

# **Pronósticos del PIB mediante modelos de factores dinámicos**

**Juan Carlos Carlo Santos\***

---

\* El contenido del presente documento es de responsabilidad del autor y no compromete la opinión del Banco Central de Bolivia.

## RESUMEN

El Instituto Nacional de Estadística de Bolivia publica los datos del Producto Interno Bruto con un retraso de entre tres y cuatro meses, reduciendo así el margen de acción de los responsables de la política económica frente a cambios imprevistos de esta variable. Ante esta necesidad, el presente documento apunta a tener estimaciones tempranas de este agregado macroeconómico mediante el uso de modelos de factores dinámicos propuestos por Stock y Watson (1988). Las series de datos incluidas en el modelo corresponden a variables relacionadas con los sectores financiero, monetario, real, y externo, incluso variables de precios. Los resultados obtenidos muestran que las estimaciones, a través de esta metodología, son más robustas en comparación con los modelos univariados y multivariados en las evaluaciones tanto dentro como fuera de la muestra.

**Clasificación JEL:** C32, C53, E23

**Palabras clave:** *Producto Interno Bruto, modelo de factores dinámicos, nowcasting*

# GDP forecasts through dynamic factor models

## ABSTRACT

The National Institute of Statistics of Bolivia publishes the Gross Domestic Product data with a lag of between three and four months, thereby reducing the margin of action of economic policy makers against unforeseen changes of this variable. Given this need, the present paper aims to have early estimates of this macroeconomic aggregate through the use of dynamic factor models proposed by Stock and Watson (1988). The data series included in the model correspond to variables related to financial, monetary, real, and external sectors, even price variables. The results obtained show that estimates through this methodology are more robust compared to univariate and multivariate models related to evaluations both inside and outside the sample.

**JEL Classification:** C32, C53, E23

**Keywords:** *Gross Domestic Product, dynamic factor model, nowcasting*

## I. Introducción

En el análisis económico, el Producto Interno Bruto (PIB) es uno de los principales agregados macroeconómicos que es constantemente monitoreado por las autoridades económicas, analistas, académicos y estudiantes, al ser considerado como uno de los principales indicadores que mide la producción de bienes y servicios de un determinado país. Sin embargo, usar el PIB para realizar un diagnóstico oportuno del *performance* de la evolución de la actividad económica, presenta un problema fundamental: el rezago con el cual se publica la información de esta variable. Por ejemplo, el Instituto Nacional de Estadística de Bolivia (INE) publica el dato del PIB trimestral con una demora de tres a cuatro meses respecto al periodo al que hace referencia, constituyéndose en una gran limitante a la capacidad de reacción por parte de las autoridades.

Bajo esta necesidad de contar con información más oportuna y con la mayor rapidez posible, es que se fue desarrollando, en los últimos años, una gran variedad de modelos, denominados en la jerga económica como *nowcasting*, que se definen como la predicción del presente, el futuro muy cercano y/o el pasado muy reciente (Bañbura et al., 2013). Entre estas metodologías, se tiene al modelo de factores dinámicos propuesto por Stock y Watson (1988).

La idea que sustenta a este tipo de modelos, se basa en el supuesto principal de que la dinámica de las variables macroeconómicas puede ser explicada por unos pocos factores inobservables, mismos que pueden ser utilizados para realizar estimaciones oportunas de la actividad económica. Es decir, consiste en hallar un factor común de toda la información disponible de distintos indicadores económicos, con el objeto de aumentar la precisión de los modelos de pronóstico de corto plazo de las principales variables macroeconómicas (como ser el PIB). Con esto, se tiene la ventaja de no descuidar información que puede ser potencialmente significativa para la realización de pronósticos y donde la dinámica de las variables pueda ser determinada por unos pocos factores inobservables.

En este sentido, bajo estas principales características mencionadas del

modelo de factores dinámicos que lo hacen uno de los más utilizados por diferentes bancos centrales (Blanco et al., 2017; Rodríguez, 2014; Cristiano et al., 2012; Aguirre y Cespedes, 2004) para la obtención de estimaciones oportunas de diferentes agregados macroeconómicos, es que en el presente documento se hará uso de esta metodología para realizar estimaciones tempranas de la actividad económica de Bolivia. Además, se evaluará las estimaciones del modelo dentro y fuera de la muestra frente a modelos tradicionales como ser: los autorregresivos integrados de medias móviles (ARIMA) y vectores autorregresivos (VAR).

Los criterios que se usaron para la selección de variables que podrían ingresar al modelo se extrajeron de Camacho y Pérez-Quirós (2008). Las series de datos incluidas en el modelo corresponden a variables relacionadas con los sectores financiero, monetario, real, y externo, incluso variables de precios.

El documento se estructura de la siguiente manera: en la sección I se presenta una breve introducción al contenido del presente documento; en la sección II se hace una breve revisión de la literatura existente sobre el uso de modelos de factores dinámicos; la sección III detalla la metodología de Stock y Watson, donde se encuentra una breve explicación del factor dinámico, la representación estado-espacio y su estimación, además de modelos alternativos para la estimación del PIB y los respectivos criterios de evaluación de proyecciones; en la sección IV se presentan los datos utilizados, los criterios para su selección, la estimación y los resultados de los modelos; y finalmente, la sección V presenta las conclusiones y recomendaciones respectivas.

## II. Revisión de la literatura

Los modelos de factores dinámicos fueron originalmente propuestos por Geweke (1977) como una extensión de series temporales de modelos de factores previamente desarrollados para datos transversales. También, un trabajo que es pertinente mencionar es el de Sargent y Sims (1977), autores que demostraron que dos factores dinámicos podrían explicar una gran porción de la varianza de las variables macroeconómicas trimestrales (producto, empleo y precios) de los Estados Unidos.

Más adelante, Stock y Watson (1988) presentaron un modelo explícito de serie de tiempo (formalmente, un análisis de factor dinámico o modelo de índice único) que define implícitamente cómo, a través de una variable que no es observable, se puede representar el estado general de una economía. Con información que data desde 1959 a 1987, la estimación del Indicador Económico Coincidente experimental (*Coincident Economic Indicator* – CEI) estuvo altamente correlacionada con la serie oficial del Departamento de Comercio de Estados Unidos, particularmente sobre los horizontes del ciclo económico. En general, este hallazgo central de que algunos factores pueden explicar una gran fracción de la varianza de muchas series macroeconómicas ha sido confirmado por muchos estudios; ver por ejemplo Giannone et al. (2004) y Watson (2004).

Posteriormente Stock y Watson (2011) en términos generales, encontraron que los pronósticos lineales de los factores funcionan muy bien en relación con muchas, pero no todas las series macroeconómicas. Por ejemplo, para las series de la actividad real de Estados Unidos, las reducciones en los errores de previsión cuadrática fuera de la muestra en el horizonte de dos a cuatro trimestres suelen estar en el rango de 20% y -40%; en cambio se observan mejoras menores o ninguna para otras series, como es el caso de la inflación después de 1990.

Camacho y Pérez-Quirós (2008), construyeron un nuevo indicador coincidente para las economías de la zona del Euro. De este documento vamos a resaltar los criterios que estos autores proponen para la selección de variables potenciales que podrían ser introducidas al modelo de factores dinámicos. Y para completar la información faltante, se recurre a modelos auxiliares, tal como lo propone en su documento Rodríguez (2014).

En el caso de Bolivia, si bien los estudios bajo este tipo de modelos son nulos, existen documentos que usaron otras metodologías para realizar proyecciones de la actividad económica. Por ejemplo, Loayza y Valdivia (2010) a través de modelos ARIMA realizaron proyecciones del Índice Global de Actividad Económica (IGAE) a once actividades, cuyos resultados estiman una tasa de crecimiento global de 3,67% para

la gestión 2010. Cáceres (2016) hace uso de modelos econométricos y redes neuronales artificiales para realizar pronósticos del IGAE, donde sus resultados muestran que ambos tipos de modelos capturan adecuadamente las fluctuaciones del IGAE. Por último, se tiene el de Aguilar y Lora (1999) quienes, a través de modelos econométricos individuales (regresiones simples) y un modelo de corrección de errores (ECM, por sus siglas en inglés) proponen que los indicadores  $M1'$ ,  $M2'$ , producción de cemento y cantidad consumida de energía eléctrica, pueden anticipar la evolución del Índice Mensual de Actividad Económica (IMAE).

Respecto a los países de la región, se pudo encontrar varios documentos que usan la metodología de factores dinámicos para las estimaciones de sus principales agregados macroeconómicos; por ejemplo, para Brasil, Branco (2018); para Argentina, Blanco et al. (2017); Ecuador, Casares (2017); Uruguay, Rodríguez (2014); Colombia, Cristiano et al. (2012); y Chile, Aguirre y Céspedes (2004).

La metodología utilizada en el presente documento es el de modelos de factores dinámicos planteada por Stock y Watson (1988) donde la dinámica de las variables se puede describir mediante dos componentes: un componente común, que es una combinación lineal de factores comunes y, por lo tanto, se correlaciona estrechamente con todas las variables; y por componentes idiosincráticos, que comprenden información específica sobre cada variable y están débilmente correlacionados con las otras variables. Además, este enfoque tiene la ventaja de no descuidar información que puede ser potencialmente significativa para la realización de pronósticos y donde la dinámica de las variables macroeconómicas está determinada por unos pocos factores inobservables.

### **III. Metodología**

El modelo de factores dinámicos planteado por Stock y Watson (1988) parte de la noción de que los coeficientes de las variables macroeconómicas tienen un elemento común que puede ser capturado por una sola variable subyacente no observada. En síntesis, esta variable representa el estado general de la economía. Esta metodología

ubica al ciclo de referencia de Burns y Mitchell (1946) en un modelo de probabilidad completamente especificado y es una versión paramétrica de los modelos de índice único de Sargent y Sims (1977).

### III.1. Modelo de factores dinámicos

Se tiene el vector de variables  $X_t$  de tamaño  $n \times 1$  que poseen la hipótesis de moverse contemporáneamente con el PIB. Además,  $X_t$  contempla el componente común no observado denominado factor dinámico  $S_t$  que ingresa a cada una de las variables de forma contemporánea, y un componente idiosincrático representado por el error de medición  $v_t$ . En ambos casos se modela de acuerdo a estructuras estocásticas lineales:

$$X_t = \beta + \gamma S_t + v_t \quad (1)$$

$$\tilde{\phi}(L)S_t = \delta + \eta_t \quad (2)$$

$$\tilde{D}(L)v_t = \epsilon_t \quad (3)$$

donde  $\tilde{\phi}(L)$  es un polinomio de rezagos escalar, y  $\tilde{D}(L)$  es un polinomio de rezagos matricial. De acuerdo con la ecuación 1, el factor dinámico  $S_t$  ingresa a cada una de las 'n' ecuaciones en 1 con rezagos y pesos variables.

Según Nelson y Plosser (1982), la gran mayoría de las series macroeconómicas se caracterizan por contener tendencias estocásticas. Una posibilidad teórica es que estas tendencias estocásticas entrarían a través de  $S_t$  a cada elemento de  $X_t$ , por lo cual cada elemento de  $X_t$  tendría en común una tendencia estocástica. Por lo tanto,  $X_t$  podría estar cointegrado. En este sentido, no podemos rechazar la hipótesis de que las series coincidentes contienen una tendencia estocástica, así como tampoco podemos rechazar la hipótesis de que no hay cointegración entre estas variables<sup>1</sup>. Para resolver este problema, las ecuaciones del 1 al 3 se reformulan en términos de variaciones. Específicamente,  $\tilde{\phi}(L) = \phi(L)\Delta$  y  $\tilde{D}(L) = D(L)\Delta$ , donde  $\Delta = 1 - L$ ,  $Y_t = \Delta X_t$  y  $u_t = \Delta v_t$ , con lo que se tiene las siguientes expresiones:

1 Para más detalle véase Stock y Watson (1988) y sobre estrategias de modelado de componentes no observados con variables cointegradas véase Fernández et al. (1987).

$$Y_t = \beta + \gamma \Delta S_t + u_t \tag{4}$$

$$\phi(L)\Delta S_t = \delta + \eta_t \tag{5}$$

$$D(L)u_t = \epsilon_t \tag{6}$$

donde los polinomios de rezago  $\phi(L)$  y  $D(L)$  tienen órdenes finitos  $p$  y  $k$ , respectivamente.

### III.2. Representación del estado – espacio

Para estimar los coeficientes de las ecuaciones 4 – 6 mediante el filtro de Kalman se las lleva a la forma de estado – espacio, donde la ecuación de estado (o de transición) describe la dinámica del vector de estado no observado, que consiste en  $\Delta S_t$ ,  $u_t$ , y sus rezagos, y la ecuación de medición relaciona las variables observadas con los elementos del vector de estado. Entonces, la ecuación de transición, que se obtiene al combinar las ecuaciones 5 y 6, es la siguiente:

$$\begin{bmatrix} S_t^* \\ u_t^* \\ S_{t-1} \end{bmatrix} = \begin{bmatrix} \Phi^* & 0 & 0 \\ 0 & D^* & 0 \\ Z_c & 0 & 1 \end{bmatrix} \begin{bmatrix} S_{t-1}^* \\ u_{t-1}^* \\ S_{t-2}^* \end{bmatrix} + \begin{bmatrix} Z_c & 0 \\ 0 & Z_u \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_t \\ \epsilon_t \end{bmatrix} \tag{7}$$

donde:

$$\begin{aligned} S_t^* &= [\Delta S_t \quad \Delta S_{t-1} \quad \dots \quad \Delta S_{t-p+1}]' \\ u_t^* &= [u_t' \quad u_{t-1}' \quad \dots \quad u_{t-k+1}']' \\ \Phi^* &= \begin{bmatrix} \phi_1 & \dots & \phi_{p-1} & \phi_p \\ & & I_{p-1} & 0 \end{bmatrix} \\ D^* &= \begin{bmatrix} D_1 & \dots & D_{k-1} & D_k \\ & & I_{n(k-1)} & 0 \end{bmatrix} \\ Z_c &= [1 \quad 0_{1 \times (p-1)}] \\ Z_u &= [I_n \quad 0_{n \times n(k-1)}] \end{aligned}$$

donde  $I_n$  denota una matriz identidad de orden  $n \times n$ ,  $0_{n \times k}$  denota una matriz de ceros de orden  $n \times k$ , y  $D_i = \text{diag}(d_{1i}, \dots, d_{ni})$  donde  $d_j(L) = 1 - \sum_{i=1}^k d_{ji} L^i$ .

En cuanto a la ecuación de medida, esta se obtiene escribiendo la ecuación 4 como una combinación lineal del vector de estado:

$$Y_t = \beta + [yZ_c \quad Z_u \quad 0] \begin{bmatrix} S_t^* \\ u_t^* \\ S_{t-1} \end{bmatrix} \quad (8)$$

El sistema 7 y 8 se puede reescribir de forma más compacta en la forma estándar:

$$\alpha_t = T_t \alpha_{t-1} + R \zeta_t \quad (9)$$

$$Y_t = \beta + Z \alpha_t + \xi_t \quad (10)$$

donde:

$$\alpha_t = [S_t^{*'} \quad u_t^{*'} \quad S_{t-1}]'$$

$$\zeta_t = [\eta_t \quad \epsilon_t']'$$

y donde las matrices  $T_t$ ,  $R$  y  $Z$  denotan la matriz de transición en la ecuación 7, y la matriz de medición en la ecuación 8. La matriz de covarianza de  $\zeta_t$  es  $E[\zeta_t \zeta_t'] = \Sigma$ . En general, se ha agregado un término de error de medición  $\xi_t$  (no correlacionado con  $\zeta_t$ ) a la ecuación de medición 10 y se permite que la matriz de transición  $T_t$  varíe con el tiempo.

### III.3. Filtro de Kalman

El filtro de Kalman es un algoritmo recursivo muy utilizado para estimar el mínimo error cuadrático medio de las series no observables. Este filtro contempla ecuaciones de predicción y de actualización. Para las ecuaciones de predicción, denotamos  $\alpha_{t|t-1}$  como la estimación de  $\alpha_t$  basado en información contenida en  $(y_1, \dots, y_{t-1})$  y  $E[\xi_t \xi_t'] = H$ , y  $E[\zeta_t \zeta_t'] = \Sigma$  son las matrices de covarianzas de  $\xi_t$  y  $\zeta_t$ , respectivamente. Además  $P_{t|t-1} = E[(\alpha_{t|t-1} - \alpha_t)(\alpha_{t|t-1} - \alpha_t)']$ . Con esta notación, las ecuaciones de predicción son:

$$\alpha_{t|t-1} = T_t \alpha_{t-1|t-1} \quad (11)$$

$$P_{t|t-1} = T_t P_{t-1|t-1} T_t' + R \Sigma R' \tag{12}$$

El pronóstico de  $Y_t$  en el tiempo  $t - 1$  es  $Y_{t|t-1} = \beta + Z \alpha_{t|t-1}$  y el error de pronóstico es  $v_t = Y_t - \beta - Z \alpha_{t|t-1}$ . Respecto a las ecuaciones de actualización del filtro se tiene:

$$\alpha_{t|t} = \alpha_{t|t-1} + P_{t|t-1} Z' F_t^{-1} v_t \tag{13}$$

$$P_{t|t} = P_{t|t-1} + P_{t|t-1} Z' F_t^{-1} Z P_{t|t-1} \tag{14}$$

donde  $F_t = E[v_t v_t'] = Z P_{t|t-1} Z' + H$ . Las ecuaciones (11) – (14) permiten el cálculo recursivo del vector de estado pronosticado  $\alpha_{t|t-1}$ , y la matriz de covarianza de esta estimación  $P_{t|t-1}$ , dados los parámetros asumidos en  $T_t$ ,  $R$ ,  $\Sigma$ ,  $H$  y  $Z$  y valores iniciales para  $\alpha_{t|t}$  y  $P_{t|t}$ . Para una estimación exacta, los valores iniciales se toman como la expectativa incondicional de  $\alpha_t$  y su matriz de covarianza,  $E[(\alpha_t - E\alpha_t)(\alpha_t - E\alpha_t)']$ ; es decir,  $\alpha_{0|0} = 0$  y  $P_{0|0} = \sum_{j=0}^{\infty} T_{t-j}^j \Sigma T_{t-j}^{j'}$ . Finalmente, este marco también permite el cálculo de estimaciones retrospectivas del estado de la economía  $C_{t|T}$ , y más generalmente  $\alpha_{t|T}$ . Para las estimaciones de  $\alpha_t$  se estimarán mediante el suavizador de Kalman y la estimación de  $\alpha_{t|t}$  es lineal tanto en las observaciones actuales como en las rezagadas de  $Y_t$ . Al sustituir la relación  $v_t = Y_t(\beta + Z \alpha_{t|t-1})$  en la ecuación 13 y luego usando la ecuación 11 se obtiene:

$$\alpha_{t|t} = (I - G_t Z) T_t \alpha_{t-1|t-1} + G_t Y_t - G_t \beta \tag{15}$$

donde  $G_t = P_{t|t-1} Z' F_t^{-1}$  es la ganancia de Kalman.

### III.4. Modelos alternativos para la proyección del PIB

#### III.4.1. Modelos ARIMA

Uno de los modelos comúnmente usados para la realización de las proyecciones de corto plazo, ya sea por su sencillez o bajo costo, son los modelos autorregresivos integrados de media móvil (ARIMA). Este tipo de modelos fueron popularizados por George Box y Gwilym Jenkins a principios de la década de los 1970. Como este tipo de modelos es bien conocido y/o utilizado, no se abundará mucho en su descripción.

El modelo general no estacional se conoce como ARIMA ( $p, d, q$ ) el cual es una combinación de tres procesos: i) procesos autorregresivos ( $p$  es el orden autoregresivo), ii) proceso de diferenciación ( $d$  es el grado de diferenciación) y iii) proceso de medias móviles ( $q$  es el orden de media móvil). Estos procesos se conocen en la literatura estadística como los principales modelos de series de tiempo univariados y se usan comúnmente en muchas aplicaciones.

El enfoque de Box y Jenkins (1976) para encontrar el mejor modelo para la realización de proyecciones sigue cuatro etapas:

- i) Identificación. En esta etapa se tiene que verificar que la serie sea estacionaria, tanto en la frecuencia regular como en la frecuencia estacional. Después, con la ayuda de las funciones de autocorrelación simple (FAS) y funciones de autocorrelación parcial (FAP) de las series en cuestión se tiene que identificar los componentes autorregresivos y/o de medias móviles del modelo.
- ii) Estimación. Una vez identificado el modelo, se procede a la estimación del mismo usando algoritmos computacionales. El método común más usado es la estimación de máxima verosimilitud.
- iii) Diagnóstico. El modelo se considera adecuado si los residuos de la estimación son simplemente ruido blanco.
- iv) Pronóstico. Una vez evaluado el modelo con las respectivas pruebas de diagnóstico, se procede a realizar las proyecciones.

En el caso de modelos de frecuencia menor a un año, se tiene que agregar ( $P, D, Q$ ), donde  $P$  es el orden del proceso autorregresivo estacional,  $D$  es el grado de diferencia estacional y  $Q$  es el orden del proceso de media móvil estacional (Novales, 1993). Entonces, la especificación quedaría como ARIMA ( $p, d, q$ )( $P, D, Q$ ).

### III.4.2. Modelos VAR

Un modelo de vectores autoregresivos (VAR) de orden  $n$ , con variables endógenas se especifica como:

$$Y_t = A_t + \sum_{s=1}^n A_s Y_{t-s} + G W_t + u_t \quad (16)$$

donde  $Y_t$  es un vector columna  $k \times 1$ ;  $n$  es el orden del modelo VAR, y  $u_t$  es un vector  $k \times 1$  de innovaciones, es decir, procesos sin autocorrelación y con  $Var(u_t) = \Sigma$ , constante.  $W_t$  es un vector de variables exógenas y  $G$  es la matriz de parámetros de sus parámetros.

Para el presente caso, no se consideran variables exógenas dentro del modelo, por lo que  $G = 0_k$ . El elemento  $(i, j)$  en la matriz  $A_s$ ,  $1 \leq s \leq n$  mide el efecto directo o parcial de un cambio unitario en  $Y_j$  en el instante  $t$  sobre los valores numéricos de  $Y_i$  al cabo de  $s$  periodos, es decir, sobre el vector  $Y_{i,t+s}$ . La columna  $j$  de la matriz  $A_s$  mide el efecto que un cambio unitario en  $Y_j$  en el instante  $t$  tiene sobre el vector  $Y_{t+s}$ . El elemento  $i$ -ésimo en  $u_t$  es el componente de  $Y_{it}$  que no puede ser previsto utilizando el pasado de las variables que integran el vector  $Y_t$ .

Según Novales (1993), sobre las estimaciones de un modelo VAR se tienen que tomar las siguientes consideraciones:

- i) Todas las variables del modelo VAR son tratadas simétricamente, siendo explicadas por el pasado de todas ellas; el modelo tienen tantas ecuaciones como variables.
- ii) Una vez estimado el modelo, se puede proceder a excluir algunas variables explicativas, en función de su significancia estadística, pero hay razones para no hacerlo. Por un lado, las estimaciones son eficientes si se mantiene el mismo conjunto de variables explicativas; por el otro, la presencia de bloques de retardos como variables explicativas conlleva a problemas de multicolinealidad.

- iii) En el modelo VAR pueden estimarse con bastante precisión los elementos globales del modelo, como el coeficiente de determinación, la desviación típica residual, y los mismos residuos, o el efecto global de una variable sobre otra, lo que se resume en los contrastes de causalidad. Sin embargo, no cabe hacer interpretaciones de coeficientes individuales en distintos retardos, ni llevar a cabo contrastes de hipótesis sobre coeficientes individuales.

### III.5. Criterios para la evaluación de proyecciones

Para la evaluación de las proyecciones, tanto dentro como fuera de la muestra, se hará uso de los estadísticos tradicionales (Pindyck y Rubinfeld, 2001). Supongamos que el pronóstico de la muestra es  $j = T + 1, T + 2, T + 3, \dots, T + h$ , y los valores observado y proyectado están denotados por  $y_t$  y  $\hat{y}_t$ , respectivamente. Las estadísticas de error de pronóstico se calculan de la siguiente manera:

**Cuadro 1: ESTADÍSTICOS DE EVALUACIÓN DE PROYECCIONES**

Raíz del error cuadrático medio (RMSE, por sus siglas en inglés)	$\sqrt{\sum_{t=T+1}^{T+h} \frac{(\hat{y}_t - y_t)^2}{h}}$
Error absoluto medio (MAE, por sus siglas en inglés)	$\sum_{t=T+1}^{T+h} \frac{ \hat{y}_t - y_t }{h}$
Error porcentual absoluto medio (MAPE, por sus siglas en inglés)	$\left( \sum_{t=T+1}^{T+h} \frac{ \hat{y}_t - y_t }{y_t} \right) \times 100$
Coefficiente de desigualdad de Theil	$\frac{\sqrt{\sum_{t=T+1}^{T+h} \frac{(\hat{y}_t - y_t)^2}{h}}}{\sqrt{\sum_{t=T+1}^{T+h} \frac{\hat{y}_t^2}{h} + \sum_{t=T+1}^{T+h} \frac{y_t^2}{h}}}$

Los dos primeros estadísticos dependen de la escala de la variable dependiente. Por lo que deben usarse como medidas relativas para comparar pronósticos para la misma serie en diferentes modelos; cuanto menor es el error, mejor es la capacidad de pronóstico de ese

modelo de acuerdo con ese criterio. Las dos estadísticas restantes son invariantes de escala. El coeficiente de desigualdad de Theil siempre se encuentra entre cero y uno, donde cero indica un ajuste perfecto.

## **IV. Selección de variables y resultados**

### ***IV.1. Criterios para la selección de variables***

Los criterios para la selección de variables que podrían ingresar al modelo, se extrajeron de Camacho y Pérez-Quirós (2008). Estos criterios se pueden resumir en lo siguiente:

1. Que la serie en cuestión tenga una alta correlación con los datos del PIB.
2. Oportunidad en la publicación de los datos, privilegiando aquellas que tuvieran un menor rezago.
3. Que refleje la lógica de las cuentas nacionales.
4. Que su inclusión genere un incremento de la varianza del PIB explicada por el factor común estimado, además de ser estadísticamente significativo.

En el caso del punto 2, si la estimación se realiza cuando se tiene información disponible para el primero y/o segundo mes del trimestre en cuestión, es necesario completar la información de los meses faltantes mediante modelos auxiliares como ser un modelo autorregresivo integrado de medias móviles (Rodríguez, 2014).

### ***IV.2. Selección de variables***

Para realizar la selección de variables se tomó en cuenta más de 50 series con frecuencia mensual, que van desde 2000 a marzo de 2018<sup>2</sup>. Los datos se clasificaron en variables correspondientes a la actividad económica, precios, financieras, monetarias y externas.

---

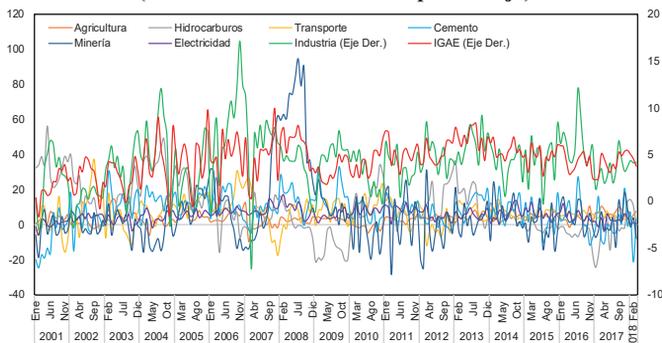
2 Para una información más detallada de las variables, sus fuentes y sus rezagos en el acceso a la información, ver Apéndice A.

## Variables de la actividad económica

Dada la necesidad de contar con una serie mensual del PIB para su inclusión en el modelo, es que se utilizó el IGAE en su lugar. En este sentido, los criterios de selección de variables potenciales para explicar la evolución del PIB se basaron en este índice.

Como se observa en el Gráfico 1, entre los indicadores de la actividad económica se tiene al índice de producción de agricultura, índice producción de cemento, índice de producción de industria, índice de producción de hidrocarburos, índice de generación de electricidad, índice de transporte y el índice de producción de minería.

**Gráfico 1: VARIABLES DE LA ACTIVIDAD ECONÓMICA,  
ENERO 2001 – MARZO 2018  
(Crecimiento interanual en porcentaje)**



Fuente: Elaboración propia con datos del INE, Yacimientos Petrolíferos Fiscales Bolivianos (YPFB) y Comité Nacional de Despacho y Carga (CNDC)

Según el Cuadro 2, entre las variables con un alto grado de correlación con el IGAE se encuentran electricidad (0,65), industria (0,60), cemento (0,55) y minería (0,42). Todas las variables son componentes importantes del PIB por el lado de la oferta.

**Cuadro 2: MATRIZ DE CORRELACIONES ENTRE EL IGAE Y VARIABLES DE LA ACTIVIDAD ECONÓMICA**

IGAE	Agricultura	Gas	Petróleo	Transporte	Cemento	Electricidad	Minería	Industria
IGAE	1							
Agricultura	0,06	1						
Gas	-0,25	-0,31	1					
Petróleo	0,19	-0,13	0,70	1				
Transporte	0,14	-0,31	-0,06	-0,08	1			
Cemento	0,55	0,06	-0,30	-0,18	0,13	1		
Electricidad	0,65	-0,29	-0,27	-0,17	0,18	0,59	1	
Minería	0,42	0,03	-0,13	-0,22	-0,22	0,53	0,52	1
Industria	0,60	0,03	-0,36	0,06	0,19	0,36	0,46	0,04

Fuente: Elaboración propia

Nota: La matriz de correlaciones se la hizo con las tasas de variación acumulada (ene 2001 – mar 2018)

Respecto al incremento de la varianza, los índices de industria (0,31pp), electricidad (0,23pp) y agricultura (0,15pp) son las variables que explican una gran parte de la varianza del IGAE, además de ser estadísticamente significativos (Cuadro 3).

**Cuadro 3: INCREMENTO DE LA VARIANZA DEL IGAE CON LAS VARIABLES DE LA ACTIVIDAD ECONÓMICA**

(En puntos porcentuales)

IGAE	Incremento	Error estándar	t-estadístico	probabilidad
Agricultura	0,15	0,0312	4,8720	0,0000 ***
Hidrocarburos	0,02	0,0055	4,2782	0,0000 ***
Transporte	0,04	0,0161	2,2013	0,0289 **
Electricidad	0,23	0,0370	6,0972	0,0000 ***
Minería	0,02	0,0056	3,1984	0,0016 ***
Industria	0,31	0,0416	7,5579	0,0000 ***
Cemento	0,02	0,0119	1,2742	0,2041

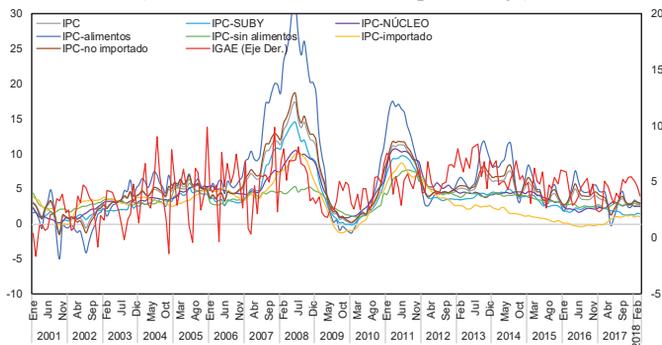
\*\*\*p<0,01, \*\*p<0,05, \*p<0,1

Fuente: Elaboración propia en base a salidas del Eviews

**Variables de precios**

En cuanto a indicadores de precios, se tiene al índice de precios al consumidor (IPC), IPC subyacente, IPC núcleo, IPC alimentos, IPC sin alimentos, IPC importado e IPC no importado a nivel nacional (Gráfico 2).

**Gráfico 2: VARIABLES DE PRECIOS, ENERO 2001 – MARZO, 2008**  
(Crecimiento interanual en porcentaje)



Fuente: Elaboración propia con datos del INE

Las variables que registraron una alta correlación con el IGAE son IPC núcleo (0,59), IPC no importado (0,59), IPC (0,58), principalmente (Cuadro 4).

**Cuadro 4: MATRIZ DE CORRELACIONES ENTRE EL IGAE Y LAS VARIABLES DE PRECIOS**

	IGAE	IPC	IPC-Núcleo	IPC-Subyacente	IPC-Alimentos	IPC-Sin alimentos	IPC-Importado	IPC-No importado
IGAE								
IPC	0,58	1						
IPC-Núcleo	0,59	0,91	1					
IPC-Subyacente	0,57	0,99	0,92	1				
IPC-Alimentos	0,57	0,99	0,87	0,97	1			
IPC-Sin alimentos	0,43	0,72	0,80	0,71	0,60	1		
IPC-Importado	0,32	0,79	0,85	0,83	0,75	0,71	1	
IPC-No importado	0,59	1,00	0,90	0,98	0,99	0,70	0,75	1

Fuente: Elaboración propia

Nota: La matriz de correlaciones se hizo con las tasas de variación acumulada (ene 2001 – mar 2018)

En cuanto al incremento de la varianza, se tiene al IPC (10,92pp), IPC sin alimentos (-3,09pp) e IPC alimentos (-2,60pp). Sin embargo, el IPC no importado (-4,33pp) y el IPC subyacente (-0,34pp) no son estadísticamente significativos al 5%, por lo cual no serían posibles candidatos a ser considerados en el modelo (Cuadro 5).

**Cuadro 5: INCREMENTO DE LA VARIANZA DEL IGAE CON LAS VARIABLES DE PRECIOS**  
(En puntos porcentuales)

IGAE	Incremento	Error estándar	t-estadístico	probabilidad
IPC	10,92	3,5680	3,0592	0,0025 ***
IPC-Núcleo	0,90	0,1402	6,4234	0,0000 ***
IPC-Subyacente	-0,34	0,2402	-1,4118	0,1596
IPC-Alimentos	-2,60	0,6054	-4,2870	0,0000 ***
IPC-Sin alimentos	-3,09	0,7202	-4,2943	0,0000 ***
IPC-Importado	-1,08	0,3620	-2,9814	0,0032 ***
IPC-No importado	-4,33	2,5874	-1,6725	0,0960 *

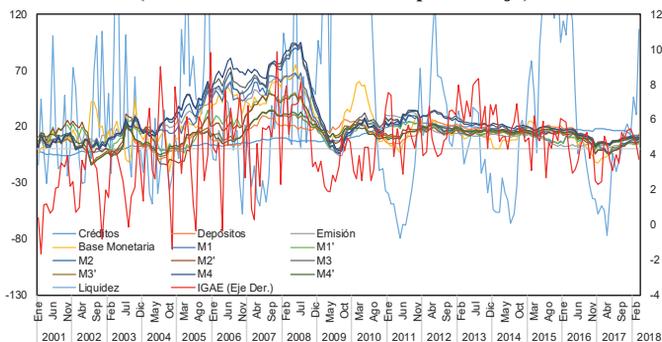
\*\*\*p<0,01, \*\*p<0,05, \*p<0,1

Fuente: Elaboración propia en base a salidas del Eviews

**VARIABLES del sector financiero y monetario**

En cuanto a los indicadores monetarios y financieros, se tiene la cartera de créditos del sistema financiero, los depósitos del sistema financiero, emisión monetaria, base monetaria, liquidez del sistema financiero y agregados monetarios (M1, M1', M2, M2', M3, M3', M4, M4') (Gráfico 3).

**Gráfico 3: VARIABLES FINANCIERAS Y MONETARIAS, ENERO 2001 – MARZO 2018**  
(Crecimiento interanual en porcentaje)



Fuente: Elaboración propia con datos de la Autoridad de Supervisión del Sistema Financiero (ASFI) y Banco Central de Bolivia (BCB)

Entre las variables financieras y monetarias que registraron una correlación alta con el IGAE se encuentran Créditos (0,61), M3' (0,47), M3 (0,46), principalmente (Cuadro 6).

### Cuadro 6: MATRIZ DE CORRELACIONES ENTRE EL IGAE, VARIABLES FINANCIERAS Y MONETARIAS

	IGAE	Créditos	Depósitos	Emisión M.	Base M.	Liquidez	M1	M1'	M2	M2'	M3	M3'	M4	M4'
IGAE	1													
Créditos	0,61	1												
Depósitos	0,39	0,57	1											
Emisión M.	0,36	-0,14	0,16	1										
Base M.	0,24	-0,13	0,45	0,79	1									
Liquidez	-0,15	-0,03	0,39	-0,09	0,45	1								
M1	0,39	-0,06	0,32	0,98	0,83	-0,05	1							
M1'	0,31	0,00	0,45	0,82	0,79	0,03	0,88	1						
M2	0,41	-0,10	0,33	0,96	0,83	-0,03	0,99	0,84	1					
M2'	0,24	0,03	0,59	0,66	0,75	0,13	0,76	0,95	0,72	1				
M3	0,46	0,00	0,39	0,95	0,83	-0,01	0,97	0,81	0,99	0,69	1			
M3'	0,47	0,46	0,92	0,51	0,67	0,28	0,64	0,72	0,64	0,78	0,68	1		
M4	0,43	-0,04	0,36	0,95	0,81	-0,06	0,97	0,80	0,99	0,69	0,99	0,66	1	
M4'	0,43	0,38	0,90	0,48	0,62	0,21	0,62	0,72	0,62	0,81	0,65	0,98	0,64	1

Fuente: Elaboración propia

Nota: La matriz de correlaciones se hizo con las tasas de variación acumulada (ene 2001 – mar 2018)

En cuanto a las variables que explican en mayor medida la varianza del IGAE se encuentran M1 (-0,49pp), M2 (0,43pp), depósitos (0,37pp), créditos (0,19pp) y emisión monetaria (0,17pp). Todas son estadísticamente significativas, con excepción de la base monetaria, la liquidez y el agregado monetario M4 (Cuadro 7).

### Cuadro 7: INCREMENTO DE LA VARIANZA DEL IGAE CON LAS VARIABLES FINANCIERAS Y MONETARIAS

(En puntos porcentuales)

IGAE	Incremento	Error estándar	t-estadístico	probabilidad
Créditos	0,19	0,0148	12,6998	0,0000 ***
Depósitos	0,37	0,0751	4,8757	0,0000 ***
Emisión M.	0,17	0,0401	4,1862	0,0000 ***
Base M.	-0,01	0,0136	-0,8833	0,3782
Liquidez	0,00	0,0016	-0,6118	0,5414
M1	-0,49	0,0475	-10,3444	0,0000 ***
M1'	0,27	0,0519	5,1782	0,0000 ***
M2	0,43	0,0368	11,7048	0,0000 ***
M2'	-0,13	0,0411	-3,2741	0,0013 ***
M3	-0,13	0,0536	-2,4973	0,0134 **
M3'	-0,31	0,1166	-2,6901	0,0078 ***
M4	0,01	0,0419	0,1453	0,8846
M4'	-0,11	0,0552	-2,0406	0,0427 **

\*\*\*p<0,01, \*\*p<0,05, \*p<0,1

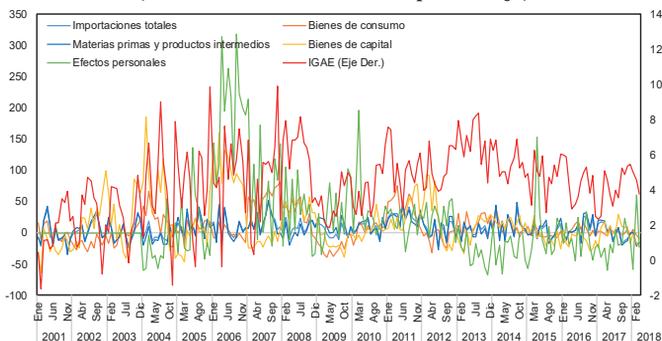
Fuente: Elaboración propia en base a salidas del Eviews

### Variables externas

Dada la gran variedad de variables externas, éstas se las clasificó en tres grupos: variables de importación, variables de exportación e indicadores externos.

Las variables correspondientes a importaciones son: importación de bienes de consumo, importación de materias primas y productos intermedios, importación de bienes de capital, importación de efectos personales e importaciones totales (Gráfico 4).

**Gráfico 4: VARIABLES DE IMPORTACIÓN, ENERO 2001 – MARZO 2018**  
(Crecimiento interanual en porcentaje)



Fuente: Elaboración propia con datos del INE

Respecto a las correlaciones, las variables que fueron más representativas son las importaciones de bienes de consumo (0,44) y las importaciones totales (0,29) (Cuadro 8).

**Cuadro 8: MATRIZ DE CORRELACIONES ENTRE EL IGAE Y VARIABLES DE IMPORTACIÓN**

	IGAE	Imp. Total	Bienes Consumo	Bienes Capital	Materias Primas	Diversos
IGAE	1					
Imp. Total	0,29	1				
Bienes Consumo	0,44	0,47	1			
Bienes Capital	0,23	0,11	0,13	1		
Materias Primas	0,07	0,87	0,09	-0,18	1	0,00
Diversos	0,01	0,05	-0,07	0,14	0,00	1

Fuente: Elaboración propia

Nota: La matriz de correlaciones se lo hizo con las tasas de variación acumulado (Ene 2001 – Mar 2018)

En cuanto a la aportación a la varianza del IGAE, se destacan solamente la correspondiente a importaciones totales (0,13), estadísticamente significativa al 5% pero no al 1%. Los demás no son estadísticamente significativos con un nivel de confianza del 95% (Cuadro 9).

### Cuadro 9: INCREMENTO DE LA VARIANZA DEL IGAE CON LAS VARIABLES DE IMPORTACIONES

(En puntos porcentuales)

IGAE	Incremento	Error estándar	t-estadístico	probabilidad
Imp. Total	0,13	0,0622	2,1675	0,0314 **
Bienes Consumo	0,01	0,0094	1,2954	0,1967
Bienes Capital	0,00	0,0040	0,3745	0,7085
Materias Primas(*)	-0,09	0,0487	-1,8881	0,0605 *
Diversos	0,00	0,0001	-0,2169	0,8285

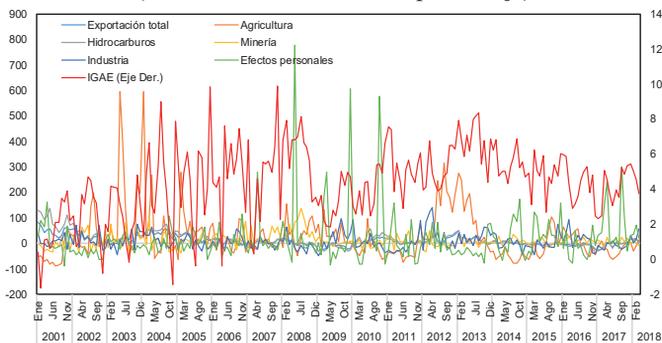
\*\*\*p<0,01, \*\*p<0,05,\*p<0,1

Fuente: Elaboración propia en base a salidas del Eviews

(\*) Más productos intermedios

Las variables correspondientes a exportaciones son: exportación de productos de agricultura, exportación de hidrocarburos, exportación de productos de minería, exportación de productos de industria, exportación de productos de efectos personales y exportaciones totales (Gráfico 5).

### Gráfico 5: VARIABLES DE EXPORTACIÓN, ENERO 2001 – MARZO 2018 (Crecimiento interanual en porcentaje)



Fuente: Elaboración propia con datos del INE

Respecto a las correlaciones de estas series con el IGAE, las variables que fueron más representativas fueron la exportación de minerales (0,42), seguida de la exportación de hidrocarburos (-0,36) (Cuadro 10). Una relación que se rescata también es la fuerte correlación existente entre exportaciones totales y la exportación de hidrocarburos (0,96) seguida posteriormente de industria (0,40).

**Cuadro 10: MATRIZ DE CORRELACIONES ENTRE EL IGAE Y VARIABLES DE EXPORTACIÓN**

	IGAE	Exp. Total	Agricultura	Hidrocarburos	Minería	Industria	Efectos Pers.
IGAE	1						
Exp. Total	-0,23	1					
Agricultura	0,13	0,13	1				
Hidrocarburos	-0,36	0,96	-0,05	1			
Minería	0,42	-0,16	0,15	-0,25	1		
Industria	-0,03	0,40	0,36	0,26	-0,14	1	
Efectos Pers.	-0,05	-0,10	-0,18	-0,01	0,17	-0,18	1

Fuente: Elaboración propia

Nota: La matriz de correlaciones se hizo con las tasas de variación acumulada (ene 2001 – mar 2018)

En cuanto a la aportación a la varianza del IGAE, se destacan las exportaciones totales (0,18pp), seguida de hidrocarburos (-0,13pp) ambos estadísticamente significativos (Cuadro 11).

**Cuadro 11: INCREMENTO DE LA VARIANZA DEL IGAE CON LAS VARIABLES DE EXPORTACIÓN (En puntos porcentuales)**

IGAE	Incremento	Error estándar	t-estadístico	probabilidad
Exp. Total	0,18	0,0261	6,7559	0,0000 ***
Agricultura	0,00	0,0014	-2,9456	0,0036 ***
Hidrocarburos	-0,13	0,0173	-7,5686	0,0000 ***
Minería	0,02	0,0063	2,7037	0,0074 ***
Industria	-0,01	0,0062	-2,3742	0,0185 **
Efectos Pers.	0,00	0,0033	0,7621	0,4469

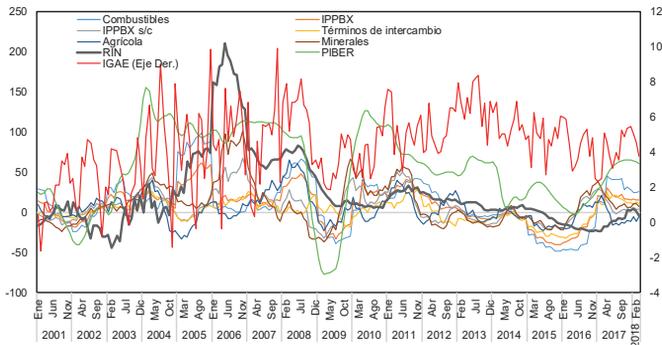
\*\*\*p<0,01, \*\*p<0,05, \*p<0,1

Fuente: Elaboración propia en base a salidas del Eviews

Finalmente, las variables correspondientes a indicadores externos son: el índice de precios de combustibles, índice de precios agrícolas, índice de precios de minerales, el índice de precios de productos básicos de exportación (IPPBX), IPPBX sin combustibles, los términos de intercambio (TI), las reservas internacionales (RIN) y el Producto Interno Bruto externo relevante (PIBER<sup>3</sup>) (Gráfico 6).

3 Dada la necesidad de contar con información mensual del PIBER, es que se cambió de frecuencia a una mensual mediante la metodología propuesta por Denton (1971).

**Gráfico 6: INDICADORES EXTERNOS, ENERO 2001 – MARZO 2018**  
(Crecimiento interanual en porcentaje)



Fuente: Elaboración propia con datos del INE y BCB

Con las correlaciones, las variables que se destacan son el PIBER (0,36), índice de precios agrícolas (0,27) y las RIN (0,24) (Cuadro 12).

**Cuadro 12: MATRIZ DE CORRELACIONES ENTRE EL IGAE E INDICADORES EXTERNOS**

	IGAE	Minerales	Agrícola	Combustible	IPPBX	IPPBX S/C	RIN	TI	PIBER
IGAE	1								
Minerales	0,06	1							
Agrícola	0,27	0,37	1						
Combustible	0,06	0,27	0,25	1					
IPPBX	0,10	0,55	0,47	0,94	1				
IPPBX S/C	0,10	0,95	0,64	0,26	0,57	1			
RIN	0,24	0,48	0,22	0,21	0,32	0,44	1		
TI	-0,13	0,59	0,51	0,45	0,65	0,67	0,28	1	
PIBER	0,36	0,72	0,44	0,53	0,70	0,71	0,45	0,45	1

Fuente: Elaboración propia

Nota: La matriz de correlaciones se lo hizo con las tasas de variación acumulado (Ene 2001 – Mar 2018)

En cuanto a la aportación a la varianza del IGAE, se destacan el IPPBX (0,49pp), PIBER (0,30pp), índice de precios de minerales (0,16pp) e índice de precios agrícolas (0,10pp) (Cuadro 13).

**Cuadro 13: INCREMENTO DE LA VARIANZA DEL IGAE CON LOS INDICADORES EXTERNOS**  
(En puntos porcentuales)

IGAE	Incremento	Error estándar	t-estadístico	probabilidad
Minerales	0,16	0,0512	3,1731	0,0017 ***
Agrícola	0,10	0,0285	3,3962	0,0008 ***
Combustible	-0,30	0,0528	-5,6419	0,0000 ***
IPPBX	0,49	0,0906	5,3527	0,0000 ***
IPPBX S/C*	-0,44	0,0862	-5,0702	0,0000 ***
RIN*	0,01	0,0024	2,2741	0,0240 **
TI*	-0,07	0,0146	-4,8732	0,0000 ***
PIBER*	0,30	0,0669	4,4648	0,0000 ***

\*\*\*p<0,01, \*\*p<0,05, \*p<0,1

Fuente: Elaboración propia en base a salidas del Eviews

En base a los resultados obtenidos mediante los criterios de selección, las variables candidatas a ingresar al modelo son las siguientes:

- Variables de la actividad económica: industria, electricidad, agricultura.
- Variables de precios: IPC, IPC sin alimentos e IPC alimentos.
- Variables financieras y monetarias: M1, M2, depósitos, créditos y emisión monetaria.
- Variables externas: Importaciones totales, exportaciones totales, exportación de hidrocarburos, IPPBX, PIBER, índice de precios de minerales e índice de precios agrícolas.

### IV.3. Resultados

#### IV.3.1. Modelos de factores dinámicos

Antes de la estimación del modelo se siguieron los siguientes pasos:

1. El primer paso fue realizar las pruebas de raíz unitaria para cada uno de los indicadores coincidentes y con ello verificar el orden de cointegración. Como las series en cuestión son de frecuencia estacional, entonces se procedió a realizar la prueba de HEGY<sup>4</sup>, cuyos resultados muestran que las series son integradas de orden 1.

4 Los resultados de la prueba de Hylleberg, Engle, Granger, and Yoo (HEGY) se encuentran en el Apéndice B.

2. El segundo paso fue realizar pruebas de cointegración. Los resultados muestran que, con un nivel de significancia del 5%, existe evidencia de cointegración, entonces, las variables que contemplan el modelo de factores dinámicos, tienen una relación común de largo plazo<sup>5</sup>.
3. Tercer paso, como las series son integradas de orden uno y tienen una relación común de largo plazo, entonces, siguiendo a Stock y Watson (1988), se expresaron las series como crecimiento interanual en porcentaje; o, siendo más preciso, como la diferencia interanual de los logaritmos de las series temporales originales con lo cual se consiguió eliminar posibles problemas de estacionalidad que pudieran generar las variables.
4. Cuarto paso, para evitar problemas de identificación se restó la media de la muestra a todas las observaciones de las series (Stock y Watson, 1988).
5. Finalmente, se estimó  $\alpha_t$  utilizando el suavizador de Kalman (Harvey, 1981).

La ecuación a estimar mediante el uso del filtro de Kalman (ecuaciones 10 y 11) es la ecuación 4 vista en el apartado metodológico<sup>6</sup>:

$$Y_t = \beta + \gamma \Delta S_t + u_t$$

Con las variables seleccionadas se realizó diferentes combinaciones de modelos y se contrastaron sus resultados con los datos observados del PIB, con el objeto de seleccionar aquel modelo que se ajuste lo más próximo posible al dato observado.

Según lo mencionado, las variables que ingresaron finalmente al modelo son el IGAE (variable *proxy* al PIB), índice de producción de

---

5 Los resultados de la prueba de Johansen se encuentran en el Apéndice C.

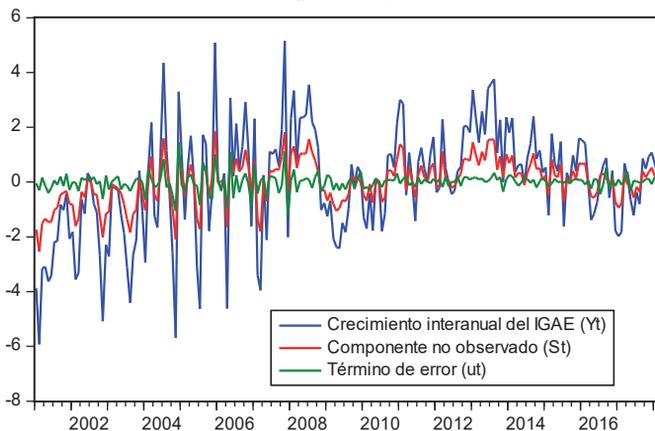
6 En el Apéndice D podrá ver la forma más detallada del modelo estado-espacio de Kalman.

agricultura, producción de gas<sup>7</sup>, índice de producción de industria, índice de producción de electricidad, índice de producción de minería, IPC, IPPBX y créditos del sistema financiero. En un principio se realizó el *nowcast* con datos a marzo de 2018 y posteriormente con datos a junio de 2018. Los resultados de la estimación econométrica se encuentran en el Apéndice E del presente documento.

**Estimación del PIB con datos a marzo de 2018**

Los resultados de las variables que conforman la ecuación 4, de crecimiento interanual del IGAE ( $Y_t$ )<sup>8</sup>, componente no observado o factor común ( $S_t$ ) y el término de error ( $u_t$ ) se observan en el Gráfico 7.

**Gráfico 7: VARIABLES DE LA ECUACIÓN DE MEDICIÓN,  
ENERO 2001 – MARZO 2018  
(En porcentaje)**



Fuente: Elaboración propia en base a salidas del Eviews

- 7 En un principio, una de las variables potenciales de la actividad económica fue el índice de producción de hidrocarburos, pero después de varias estimaciones en base a diferentes modelos, la variable que se ajustó más a los datos observados fue la producción de gas natural, dato que no se aleja de la realidad, dado que, aproximadamente más del 70% de la producción de hidrocarburos lo representa este energético.
- 8 Como se mencionó en la parte metodológica, el crecimiento interanual es igual a la diferencia interanual del logaritmo del IGAE.

La estimación de la ecuación 4 es la siguiente:

$$g\_IGAE_t = \beta + 2,212 \Delta S_t + u_t$$

En base a este resultado, la tasa de crecimiento interanual del PIB al primer trimestre de 2018 es de:

$$g\_IGAE_{I-18} = 4,7930 + 2,212(-0,0299) + 0,0646 = 4,7915$$

Finalmente, recuperando las series originales a través de esta tasa de crecimiento y sacando el antilogaritmo (para una estimación más exacta), se tiene que el valor del PIB del primer trimestre de 2018 es de Bs10.919 millones, con una tasa de crecimiento de 4,89% respecto a similar periodo en 2017<sup>9</sup>. Posteriormente, en julio de 2018 se publicó el dato oficial del PIB con un valor a precios constantes de Bs10.977 millones, con un crecimiento de 4,93%.

### ***Estimación del PIB con datos a junio de 2018***

Posteriormente, con datos a junio de 2018 se realizó el *nowcast* del PIB. La ecuación de estimación cambió ligeramente respecto a la obtenida al primer trimestre de 2018:

$$g\_IGAE_t = \beta + 2,329 \Delta S_t + u_t$$

En base a esta nueva estimación, el logaritmo de la tasa de crecimiento del PIB al segundo trimestre de 2018 es de:

$$IGAE_{II-1} = 4,7701 + 2,329(-0,0767) + 0,1782 = 4,7697$$

Recuperando las series originales como se lo hizo anteriormente para el primer trimestre, se tiene que el valor del PIB del segundo trimestre de 2018 es de Bs12.128 millones, con una variación acumulada de 4,86% respecto a similar periodo 2017. El dato estimado se encuentra ligeramente por debajo del dato observado de 4,85%.

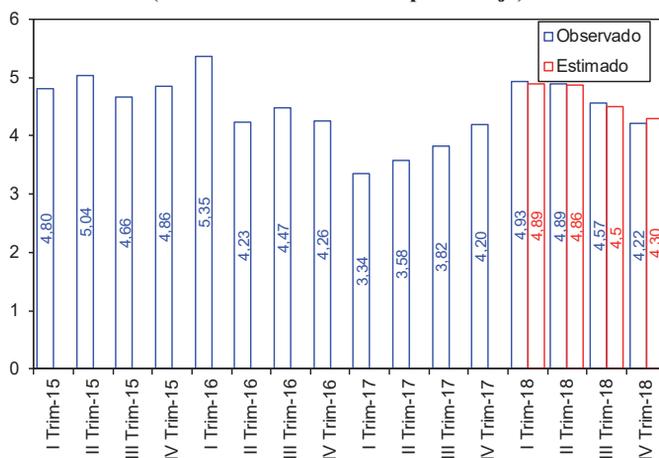
---

9 El resto de las proyecciones de las demás variables que comprenden el modelo se encuentran en el Apéndice F del presente documento.

Posteriormente se realizó proyecciones de corto plazo para el tercer y cuarto trimestres de 2018. Los resultados indican un crecimiento de 4,50% y 4,30%, respectivamente.

Como se observa en el Gráfico 8, las estimaciones en base al modelo de factores dinámicos subestiman el dato oficial publicado por el INE, donde la raíz del error cuadrático medio (RMSE) de las cuatro estimaciones del modelo es de 0,40%.

**Gráfico 8: PIB OBSERVADO Y ESTIMADO, 2013 – IV TRIMESTRE 2018**  
(Variación acumulada en porcentaje)



Fuente: Elaboración propia en base a los resultados del modelo

### IV.3.2 Modelo ARIMA

Para realizar la proyección mediante el modelo ARIMA se tomó la serie del IGAE en la misma que, al ser una serie de frecuencia mensual, se procedió a realizar pruebas de raíz unitaria en la frecuencia estacional<sup>10</sup>. Analizando las FAS y FAP se determinó la especificación del siguiente modelo:

$$\Delta^d \Delta_s^D IGAE_t = \alpha_0 + \alpha_1 AR(1) + \alpha_2 SAR(12) + \alpha_3 MA(1) + \varepsilon_t$$

donde  $d$  es el grado de diferenciación de la frecuencia regular,  $D$  es el orden de diferenciación de la frecuencia estacional y  $s$  representa el número de frecuencia estacional (por ejemplo, en nuestro caso que son datos mensuales  $s = 12$ ). Entonces, se dice que la serie es integrada de órdenes  $(d, D)$  lo que se representa por  $IGAE_t \sim (1, 1)$ . Es decir, para que la serie sea estacionaria se debe realizar una diferencia tanto en la frecuencia regular como en la estacional. Los resultados del modelo y las pruebas de diagnóstico se pueden ver en el Apéndice G.

$$\Delta^1 \Delta_{12}^1 IGAE_t = 0,0526 + [AR(1) = 0,283, SAR(12) = -0,187, MA(1) = -0,901]$$

### IV.3.3 Modelo VAR

Para la estimación del modelo VAR se trabajó con las series en diferencias estacionales, dado que son integradas de orden uno<sup>11</sup>. El modelo estimado es un VAR de orden 12, con 9 variables explicativas: IGAE, agricultura, gas, industria, electricidad, minería, IPC, IPPBX y créditos.

Con 12 rezagos, los residuos son ruido blanco, es decir, no sufren problemas de autocorrelación, tienen distribución normal y no hay evidencia de heteroscedasticidad (ver Apéndice H).

10 Los resultados de la prueba de HEGY indican que el modelo no contempla problemas de raíz unitaria en la frecuencia estacional, pero sí en la frecuencia regular, por lo que se procedió a realizar una diferencia en esta frecuencia (ver Apéndice G1).

11 Los resultados de las pruebas de raíz unitaria se encuentran en el Apéndice B.

### ***Evaluación dentro y fuera de la muestra***

Para las evaluaciones, tanto dentro como fuera de la muestra, se compararon las proyecciones de los diferentes modelos (ARIMA, VAR y Factores Dinámicos) con los datos observados del IGAE. Para la evaluación dentro de la muestra, se realizó la estimación para el periodo 2003M1 – 2018M3. Respecto a las estimaciones fuera de la muestra, se tomó el periodo 2015M1 – 2018M3

Los resultados dentro de la muestra exponen que las estimaciones del modelo de factores dinámicos incurren en un menor error de pronóstico como puede observarse, por ejemplo, con el coeficiente de Theil, el mismo que es cercano a cero. Pero le siguen muy de cerca los modelos univariados, donde las diferencias en los errores de pronóstico son mínimas (Cuadro 14).

**Cuadro 14: ESTADÍSTICOS DE EVALUACIÓN DENTRO DE LA MUESTRA  
PARA EL PERÍODO 2003M1 – 2018M3**

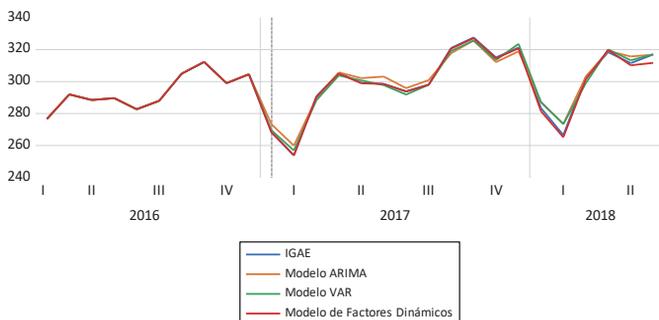
Proyección	RMSE	MAE	MAPE	THEIL
ARIMA	3,3274	2,5642	1,2478	0,0074
VAR	8,5732	6,8165	3,1056	0,0189
Factores Dinámicos	3,2206	2,5151	1,2285	0,0071

De similar forma, las proyecciones fuera de la muestra dan una sólida evidencia de que el modelo de factores dinámicos es el que mejor estima la evolución del IGAE, es decir, son los más cercanos al dato observado, seguida de los modelos VAR (Cuadro 15).

**Cuadro 15: ESTADÍSTICOS DE EVALUACIÓN FUERA DE LA MUESTRA  
PARA EL PERÍODO 2015M1 – 2018M06**

Proyección	RMSE	MAE	MAPE	THEIL
ARIMA	3,4415	2,8365	0,9833	0,0057
VAR	2,6450	2,1449	0,7363	0,0044
Factores Dinámicos	1,5083	0,8894	0,2922	0,0025

En efecto, como se puede observar en el Gráfico 9, las proyecciones del modelo de factores dinámicos (línea roja) son más próximos al dato observado (lineal azul).

**Gráfico 9: COMPARACIÓN DE PRONÓSTICOS FUERA DE LA MUESTRA**

Fuente: Elaboración propia

## V. Conclusiones

Con el objetivo de realizar estimaciones tempranas del PIB, se utilizó el enfoque de modelos de factores dinámicos planteado por Stock y Watson (1988). El cual tiene la ventaja de usar información de diferentes indicadores que pueda ser potencialmente significativa para la realización de pronósticos. Según este modelo, las variables se pueden describir mediante dos componentes, el factor común a todas las series y los factores idiosincrásicos.

Las variables que se utilizaron para la realización de este modelo fueron el índice de producción de agricultura, producción de gas, índice de producción de industria, índice de producción de electricidad, índice de producción de minería, IPC, IPPBX y créditos del sistema financiero. Los resultados del *nowcast* estiman para el primer trimestre de 2018 un crecimiento del PIB de 4,89% y con datos a junio de 2018, se estimó una tasa de crecimiento acumulada de 4,86%; en ambos casos las estimaciones se encontraron ligeramente por debajo de los datos observados. Luego, se realizaron proyecciones de corto plazo para el tercer y cuarto trimestre de 2018, donde los resultados indican un crecimiento de 4,50% y 4,30%, respectivamente.

Una vez obtenidas las estimaciones del PIB boliviano, se tiene que advertir a los hacedores de política económica que se espera un crecimiento de 4,89% para el primer trimestre y que posteriormente

se irá desacelerando hasta llegar a 4,30% en el acumulado al cuarto trimestre de 2018; con estos resultados, las autoridades del BCB podrían efectuar recomendaciones al Órgano Ejecutivo sobre políticas económicas con el objetivo de prever la desaceleración de la actividad económica boliviana.

Por otra parte, se compararon las estimaciones del modelo de factores dinámicos frente a modelos ARIMA y VAR. Los resultados de los criterios de evaluación, tanto dentro como fuera de la muestra, reflejan que los estimadores del modelo de factores dinámicos son más precisos que los modelos tradicionales.

Se sugiere realizar un contraste de las estimaciones de este modelo con otras metodologías como por ejemplo, utilizando modelos de frecuencia mixta (MiDas y Modelos Puente).

## Referencias bibliográficas

AGUILAR, M. A. y O. LORA (1999). "Indicadores para anticipar la evolución de la actividad económica", Banco Central de Bolivia, *Revista de Análisis*, 3 (1), pp. 87 - 119

AGUIRRE, A. y L. CÉSPEDES (2004). "Uso de análisis factorial dinámico para proyecciones macroeconómicas", Banco Central de Chile, Documento de trabajo No. 274, noviembre.

AJEVSKIS, V. and G. DĀVIDSONS (2008). "Dynamic Factor Models in Forecasting Latvia's Gross Domestic Product", Latvijas Banka, Working Paper 2•2008

BAÑBURA, M., D. GIANNONE, M. MODUGNO, L. REICHLIN "Now-Casting and the Real-Time Data Flow" in ELLIOT, G. and A. TIMMERMANN (Eds.) (2013) *Handbook of Economic Forecasting*, Volume 2A, Elsevier B.V., The Netherlands, pp. 195 - 237

BLANCO, E., L. D'AMATO, F. DOGLIOLO, L. GAREGNANI (2017). "Nowcasting GDP in Argentina: Comparing the Predictive Ability of Different Models", Banco Central de la República Argentina, Economic Research Working Paper No. 74, December

BRANCO, G. (2018). "Nowcasting Brazilian GDP: A Performance Assessment of Dynamic Factor Models", Fundação Getulio Vargas, Escola de Pós-Graduação em Economia, Dissertação para obtenção do grau de mestre apresentada à Escola de Pós-Graduação em Economia, março

BOX, G. E. P. and G. M. JENKINS (1976). *Time Series Analysis: Forecasting and Control*, Revised edition, Holden Day, San Francisco

BURNS, A. and W. MITCHELL (1946). *Measuring Business Cycles*, National Bureau of Economic Research, Studies in Business Cycles No. 2, New York

CÁCERES, J. (2016). "Pronóstico de la actividad económica con base en el volumen transaccional - caso boliviano", Banco Central de Bolivia, *Revista de Análisis*, 24, pp. 115 - 145

CAMACHO, M. and J. MARTÍNEZ-MARTÍN (2014). “Real-time forecasting US GDP from small-scale factor models”, Banco de España, Documento de Trabajo No. 1425

CAMACHO, M. y G. PÉREZ-QUIRÓS (2011). “Latin STINGS: Indicadores de crecimiento a corto plazo de los países de América Latina”, Comisión Económica para América Latina y el Caribe, Serie Macroeconomía del desarrollo No. 108, enero

CAMACHO, M. y G. PÉREZ-QUIRÓS (2008). “Introducing the euro-sting: Short term indicator of euro area growth”, Banco de España, Documento de Trabajo No. 0807

CASARES, F. (2017). “*Nowcasting*: Modelos de factores dinámicos y ecuaciones puente para la proyección del PIB de Ecuador”, *Compendium*, 4 (8), pp. 45 - 66

CHEUNG, C. and F. DEMERS (2008). “Evaluating Forecasts from Factor Models for Canadian GDP Growth and Core Inflation”, Bank of Canada, Working Paper 2007-8, February

CRISTIANO, D., M. D. HERNÁNDEZ, J. D. PULIDO (2012). “Pronósticos de corto plazo en tiempo real para la actividad económica colombiana”, Banco de la República – Colombia, Borradores de Economía No. 724, julio

DENTON, F. (1971). “Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization”, *Journal of the American Statistical Association*, 66 (333), pp. 99 - 102

DIAS, F., M. PINHEIRO, A. RUA (2014). “Forecasting Portuguese GDP with factor models”, Banco de Portugal, *Economic Bulletin*, June, pp. 85 - 100

EMIRIS, M. (2016). “A dynamic factor model for forecasting house prices in Belgium”, National Bank of Belgium, Working Paper Research No. 313, November

ENGLE, R. F. and C. W. J. GRANGER (1987). “Co-Integration and Error Correction: Representation, Estimation and Testing”, *Econometrica*, 55 (2), pp. 251 - 276

GEWEKE, J. "The Dynamic Factor Analysis of Economic Time Series Models" in AIGNER, D. and A. GOLDBERGER (Eds.) (1977) *Latent Variables in Socio-Economic Models*, North Holland, New York, pp. 365 - 383

GIANNONE, D., L. REICHLIN, L. SALA "Monetary Policy in Real Time" in GERTLER, M. and K. ROGOFF (Eds.) (2004) *NBER Macroeconomics Annual 2004*, Number 19, The MIT Press, London, pp. 161 - 200

GONZÁLEZ, M. (2017). "Un modelo de *nowcasting* para la tasa de crecimiento del PIB real de Ecuador", Corporación de Estudios para el Desarrollo, noviembre

HAMILTON, J. (1994). *Time Series Analysis*, Princeton University Press, New Jersey

HARVEY, A. (1981). *Time Series Models*, first edition, John Wiley & Sons Inc., United States of America

FERNÁNDEZ, F. J., A. C. HARVEY, J. H. STOCK (1987). "Forecasting and Interpolation Using Vector Autoregressions with Common Trends", *Annales d'Économie et de Statistique*, 6/7, pp. 279 - 287

LIU, P., T. MATHESON, R. ROMEU (2011). "Real-time Forecasts of Economic Activity for Latin American Economies", International Monetary Fund, Working Paper WP/11/98, April

LOAYZA, L. y D. VALDIVIA (2010). "Un método simple para proyecciones de la actividad económica: Una aproximación mediante utilización de ARIMA" paper presented at the Second Bolivian Conference on Development Economics (BCDE 2010), November 18th – 19th, 2010, organized by Institute for Advanced Development Studies, La Paz, Bolivia

NELSON, C. and C. R. PLOSSER (1982). "Trends and random walks in macroeconomic time series: Some evidence and implications", *Journal of Monetary Economics*, 10 (2), pp. 139 - 162

NOVALES, A. (1993). *Econometría*, McGraw-Hill, Madrid, España

PINDYCK, R. S. and D. L. RUBINFELD (2001). *Econometría: Modelos y pronósticos*, McGraw-Hill, México D.F.

RODRÍGUEZ, H. (2014). “Un indicador de la evolución del PIB uruguayo en tiempo real”, Banco Central del Uruguay, Documento de trabajo No. 009-2014

SARGENT, T. and C. SIMS (1977). “Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory”, Federal Reserve Bank of Minneapolis, Working Paper # 55, January

SOLBERGER, M. and E. SPÅNBERG (2017). “Estimating a Dynamic Factor Model in EViews Using the Kalman Filter and Smoother”, Uppsala Universiteit, Working Paper 2017:2, April

STOCK, J. and M. WATSON “Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics” in TAYLOR, J. B. and H. UHLIG (Eds.) (2016) *Handbook of Macroeconomics*, Volume 2A, North Holland, The Netherlands, pp. 415 – 525

STOCK, J. and M. WATSON “Dynamic Factor Models” in CLEMENTS, M. P. and D. F. HENDRY (Eds.) (2011) *The Oxford Handbook of Economic Forecasting*, Oxford University Press, New York, pp. 35 - 59

STOCK, J. and M. WATSON (2005). “Implications of Dynamic Factor Models for VAR Analysis”, National Bureau of Economic Research, Working Paper, 11467, June

STOCK, J. and M. WATSON (2002). “Forecasting Using Principal Components from a Large Number of Predictors”, *Journal of the American Statistical Association*, 97 (460), pp. 1167 - 1179

STOCK, J. and M. WATSON “New Indexes of Coincident and Leading Economic Indicators” in BLANCHARD, O. J. and S. FISCHER (Eds.) (1989) *NBER Macroeconomics Annual 1989*, The MIT Press, London, pp. 351 - 394

STOCK, J. and M. WATSON (1988). "A Probability Model of the Coincident Economic Indicators", National Bureau of Economic Research, Working Paper No. 2772, November

WATSON, M. "Comment" [to GIANNONNE, D., L. REICHLIN, L. SALA "Monetary Policy in Real Time"]in GERTLER, M. and K. ROGOFF (Eds.) (2004) *NBER Macroeconomics Annual 2004*, Number 19, The MIT Press, London, pp. 216 - 221

## APÉNDICES

### Apéndice A: Variables potenciales para el modelo de factores dinámicos

Variable	Fuente	Frecuencia	Rezag
<b>Variables de la actividad económica</b>			
Índice Global de Actividad Económica	Instituto Nacional de Estadística	Mensual	45 días
Índice de producción de agricultura	Instituto Nacional de Estadística	Mensual	30 días
Índice de producción de hidrocarburos	Instituto Nacional de Estadística	Mensual	30 días
Índice de transporte	Instituto Nacional de Estadística	Mensual	30 días
Índice de producción de cemento	Instituto Nacional de Estadística	Mensual	30 días
Índice de generación de energía eléctrica	Comité Nacional de Despacho y Carga	Mensual	20 días
Índice de producción de minería	Instituto Nacional de Estadística	Mensual	30 días
Índice de producción industrial	Instituto Nacional de Estadística	Mensual	30 días
Producción de petróleo crudo y condensados	Yacimientos Petrolíferos Fiscales Bolivianos	Mensual	15 días
Producción de gas natural	Yacimientos Petrolíferos Fiscales Bolivianos	Mensual	15 días
<b>Variables de precios</b>			
IPC	Instituto Nacional de Estadística	Mensual	5 días
IPC subyacente	Instituto Nacional de Estadística	Mensual	5 días
IPC núcleo	Instituto Nacional de Estadística	Mensual	5 días
IPC sin alimentos	Instituto Nacional de Estadística	Mensual	5 días
IPC alimentos	Instituto Nacional de Estadística	Mensual	5 días
IPC importado	Instituto Nacional de Estadística	Mensual	5 días
IPC no importado	Instituto Nacional de Estadística	Mensual	5 días
<b>Variables financieras y monetarias</b>			
Créditos del sistema financiero	Autoridad de Supervisión del Sistema Financiero	Mensual	1 día
Depósitos del sistema financiero	Autoridad de Supervisión del Sistema Financiero	Mensual	1 día
Emisión monetaria	Banco Central de Bolivia	Mensual	1 día
Base monetaria	Banco Central de Bolivia	Mensual	1 día
Agregados monetarios	Banco Central de Bolivia	Mensual	1 día
Liquidez del sistema financiero	Banco Central de Bolivia	Mensual	1 día
<b>Variables externas</b>			
Reservas internacionales netas	Banco Central de Bolivia	Mensual	1 día
Importaciones totales	Instituto Nacional de Estadística	Mensual	30 días
Importación de bienes de consumo	Instituto Nacional de Estadística	Mensual	30 días
Importación de bienes de capital	Instituto Nacional de Estadística	Mensual	30 días
Importación de materias primas y pptos. intermedios	Instituto Nacional de Estadística	Mensual	30 días
Importación de pptos. diversos	Instituto Nacional de Estadística	Mensual	30 días
Exportaciones totales	Instituto Nacional de Estadística	Mensual	30 días
Exportaciones de agricultura, ganadería y otros	Instituto Nacional de Estadística	Mensual	30 días
Exportaciones de hidrocarburos	Instituto Nacional de Estadística	Mensual	30 días
Exportaciones de minería	Instituto Nacional de Estadística	Mensual	30 días
Exportaciones de la industria manufacturera	Instituto Nacional de Estadística	Mensual	30 días
Exportación de efectos personales	Instituto Nacional de Estadística	Mensual	30 días
Términos de intercambio	Banco Central de Bolivia	Mensual	90 días
IPPBX minerales	Banco Central de Bolivia	Mensual	1 día
IPPBX agrícola	Banco Central de Bolivia	Mensual	1 día
IPPBX combustibles	Banco Central de Bolivia	Mensual	1 día
IPPBX ponderado	Banco Central de Bolivia	Mensual	1 día
IPPBX ponderado sin combustibles	Banco Central de Bolivia	Mensual	1 día
Producto Interno Bruto externo relevante	Banco Central de Bolivia	Mensual	1 día

## Apéndice B: Pruebas de raíz unitaria:

### *Traditional Hylleberg, Engle, Granger, and Yoo (HEGY) test*

	Frecuencia anual	Frecuencia estacional	Todas las frecuencias
d(log(igae))	-3,1356 ***	49,9839 ***	48,4735 ***
d(log(agricultura))	-1,6412 *	21,5718 **	19,9604 **
d(log(gas))	-3,1356 ***	49,9839 ***	48,4735 ***
d(log(industria))	-1,5655 *	32,9853 ***	30,4083 ***
d(log(electricidad))	-1,6355 *	56,0620 ***	51,6816 ***
d(log(ipc))	-1,5932 *	33,5404 ***	30,7738 ***
d(log(ipbbx))	-2,1363 **	44,9834 ***	42,3490 ***
d(log(creditos))	-1,1943 *	44,4003 ***	40,9903 ***

\*\*\*p<0,01, \*\*p<0,05, \*p<0,1

Nota: las pruebas se realizaron a las series en logaritmos con diferencia interanual para 2001m1 - 2018m6

## Apéndice C: Prueba de cointegración de Johansen

Date: 09/23/19 Time: 10:51  
 Sample: 2000M01 2018M06  
 Included observations: 210  
 Series: IGAE AGRICULTURA GAS INDUSTRIA ELECTRICIDAD MINERIA IPC ...  
 Lags interval: 1 to 8

Selected (0.05 level\*) Number of Cointegrating Relations by Model

Data Trend:	None	None	Linear	Linear	Quadratic
Test Type	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Trace	4	5	4	4	4
Max-Eig	2	3	3	2	2

\*Critical values based on MacKinnon-Haug-Michelis (1999)

Information Criteria by Rank and Model

Data Trend:	None	None	Linear	Linear	Quadratic
Rank or No. of CEs	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Log Likelihood by Rank (rows) and Model (columns)					
0	-5678.664	-5678.664	-5670.126	-5670.126	-5661.594
1	-5624.743	-5618.439	-5610.275	-5571.059	-5562.612
2	-5595.984	-5574.308	-5566.236	-5525.977	-5518.797
3	-5576.703	-5549.421	-5541.352	-5500.819	-5497.103
4	-5563.327	-5530.144	-5522.340	-5479.775	-5476.301
5	-5550.896	-5516.774	-5511.101	-5468.002	-5464.726
6	-5541.675	-5506.065	-5501.643	-5457.532	-5454.467
7	-5537.125	-5496.847	-5493.631	-5448.356	-5447.577
8	-5534.945	-5492.300	-5490.784	-5442.181	-5442.158
9	-5534.253	-5490.400	-5490.400	-5439.476	-5439.476

Akaike Information Criteria by Rank (rows) and Model (columns)

0	60.25394	60.25394	60.25834	60.25834	60.26280
1	59.91184	59.86133	59.85976	59.49580	59.49154
2	59.80937	59.62198	59.61177	59.24740	59.24569
3	59.79717	59.56592	59.54621	59.18876	59.21051
4	59.84121	59.56328	59.53657	59.16929*	59.18382
5	59.89425	59.61689	59.60096	59.23811	59.24501
6	59.97786	59.69586	59.68231	59.31935	59.31873
7	60.10596	59.78902	59.77744	59.41291	59.42454
8	60.25662	59.92667	59.92175	59.53506	59.54436
9	60.42146	60.08952	60.08952	59.69025	59.69025

Schwarz Criteria by Rank (rows) and Model (columns)

0	70.58216	70.58216	70.73001	70.73001	70.87792
1	70.52695	70.49238	70.61832	70.27030*	70.39355
2	70.71138	70.55586	70.65723	70.32473	70.43459
3	70.98607	70.80264	70.87855	70.56892	70.68631
4	71.31701	71.10283	71.15582	70.85229	70.94651
5	71.65694	71.45928	71.50710	71.22394	71.29460
6	72.02744	71.84107	71.87535	71.60802	71.65521
7	72.44244	72.23707	72.25737	72.00441	72.04792
8	72.88000	72.67756	72.68858	72.42940	72.45463
9	73.33173	73.14324	73.14324	72.88742	72.88742

## Apéndice D: Modelo de estado – espacio

Las ecuaciones de medición definen las relaciones entre las variables observables, como el crecimiento del PIB ( $g\_IGAE$ ), y un factor común real no observable  $S_t$ ; por analogía con Stock y Watson (1988), el factor común puede llamarse “índice de indicadores coincidentes”. Las ecuaciones de medición del modelo están dadas por:

$$\begin{aligned}
 g\_IGAE_t &= \lambda_{IGAE} S_t + e_{IGAE_t} \\
 g\_AGRICULTURA_t &= \lambda_{AGRICULTURA} S_t + e_{AGRICULTURA_t} \\
 g\_GAS_t &= \lambda_{GAS} S_t + e_{GAS_t} \\
 g\_INDUSTRIA_t &= \lambda_{INDUSTRIA} S_t + e_{INDUSTRIA_t} \\
 g\_ELECTRICIDAD_t &= \lambda_{ELECTRICIDAD} S_t + e_{ELECTRICIDAD_t} \\
 g\_MINERIA_t &= \lambda_{MINERIA} S_t + e_{MINERIA_t} \\
 g\_IPC_t &= \lambda_{IPC} S_t + e_{IPC_t} \\
 g\_IPPBX_t &= \lambda_{IPPBX} S_t + e_{IPPBX_t} \\
 g\_CREDITOS_t &= \lambda_{CREDITOS} S_t + e_{CREDITOS_t}
 \end{aligned}$$

Las ecuaciones de estado (transición) describen el movimiento del factor real y de los términos de error (factores idiosincráticos) en ecuaciones indicativas como procesos AR(1):

$$\begin{aligned}
 S_t &= \rho_S S_{t-1} + \varepsilon_{S_t} \\
 e_{IGAE_t} &= \rho_{IGAE} e_{IGAE_{t-1}} + \varepsilon_{IGAE_t} \\
 e_{AGRICULTURA_t} &= \rho_{AGRICULTURA} e_{AGRICULTURA_{t-1}} + \varepsilon_{AGRICULTURA_t} \\
 e_{GAS_t} &= \rho_{GAS} e_{GAS_{t-1}} + \varepsilon_{GAS_t} \\
 e_{INDUSTRIA_t} &= \rho_{INDUSTRIA} e_{INDUSTRIA_{t-1}} + \varepsilon_{INDUSTRIA_t} \\
 e_{ELECTRICIDAD_t} &= \rho_{ELECTRICIDAD} e_{ELECTRICIDAD_{t-1}} + \varepsilon_{ELECTRICIDAD_t} \\
 e_{MINERIA_t} &= \rho_{MINERIA} e_{MINERIA_{t-1}} + \varepsilon_{MINERIA_t} \\
 e_{IPC_t} &= \rho_{IPC} e_{IPC_{t-1}} + \varepsilon_{IPC_t} \\
 e_{IPPBX_t} &= \rho_{IPPBX} e_{IPPBX_{t-1}} + \varepsilon_{IPPBX_t} \\
 e_{CREDITOS_t} &= \rho_{CREDITOS} e_{CREDITOS_{t-1}} + \varepsilon_{CREDITOS_t}
 \end{aligned}$$

Las perturbaciones restantes son variables aleatorias independientes e idénticamente distribuidas:

$$\begin{aligned}\varepsilon_{IGA E_t} &\sim iid(0, \sigma_{IGA E}^2), \varepsilon_{AGRICULTURA_t} \sim iid(0, \sigma_{AGRICULTURA}^2), \varepsilon_{GAS_t} \sim iid(0, \sigma_{GAS}^2), \varepsilon_{INDUSTRIA_t} \sim iid(0, \sigma_{INDUSTRIA}^2), \\ \varepsilon_{ELECTRICIDAD_t} &\sim iid(0, \sigma_{ELECTRICIDAD}^2), \varepsilon_{MINERIA_t} \sim iid(0, \sigma_{MINERIA}^2), \varepsilon_{IPC_t} \sim iid(0, \sigma_{IPC}^2), \varepsilon_{IPPBX_t} \sim iid(0, \sigma_{IPPBX}^2), \\ \varepsilon_{CREDITOS_t} &\sim iid(0, \sigma_{CREDITOS}^2) \text{ y } \varepsilon_{s_t} \sim iid(0, 1)\end{aligned}$$

# Apéndice E: Resultados de la estimación usando el Filtro de Kalman

## A. Con información al primer trimestre de 2018

Sspace: SS\_9  
 Method: Maximum likelihood (BFGS / Marquardt steps)  
 Date: 09/14/18 Time: 10:22  
 Sample: 2000M01 2018M03  
 Included observations: 219  
 Valid observations: 207  
 Partial observations: 12  
 Estimation settings: tol= 1.0e-09, derivs=accurate numeric  
 Initial Values: C(1)=1.00000, C(2)=1.97898, C(3)=1.61630,  
 C(4)=3.19862, C(5)=7.53224, C(6)=14.9062, C(7)=1.50076,  
 C(8)=2.96998, C(9)=1.71926, C(10)=3.40238, C(11)=8.18558,  
 C(12)=16.1991, C(13)=1.61935, C(14)=3.20468, C(15)=11.10633,  
 C(16)=21.9792, C(17)=1.00000, C(18)=7.51076, C(19)=0.90000,  
 C(20)=0.70000, C(21)=0.70000, C(22)=0.70000, C(23)=0.70000,  
 C(24)=0.70000, C(25)=0.70000, C(26)=0.70000, C(27)=0.70000,  
 C(28)=0.70000  
 Convergence achieved after 84 iterations  
 Coefficient covariance computed using outer product of gradients

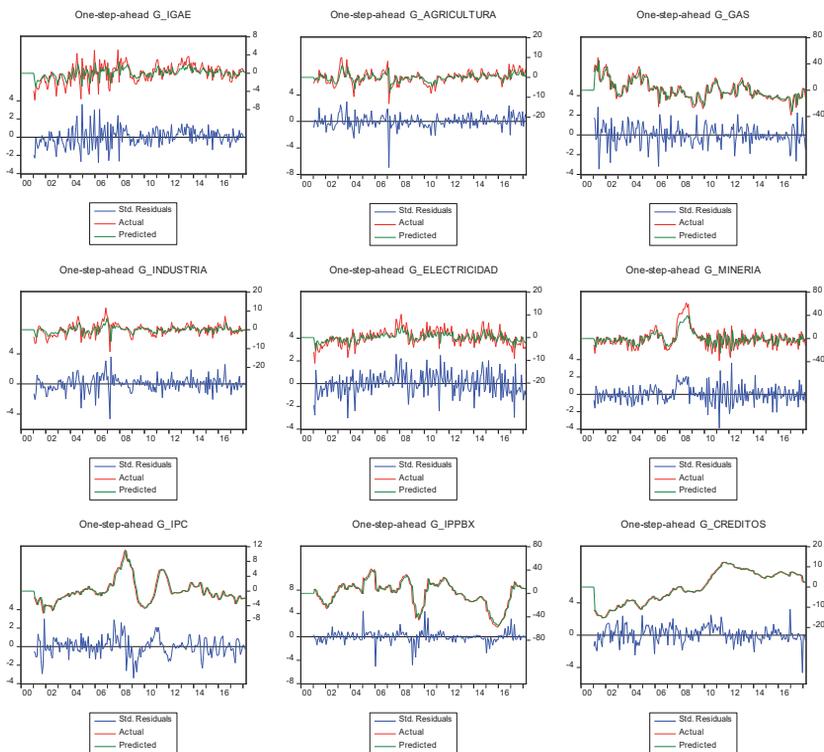
	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	2.212174	0.258956	8.542649	0.0000
C(2)	0.623201	0.331046	1.882523	0.0598
C(3)	0.841238	0.339713	2.476317	0.0133
C(4)	2.537500	0.097023	26.15353	0.0000
C(5)	3.567839	1.118992	3.188442	0.0014
C(6)	-7.294859	0.341554	-21.35784	0.0000
C(7)	1.732346	0.338970	5.142480	0.0000
C(8)	2.187548	0.100742	21.71437	0.0000
C(9)	1.702193	0.413064	4.120899	0.0000
C(10)	2.616030	0.136592	19.15220	0.0000
C(11)	6.265829	1.723719	3.646667	0.0003
C(12)	11.33127	0.603719	18.76912	0.0000
C(13)	-0.066390	0.109055	-0.608780	0.5427
C(14)	-0.727152	0.031153	-23.34163	0.0000
C(15)	0.943456	0.746109	1.264501	0.2061
C(16)	6.282164	0.153315	40.97555	0.0000
C(17)	-0.034216	0.067841	-0.504354	0.6140
C(18)	0.634609	0.023201	27.35279	0.0000
C(19)	0.555960	0.089594	6.205351	0.0000
C(20)	-0.013308	0.499245	-0.026657	0.9787
C(21)	0.594116	0.050200	11.87754	0.0000
C(22)	0.886858	0.024591	36.03636	0.0000
C(23)	0.528558	0.055335	9.551920	0.0000
C(24)	0.483133	0.065069	7.578576	0.0000
C(25)	0.662775	0.044180	15.00154	0.0000
C(26)	0.971798	0.011888	81.74421	0.0000
C(27)	0.954634	0.023688	40.30090	0.0000
C(28)	0.996895	0.004875	204.4850	0.0000
	Final State	Root MSE	z-Statistic	Prob.
S1	-0.215943	0.767487	-0.281363	0.7784
E_IGAE	8.19E-07	0.623257	1.31E-06	1.0000
E_AGRICULTURA	2.599658	2.551639	1.018919	0.3083
E_GAS	-8.930944	7.489926	-1.192394	0.2331
E_INDUSTRIA	0.298588	2.242609	0.133175	0.8941
E_ELECTRICIDAD	-2.119731	2.654359	-0.798531	0.4246
E_MINERIA	-7.060030	11.54977	-0.611270	0.5410
E_IPC	-2.076872	0.727976	-2.852938	0.0043
E_IPPBX	8.544050	6.300734	1.356040	0.1751
E_CREDITOS	2.074163	0.634873	3.270504	0.0011
Log likelihood	-4462.709	Akaike info criterion	43.38849	
Parameters	28	Schwarz criterion	43.83930	
Diffuse priors	0	Hannan-Quinn critier.	43.57079	

## B. Con información al segundo trimestre de 2018

Sspace: SS\_9  
 Method: Maximum likelihood (BFGS / Marquardt steps)  
 Date: 09/14/18 Time: 15:19  
 Sample: 2000M01 2018M06  
 Included observations: 222  
 Valid observations: 210  
 Partial observations: 3  
 Estimation settings: tol= 1.0e-09, derivs=accurate numeric  
 Initial Values: C(1)=1.00000, C(2)=1.97462, C(3)=1.63021,  
 C(4)=3.21905, C(5)=7.56246, C(6)=14.9330, C(7)=1.49354,  
 C(8)=2.94917, C(9)=1.72388, C(10)=3.40401, C(11)=8.15253,  
 C(12)=16.0381, C(13)=1.61470, C(14)=3.18842, C(15)=11.0720,  
 C(16)=21.8629, C(17)=1.00000, C(18)=7.53896, C(19)=0.90000,  
 C(20)=0.70000, C(21)=0.70000, C(22)=0.70000, C(23)=0.70000,  
 C(24)=0.70000, C(25)=0.70000, C(26)=0.70000, C(27)=0.70000,  
 C(28)=0.70000  
 Convergence achieved after 152 iterations  
 Coefficient covariance computed using outer product of gradients

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	2.328649	0.142791	16.30809	0.0000
C(2)	0.146576	0.118397	1.238000	0.2157
C(3)	0.880327	0.305400	2.882539	0.0039
C(4)	2.518646	0.100022	25.18087	0.0000
C(5)	3.204886	0.788169	4.066240	0.0000
C(6)	-7.442269	0.337868	-22.02716	0.0000
C(7)	1.570976	0.288079	5.453288	0.0000
C(8)	2.219327	0.083467	26.58809	0.0000
C(9)	1.395752	0.459843	3.035278	0.0024
C(10)	2.719550	0.137656	19.75615	0.0000
C(11)	5.263848	1.563158	3.367445	0.0008
C(12)	11.51270	0.558093	20.62865	0.0000
C(13)	-0.032395	0.097792	-0.331265	0.7404
C(14)	0.723478	0.031363	23.06821	0.0000
C(15)	0.759835	0.577167	1.315971	0.1882
C(16)	6.250758	0.147315	42.43128	0.0000
C(17)	-0.019511	0.056845	-0.343231	0.7314
C(18)	0.630538	0.022656	27.83134	0.0000
C(19)	0.358354	0.087963	4.073930	0.0000
C(20)	0.981819	0.023180	42.35704	0.0000
C(21)	0.606024	0.049160	12.32771	0.0000
C(22)	0.871774	0.026143	33.34697	0.0000
C(23)	0.535350	0.052833	10.13295	0.0000
C(24)	0.599406	0.067353	7.563200	0.0000
C(25)	0.648599	0.041356	15.68316	0.0000
C(26)	0.971188	0.011808	82.24890	0.0000
C(27)	0.954807	0.023599	40.01867	0.0000
C(28)	0.996921	0.004656	214.1359	0.0000
	Final State	Root MSE	z-Statistic	Prob.
S1	-0.008324	0.740622	-0.011239	0.9910
E_IGAE	0.171767	0.536597	0.320105	0.7489
E_AGRICULTURA	1.340498	2.539305	0.527775	0.5977
E_GAS	3.025614	7.637851	0.396134	0.6920
E_INDUSTRIA	-0.124345	2.278653	-0.054570	0.9565
E_ELECTRICIDAD	-2.325700	2.754444	-0.844345	0.3985
E_MINERIA	-2.049435	11.70241	-0.175129	0.8610
E_IPC	-1.608812	0.723736	-2.222927	0.0282
E_IPPBX	11.46212	6.266633	1.829072	0.0674
E_CREDITOS	2.182998	0.630651	3.461500	0.0005
Log likelihood	-4520.399	Akaike info criterion	43.31808	
Parameters	28	Schwarz criterion	43.76437	
Diffuse priors	0	Hannan-Quinn critier.	43.49850	

## Apéndice F: One-step-ahead de las series del modelo



## Apéndice G: Pruebas de identificación, estimación, diagnóstico del modelo ARIMA

### Prueba de HEGY, en niveles en el lado izquierdo y con una diferencia en la frecuencia estacional y regular en el lado derecho

Seasonal Unit Root test   HEGY			Seasonal Unit Root test   HEGY		
Null	Simulated P-value*	Statistical	Null	Simulated P-value*	Statistical
Nonseasonal unit root (Zero frequency)	0.559768	-2.043752	Nonseasonal unit root (Zero frequency)	0.005643	-5.697132
Seasonal unit root (2 months per cycle)	0.107172	-2.474020	Seasonal unit root (2 months per cycle)	0.005643	-5.199828
Seasonal unit root (4 months per cycle)	0.541115	2.270080	Seasonal unit root (4 months per cycle)	0.000000	19.81589
Seasonal unit root (2.4 months per cy...	0.031221	6.475883	Seasonal unit root (2.4 months per cy...	0.000000	24.33975
Seasonal unit root (12 months per cycle)	0.965724	0.230242	Seasonal unit root (12 months per cycle)	0.000000	17.47571
Seasonal unit root (3 months per cycle)	0.612809	1.913157	Seasonal unit root (3 months per cycle)	0.000000	22.73105
Seasonal unit root (6 months per cycle)	0.769914	1.278961	Seasonal unit root (6 months per cycle)	0.000000	35.13258

\*Monte Carlo Simulations: 1000  
Selected lag using aic criteria: 2

\*Monte Carlo Simulations: 1000  
Selected lag using aic criteria: 0

### Estimación del modelo ARIMA por Mínimos Cuadrados Generalizados

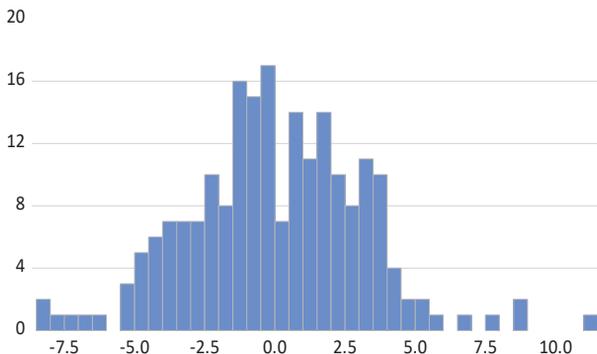
Dependent Variable: D(I(AE,1,12)  
Method: ARMA Generalized Least Squares (Gauss-Newton)  
Date: 09/24/19 Time: 11:07  
Sample: 2001M02 2018M03  
Included observations: 206  
Convergence achieved after 16 iterations  
Coefficient covariance computed using outer product of gradients  
d.f. adjustment for standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.052562	0.026918	1.952669	0.0522
AR(1)	0.282963	0.082634	3.424286	0.0007
SAR(12)	-0.186952	0.071894	-2.600375	0.0100
MA(1)	-0.901061	0.038125	-23.63428	0.0000
R-squared	0.319121	Mean dependent var	0.050460	
Adjusted R-squared	0.309009	S.D. dependent var	3.785851	
S.E. of regression	3.147022	Akaike info criterion	5.158254	
Sum squared resid	2000.557	Schwarz criterion	5.222873	
Log likelihood	-527.3002	Hannan-Quinn criter.	5.184388	
F-statistic	31.55844	Durbin-Watson stat	2.019212	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.84+.23i .28 -.23-.84i -.84+.23i	.84-.23i .23+.84i -.61+.61i -.61+.61i	.61-.61i .23-.84i -.61+.61i -.61+.61i	.61-.61i -.23+.84i -.84-.23i
Inverted MA Roots	.90			

### Pruebas de diagnóstico del modelo ARIMA

Date: 09/24/19 Time: 11:23  
 Sample: 2000M01 2018M06  
 Included observations: 206  
 Q-statistic probabilities adjusted for 3 ARMA terms

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.011	-0.011	0.0273	
		2 0.018	0.018	0.0987	
		3 0.107	0.108	2.5332	
		4 -0.059	-0.057	3.2704	0.071
		5 -0.024	-0.030	3.3935	0.183
		6 -0.073	-0.084	4.5420	0.209
		7 -0.041	-0.030	4.9082	0.297
		8 0.077	0.084	6.2066	0.287
		9 0.086	0.107	7.8123	0.252
		10 0.014	0.012	7.8522	0.346
		11 -0.049	-0.084	8.3859	0.397
		12 0.007	-0.021	8.3966	0.495
		13 0.106	0.121	10.904	0.365
		14 -0.034	0.008	11.157	0.430
		15 0.033	0.044	11.397	0.495
		16 -0.130	-0.171	15.207	0.295
		17 0.083	0.069	16.768	0.269
		18 -0.044	-0.056	17.218	0.306
		19 -0.201	-0.146	26.446	0.048
		20 -0.014	-0.033	26.493	0.066
		21 -0.043	-0.040	26.919	0.081



Series: Residuals	
Sample 2001M02 2018M03	
Observations 206	
Mean	0.041396
Median	-0.077027
Maximum	11.47816
Minimum	-8.496063
Std. Dev.	3.123635
Skewness	0.131208
Kurtosis	3.779223
Jarque-Bera	5.802777
Probability	0.054947

## Apéndice H: Pruebas de diagnóstico del modelo VAR

VAR Residual Serial Correlation LM Tests

Date: 09/24/19 Time: 19:28

Sample: 2000M01 2018M03

Included observations: 195

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	116.1834	81	0.0063	1.489873	(81, 435.5)	0.0067
2	96.02226	81	0.1218	1.204530	(81, 435.5)	0.1254
3	91.12372	81	0.2071	1.137010	(81, 435.5)	0.2120
4	92.21402	81	0.1853	1.151978	(81, 435.5)	0.1899
5	113.4914	81	0.0100	1.451068	(81, 435.5)	0.0106
6	104.3383	81	0.0415	1.320763	(81, 435.5)	0.0432
7	80.58730	81	0.4920	0.994117	(81, 435.5)	0.4981
8	90.80279	81	0.2139	1.132610	(81, 435.5)	0.2188
9	98.84315	81	0.0866	1.243729	(81, 435.5)	0.0895
10	79.28341	81	0.5332	0.976653	(81, 435.5)	0.5392
11	86.55818	81	0.3159	1.074703	(81, 435.5)	0.3216
12	89.00907	81	0.2541	1.108076	(81, 435.5)	0.2594
13	92.65096	81	0.1771	1.157986	(81, 435.5)	0.1815

## Apéndice H: Pruebas de diagnóstico del modelo VAR (Cont.)

VAR Residual Normality Tests  
 Orthogonalization: Cholesky (Lutkepohl)  
 Null Hypothesis: Residuals are multivariate normal  
 Date: 09/24/19 Time: 19:28  
 Sample: 2000M01 2018M03  
 Included observations: 195

Component	Skewness	Chi-sq	df	Prob.*
1	0.046427	0.070052	1	0.7913
2	0.088807	0.256320	1	0.6127
3	0.189560	1.167817	1	0.2799
4	0.009566	0.002974	1	0.9565
5	0.024559	0.019602	1	0.8887
6	0.303522	2.994079	1	0.0836
7	0.293917	2.807591	1	0.0938
8	-0.224780	1.642091	1	0.2000
9	-0.158553	0.817016	1	0.3661
Joint		9.777541	9	0.3688

Component	Kurtosis	Chi-sq	df	Prob.
1	3.253732	0.523086	1	0.4695
2	3.485217	1.912912	1	0.1666
3	2.952556	0.018289	1	0.8924
4	2.874100	0.128787	1	0.7197
5	2.861449	0.155970	1	0.6929
6	3.188302	0.288093	1	0.5914
7	3.405443	1.335619	1	0.2478
8	3.546831	2.429574	1	0.1191
9	3.337991	0.928181	1	0.3353
Joint		7.720512	9	0.5625

Component	Jarque-Bera	df	Prob.
1	0.593138	2	0.7434
2	2.169232	2	0.3380
3	1.186107	2	0.5526
4	0.131761	2	0.9362
5	0.175572	2	0.9160
6	3.282172	2	0.1938
7	4.143210	2	0.1260
8	4.071665	2	0.1306
9	1.745197	2	0.4179
Joint	17.49805	18	0.4891

\*Approximate p-values do not account for coefficient estimation

## Apéndice I: Estadísticos de evaluación de la proyección

### Evaluación dentro de la muestra para el periodo 2003M1-2018M3

Forecast Evaluation  
Date: 09/24/19 Time: 18:12  
Sample: 2003M01 2018M03  
Included observations: 183  
Evaluation sample: 2003M01 2018M03  
Number of forecasts: 3

Combination tests						
Null hypothesis: Forecast i includes all information contained in others						
Forecast	F-stat	F-prob				
Modelo ARIMA	0.006040	0.9940				
Modelo VAR	2.724755	0.0683				
Factores Dinámicos	0.314092	0.7309				
Evaluation statistics						
Forecast	RMSE	MAE	MAPE	SMAPE	Theil U1	Theil U2
Modelo ARIMA	3.327370	2.564187	1.247795	1.248315	0.007383	0.261015
Modelo VAR	8.573195	6.816518	3.105613	3.051995	0.018870	0.585016
Factores Dinámicos	3.220648	2.515127	1.228527	1.229253	0.007145	0.254266

### Evaluación fuera de la muestra para el periodo 2015M1 – 2018M12

Evaluation sample: 2017M01 2018M06  
Number of forecasts: 3

Combination tests						
Null hypothesis: Forecast i includes all information contained in others						
Forecast	F-stat	F-prob				
Modelo ARIMA	16.49731	0.0002				
Modelo VAR	4.673486	0.0264				
Factores Dinámicos	0.084545	0.9194				
Evaluation statistics						
Forecast	RMSE	MAE	MAPE	SMAPE	Theil U1	Theil U2
Modelo ARIMA	3.441480	2.836473	0.983272	0.977369	0.005717	0.171639
Modelo VAR	2.645020	2.144891	0.736270	0.734247	0.004403	0.141995
Factores Dinámicos	1.508327	0.889448	0.292182	0.293260	0.002515	0.076092