

Cambios absolutos y relativos del índice de precios al consumidor

Absolute and relative changes in the consumer price index

J. Cerda-Hernández^{1*} , J. Fernández-Villarreal¹

Departamento de Ing. Económica y Estadística, Universidad Nacional de Ingeniería, Lima, Perú¹

*E-mail: icerdah@uni.edu.pe

Recibido (Received): 23/06/2020 Aceptado (Accepted): 07/10/2020 Publicado (Published): 15/12/2020

RESUMEN

El presente trabajo estima una descomposición factorial dinámica para la inflación peruana, usando datos desagregados mensuales del índice de precios al consumidor de Lima Metropolitana de enero de 2000 hasta diciembre de 2019¹. Uno de los objetivos de la macroeconomía es explicar las fuentes agregadas de cambios en los precios de los bienes de una economía. Este modelo permite identificar precios relativos idiosincráticos, precios relativos agregados y precios absolutos, que representan los choques de oferta y demanda que pueden aparecer en una economía. La componente común a todos los rubros del IPC sirve como una medida alternativa de la inflación subyacente, y nuestros resultados muestran que esta componente está altamente correlacionada con la inflación subyacente publicada por el BCRP. En general, los resultados obtenidos en el presente trabajo son robustos para diferentes estructuras de dependencia intertemporal para los factores considerados.

Palabras clave: Descomposición factorial, filtro de Kalman, inflación pura, inflación subyacente, IPC, análisis de componentes principales.

ABSTRACT

This work estimates a dynamic factor decomposition for Peruvian inflation, using monthly disaggregated data from the consumer price index of Metropolitan Lima from January 2000 to December 2019. One of the objectives of macroeconomics is to explain the aggregate sources of changes in the prices of goods in an economy. This model makes it possible to identify idiosyncratic relative prices, aggregate relative prices and absolute prices, which represent the supply and demand shocks that can appear in an economy. The component common to all items of the CPI serves as an alternative measure of core inflation, and our results show that this component is highly correlated with core inflation published by the BCRP. In general, the results obtained in the present work are robust for different intertemporal dependency structures for the factors considered.

Keywords: Factor decomposition, Kalman filter, pure inflation, core inflation, CPI, principal components analysis.

¹ Estos datos son medidos por el INEI.

1. INTRODUCCIÓN

Uno de los principales objetivos de los modelos econométricos es entender y analizar las características estadísticas de la variable estudiada. Particularmente, el análisis factorial tiene como objetivo primario encontrar componentes o variables no observadas que estén altamente correlacionadas y puedan explicar la variable de interés. *Bai (2003)* y *Bai (2004)* ofrecen una excelente revisión sobre los modelos factoriales dinámicos. El objetivo del presente trabajo es estudiar la inflación peruana usando los modelos factoriales dinámicos. Usaremos el IPC de Lima Metropolitana proporcionado por el Instituto Nacional de Estadística e Informática (INEI), de enero del 2000 a diciembre del 2019. Dichos datos son publicados por el Banco Central de Reserva (BCRP). La inflación se entiende como el aumento generalizado del nivel de precios en una economía, y su control es de vital importancia para la política monetaria de un país. El análisis del incremento de los precios se hace a través del Índice de Precios al Consumidor (IPC), el cual muestra las fluctuaciones que tienen los precios mensualmente producidos por choques internos o externos. *Armas et al. (2011)* mencionan que el IPC de Lima es una buena variable instrumental de la inflación nacional dado que el gasto de Lima representa el 70 por ciento del gasto nacional. Esto fue corroborado en *Monge y Winkelried (2009)*, donde los autores demuestran que el uso del IPC de Lima para el cálculo de la meta inflacionaria del Banco Central garantiza un ancla para el IPC nacional en su conjunto.

Desde el año 2002, el Banco Central de Reserva del Perú (BCRP) viene dirigiendo formalmente la política monetaria del país a través de metas de inflación (*target inflation*), tomando medidas pertinentes para que en el mediano plazo se mantenga dicha meta. La inflación del IPC a 12 meses es la meta numérica oficial del Banco Central.

La evolución de la inflación en el Perú se ha visto marcada por hechos relevantes. En el 2010, el Instituto Nacional de Estadísticas (INEI) actualizó la canasta de consumo utilizada para medir el IPC de Lima Metropolitana. Por otro lado, en el año 2008, ocurrió un aumento general de los precios internacionales de alimentos, lo cual hizo que se alcanzara una inflación de 6,65%. Esto es relativamente alto considerando que el BCRP toma como meta inflacionaria 2% con umbrales de tolerancia del +1% y -1%. Entre el periodo 2002-2009, se llegó a una inflación promedio de 2,6% y siendo las posteriores, al año 2009, crecientes.

En este contexto, entendiendo los posibles impactos que tiene una economía frente a fluctuaciones grandes de la inflación y su importancia por entender su dinámica. W. Phillips estudia el vínculo que existe entre la inflación y el desempleo (ver Phillips, 1958). Este investigador plasmó su investigación en la relación de la tasa de desempleo y la tasa de inflación en el Reino Unido (1861-1957) publicada en 1958. En ese trabajo se encontró que estas tasas se relacionan de manera inversa. Para su estudio, se partió con observar que de manera general una mayor demanda causa aumento de precios y que a una demanda

menor los precios disminuyen. Esto fue inferido al mercado de trabajo de la siguiente manera: cuando la demanda de trabajo aumenta y hay presencia de poco desempleo se espera que el salario aumente ya que las empresas ofrecen mayores salarios buscando atraer personal más calificado; de otro lado, cuando la demanda de mano de obra baja, los empleados están dispuestos a ofertar su trabajo a salarios vigentes (salario básico). Este tipo de relaciones es conocido en la literatura como Curva de Phillips. Basados en este tipo de análisis, que relaciona algunas variables macroeconómicas con la inflación, el BCRP propone el Modelo de Proyección Trimestral (2009) que es utilizado para fines de simulación de política monetaria y de proyección de las principales variables macroeconómicas; y la inflación es descrita o representada por una Curva de Phillips del tipo nekeynesiano que considera la existencia de rigideces en la formación de precios de la economía.

En ese sentido, la inflación constituye uno de los pilares sobre los que se construye la política económica de los Bancos Centrales, pues es una medida directa del nivel de bienestar de una sociedad. En países emergentes como el Perú, con un programa de metas explícitas desde el año 2002 orientado a mantener la estabilidad, adquiere vital importancia la comprensión de la dinámica, los mecanismos y fuerzas propios de dicha variable, como se puede ver en algunos estudios realizados por el BCRP (ver, por ejemplo, *Rossini et al. (2016)*, *Armas et al. (2011)*, *Winkelried (2012)* y *Winkelried y Ramos (2016)*).

Existen muchas formas de medir y analizar la inflación de un país, siendo los más frecuentes el índice de precios al consumidor (ver, por ejemplo, *Winkelried y Ramos, 2016*; *Armas et al., 2011*; *Reis y Watson, 2010*; y sus referencias dentro) y el deflactor del PBI (ver, por ejemplo, *Bollerslev, 1986*; *Engle, 1982* y sus referencias dentro), siendo lo más común los cambios en el IPC, en la actualidad. En esa dirección, algunos trabajos, como los de *Samuelson (1960)*, *Robert F. Engle (1982)*, *Tim Bollerslev (1986)*, *Bryan y Cechetti (1993)*, *Bai y Ng (2002)*, *Bai (2004)*, *Phillips y Sul (2007)*, *Stock y Watson (2002, 2007)* y *Reis and Watson (2007)* constituyen un aporte sustancial en el sentido de ampliación de la perspectiva para la selección de modelos que se ajusten a los datos disponibles. Considerar construcciones econométricas más robustas permite obtener horizontes predictivos más fiables y realistas, permitiendo implementar mejores políticas monetarias.

2. MODELO FACTORIAL DINÁMICO PARA LA INFLACIÓN

Sea $\pi_{i,t}$ la tasa de cambio para un ítem o rubro i del IPC entre el tiempo $t - 1$ y t , para $i = 1, \dots, N$ y $t = 1, \dots, T$, donde el índice t denota el tiempo. Denotamos por π_t el vector que contiene las tasas cambio de los N ítems o rubros del IPC. Suponemos que cada ítem del IPC es la suma de un efecto fijo $\lambda_{i,t}$ para cada ítem i , k , factores o componentes ortogonales no observables que captura las fuentes comunes de la variación en los precios y una componente idiosincrática $\varepsilon_{i,t}$ que representan los precios relativos idiosincráticos que

captura la variabilidad de los precios relativos de un bien específico ocasionados por información no anticipada en la economía. Las componentes de $\pi_{i,t}$ son obtenidas extrayendo un número pequeño $k \geq 1$ de factores comunes $\{F_{jt}\}$, $j = 1, \dots, k$. La componente idiosincrática ε_{it} cubre choques específicos en cada uno de los ítems. Así, el modelo factorial del vector $\pi_t = (\pi_{1t}, \dots, \pi_{Nt})$ puede escribirse de la siguiente forma:

$$\pi_{i,t} = \lambda_i + \theta_{i1}F_{1t} + \dots + \theta_{ik}F_{kt} + \varepsilon_{it} \tag{1}$$

donde $F_t = (F_{1t}, \dots, F_{kt})^T$ es el vector de los k factores que explican las fuentes comunes de los cambios en los precios para $i = 1, \dots, N$ y $t = 1, \dots, T$. Estos factores son variables no observadas². El lector interesado puede revisar *Wansbeek y Meijer (2000)* y *Bai (2003)* para obtener una perspectiva econométrica más amplia y moderna de los modelos de factoriales. Los parámetros θ_{ij} , para $i = 1, \dots, N$ y $j = 1, \dots, k$ representa la contribución del factor común F_j , y mide la i -ésima respuesta del precio al choque. El vector $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})^T$ es el vector que captura la variabilidad del precio relativo asociada a eventos sectoriales idiosincráticos o errores de medida de π_t . Como propone *Reis y Watson (2007)*, puede deberse a choques agregados que afectan a todos los sectores, como cambios en la productividad agregada, el gasto público o la política monetaria, o pueden deberse a choques que afectan a muchos, pero no a todos los sectores, como cambios en los precios de la energía, eventos climáticos en agricultura, o las fluctuaciones del tipo de cambio y el precio de los comerciantes.

El modelo factorial puede ser representado de forma matricial como sigue

$$\pi_t = \lambda + \Theta F_t + \varepsilon_t \tag{2}$$

para todo $t = 1, \dots, T$, donde $\lambda = (\lambda_1, \dots, \lambda_N)^T$, $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})$, y

$$\Theta = \begin{bmatrix} \theta_{11} & \dots & \theta_{1k} \\ \vdots & & \vdots \\ \theta_{N1} & \dots & \theta_{Nk} \end{bmatrix}$$

Para la ecuación de estados o componentes no observadas, podemos proponer un vector autorregresivo (VAR),

$$\Phi(L)F_t = \Omega(L)w_t \tag{3}$$

² Note que N representa el número de variables y T representa el número de observaciones, una interpretación utilizada por el análisis factorial clásico.

donde $w_t \sim N(0, Q)$, y Q es la matriz correlaciones.

Usando la descomposición (1) y (3) para la inflación, *Reis y Watson (2010)* proponen reducir el número de factores mediante la aplicación de la metodología de *Bai (2004)* para explicar las fuentes agregadas de cambios en los precios de los bienes. Para esto usan una descomposición factorial y separan empíricamente las fluctuaciones de los precios en tres componentes: inflación pura, un índice de precios agregados relativos y un precio relativo idiosincrático. Esta descomposición proporciona flexibilidad en el comportamiento idiosincrático sobre el tiempo y el *cross section*. Aplicado a la inflación, el modelo permite analizar el comportamiento de largo plazo de los rubros que componen el índice de precios al consumidor. En dicho artículo suponen que los ítems que componen el IPC se mueven usando un factor común, una inflación pura, que se interpreta como la componente que representa el cambio absoluto de los precios, sin presentar alta volatilidad ni estar influenciado por choques de oferta. El resto de factores en el modelo representa las componentes de precios relativos.

Para ajustar los datos al modelo anterior, primero tenemos que estimar el número de factores que se incluirían, y analizar la evidencia empírica para el caso del Perú. Para esto, seguimos el método propuesto por *Reis y Watson (2010)*, quienes utilizan los resultados presentados en *Bai y Ng (2002)* para el cálculo del número de factores en una descomposición factorial dinámica. Según los resultados obtenidos en *Reis y Watson (2010)*, la inflación -para el caso de los Estados Unidos de Norteamérica- se puede explicar usando solamente dos factores. En el presente trabajo asumimos este resultado, es decir, asumimos que el número de factores es $k = 2$. En el paper *Humala y Rodríguez (2012)*, los autores comprueban que dicha descomposición es factible, por lo menos para los datos observados entre enero de 1995 y julio del 2008. Además, verifican a través de una prueba t que la hipótesis nula $\theta_{i,1} = 1$ no puede ser rechazada para todo $i = 1, \dots, N$, con un p-valor de 11,1% y 8,9% para los niveles de significancia de 5% y 1%, respectivamente. Lo que los autores no testan es si los efectos fijos de cada rubro son iguales a cero. De esta forma, el modelo para la inflación peruana se escribiría como

$$\pi_{i,t} = \lambda_i + F_{1t} + \theta_{12} F_{2t} + \varepsilon_{it} \quad (4)$$

donde $\varepsilon_{it} \sim N(0, \Sigma)$, con $\Sigma = \text{diag}[\sigma_1^2, \dots, \sigma_N^2]$, y por Θ la matriz de coeficiente que mide la sensibilidad (sensitivity) de π_t a cambios del factor F_t

$$\Theta = \begin{bmatrix} 1 & \dots & \theta_{12} \\ \vdots & & \ddots \\ 1 & \dots & \theta_{N2} \end{bmatrix}$$

donde N es el número de rubros que forman el IPC, k el número de factores, y $0 \leq t \leq T$ es el número de observaciones en cada serie.

Para la ecuación de estados usamos diferentes combinaciones, donde se observa que valores altos en los rezagos para el modelo VAR no cambian significativamente los resultados. Otro inconveniente en considerar estructuras autorregresivas con errores Gaussianos más complejos hace la estimación muy complicada, puesto que el número de parámetros a estimar crece de forma exponencial, haciendo que el modelo sea no identificable. Así, tenemos el siguiente modelo para los estados del proceso

$$\Phi(L)F_t = \mu + w_t \quad (5)$$

donde $w_t \sim N(0, Q)$ y $\Phi(L) = 1 - BL$.

3. ESTIMANDO EL MODELO FACTORIAL

El modelo anterior no puede ser estimado usando métodos de regresión desde que las variables explicativas en el modelo (los factores) son variables no observables. Más aún, de modo general, el modelo es no identificable, por ejemplo *Reis y Watson (2010)* y *Lee and Carter (1992)*. La no identificabilidad se debe a que cambios en los precios absolutos no pueden ser distinguidos de los cambios en los precios relativos medios. Para resolver este modelo, se tiene que introducir algunas restricciones en los parámetros o en los factores. Una condición natural que se puede asumir es que los cambios en los precios relativos deben sumar cero³.

Para estimar el modelo, usaremos el filtro de Kalman y la función de máxima verosimilitud para realizar las estimaciones de los parámetros y los componentes no observados⁴. Dado que por definición los residuos tienen distribución normal y son independientes, podemos obtener el logaritmo de la función de verosimilitud, que es dado por:

$$\ln L(\Xi|\pi) = -\frac{1}{2} \sum_{t=1}^T \ln |\Sigma_t| - \frac{1}{2} \sum_{t=1}^T e^T \Sigma_t^{-1} e_t \quad (6)$$

³ Esta condición fue introducida en *Reis y Watson (2007)*.

⁴ *Reis y Watson (2010)* utilizan otro abordaje para estimar el modelo factorial dinámico. Ellos utilizan estimadores de mínimos cuadrados para modelos factoriales, $\sum_{i=1}^N w_i \sum_{t=1}^T (\pi_{it} - \theta F_t)^2$, desarrollados por *Stock y Watson (2002)*, *Bai (2003)* y *Bai y Ng (2006)*.

que es una función no lineal de los parámetros Ξ . Aquí Σ_t es igual a

$$\Sigma_t = \Theta P_{t|t-1} \Theta^T + \Sigma$$

donde es la matriz de covarianza del error de estimación. La forma de obtener el valor de los parámetros que maximiza la función de verosimilitud es usar el algoritmo de optimización de Newton-Raphson de manera recursiva, hasta cierto punto en que la función de verosimilitud sea maximizada por completo. Ver Anexo A para la descripción del algoritmo.

4. LOS DATOS

Los datos utilizados en este trabajo es la variación porcentual mensual de los índices de precios al consumidor de Lima Metropolitana: clasificación transables - no transables, desde enero del 2000 hasta diciembre de 2019. Los datos fueron tomados de la página web del Banco Central de Reserva del Perú⁵. Estos datos corresponden a diferentes rubros de productos transables y no transables, que ya fueron agrupados por sectores y reducidos solamente en 16 series de tiempo. Ver Anexo B para el detalle de cada serie.

Como estos datos ya son variaciones porcentuales, no realizamos la transformación usada por *Reis y Watson (2010)* y *Humala y Rodríguez (2012)* $\pi_{it} = 1200 \times \ln(P_{it}/P_{it-1})$. Para este conjunto de series no excluimos ninguna serie por datos faltantes, ni por problemas de alta correlación. Las mayores correlaciones encontradas en los datos son $Cor(V_7, V_{11}) = 0.637$, $Cor(V_7, V_{12}) = 0.573$, $Cor(V_4, V_{14}) = 0.48$ y $Cor(V_7, V_8) = 0.464$. Observemos que el criterio usado por *Reis y Watson* para eliminar series correlacionadas es $Cor(\pi_{it}, \pi_{jt}) > 0.99$. Entonces, para ver la performance del algoritmo propuesto usaremos $N = 16$ series no correlacionadas. En las series evidentemente se observaron outliers grandes en algunas series de la inflación, como por ejemplo la serie V_1, V_2, V_8, V_{11} y V_{16} y esas observaciones son reemplazadas por la mediana de los 6 valores más próximos al outlier.

En la Figura 6, ver Anexo B, mostramos la dinámica de las series que usamos en esta sección. En los gráficos observamos diferentes comportamientos para los 16 rubros seleccionados del índice de precios. Notemos que cada una de las 16 variables es, a su vez, un conglomerado de diferentes productos de la canasta básica familiar que se incorporan para el cálculo de la inflación (variación porcentual mensual). En la gráfica podemos observar que cada rubro tiene un comportamiento relativamente diferente; lo cual significa que cada rubro tiene una componente dinámica idiosincrática, propio del sector, como propone *Reis y Watson (2010)*. Además, eso se puede observar al calcular las correlaciones,

⁵ www.bcrp.gob.pe

que no pasan de 0,637 en el caso de la serie V_7 y V_{11} . En el modelo propuesto se espera capturar los choques agregados más importantes en la inflación.

Como los datos son mensuales, esto implica que cambios equi-proporcionales en todos los rubros V_i , $1 \leq i \leq 16$, son incluidos en la variable latente F_{1t} , incluso si ocurren en diferentes instantes de tiempo. Por otro lado, el factor F_{2t} captura solo los cambios relativos que persisten por al menos un mes. La mayoría de los análisis macroeconómicos centrados en los shocks agregados utilizan datos trimestrales. En nuestro caso, nos alejaremos un poco de lo tradicional y analizaremos datos mensuales. Los datos que consideraremos en esta sección fueron tratados y publicados por el BCRP, es por eso que creemos que las series tienen una correlación muy baja.

En la gráfica 6 también observamos que muchas de las series tienen mayor volatilidad durante la crisis del 2008, y que algunos picos de volatilidad, ocasionados por algún choque interno o externo, no fueron observados en las otras series en el mismo periodo. Para obtener un mejor ajuste del modelo propuesto reemplazamos estos picos o datos atípicos por la mediana de los 6 valores más próximos, que es una técnica estándar cuando se trabaja con datos atípicos o outliers⁶. Las series asociadas con los productos no transables servicios públicos, servicios de transporte y servicios de educación presentan una componente estacional. En particular, la serie asociada con los servicios de educación tiene una componente estacional fuerte en el mes de marzo, que lo traslada a la inflación subyacente posteriormente, como lo muestra el gráfico 6. Esta componente estacional no es considerada en el modelo factorial dinámico, lo que creemos que afectará en las estimaciones (Ver sección Discusión). En la Figura 6 mostramos las fluctuaciones de los 16 rubros considerados del IPC.

5. ESTIMANDO LOS PARÁMETROS DEL MODELO

En esta sección, presentamos los resultados de los diferentes modelos ajustados a la inflación, derivados de (4) y (5). Comparamos diversos modelos para diferentes estructuras de dependencia temporal para la ecuación de los estados, y para diferentes estructuras de dependencia en los residuos.

Primero observamos que si colocamos una estructura de dependencia en los residuos, el factor F_{1t} modelo ajustado sobreestima la inflación subyacente en diferentes intervalos de tiempo. Los modelos más adecuados se obtuvieron cuando los residuos no tienen correlación entre cada uno de los rubros. Siguiendo *Reis y Watson (2010)*, consideraremos los siguientes modelos

⁶ Esta metodología para tratar outliers fue utilizada por *Reis y Watson (2010)*.

- **Modelo 1:** $F_t = BF_{t-1} + u + w_t$, $w_t \sim N(0, Q)$, donde $Q = \text{diag}[q_{11}, q_{22}]$, $u = [u_1, u_2]^T$
- **Modelo 2:** $\lambda = 0$, $F_t = BF_{t-1} + w_t$, $w_t \sim N(0, Q)$, donde $Q = \text{diag}[q_{11}, q_{22}]$
- **Modelo 3:** $\lambda = 0$, $F_t = F_{t-1} + w_t$, $w_t \sim N(0, Q)$, donde $Q = \text{diag}[q_{11}, q_{22}]$

La Tabla 1 muestra un resumen de la estimación de los parámetros asociados a los factores de los tres modelos considerados. En esta tabla, podemos observar que el segundo factor actúa de forma positiva y negativa para los diferentes rubros considerados del IPC.

Tabla 1. Estimaciones de la matriz Θ de los modelos considerados

	Modelo 1	Modelo 2	Modelo 3
$\theta_{1,2}$	-1.942	1.471	1.175
$\theta_{2,2}$	-0.098	-0.125	0.252
$\theta_{3,2}$	2.052	-0.269	0.148
$\theta_{4,2}$	1.410	-3.611	-1.708
$\theta_{5,2}$	0.692	-0.461	0.050
$\theta_{6,2}$	-1.774	2.288	1.597
$\theta_{7,2}$	0.110	0.638	0.687
$\theta_{8,2}$	0.918	-0.980	-0.130
$\theta_{9,2}$	0.392	-0.955	-0.241
$\theta_{10,2}$	0.258	0.624	0.635
$\theta_{11,2}$	0.655	0.043	0.366
$\theta_{12,2}$	0.203	2.030	1.431
$\theta_{13,2}$	-1.096	2.249	1.589
$\theta_{14,2}$	1.310	-2.559	-1.105
$\theta_{15,2}$	0.731	-0.486	0.027
$\theta_{16,2}$	0.135	-1.444	-0.513

En esta tabla, observamos que la componente asociada a los cambios relativos de algunos de los sectores considerados tiene mayor impacto sobre la inflación. Por ejemplo, según la Tabla 1, choques en la componente de cambios relativos asociada a Productos Transables - Alimentos, Productos Transables - Aparatos Electrodomésticos, Productos No Transables - Alimentos, Productos No Transables - Servicios - Comida Fuera del Hogar y Productos No Transables - Servicios - Alquileres generan fluctuaciones más fuerte en la inflación.

Uno de los criterios más comunes utilizados como herramienta de selección de un modelo en el contexto de máxima verosimilitud es la construcción de funciones penalizadoras como, por ejemplo, el criterio de Akaike (AIC), *Akaike (1973, 1974)*, o correcciones de esta (AICc), *Akaike (1979)*, *Hurvich y Tsai (1989)*. Como el objetivo de aplicar estos modelos a la inflación desagregada es estimar la inflación subyacente, usamos como criterio de selección el nivel de ajuste y la capacidad predictiva del factor asociado a cambios absolutos en la inflación. Para estos criterios, calculamos el Mean squared error (MSE) entre el trend de la inflación subyacente y el factor asociado a cambios absolutos, así como la correlación entre ellos. En la Tabla 2, resumimos nuestro análisis

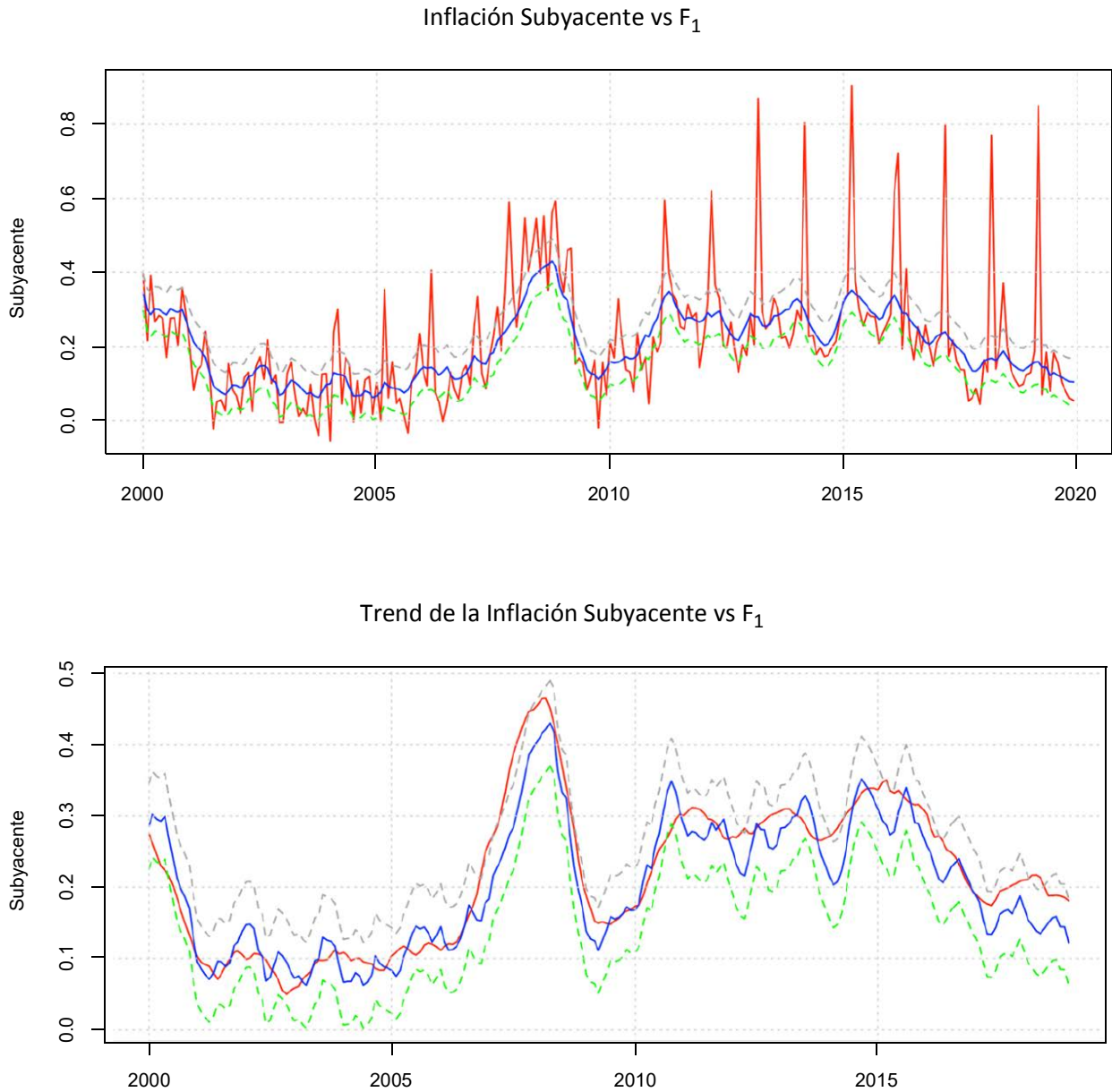
Tabla 2. Criterios de selección para los modelos ajustados.

	Modelo 1	Modelo 2	Modelo 3
AIC	3935.975	3968.696	3980.074
AICc	3937.785	3969.559	3980.775
MSE	0.03732	0.04823	0.07754
Correlación	0.93801	0.93582	0.88398

En la Figura 1, graficamos la inflación subyacente con el factor asociado a los cambios absolutos para el modelo que mejor se ajusta a nuestros datos, el modelo 1. Observamos que el factor F_{1t} asociado a los cambios absolutos de la inflación sigue el patrón de la inflación subyacente, como propone *Reis y Watson (2010)*. Desestacionalizamos la inflación subyacente y lo graficamos junto con el factor F_{1t} para compararlas. Observamos que ambas series siguen el mismo patrón de comportamiento, como se puede observar en la Figura inferior de la Figura 1. Además, la Tabla 2 muestra que la correlación calculada entre la inflación subyacente desestacionalizada y el factor común asociado a los cambios absolutos es de 0,93801. La medida de los cambios absolutos es cercana a la medida estándar de la inflación subyacente del BCRP, y la correlación es altamente significativa.

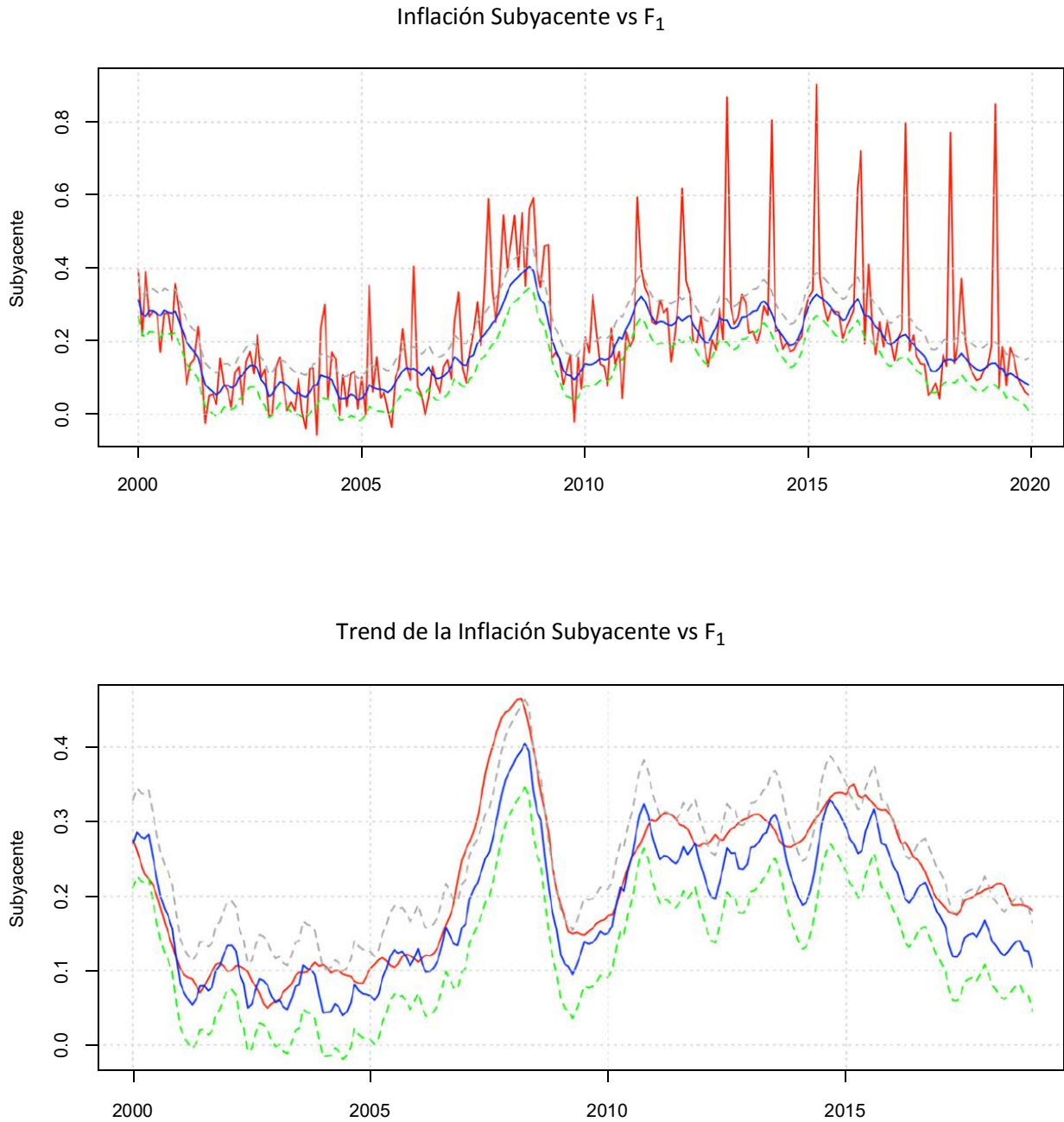
El modelo factorial dinámico propuesto por *Reis y Watson (2010)* no incluye una componente estacional para ningún rubro del índice de precios al consumidor. En la gráfica de la inflación subyacente variación mensual, observamos que en el caso peruano existe una componente estacional muy marcada, por lo menos del 2010 hacia adelante. Esta componente observada empíricamente puede hacer que el nivel de la inflación subyacente sea súper-estimada en algunas regiones, haciendo que exista un gap mayor entre el factor F_{1t} y el trend de la inflación subyacente.

Figura 1. Modelo 1. Figura superior: inflación subyacente variación mensual (curva roja) vs factor asociado a los cambios absolutos en la descomposición factorial del modelo 1 (curva azul). Las curvas ploma y verde representan el intervalo de confianza de la estimación para $\alpha = 5\%$. Figura inferior: inflación subyacente desestacionalizada variación mensual (curva roja) vs factor asociado a los cambios absolutos en la descomposición factorial del modelo 1 (curva azul). Las curvas ploma y verde representan el intervalo de confianza de la estimación para $\alpha = 5\%$.



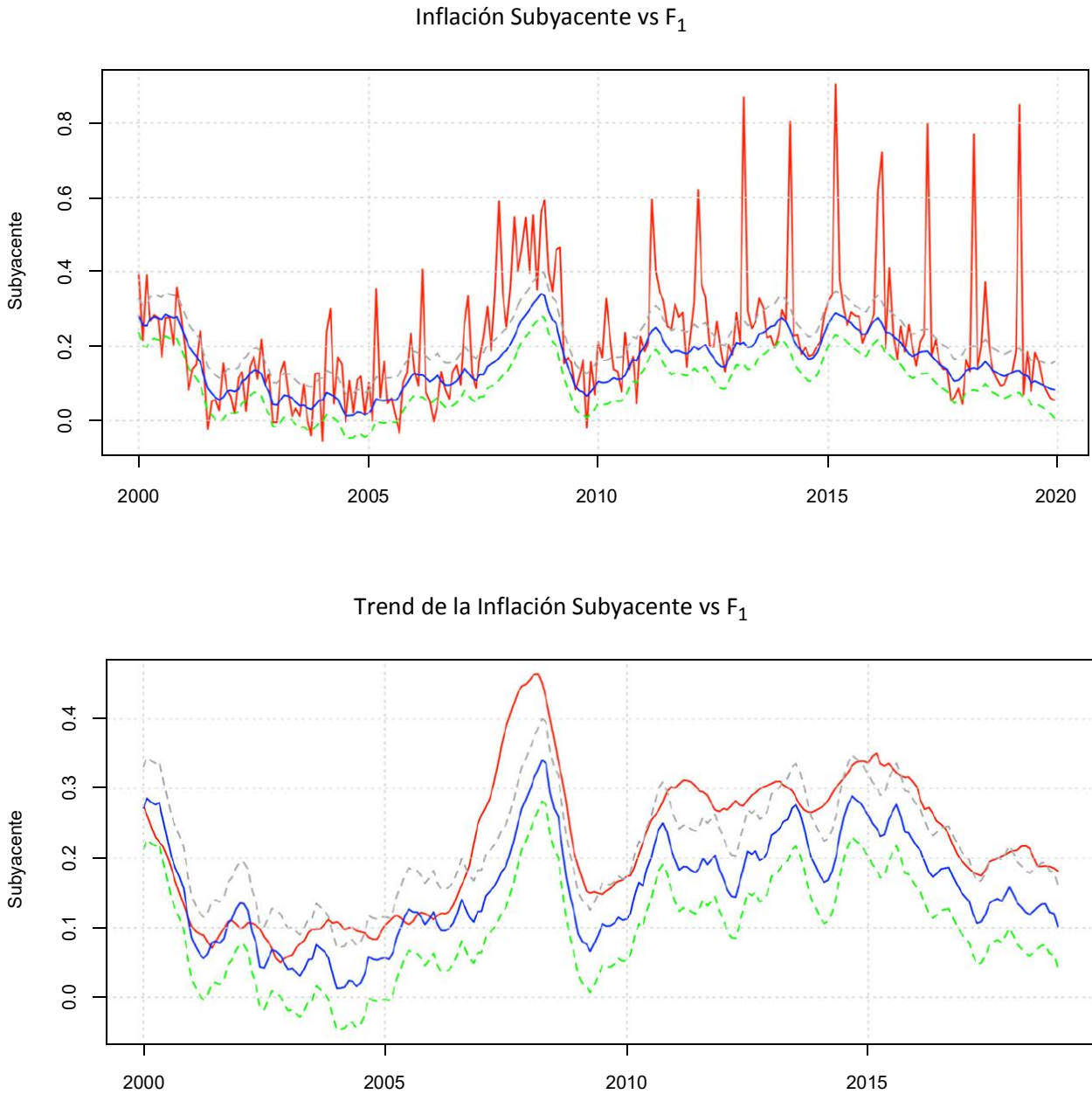
En la Figura 1, observamos que la inflación subyacente desestacionalizada solo se escapa del intervalo de confianza con $\alpha = 5\%$ entre 2007 y 2008, periodo de alta volatilidad en los mercados por la crisis de las hipotecas de los EE.UU.

Figura 2. Modelo 2. Figura superior: inflación subyacente variación mensual (curva roja) vs factor asociado a los cambios absolutos en la descomposición factorial del modelo 2 (curva azul). Las curvas ploma y verde representan el intervalo de confianza de la estimación para $\alpha = 5\%$. Figura inferior: inflación subyacente desestacionalizada variación mensual (curva roja) vs factor asociado a los cambios absolutos en la descomposición factorial del modelo 1 (curva azul). Las curvas ploma y verde representan el intervalo de confianza de la estimación para $\alpha = 5\%$.



Por otro lado, en la Figura 2 y 3, presentamos el ajuste de los modelos 2 y 3, respectivamente.

Figura 3. Modelo 3. Figura superior: inflación subyacente variación mensual (curva roja) vs factor asociado a los cambios absolutos en la descomposición factorial del modelo 3 (curva azul). Las curvas ploma y verde representan el intervalo de confianza de la estimación para $\alpha = 5\%$. Figura inferior: inflación subyacente desestacionalizada variación mensual (curva roja) vs factor asociado a los cambios absolutos en la descomposición factorial del modelo 1 (curva azul). Las curvas ploma y verde representan el intervalo de confianza de la estimación para $\alpha = 5\%$.



CONCLUSIONES

- La información desagregada por categorías de precios en el IPC se puede reducir a unos pocos factores con interpretación económica. Según *Reis y Watson (2010)*, la inflación podría descomponerse en una medida de cambio absoluto F_{1t} (inflación pura) y una parte de cambio relativo F_{2t} (factores internos y externos). La medida de inflación pura se estima como una representación del efecto común al que todos los cambios en los precios están expuestos, y que no están correlacionados con los cambios en los precios relativos. Dado que esta medida de inflación pura aísla la inflación de los precios con alta volatilidad, proporciona una medida alternativa de la inflación subyacente. En este trabajo, mostramos que, en el caso de Perú, las estimaciones de inflación pura están altamente correlacionadas con las medidas estándar de inflación subyacente, a pesar de que se calculan de maneras muy diferentes. Nuestros resultados empíricos consiguen capturar y reproducir los movimientos de la inflación subyacente calculada por el BCRP. Además, este trabajo muestra evidencia fuerte de que la inflación se puede descomponer como propone *Reis y Watson (2010)*, usando metodologías desarrolladas por *Stock y Watson (2002)*, *Bai y Ng (2002)*, *Bai (2003)*, *Bai (2004)*, y *Bai y Ng (2007)*.
- El Modelo 1 se ajustó mejor a los datos de la inflación; obtuvimos una alta correlación entre la inflación subyacente y la inflación pura del modelo. Además, observamos que al eliminar el ruido y la componente estacional de la inflación subyacente se obtiene una alta correlación con la inflación pura, 93,8%. El modelo original propuesto por *Reis y Watson (2010)* no considera una componente estacional para los cambios en los precios. Incorporar una componente estacional en el modelo explicaría mejor los picos estacionales de marzo, fenómeno común observado en la inflación peruana. Creemos que si se redefine el modelo de *Reis y Watson* incorporando componentes estacionales, se obtendrá un mejor ajuste y explicación de los datos. Esto requiere más análisis, y puede ser una línea de investigación futura.

AGRADECIMIENTOS

El presente trabajo fue financiado por el Instituto de Investigación de la Facultad de Ingeniería Económica y Estadística de la Universidad Nacional de Ingeniería (IECOS-UNI). El punto de vista presentado en este trabajo corresponde a los autores y no representa ninguna posición oficial de la institución a la que los autores están afiliados. Los autores agradecen los comentarios y las sugerencias de la Profesora A. Sikov durante el desarrollo de este trabajo.

REFERENCIAS

- Armas, A., L. Vallejos y M. Vega (2011). Indicadores tendenciales de inflación y su relevancia como variables indicativas de política monetaria. Banco Central de Reserva del Perú, *Revista Estudios Económicos*, 20, 27-56.
- Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometría*, 71, 135-71.
- Bai, J. and Ng S. (2007) Determining the number of primitive shocks in factor models. *Journal of Business and Economic Statistics*, 25, 52-60.
- Bai, J. and Ng, S. (2002) Determining the number of factors in approximate factor models. *Econometrica*, 70, 191- 221.
- Bai, J. (2004) Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics*, 122, 137-83.
- Bollerslev T., Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*. Vol. 31, Pag 307- 327 (1986).
- Bryan, M. F. and Cechetti, S. G. (1993) The consumer price index as a measure of inflation. *Federal Reserve Bank of Cleveland Economic Review*, 29, 15-24.
- Chan, J. C., Koop, G and Potter, S. M. (2013). A new model of trend inflation. *Journal of Business and Economic Statistics* 31, 94-106.
- Clark, T. E. and Doh, T. (2014), Evaluating alternative models of trend inflation. *International Journal of Forecasting* 30, 426-448.
- Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometría*, Vol. 50, pp. 987-1007.
- Evans, M. (1991). Discovering the link between inflation Rates and inflation Uncertainty. *Journal of Money, Credit and Banking* 23, 169-184.
- Faust, J and Wright J. (2011). Forecasting inflation. Manuscript, Johns Hopkins University.
- Friedman, M. (1976). Inflation and Unemployment. Nobel memorial lecture.
- Harvey, A. C. (1989). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.
- Humala A. y G. Rodríguez (2012), A factorial decomposition of inflation in Peru: An alternative measure of core inflation. *Applied Economics Letters*, 19(14), 1331-1334.
- Mishkin, F. (2007). Inflation Dynamics. *International Finance*, 10(3), 317-334.

- Modelo de Proyección Trimestral del BCRP. Banco Central de Reserva del Perú, Documento de Trabajo DT. N 2009-006.
- Phillips, A. W. (January, 1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom. *Económica* 25(100).
- Phillips, P. C. B. y D. Sul (2007a). Transition modeling and econometric convergence tests. *Econometría*, 75(6), 1771-1855.
- Ramos, M., Winkelried, D. (2016). Tendencias comunes en el índice de precios al consumidor. Banco Central de Reserva del Perú. *Revista Estudios Económicos* 31, 29 – 44.
- Reis, R. and Watson, M. W. (2007). Relative goods' prices, pure inflation and the Phillips correlation. *American Economic Journal: Macroeconomics*, 2, 128-57.
- Rodríguez, G. (2010). Estimating Output Gap, Core Inflation, and the NAIRU for Perú. *Applied Econometrics and International Development* 10(1), 149-160.
- Rossini, R.; Vega, M.; Quispe, Z.; Pérez, F. (2016). Expectativas de inflación y dolarización en el Perú. Banco Central de Reserva del Perú. *Revista Estudios Económicos* 31, 71 – 84.
- Samuelson, P.A. and R. Solow (1960). Analytical Aspects of Anti-Inflation Policy. *American Economic Review Papers and Proceedings* 50(2), 177-194.
- Samuelson, P. A. (1960). Reflections on Monetary Policy. *The Review of Economics and Statistics*, 42(3), 263-269.
- Solow, Robert M. A. (1962). Policy for Full Employment. *Industrial Relations: A Journal of Economy and Society*, 2(1), 1-14.
- Stock, J. H. and M. Watson (2007). Why has U.S. Inflation become harder to forecast? *Journal of Money, Credit and Banking* 39, 30-33.
- Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97, 1167- 79.

ANEXOS

A. Estimación de los parámetros usando el Filtro de Kalman. Definamos la variable aleatoria

$$F_{t|s} = E(\pi_0, \pi_1, \dots, \pi_s) \quad (7)$$

que se interpreta como la estimación de la variable X en el tiempo t usando la información disponible solo hasta el tiempo s , con $s < t$. Definamos las matrices de covarianza

$$P_{t_1 t_2 | s} = E[(F_{t_1} - F_{t_1 | s})(F_{t_2} - F_{t_2 | s})] \tag{8}$$

Cuando $t_1 = t_2$, escribimos $P_{t | s}$. Este número se interpreta como error cometido de aproximar F_t por $F_{t | s}$. Note que los factores F_t son variables no observadas que se tienen que estimar. Estos errores se usarán en el algoritmo del filtro de Kalman y serán actualizados en cada etapa.

Consideremos las siguientes condiciones iniciales $F_0 = \mu$ y $P_0 = \Sigma_0$. Usando la notación $t | s$ en F y P , para todo $t = 1, \dots, N$ tenemos que $F_{t | t-1}$ sería la estimación de F_t usando las observaciones $\{\pi_0, \pi_1, \dots, \pi_{t-1}\}$, y $P_{t | t-1}$ sería la matriz de covarianza del error de estimar F_t por $F_{t | t-1}$, es decir,

$$F_{t | t-1} = E[F_t | \pi_0, \pi_1, \dots, \pi_{t-1}]$$

$$P_{t | t-1} = E[(F_t - F_{t | t-1})(F_t - F_{t | t-1})^T]$$

Queremos escribir las expresiones anteriores en términos $\pi_t, \Theta, G, \sigma^2, \Sigma_1$ y Q_t . Primero sea $F_{t-1 | t-1}$ el estimador óptimo de F_{t-1} basado en las observaciones $\{\pi_0, \pi_1, \dots, \pi_{t-1}\}$. Sea $P_{t-1 | t-1}$ la matriz de covarianza del error de estimar F_{t-1} por $F_{t-1 | t-1}$, es decir,

$$P_{t-1 | t-1} = E[(F_{t-1} - F_{t-1 | t-1})(F_{t-1} - F_{t-1 | t-1})^T]$$

Dado $F_{t-1 | t-1}$ y $P_{t-1 | t-1}$ el estimador óptimo de F_t es dado por

$$F_{t | t-1} = GF_{t-1 | t-1} \tag{9}$$

mientras que la matriz de covarianza del error de estimación es dada por

$$P_{t | t-1} = E[(F_t - F_{t | t-1})(F_t - F_{t | t-1})^T]$$

$$= GP_{t-1 | t-1}G^T + Q \tag{10}$$

Las ecuaciones (9) y (10) son llamadas ecuaciones de predicción. El *segundo* paso consiste en actualizar las ecuaciones (9) y (10) usando la observación disponible π_t . Dichas actualizaciones son

$$F_{t|t} = F_{t|t-1} + K_t(\pi_t - \Theta F_{t|t-1}) \quad (11)$$

$$P_{t|t} = [I - K_t \Theta] P_{t|t-1} \quad (12)$$

donde

$$K_t = P_{t|t-1} \Theta^T [\Theta P_{t|t-1} \Theta^T + \Sigma]^{-1} \quad (13)$$

que es denominada la ganancia o corrección de Kalman. La ecuación recursiva para la matriz de covarianza del error es

$$P_{t+1|t} = G[I - K_t \Theta] P_{t|t-1} G^T + Q \quad (14)$$

para $t = 1, \dots, T$. La ecuación anterior es conocida como Ecuación de Riccati.

Este algoritmo recursivo recorre hasta un instante t usando los valores iniciales en el cual estos valores se irán actualizando con la información que se tenga hasta el tiempo $t - 1$. Pero una vez actualizados estos, se utiliza las funciones de filtro con el fin de ser más precisos en los coeficientes y reducir el error de estimación. Dado que solo se tiene información hasta el tiempo t de ahí en adelante se usarán las ecuaciones de predicción. Los valores iniciales del algoritmo pueden ser dados en términos de F_0 y P_0 ó $F_{1|0}$ y $P_{1|0}$.

A1. Estimador de Máxima Verosimilitud. Para estimar el modelo factorial dinámico, *Reis y Watson (2010)* usaron componentes principales restringidos. El método consiste en resolver el siguiente problema de optimización

$$\min \sum_{i=1}^N w_i \sum_{t=1}^T (\pi_{it} - \Theta F_t)^2 \quad (15)$$

donde w_i es el igual a la inversa de la varianza muestral de π_{it} . Las propiedades estadísticas de este método son estudiadas en *Stock y Watson (2002)*, *Bai (2003)* y *Bai y Ng (2006)*.

El método que usaremos hace algunas suposiciones paramétricas sobre los parámetros del modelo y las variables latentes, y los parámetros son estimados a través de la máxima verosimilitud. Observemos que los parámetros de la matriz Θ no son conocidos, así como las varianzas de los residuos. Entonces, primero se tiene que estimar los parámetros del modelo

$$\Xi = \{\sigma^2, \sigma_1^2, \dots, \sigma_k^2, \Theta, \Sigma, Q\} \tag{16}$$

Además, suponemos que el estado inicial tiene una distribución gaussiana o normal, $F_0 \sim N(\mu_0, C_0)$, y que los ruidos $\eta_1, \dots, \eta_T, \varepsilon_1, \dots, \varepsilon_T$, y w_1, \dots, w_t son variables no correlacionadas y conjuntamente normales. Sobre esta suposición, podemos calcular la función de verosimilitud utilizando las innovaciones, dadas por

$$e_t = \pi_t - E(\pi|\pi_{t-1}) = \pi_t - \Theta F_{t|t-1} \tag{17}$$

De lo anterior, note que

$$E(e_t) = E(Z_t) - E(Z_t) = 0 \tag{18}$$

Denotemos por Σ_t la varianza del error e_t , entonces la varianza se puede escribir de la siguiente forma

$$\begin{aligned} \Sigma_t &= Var(e_t) = Var(\pi_t - \Theta F_{t|t-1}) \\ &= Var(\Theta F_t + [\varepsilon_p, w_t] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^T - \Theta F_{t|t-1}) \\ &= Var\Theta(F_t - F_{t|t-1}) + [\varepsilon_p, w_t] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^T (F_t - F_{t|t-1})^T \\ &= \Theta E[(F_t - F_{t|t-1})(F_t - F_{t|t-1})^T] \Theta^T + \Sigma \end{aligned}$$

de la ecuación anterior obtenemos, finalmente, la siguiente identidad

$$\Sigma_t = \Theta P_{t|t-1} \Theta^T + \Sigma$$

Dado que por definición los residuos tienen distribución normal y son independientes, podemos obtener el logaritmo de la función de verosimilitud, que es dado por:

$$\ln L(\Xi, \pi) = -\frac{1}{2} \sum_{T=1}^T \ln |\Sigma_t| - \frac{1}{2} \sum_{T=1}^T e^T \Sigma_t^{-1} e_t \tag{19}$$

que es una función no lineal de los parámetros Ξ . La forma de obtener el valor de los parámetros que maximiza la función de verosimilitud es usar el algoritmo de optimización de Newton-Raphson de manera recursiva, hasta cierto punto en que la función de verosimilitud sea maximizada por completo. Los pasos para la estimación son:

1. Seleccionar los valores iniciales para los parámetros, denotados por $\Xi^{(0)}$.
2. Utilizar el filtro de Kalman con los valores iniciales $\Xi^{(0)}$, para obtener un nuevo conjunto de parámetros, que incluye a las matrices de covarianzas.

$$\{e_t^{(0)}, \Sigma_t^{(0)}, t = 1, \dots, T\} \quad (20)$$

3. Ejecutar una iteración del algoritmo de Newton-Raphson utilizando los valores obtenidos en el paso 2 y obteniendo con esto un nuevo conjunto de parámetros $\Xi^{(1)}$.
4. Para cada iteración repetimos el paso 2 utilizando $\Xi^{(j)}$ para obtener un nuevo conjunto de innovaciones y sus respectivas matrices de covarianzas, $\{e_t^{(j)}, \Sigma_t^{(j)}, t = 1, \dots, T\}$. Repetir el paso 3, para obtener una nueva estimación $\Xi^{(j+1)}$. Cuando las estimaciones se estabilizan o cuando la función de verosimilitud se estabiliza, el algoritmo para estabilizar significa que

$$\|\Xi^{(j+1)} - \Xi^{(j)}\| < \epsilon \quad \|\ln L(\Xi^{(j+1)} | \pi) - \ln L(\Xi^{(j)} | \pi)\| < \epsilon$$

donde ϵ es una cantidad pequeña preestablecida.

B. Clasificación de las series utilizadas. Las series utilizadas en el presente trabajo fueron tomadas de la página del BCRP, y son:

- V1: Productos Transables - Alimentos
- V2: Productos Transables - Textil y Calzado
- V3: Productos Transables - Combustibles
- V4: Productos Transables - Aparatos Electrodomésticos
- V5: Productos Transables - Otros Transables
- V6: Productos No Transables - Alimentos
- V7: Productos No Transables - Servicios
- V8: Productos No Transables - Servicios - Servicios Públicos
- V9: Productos No Transables - Servicios - Otros Servicios Personales
- V10: Productos No Transables - Servicios - Servicios de Salud
- V11: Productos No Transables - Servicios - Servicios de Transporte
- V12: Productos No Transables - Servicios - Servicios de Educación
- V13: Productos No Transables - Servicios - Comida Fuera del Hogar
- V14: Productos No Transables - Servicios - Alquileres
- V15: Productos No Transables - Servicios - Otros Servicios
- V16: Productos No Transables - Otros No Transables

“La información del IPC se puede clasificar en sus componentes transables y no transables. Los productos transables son aquellos bienes susceptibles de ser comercializados en el mercado internacional, por lo que sus precios se ven influenciados por la evolución de los precios internacionales, aranceles, costos de transporte y el tipo de cambio. Comprenden el 37 por ciento de la canasta de consumo. Los productos no transables son aquellos bienes y servicios que por su naturaleza no son susceptibles de ser comercializados en el mercado internacional, por lo que su precio se determina por las condiciones de oferta y demanda en el mercado interno. Comprenden el 59 por ciento de la canasta de consumo”⁷.

Figura 4. Correlaciones entre las variables seleccionadas.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16
V1	1.000	0.240	0.024	0.048	0.031	0.026	0.034	-0.045	-0.034	0.003	-0.072	0.053	0.287	-0.045	0.033	0.050
V2	0.240	1.000	-0.174	0.071	0.107	-0.026	0.034	-0.001	0.164	0.036	-0.019	-0.072	0.353	0.081	0.099	0.010
V3	0.024	-0.174	1.000	0.090	-0.005	0.032	0.121	0.064	-0.076	-0.020	0.109	0.061	-0.088	-0.012	-0.019	-0.017
V4	0.048	0.071	0.090	1.000	0.162	-0.051	0.056	0.127	0.029	0.101	-0.075	-0.028	0.031	0.480	0.180	0.108
V5	0.031	0.107	-0.005	0.162	1.000	0.045	0.195	0.106	0.207	0.149	0.008	0.108	0.244	0.309	0.167	0.225
V6	0.026	-0.026	0.032	-0.051	0.045	1.000	0.144	0.057	-0.037	-0.040	-0.025	0.238	0.075	0.003	0.012	0.101
V7	0.034	0.034	0.121	0.056	0.195	0.144	1.000	0.464	0.193	0.185	0.637	0.573	0.264	0.197	0.086	0.084
V8	-0.045	-0.001	0.064	0.127	0.106	0.057	0.464	1.000	0.104	0.108	0.000	-0.022	0.131	0.139	0.130	0.074
V9	-0.034	0.164	-0.076	0.029	0.207	-0.037	0.193	0.104	1.000	0.194	0.075	0.001	0.254	0.025	0.103	-0.039
V10	0.003	0.036	-0.020	0.101	0.149	-0.040	0.185	0.108	0.194	1.000	0.017	0.024	0.266	0.070	0.101	0.097
V11	-0.072	-0.019	0.109	-0.075	0.008	-0.025	0.637	0.000	0.075	0.017	1.000	0.072	-0.101	0.104	-0.116	0.011
V12	0.053	-0.072	0.061	-0.028	0.108	0.238	0.573	-0.022	0.001	0.024	0.072	1.000	0.057	-0.037	0.017	0.008
V13	0.287	0.353	-0.088	0.031	0.244	0.075	0.264	0.131	0.254	0.266	-0.101	0.057	1.000	0.056	0.134	0.144
V14	-0.045	0.081	-0.012	0.480	0.309	0.003	0.197	0.139	0.025	0.070	0.104	-0.037	0.056	1.000	0.085	0.107
V15	0.033	0.099	-0.019	0.180	0.167	0.012	0.086	0.130	0.103	0.101	-0.116	0.017	0.134	0.085	1.000	0.092
V16	0.050	0.010	-0.017	0.108	0.225	0.101	0.084	0.074	-0.039	0.097	0.011	0.008	0.144	0.107	0.092	1.000

⁷ Esta definición fue tomada de la *Guía metodológica de la inflación del BCRP*.

Figura 5. Gráficos BoxPlot de las fluctuaciones de los 16 rubros considerados del IPC.

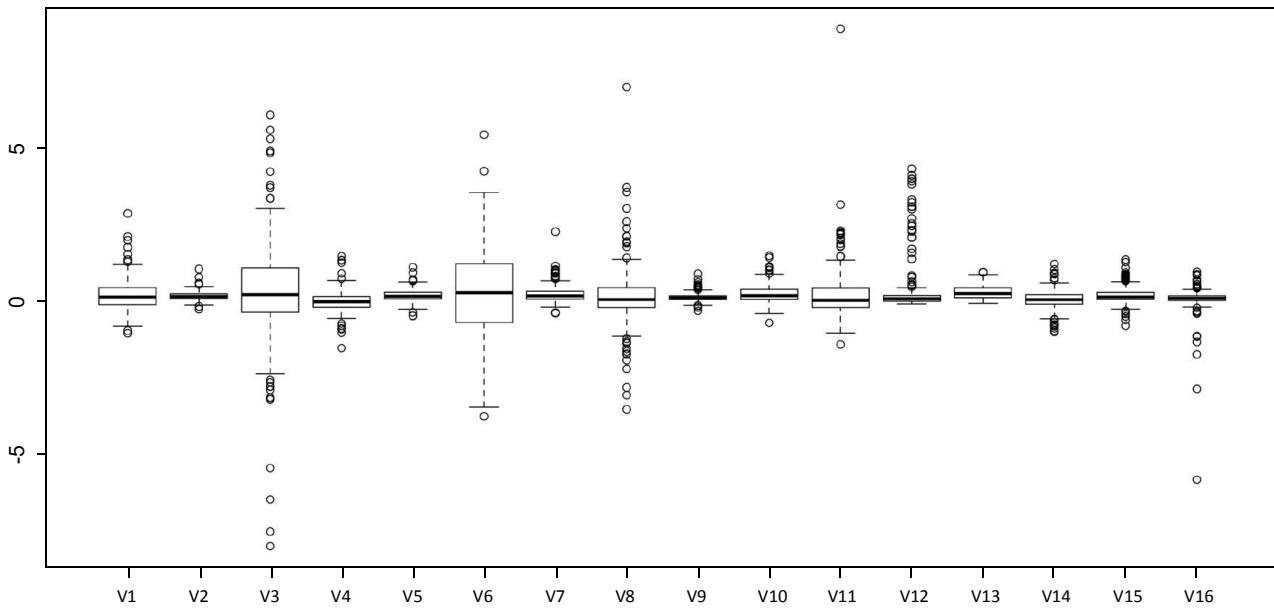
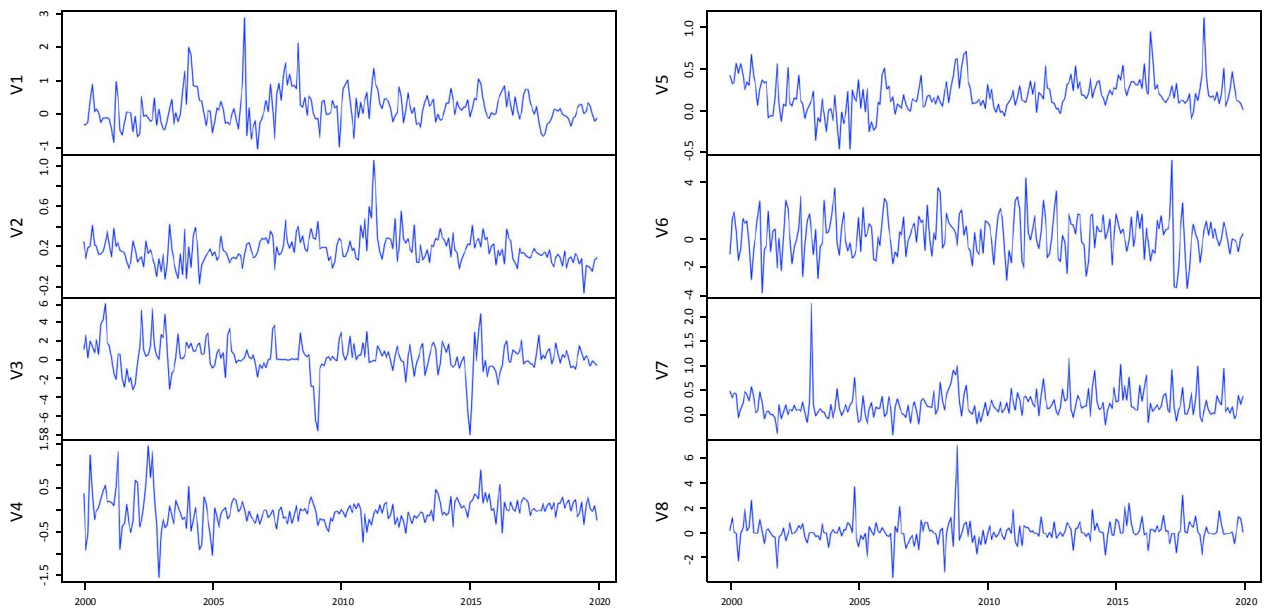
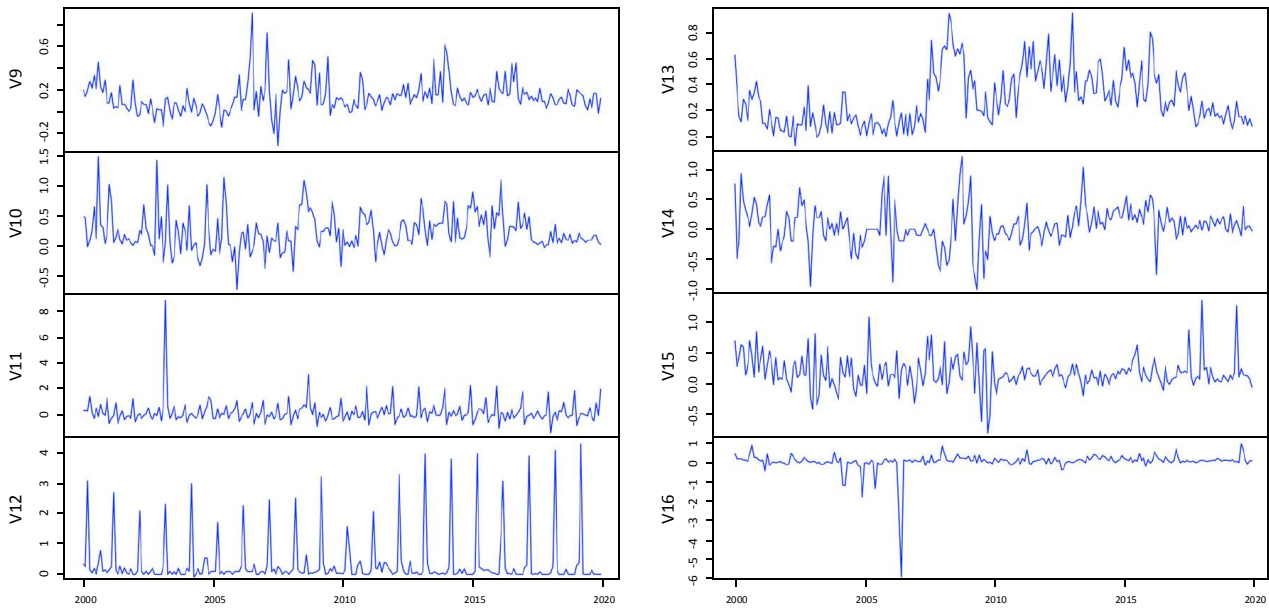


Figura 6. En esta figura mostramos la dinámica de la inflación para los rubros \$V_1, \dots, V_{16}\$ entre enero del 2000 y diciembre de 2019.





Los artículos publicados por IECOS pueden ser compartidos a través de la licencia Creative Commons: CC BY 4.0 Perú. Permisos lejos de este alcance pueden ser consultados a través del correo revistas@uni.edu.pe.



