

CAPÍTULO VI

Análisis de factores comunes estacionales en datos masivos

Fabio H. Nieto*
Daniel Peña
Stevenson Bolívar

Las variables que se estudian en economía y finanzas tienen con frecuencia un comportamiento estacional. Al analizar este tipo de variables para hacer inferencia sobre su dinámica no conviene desestacionalizarlas, ya que esto supone aplicar filtros de corrección estándar de la estacionalidad que pueden no ser adecuados para las variables estudiadas. En la actualidad podemos analizar conjuntos muy grandes de variables, dada la presencia cada vez más frecuente de datos masivos, y los métodos tradicionales para tratar la estacionalidad son difíciles de aplicar en estos casos. Una manera efectiva de hacerlo es mediante el análisis factorial, que resume todas las relaciones de dependencia comunes, incluyendo la estacional, en un conjunto pequeño de factores, permitiendo además una dinámica específica de cada serie. En este capítulo presentamos la metodología para construir un modelo factorial estacional e ilustramos su aplicación en un conjunto de variables macroeconómicas estacionales que corresponden al dinero en efectivo en circulación en algunos países latinoamericanos.

Palabras clave: datos masivos, estacionalidad, modelo factorial dinámico, series temporales multivariadas; procesos integrados.

* Los autores del presente trabajo agradecen a Luis Fernando Melo, investigador en econometría del Banco de La República, el Banco Central de Colombia, por sus sugerencias sobre el tipo de variables a utilizar en este estudio, y a Fabio D. Nieto-Mosquera, economista jefe del Banco Agrario de Colombia, por su ayuda en la interpretación económica de los resultados que se presentan en este documento. Daniel Peña agradece el apoyo parcial de la Agencia Nacional de Evaluación de la Calidad y Acreditación con referencia PID2019-109196GB-I00.

1. INTRODUCCIÓN

El análisis de factores comunes (dinámicos) en series temporales multivariadas ha recibido mucha atención en los últimos años. Este método puede considerarse como una forma de reducir la dimensión de un vector de variables y , cuando el número de series analizado es grande, es una herramienta muy utilizada para el manejo de *big data*. La idea central detrás del modelo factorial es explicar los comovimientos entre las variables observadas por un número muy pequeño de factores subyacentes latentes. Este ha sido un objetivo importante de los macroeconomistas en décadas recientes, como lo podemos consultar en los trabajos de Stock y Watson (2016, 2017) y Diebold (2003), entre otros.

Para series temporales multivariantes se han considerado tres tipos de modelos factoriales: con factores comunes dinámicos exactos (FCDE), factores comunes dinámicos aproximados (FCDA) y factores comunes dinámicos generalizados (FCDG). Entre muchos otros, los trabajos de Peña y Box (1987), Peña y Poncela (2006) y Lam y Yao (2012) están en la órbita de los FCDE; Bai (2004), Bai y Ng (2002, 2004) y Ahn y Horenstein (2013) han trabajado en el campo de los FCDA y Forni *et al.* (2000) han sido pioneros de los FCDG. En este último caso, el análisis se ha realizado bajo la premisa de que el número de variables es muy grande. En este libro y en Peña, Poncela y Ruiz (2021) pueden encontrarse varios trabajos panorámicos sobre las aplicaciones de los modelos factoriales a la predicción y al análisis económico.

Los métodos citados han supuesto, explícita o tácitamente, que las variables analizadas no son estacionales. Sin embargo, y de manera reciente, el problema de analizar factores comunes en presencia de estacionalidad ha empezado a ser considerado, ya que la desestacionalización de series temporales puede inducir comportamientos ficticios en las variables bajo estudio, como se puede consultar en la literatura (ver, por ejemplo, Bell y Hillmer, 1984). Es frecuente que las variables macroeconómicas tengan estacionalidad, y un análisis de factores comunes dinámicos deberá tener en cuenta explícitamente esta característica. Modelos factoriales con estacionalidad han sido propuestos, entre otros, por Melo *et al.* (2001), Busetti (2006), Alonso *et al.* (2011), García-Martos, Rodríguez y Sánchez (2011), Camacho, Lovcha y Perez-Quiroz (2015) y Nieto, Peña y Saboyá (2016). El objetivo de este capítulo es explicar este último enfoque y destacar sus potenciales ventajas en el análisis de datos masivos de series económicas.

El resto de este capítulo está organizado como sigue: en la sección segunda presentaremos una descripción del concepto de estacionalidad y su modelización en series económicas, tanto en series univariantes como multivariantes. La sección tercera introduce el modelo factorial dinámico estacional propuesto por Nieto, Peña y Saboyá (2016) y sus propiedades principales. La sección cuarta aplica el análisis factorial estacional a datos del dinero en efectivo en circulación en 15 países latinoamericanos. Las conclusiones de este trabajo se presentan en la sección quinta.

2. ESTACIONALIDAD EN SERIES TEMPORALES

2.1. Estacionalidad univariante

Históricamente, el concepto de estacionalidad surge al tener en cuenta el efecto de las estaciones climatológicas sobre la variabilidad de una serie temporal. Jevons (1862), citado por Klein (1997), definió este efecto al analizar series temporales comerciales (precios, ventas, inventarios, etc.), comprobando que las observaciones mensuales o trimestrales tienden a tener un valor medio diferente en distintos meses o trimestres. Se induce entonces un comportamiento periódico a través del tiempo, con periodo $s = 12$ si los datos son mensuales, o $s = 4$ si son trimestrales. La práctica inicial fue desestacionalizar este tipo de variables, esto es, estimar el ciclo estacional y restarlo a la serie, pero para ello hay que suponer un cierto modelo estacional que puede no ser adecuado para la serie, lo que puede introducir ciclos o comportamientos ficticios en la variable desestacionalizada; véase por ejemplo Bell y Hillmer (1984).

La estacionalidad de una serie temporal se ha modelado con varios tipos de modelos estadísticos. Inicialmente el ciclo estacional, generalmente ligado a la temperatura, se representaba por medio de una función determinista periódica, como por ejemplo, las funciones seno y coseno. Es claro que esta formulación de la estacionalidad es muy restrictiva: una serie mensual de ventas, puede tener estacionalidad por unas ventas promedio más altas en diciembre, efecto que se representará mal con una función sinusoidal. Una alternativa más flexible es estimar un modelo con s variables ficticias, una para cada periodo (y omitir la constante en el modelo), o introducir $s - 1$ variables ficticias más la constante. Por ejemplo, con datos mensuales los coeficientes de las 12 variables ficticias estiman la media de cada mes, mientras que con 11 y la constante los coeficientes son el incremento (o decremento si es negativo) de cada mes respecto a la media del que tomamos como referencia, cuya media se estima con la constante. En lugar de suponer una estacionalidad determinista, un modelo más general es permitir que los efectos evolucionen con el tiempo. Por ejemplo, en los modelos estructurales o de componentes no observables (Harvey, 1989), la serie se representa como:

$$y_t = T_t + S_{t,s} + I_t, \quad [1]$$

donde T_t representa la tendencia de largo plazo de la variable, $S_{t,s}$ es la estacionalidad que tiene una dinámica con periodo s y seguirá un proceso definido, en general no determinista, e I_t denota la componente irregular. El modelo que se establezca para la evolución de $S_{t,s}$ determina el comportamiento estacional.

Algunos autores consideran que la estacionalidad aleatoria puede tener efectos permanentes o transitorios. En el primer caso la estacionalidad permanece en la serie, por ejemplo, si $S_t^{(i)} = S_{t-s}^{(i)} + \eta_t$ para $i = 1, \dots, s$, donde η_t es un proceso estacionario, la estacionalidad es permanente y la serie no estacionaria. Sin embargo, si $S_t^{(i)} = \phi S_{t-s}^{(i)} + \eta_t$, con $\phi < 1$, la estacionalidad prevista desaparece a largo plazo y el proceso es estacionario. Por ejemplo,

Hylleberg *et al.* (1990) definen la estacionalidad como picos en el espectro de la serie en las frecuencias estacionales, que pueden estar causados por efectos permanentes o transitorios. Otros autores, como Box y Jenkins (1970), suponen que la estacionalidad es una característica permanente en la serie. Entonces, para conseguir una serie estacionaria es necesario aplicar diferencias estacionales, $\nabla_s = (1-B^s)$ donde B es el operador de retardo y $B^s y_t = y_{t-s}$. Esto es debido a que con estacionalidad estable los valores medios de la serie en distintos puntos del ciclo estacional serán diferentes.

Una forma general de modelar la estacionalidad es redefinir los valores temporales como $t = i + js$, donde $i = 1, \dots, s$ son los instantes del ciclo estacional y $j = 0, \dots, L-1$ representa los ciclos completos de periodo s , y supondremos para simplificar que tenemos $N = Ls$ datos. Entonces, $y_t - y_{i+js}$. Para cada valor de i podemos construir una serie temporal que relaciona los valores de los mismos periodos a lo largo del ciclo estacional, por ejemplo los meses dentro del año. Supongamos que estas series son estacionarias. Llamando c_i a la media de cada una de las s series, su estructura será del tipo:

$$y_{i+js} = c_i + \Phi_i^{-1}(B^s)\Theta_i(B^s)u_{i+js} \quad [2]$$

donde en esta ecuación i es fijo, representa el periodo, $u_{i+js} = u_t$ es un proceso estacionario de media cero. Los ruidos de estas s ecuaciones pueden unirse en un proceso único de ruido que estará incorrelado para los retardos múltiplos de s pero puede tener dependencia a retardos $1, 2, \dots$. Esta serie de ruido común, u_t tal que $u_{i+js} = u_t$, seguirá un cierto proceso ARMA del tipo:

$$u_t = \phi^{-1}(B)\theta(B)a_t, \quad [3]$$

donde a_t es ruido blanco. Las $s+1$ ecuaciones [2] y [3] representan la modelización del proceso estacional estacionario. En el caso particular de que los modelos [2] sean todos idénticos y con los mismos parámetros, es decir todos los periodos estacionales tienen la misma media y la misma dinámica, sustituyendo [3] en [2] tenemos un modelo para la serie completa y_t dado por:

$$\Phi(B^s)\phi(B)(y_t - c) = \theta(B)\Theta(B^s)a_t$$

Este es un modelo ARMA donde los operadores AR y MA factorizan en dos bloques de retardos B y B^s , pero las medias de las observaciones, tanto marginales como condicionadas al periodo, son las mismas. Por ejemplo, si la serie es mensual este modelo implica que todos los meses tienen la misma media y estructura de dependencia de los meses pasados. En este sentido son intercambiables. La estacionalidad es solo una dependencia transitoria que hace que si un mes el valor es especialmente alto o bajo es previsible que este efecto se extienda a corto plazo a los años futuros.

Supongamos ahora que las series y_{i+js} son no estacionarias. Por ejemplo siguen un paseo aleatorio $y_{i+js} = y_{i+(j-1)s} + u_{i+js}$. Entonces, suponiendo que $\nabla_s y_{i+js}$ es estacionario para $i = 1, \dots, s$, si las s series siguen el mismo modelo estacionario utilizando el mismo razonamiento anterior podemos escribir el modelo común para todas las series:

$$\Phi(B^s)\phi(B)\nabla_s y_t = \theta(B)\Theta(B^s)a_t. \quad [4]$$

Observemos que la diferencia estacional es imprescindible para permitir que la media de distintos periodos estacionales no sea la misma y exista un efecto estacional permanente. El modelo [4] puede también incluir diferencias regulares $\nabla = (1-B)$.

La recogida de datos masivos con alta frecuencia hace que muchas series actuales tengan varios tipos de estacionalidad. Por ejemplo, si medimos el consumo de electricidad cada hora tendremos una estacionalidad horaria dentro de cada día, con $s_1 = 24$ datos, diaria dentro de cada semana, con $s_2 = 7$ días x 24 horas=168 datos, y mensual dentro del año, con $s_3 = 30$ días x 24 horas=888 datos. Este modelo, suponiendo que cada tipo de dependencia es el mismo en todos los periodos de cada ciclo estacional, puede modelarse como:

$$\Phi_1(B^{s_1})\Phi_2(B^{s_2})\Phi_3(B^{s_3})\phi(B)\nabla_{s_1}\nabla_{s_2}\nabla_{s_3}y_t = \Theta_1(B^{s_1})\Theta_2(B^{s_2})\Theta_3(B^{s_3})\theta(B)a_t.$$

Otra alternativa es utilizar un modelo estructural generalizando [1], por ejemplo para tres estacionalidades, del tipo:

$$y_t = T_t + \sum_{j=1}^3 S_{t,s_j} + I_t$$

La forma más efectiva de modelar muchas estacionalidades es un tema actual de investigación. Véase por ejemplo De Livera, Hyndman y Snyder (2011).

2.2. Estacionalidad multivariante

En contraste con los muchos trabajos sobre la estacionalidad en series univariantes, esta característica ha sido poco estudiada con series multivariantes (véase por ejemplo Reinsel, 1997; Lütkepohl, 2013 o Tsay, 2014). Dado un conjunto de m series con estacionalidad del mismo periodo, s , el correspondiente vector de series y_t de dimensión m con la notación de la sección anterior, puede escribirse de forma estructural, suponiendo el caso mas interesante de no estacionariedad, como:

$$\mathbf{y}_{i+js} = \mathbf{T}_{i+js} + \mathbf{S}_i^{(j)} + \mathbf{n}_{i+js}, \quad [5]$$

o en forma reducida VARMA como:

$$\Phi(B^s)\phi(B)\nabla\nabla_s y_t = \theta(B)\Theta(B^s)a_t \quad [6]$$

que generaliza [4] con operadores matriciales. Esta representación tiene sin embargo dos problemas. El primero es que ahora el orden de los operadores es importante porque las matrices $\Phi(B^s)$ y $\phi(B)$ no conmutan en general con lo que $\Phi(B^s)\phi(B) \neq \phi(B)\Phi(B^s)$. En segundo lugar, es frecuente que no todas las m series necesiten diferencias regulares o estacionales. Es bien conocido, véase por ejemplo Hylleberg (1990) y Buseti (2006), que un vector de series no estacionarias está cointegrado si existe al menos una combinación lineal $\alpha'y_t$ estacionaria.

En estos casos no conviene diferenciar el vector de series porque introducimos términos de media móvil no invertible. Lo mismo ocurre con la diferenciación estacional, ya que las series pueden compartir estacionalidades comunes.

Para ilustrar este último aspecto consideremos la formulación [5]. Si el vector de series está cointegrado y $\alpha' y_{i+j_s}$ es estacionario esto implica que $\alpha' T_{i+j_s}$ y $\alpha' S_i^{(j)}$ deben serlo también. Por ejemplo, supongamos que T_{i+j_s} es una tendencia determinista lineal común, $T_{i+j_s} = a1 + b1t$, donde $1' = (1, \dots, 1)$, y $S_i^{(j)} = 0$. Entonces, los $m-1$ vectores α_j que verifican $\alpha_j' T_{i+j_s} = 0$ definen relaciones de cointegración. Lo mismo ocurre si $T_{i+j_s} = p r_{i+j_s}$, donde r_t es un paseo aleatorio univariante. Los $m-1$ vectores ortogonales al vector p definen relaciones de cointegración. Análogamente, si los procesos $S_i^{(j)}$ no tienen media cero, y las series tienen estacionalidad permanente, pero existe un vector $\delta' S_i^{(j)} = 0$, que conduce a la serie univariante $\delta' y_{i+j_s}$ no estacional diremos que existe cointegración estacional de periodo s . Por ejemplo, supongamos que existen solo dos tipos de estacionalidad en las series de manera que para las $m/2$ primeras (suponemos m par) hay una estacionalidad determinista con coeficientes $c_i^{(1)}$ y para las $m/2$ restantes coeficientes $c_i^{(2)}$. Definiendo los vectores $u_1 = (1, \dots, 1, 0, \dots, 0)$ y $u_2 = (0, \dots, 0, 1, \dots, 1)$, el modelo [5] puede escribirse como:

$$y_{i+j_s} = T_{i+j_s} + u_1 c_i^{(1)} + u_2 c_i^{(2)} + n_{i+j_s}$$

y los $m-2$ vectores δ_j que verifican $\delta_j' u_1 = \delta_j' u_2 = 0$ conducen a series univariantes no estacionales. Tenemos $m-2$ relaciones de cointegración estacional o dos factores comunes estacionales. Lo mismo ocurre con estacionalidades no deterministas. Por ejemplo, si $S_i^{(j)} = b c_i^{(j)} + v_i^{(j)}$, donde (\cdot) es un componente estacional estocástico común y $v_i^{(j)}$ un vector de ruido blanco, las combinaciones ortogonales a b definen $m-1$ relaciones de cointegración estacional.

Una forma alternativa de modelar la cointegración es con modelos factoriales (véase Escribano y Peña, 1994; Peña y Poncela, 2006; y Peña y Tsay, 2021). Con muchas series es más interesante estudiar los factores comunes, que suelen ser pocos, que las relaciones de cointegración, que pueden ser muchas. Esta es una ventaja clara para modelar la estacionalidad multivariante de muchas series con un modelo factorial estacional, como el que presentamos a continuación.

3. UN MODELO FACTORIAL DINÁMICO ESTACIONAL

Muchas series temporales relacionadas suelen compartir varios grupos de estacionalidad, o factores estacionales comunes. Siguiendo el enfoque de Peña y Poncela (2006) de analizar factores de tendencia común, Alonso *et al.* (2011), García-Martos, Rodríguez y Sánchez (2011) y Nieto, Peña y Saboyá (2016) han estudiado modelos factoriales estacionales. El enfoque de estos últimos autores parte del modelo factorial de Peña y Box (1987):

$$y_t = P f_t + e_t, \quad [7]$$

donde y_t es el vector de variables observables de dimensión m , f_t un vector de factores no observables (o latentes) de dimensión r , con $r < m$, P es una matriz de dimensión $m \times r$,

llamada la matriz de pesos del vector de factores f_t y $\{e_t\}$ es una sucesión de vectores idéntica e independientemente distribuidos, con media el vector 0, matriz de varianzas y covarianzas Σ_e y distribución multinormal cada uno. Conceptualmente, el modelo indica que cada variable depende de una combinación lineal de los factores f_{1t}, \dots, f_{rt} y de un ruido intrínscico a esa variable. Los pesos o ponderaciones de los factores en una variable, son las entradas de la fila de P correspondiente a ella. Así, por ejemplo, $y_{1t} = p_{11}f_{1t} + \dots + p_{1r}f_{rt} + e_{1t}$.

Con el fin de extraer los ciclos estacionales comunes del conjunto de variables observables, y además otras características comunes a las variables, como tendencias de largo plazo y movimientos de corto plazo, se supone $f_t = (f'_{1t}, f'_{2t}, f'_{3t})'$, donde f_{1t} es un vector de dimensión r_1 y representa las tendencias comunes (movimientos de largo plazo), f_{2t} es de dimensión r_2 y especifica la estacionalidad común (ciclos) de las variables observadas y f_{3t} de dimensión r_3 representa los factores comunes estacionarios (movimientos comunes de corto plazo). Nótese que $r_1+r_2+r_3 = r$ y con $P = [P_1, P_2, P_3]$, donde P_i es una submatriz de dimensión $m \times r_i$, $i = 1, 2, 3$, se obtiene que,

$$y_t = P_1 f_{1t} + P_2 f_{2t} + P_3 f_{3t} + e_t,$$

la cual es una expresión análoga al modelo de componentes no observables multivariado (Harvey, 1989).

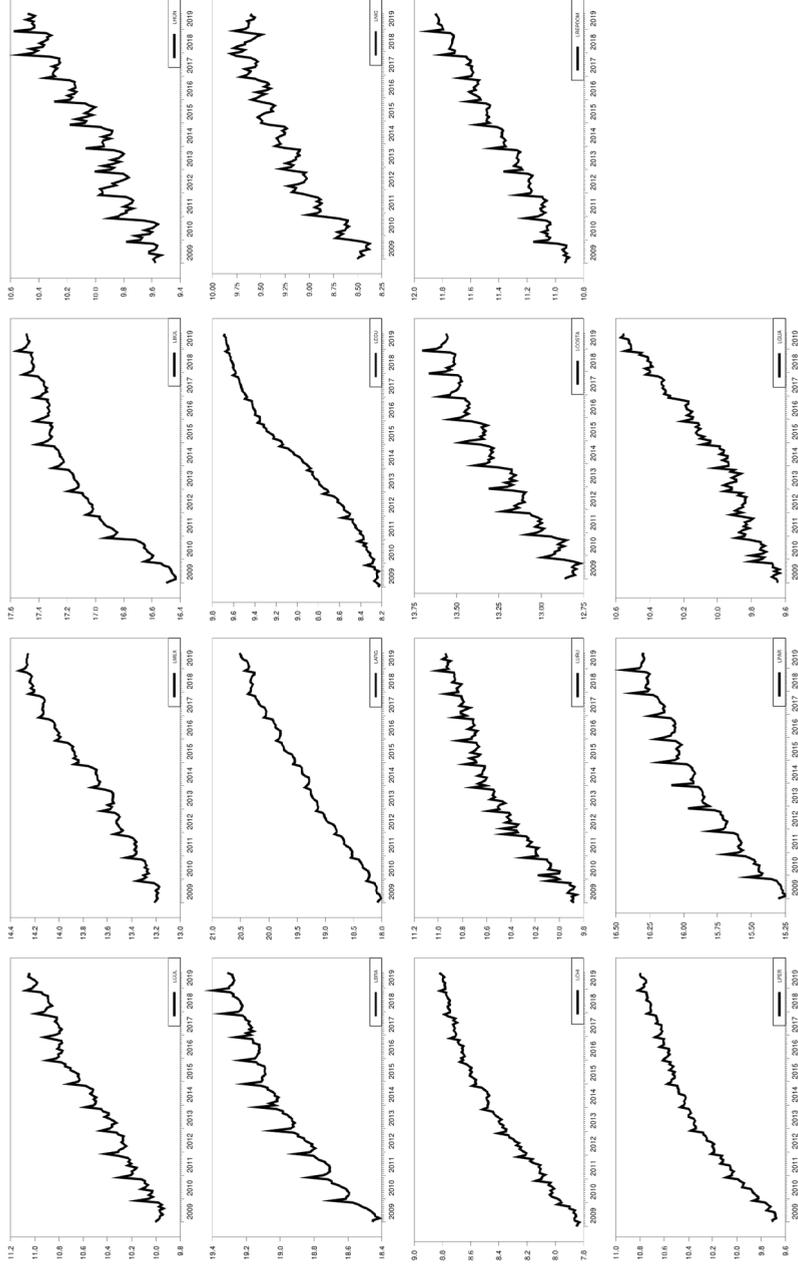
3.1. Un ejemplo de estacionalidad común

Antes de establecer las propiedades principales del modelo factorial estacional de Nieto, Peña y Saboyá (2016), consideramos el siguiente ejemplo empírico, el cual nos permite motivar la aplicación de esa metodología. Las variables de interés corresponden al dinero en efectivo en circulación, de 15 países de América Latina. Estos son: Colombia (1), Brasil (2), Chile (3), Perú (4), México (5), Argentina (6), Uruguay (7), Paraguay (8), Bolivia (9), Ecuador (10), Costa Rica (11), Guatemala (12), Honduras (13), Nicaragua (14) y República Dominicana (15). La fuente de los datos es Bloomberg y el período muestral considerado es enero, 2009-septiembre, 2019, para un total de 129 observaciones por país. En la figura 1 presentamos las gráficas de las series temporales en logaritmos, donde comenzando por Colombia, observamos de arriba hacia abajo columna por columna, en el orden establecido antes para los países. Como podemos observar, hay una tendencia creciente de tipo lineal en algunos países y de tipo curvilíneo en otros, notando que la tendencia de Ecuador difiere un poco de las demás (cambio de concavidad). También, es muy clara la presencia de estacionalidad en cada una de ellas, aunque la estacionalidad de las variables de Argentina y Ecuador no es tan marcada como en los otros países.

Es bien sabido que la inflación induce la tendencia creciente de este agregado monetario, mientras que la estacionalidad tiende a ser muy típica en épocas del año donde la demanda de liquidez de los agentes de la economía, tiende a elevarse por motivo de transacciones de navidad, verano, vacaciones colectivas, etcétera.

Figura 1.

Series temporales en logaritmos del dinero en circulación, en algunos países latinoamericanos



Fuente: Elaboración propia.

A simple vista, se intuye la presencia de, al menos, una tendencia común y un ciclo estacional común, los cuales inducen de alguna manera la tendencia y la estacionalidad de cada serie temporal. La pregunta es: ¿cuántas tendencias comunes y cuántos ciclos estacionales comunes están detrás del origen de los datos? ¿Hay movimientos de corto plazo comunes? ¿Es el número total de factores comunes mucho menor que el número de variables observadas?

Con el propósito de responder las preguntas anteriores, vamos a describir las propiedades del modelo factorial estacional de Nieto, Peña y Saboyá (2016). La idea básica es extraer la información presente en los datos, acerca del número y el tipo de los factores comunes, a partir de la estructura de asociación estadística conjunta de las variables observables. Es bien conocido que la estructura de autocorrelación conjunta de los factores comunes determina aquella de las variables observables (ver, por ejemplo, Peña y Box, 1987).

3.2. Propiedades del modelo factorial estacional

Sea $N = Ls$ el tamaño de muestra, definimos las matrices de autocovarianza muestral generalizadas (CVMG) $C(k, N)$ como:

$$C(k, N) = \frac{s^{2d}}{N^{2d}} \sum_{t=k+1}^N (\mathbf{y}_{t-k} - \bar{\mathbf{y}})(\mathbf{y}_t - \bar{\mathbf{y}})', \quad k = 0, 1, 2, \dots, \quad [8]$$

donde $\bar{\mathbf{y}} = \frac{1}{T} \sum_{t=1}^N \mathbf{y}_t$ y d es el orden de integración de las tendencias estocásticas comunes.

Nótese que $L = N/s$ es el número de ciclos estacionales en la muestra que supondremos un número entero y N múltiplo de s . Esta sucesión de matrices cuantifica la estructura de autocovarianzas conjunta de las variables. Vamos a ver que los valores y vectores propios de estas matrices son informativos sobre la presencia de factores comunes, regulares o estacionales en las series. Cuando N es muy grande y se cumplen ciertas condiciones sobre el modelo factorial estacional se comprueba que los m valores propios de las matrices $C(k, N)$, ordenados de mayor a menor, poseen la siguiente propiedad: (i) si el retardo k es estacional, es decir tiene la forma js , para algún número natural j , entonces los últimos $m - (r_1 + r_2)$ valores propios son muy cercanos a cero. (ii) Si k no es estacional los últimos $m - r_1$ valores propios son también muy cercanos a cero.

Para examinar esta característica del modelo en la práctica, estimamos las matrices $C(k, N)$ con la serie temporal multivariante, hasta un retardo K previamente establecido, y computamos los valores propios de cada una de ellas. Obtenemos las m sucesiones de valores propios, indexadas por el retardo $k = 0, 1, \dots, K$, y observamos el comportamiento de las sucesiones tanto en los retardos estacionales como en los no estacionales. Determinamos entonces un valor para r_1 a partir de los retardos no estacionales y un valor para r_2 a partir de los retardos estacionales, siguiendo las propiedades anotadas antes.

Con base en las propiedades y características del modelo [1], citadas anteriormente, la propuesta metodológica para especificar el número y tipo de factores comunes, en el vector de

variables observables, sigue los siguientes pasos: 1) graficar las sucesiones de valores propios de las matrices CVMG para identificar el número de tendencias comunes r_1 y el número r_2 de ciclos estacionales comunes; 2) por medio de algún contraste o prueba estadística, determinar el número total r de factores comunes. Algunas alternativas son la prueba de Peña y Poncela (2006), extendida por Nieto, Peña y Saboyá (2016) y Bolívar, Nieto y Peña (2020), y los contrastes de Ahn y Horenstein (2013), Lam y Yao (2012) y Caro y Peña (2020); 3) tomar $r_3 = r - (r_1 + r_2)$.

Una vez identificados el número y tipo de factores comunes, esto es, r_1 , r_2 y r_3 , se debe especificar un modelo estadístico ARIMA/SARIMA para los factores comunes. Al hacerlo, los parámetros poblacionales fijos del modelo factorial estacional pueden ser estimados, usando el método de máxima verosimilitud y los factores comunes son estimados usando el denominado suavizador de punto fijo (Harvey, 1989). La idea es transformar el modelo factorial en un modelo de estados (Harvey, 1989; Gómez y Maravall, 1994). En muchos paquetes estadísticos se encuentran códigos para la estimación de un modelo de estados. Sin embargo, cuando la dimensión del vector de variables observables es muy grande (big data), pueden ocurrir problemas computacionales complejos. En ese caso, recomendamos utilizar el método de reducción de la dimensión de un modelo de estados de Jungbaker y Koopman (2015), veáse también Poncela, Ruiz y Miranda (2020).

3.3. Un ejemplo simulado

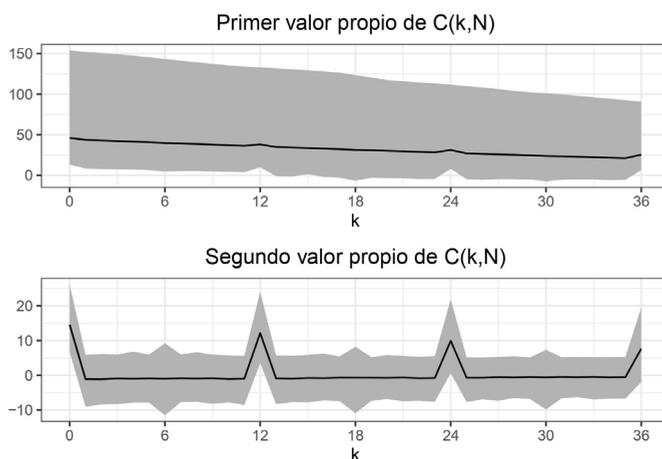
Con el fin de ilustrar el procedimiento propuesto antes, para especificar el número total de factores comunes, el número de tendencias comunes y el número de ciclos estacionales comunes, utilizaremos un modelo factorial estacional con $m = 20$, $r_1 = 1$, $r_2 = 1$ y $r = r_1 + r_2 = 2$. Las entradas de la matriz P son generadas aleatoriamente, a partir de una distribución uniforme sobre el intervalo $[-1,1]$, fijando $p_{1,2} = 0$. Además, se garantiza que P sea ortonormal. La matriz de varianzas y covarianzas de la sucesión de vectores de ruido es $\Sigma_e = I_{20}$, la matriz identidad de orden 20. Para el primer factor, esto es, $f_{11,t}$, consideramos el modelo ARIMA(0,1,1) dado por $f_{11,t} = f_{11,t-1} + a_{11,t} + 0,4a_{11,t-1}$ y para el segundo, es decir, $f_{21,t}$, el modelo estacional puro SARIMA (0,1,1)₁₂ dado por $f_{21,t} = f_{21,t-12} + a_{21,t} + 0,6a_{21,t-12}$. Acá, $\{a_{11,t}\}$ y $\{a_{21,t}\}$ son, cada una, sucesiones de variables aleatorias idéntica e independientemente distribuidas, con media cero, varianza 1 y distribución normal. Este tipo de modelos se encuentra con mucha frecuencia en el análisis de series temporales económicas, ya que determinan el llamado proceso de suavizamiento exponencial simple, un método muy usado para obtener, de manera recurrente, pronósticos de una variable de interés (ver, por ejemplo, Peña, 2005).

El diseño del experimento de simulación es el siguiente: simulamos 1.000 series temporales multivariantes de longitud 300 cada una, luego, para cada serie, calculamos los valores propios de las matrices CVMG para retardos $k = 0,1, \dots, 36$. Obtenemos así 1.000 primeros valores propios, 1.000 segundos valores propios, etc. Calculamos el promedio aritmético y los cuantiles 0.025 y 0.975 de cada muestra de valores propios, de cada retardo k , y construimos las 20 sucesiones de los valores promedio junto con intervalos del 95 % de confianza. En la

figura 2 observamos las sucesiones de los dos primeros valores propios con sus bandas de confianza resaltadas en color gris (que resultan al unir los cuantiles, tanto inferiores como superiores, de cada retardo). Observamos que la banda de confianza del primer valor propio (parte superior), prácticamente no contiene al cero, con lo que podemos validar que el primer valor propio es positivo en todo retardo k . En la parte inferior de la figura 2, observamos que la banda de confianza no contiene al cero en los retardos estacionales (0,12,24,36) mientras que sí lo incluye en los retardos no estacionales. En la figura 3, presentamos la gráfica del tercer valor propio y observamos que, excepto por el retardo $k = 0$, la banda de confianza incluye a cero en todos los demás retardos. Este comportamiento se repite en el resto de sucesiones de valores propios medios. De esta manera, concluimos que en los retardos estacionales hay

Figura 2.

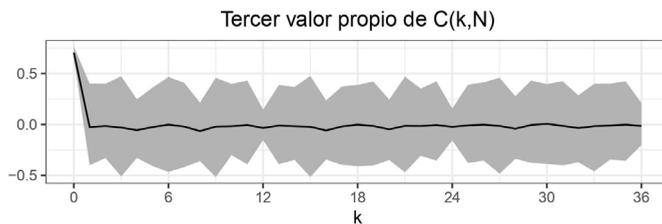
Sucesiones del primer valor propio promedio (arriba) y segundo valor propio promedio (abajo) de las matrices SGCV del ejemplo simulado, con bandas del 95 % de confianza



Fuente: Elaboración propia.

Figura 3.

Sucesión del tercer valor propio medio de las matrices SGCV del ejemplo simulado



Fuente: Elaboración propia.

dos valores propios positivos y en los no estacionales un valor propio positivo, como lo indica el modelo simulado.

4. ANÁLISIS DEL EFECTIVO EN CIRCULACIÓN EN LATINOAMÉRICA

En esta sección aplicamos la metodología del análisis factorial en presencia de estacionalidad, en el grupo de variables determinadas por el efectivo en circulación de 15 países latinoamericanos, el cual fue descrito en la sección anterior. Como es sabido en el contexto económico, el comportamiento de este tipo de variables le permite a la autoridad monetaria, evaluar la liquidez agregada de la economía.

Un análisis exploratorio preliminar de los datos, en donde se aplican pruebas de raíz unitaria de Dickey-Fuller alrededor de tendencia lineal, sugiere la presencia de una tendencia estocástica en la mayoría de las series temporales. Adicionalmente, los correlogramas de la primera diferencia finita de cada serie temporal, exhiben autocorrelaciones significativas al 5 % en los retardos 12, 24 y 36, con un lento decrecimiento. Este hecho indica la necesidad de una diferencia estacional en cada serie temporal. En consecuencia, es plausible considerar que exista, al menos, una tendencia estocástica común y un ciclo estacional común.

Las etapas del análisis son las siguientes:

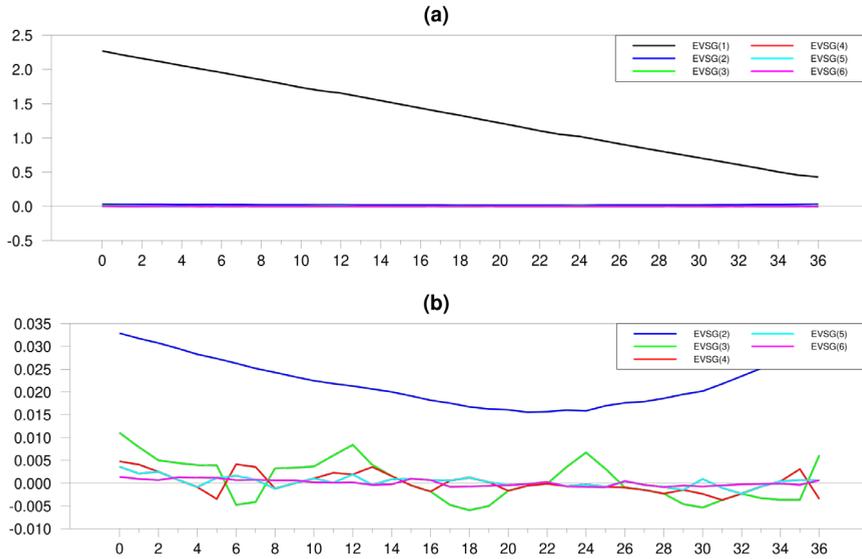
1. Identificación del número y del tipo de factores

Se computan las 15 sucesiones de valores propios de las matrices de covarianza muestral generalizada, para retardos $k = 0, 1, \dots, 36$. Las seis primeras sucesiones se presentan en la figura 4(a) y se observa que la primera sucesión es muy preponderante, en el sentido que sus valores son mucho mayores que aquellos de las restantes sucesiones, en todos los retardos considerados. Esta sucesión señala la presencia de un factor no estacionario y no estacional muy fuerte. En la figura 4(b) presentamos las sucesiones del 2° al 6° valor propio. Podemos observar que la segunda sucesión indica también la presencia de otro factor no estacionario y no estacional, menos fuerte que el anterior. La tercera sucesión señala un factor estacional, pues sus valores en los retardos estacionales son relativamente grandes (son los terceros en magnitud) y en los no estacionales son relativamente pequeños, en comparación con las dos primeras sucesiones. El resto de sucesiones señalan valores muy pequeños en todos los retardos. En consecuencia, podemos especificar que $r_1 = 2$ y $r_2 = 1$.

Con el propósito de identificar el número total de factores r , intentamos utilizar la prueba de Peña y Poncela (2006). Sin embargo, como lo destacan Nieto, Peña y Saboyá (2016), dado que $N/m = 129/15 \approx 8,6 < 20$, no tenemos soporte metodológico para realizarla. De hecho, un intento de su aplicación nos arroja el valor $r = 11$, al nivel 5 %, lo cual, desde el punto de vista de reducción de dimensión, no es un valor razonable. En su lugar, podemos considerar métodos alternativos, como los de Ahn y Horenstein (2013), Lam y Yao (2012) y Caro y Peña (2020). Al hacerlo así, se identifica un solo factor común. Por la concepción y filosofía de esos métodos, podríamos decir que ellos solo identifican el factor no estacionario y no estacional fuerte, el que induce la tendencia creciente de las variables.

Figura 4.

- (a) Primeras seis sucesiones de valores propios de las matrices de CVMG
 (b) Cinco últimas sucesiones (EVSG(i): i-ésimo valor propio)



Fuente: Elaboración propia.

Con el fin de explorar por la presencia de factores comunes adicionales, una práctica usual es corregir los datos originales por estimaciones iniciales de los factores detectados. Para obtener estimaciones iniciales de los factores, podemos proceder de la siguiente manera: con la filosofía del análisis de componentes principales y soportado también por los resultados de Nieto, Peña y Saboyá (2016), utilizamos los tres primeros vectores propios de $C(0, N)$, esto es, los asociados a los tres primeros valores propios, los cuales denotamos \hat{P}_1^0 , \hat{P}_2^0 y \hat{P}_3^0 , respectivamente. Luego, tomamos $f_{11,t}^0 = (\hat{P}_1^0)' \mathbf{y}_t$, $f_{12,t}^0 = (\hat{P}_2^0)' \mathbf{y}_t$ y $f_{21,t}^0 = (\hat{P}_3^0)' \mathbf{y}_t$, como las estimaciones iniciales de los factores. Sea $\mathbf{f}_t^{(0)} = (f_{11,t}^0, f_{12,t}^0, f_{21,t}^0)'$. En la figura 5 podemos observar dichas estimaciones iniciales. Es importante señalar en este punto que los dos primeros vectores propios de las matrices $C(0, N)$, $C(1, N)$ y $C(12, N)$ son muy parecidos, luego el primer par de vectores propios de $C(1, N)$ y $C(12, N)$ pueden ser utilizados también para computar los dos primeros factores comunes preliminares.

Ahora, la corrección por el primer factor común la realizamos a través de la expresión $\hat{\mathbf{y}}_t^{(1)} = \mathbf{y}_t - \hat{P}_1^0 f_{11,t}^0$ y aplicamos entonces los contrastes a la serie temporal $\hat{\mathbf{y}}_t^{(1)}$. Los resultados de este proceso a dos etapas se presentan en la tabla 1, donde las celdas de la tabla indican el número de factores identificado en cada etapa. Como podemos observar, dichos contrastes detectan globalmente un total de dos factores comunes, los cuales corresponderían a los dos factores no estacionarios y no estacionales identificados con los valores propios de las matrices CVMG, pero no detectan el factor común estacional.

La decisión final acerca del número r la tomaremos de la siguiente forma: sea $\hat{y}_t = \hat{P}^{(0)} \hat{f}_t^{(0)}$, donde $\hat{P}^{(0)}$ es la estimación inicial de la matriz P que se obtiene al utilizar como sus columnas, los vectores propios de $C(0, N)$ descritos antes. La serie temporal \hat{y}_t representa los datos ajustados por los tres factores identificados. Al repetir el paso 1 de la metodología propuesta en este capítulo, las sucesiones de valores propios de las matrices CVMG no señalan factores comunes adicionales. Combinando este resultado con aquellos presentados en la tabla 1, podemos concluir que una adecuada especificación del número total de factores comunes es $r = 3$.

Tabla 1.

AH.GR y AH.ER : Ahn y Horenstein (con las dos alternativas de razón de valores propios). LY(K): Lam y Yao con K retardos. CP(K): Caro y Peña con K retardos

<i>Contraste</i>	<i>Datos originales</i>	<i>Ajuste por el factor 1</i>
AH.GR	1	1
AH.ER	1	1
LY(9)	1	1
LY(12)	1	1
CP(9)	1	1
CP(12)	1	1

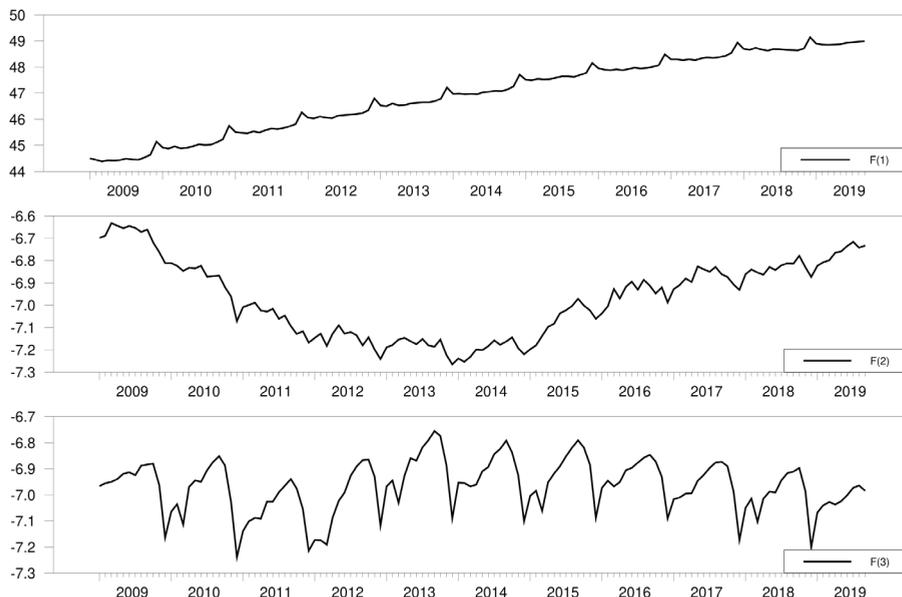
Fuente: Elaboración propia.

2. Identificación de los modelos para los factores comunes

Antes de realizar la estimación de los parámetros fijos del modelo, debemos identificar modelos ARIMA/SARIMA para los tres factores comunes. Con este fin, utilizamos las estimaciones iniciales de los factores como series temporales de los mismos e identificamos sus modelos vía, por ejemplo, el método clásico de Box y Jenkins (1976). Procedimientos automáticos de identificación también pueden ser usados para ese propósito como, por ejemplo, el paquete estadístico TSW (Caporello y Maravall, 2008).

Al hacerlo así, encontramos que los siguientes modelos son adecuados para especificar la dinámica de los factores $f_{11,t}$, $f_{12,t}$ y $f_{21,t}$: $(1-\phi B)(1-B)^2 f_{11,t} = a_{11,t}$, $(1-B)^2 f_{12,t} = a_{12,t}$ y $(1-B^{12}) f_{21,t} = (1+\Theta B^{12}) a_{21,t}$. Nótese que (i) el orden de diferenciación ordinaria especificado es $d = 2$, que concuerda con la observación empírica de una tendencia estocástica alrededor de tendencias de tipo lineal en las variables observables, (ii) el segundo factor común propuesto es una caminata aleatoria integrada, un modelo muy frecuente en análisis de series temporales económicas y (iii) el modelo estacional puro para el tercer factor representa un suavizamiento exponencial simple para las subseries mensuales.

Figura 5.

Factores comunes iniciales del efectivo en circulación

Fuente: Elaboración propia.

3. Estimación del modelo factorial

Para la estimación de los parámetros fijos del modelo factorial propuesto, utilizamos el método de máxima verosimilitud. Por razones de identificabilidad del modelo, restringimos a la matriz P de tal forma que $P'P = I_3$. En la tabla 2 presentamos las estimaciones de las componentes de P , redondeadas a dos cifras decimales, y de las varianzas de los ruidos intrínsecos ($\times 10^4$). Adicionalmente, $\hat{\phi} = -0,57$ y $\hat{\Theta} = -0,06$. En la figura 6 presentamos la gráfica de los pesos de cada factor en las diferentes variables, para una interpretación más cualitativa de los diferentes factores.

Tabla 2.

\hat{P}_i es la estimación final de la columna i de P , $i = 1, 2, 3$; $\hat{\sigma}_i^2$ estimación de la varianza intrínseca del país i , $i = 1, \dots, 15$

País	\hat{P}_1	\hat{P}_2	\hat{P}_3	$10^4 \hat{\sigma}_i^2$
Colombia	0.09	0.05	-0.03	46.9
Brasil	0.07	-0.06	-0.01	49.6

Tabla 2. (continuación)

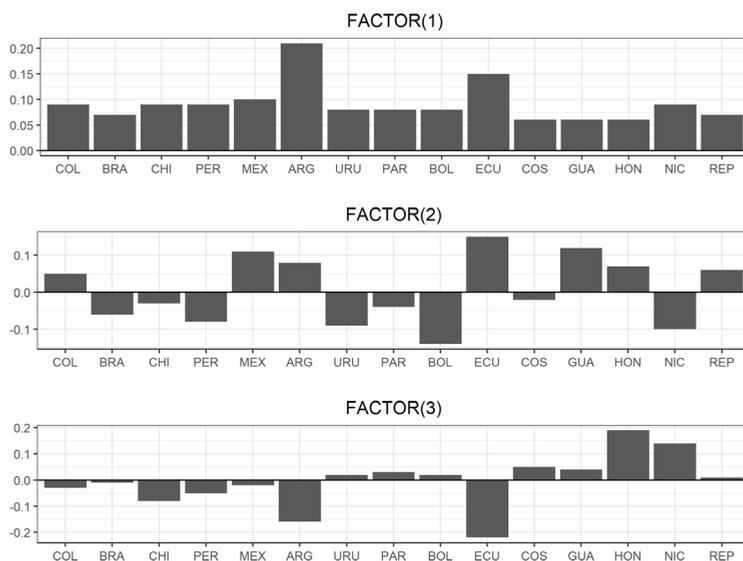
\hat{P}_i es la estimación final de la columna i de P , $i = 1, 2, 3$; $\hat{\sigma}_i^2$ estimación de la varianza intrínseca del país i , $i = 1, \dots, 15$

<i>País</i>	\hat{P}_1	\hat{P}_2	\hat{P}_3	$10^4 \hat{\sigma}_i^2$
Chile	0.09	-0.03	-0.08	14.6
Perú	0.09	-0.08	-0.05	16.3
México	0.10	0.11	-0.02	5.8
Argentina	0.21	0.08	-0.16	16.5
Uruguay	0.08	-0.09	0.02	68.5
Paraguay	0.08	-0.04	0.03	39.6
Bolivia	0.08	-0.14	0.02	14.0
Ecuador	0.15	0.15	-0.22	7.3
Costa Rica	0.06	-0.02	0.05	71.4
Guatemala	0.06	0.12	0.04	71.8
Honduras	0.06	0.07	0.19	9.0
Nicaragua	0.09	-0.10	0.14	133.3
R. Dominicana	0.07	0.06	0.01	58.3

Fuente: Elaboración propia.

Figura 6.

Pesos de cada factor en las 15 variables latinoamericanas



Fuente: Elaboración propia.

Como observamos en la tabla 2, 1ª columna de \hat{P} (o en la parte superior de la figura 6), los pesos del primer factor son todos positivos, indicando que este factor es un promedio ponderado de todas las variables. Además, Argentina tiene un peso relativamente grande en el factor 1, seguido por Ecuador. Estas dos variables indicarían una alta influencia en esta tendencia común. Es conocido en el medio latinoamericano que la economía de Argentina ha padecido de una profunda estanflación en los últimos 10 años, con crecimientos anuales del PIB real de apenas 1,4 % (uno de los más bajos de Latinoamérica) y una alta inflación que se ha mantenido en promedio por encima del 21 %.

Sobre el segundo factor no estacionario y no estacional (segunda columna de \hat{P}), podríamos decir que es un contraste entre el grupo de Brasil, Chile, Perú, Uruguay, Paraguay, Bolivia y, curiosamente, Nicaragua y el resto de países. Bolivia recibe el mayor peso negativo y Ecuador, el mayor positivo.

Observando la tercera columna de \hat{P} (o la parte inferior de la figura 6), podemos decir que el factor común estacional contrasta, esencialmente, al grupo de países centroamericanos (excepto México), con Chile, Argentina y Ecuador. Como podemos observar, Ecuador recibe el mayor peso negativo mientras que Honduras el mayor positivo.

De lo observado en la composición de los tres factores, podemos interpretar que Ecuador tiene un desempeño relativamente importante en este análisis factorial. Cabe notar que la economía de Ecuador está “dolarizada” desde hace 20 años, aproximadamente, y esto ha hecho que el Banco Central pierda la potestad de hacer emisión primaria y controlar la base monetaria de la economía.

Otro hecho interesante lo constituye la varianza del ruido intrínscico de Nicaragua, pues su valor es relativamente grande, en comparación con el resto de países. Así mismo, su peso en los tres factores es relativamente relevante. La economía de Nicaragua se ha contraído a tasas anuales del orden del -4 %, desde 2017 aproximadamente, el déficit fiscal se triplicó hasta niveles del 3 % del PIB, y la deuda pública se disparó hasta máximos históricos del 42 % del PIB. La independencia del Banco Central ha permitido actuar al emisor en función de estabilizar los medios de pago de la economía y el sistema financiero. No obstante, la crisis de los últimos años ha tenido efectos significativos sobre los ciclos naturales de varios de los principales agregados monetarios.

A la luz de los resultados anteriores, el efectivo en circulación en la muestra de 15 países latinoamericanos considerada, es explicado por tres variables comunes subyacentes (latentes). Una, explica la fuerte tendencia de tipo aproximadamente lineal de las series temporales. Otra, explica la presencia de una tendencia estocástica común de tipo caminata aleatoria integrada, que se superpone a la tendencia fuerte de las variables observadas. Y la otra, indica que las variables comparten una estacionalidad común. Al obtener esta reducción de dimensión, emergen características interesantes de las variables consideradas. Por ejemplo, y de alguna forma, el agregado monetario de Argentina, Ecuador y Nicaragua se separa del resto de variables consideradas. No obstante, el agregado monetario de Honduras y Bolivia también tiene cierto desempeño relevante en estas economías.

5. CONCLUSIONES

Con datos masivos, mucho más eficaz que investigar la cointegración estacional es construir un modelo que, además de los factores habituales integrados y estacionarios, pueda incluir factores comunes estacionales. El modelo propuesto por Nieto, Peña y Saboyá (2016) es un buen punto de partida con este objetivo.

Este modelo se ha aplicado para analizar el efectivo en circulación, como un agregado monetario, en una muestra de 15 países de América Latina. La metodología empleada permite ilustrar potencialmente, cómo podemos reducir un conjunto relativamente grande de variables observables, en un conjunto de variables latentes no observables de dimensión menor, en presencia de estacionalidad.

El método utilizado nos facilita un procedimiento numérico y gráfico para detectar los ciclos estacionales comunes, además de eventuales tendencias estocásticas comunes y comovimientos de corto plazo. Además, ha permitido separar los agregados monetarios de Argentina, Ecuador y Nicaragua, cuyas economías tienen características particulares en relación con el resto de economías consideradas.

Referencias

- AHN, S. y HORENSTEIN, A. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3), pp. 1203-1227.
- ALONSO, A. M., RODRÍGUEZ, J., GARCÍA-MARTOS, C. y SÁNCHEZ, M. J. (2011). Seasonal dynamic factor analysis and bootstrap inference: Application to electricity market forecasting. *Technometrics*, 53(2), pp. 137-151.
- BAI, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics*, 122(1), pp. 137-183.
- BAI, J. y NG, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), pp. 191-221.
- BAI, J. y NG, S. (2004). A PANIC attack on unit roots and cointegration. *Econometrica*, 72(4), pp. 1127-1177.
- BELL, W. R. y HILLMER, S. C. (1984). Issues involved with the seasonal adjustment of economic time series. *Journal of Business and Economic Statistics*, 2(4), pp. 291-320.
- BOLÍVAR, S., NIETO, F. H. y PEÑA, D. (2021). On a new procedure for identifying a dynamic common factor model. *Colombian Journal of Statistics*, 44(1), pp. 1-21.
- BOX, G. E. P. y JENKINS, G. M. (1976). *Time Series Analysis Forecasting and Control*. California: Holden-Day.
- BUSETTI, F. (2006). Tests of seasonal integration and cointegration in multivariate unobserved component models. *Journal of Applied Econometrics*, 21(4), pp. 419-438.
- CAMACHO, M., LOVCHA, Y. y PÉREZ-QUIROZ, G. (2015). Can we use seasonally adjusted indicators in dynamic factor models? *Studies in Nonlinear Dynamics and Econometrics*, 19, pp. 377-391.
- CARO, A. y PEÑA, D. (2020). A Test for the Number of Factors in Dynamic Factor Models. *Working Paper Universidad Carlos III de Madrid*.

- CAPORELLO, G. y MARAVALL, A. (2008). *Program TSW, Version Beta 1.0.4*. Banco de España.
- DE LIVERA, A. M., HYNDMAN, R. J. y SNYDER, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496), pp. 1513-1527.
- DIEBOLD, F. X. (2003). Big data dynamic factor models for macroeconomic measurement and forecasting. En: M. DEWATRIPOINT y S. TURNOVSKY (eds.), *Advances in Economics and Econometrics: Theory and Applications* (pp. 115-122). Eighth World Congress of the Econometric Society.
- ESCRIBANO, A. y PEÑA, D. (1994). Cointegration and common factors. *Journal of Time Series Analysis*, 15(6), pp. 577-586.
- FORNI, M., HALLIN, M., LIPPI, M. y REICHLIN, L. (2000). The generalized dynamic factor model: Identification and estimation. *The Review of Economic and Statistics*, 82(4), pp. 540-554.
- GÓMEZ, V. y MARAVALL, A. (1994). Estimation, prediction, and interpolation for nonstationary series with the Kalman Filter. *Journal of the American Statistical Association*, 89(426), pp. 611-624.
- GARCÍA-MARTOS, C., RODRÍGUEZ, J. y SÁNCHEZ, M. J. (2011). Forecasting electricity prices and their volatilities using unobserved components. *Energy Economics*, 33(6), pp. 1227-1239.
- HARVEY, A. C. (1989). *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge: Cambridge University Press.
- HYLLEBERG, S., ENGLE, R., GRANGER, C. W. y YOO, B. (1990). Seasonal Integration and Cointegration. *Journal of Econometrics*, 44(1-2), pp. 215-238.
- JEVONS, W. S. (1862). On the Study of Periodic Commercial Fluctuations. En: H. S. FOXWELL (ed.), *Investigations in Currency and Finance* (pp. 13-118). London: Macmillan.
- JUNGBACKER, B. y KOOPMAN, S. J. (2015). Likelihood-based dynamic factor analysis for measurement and forecasting estimation. *Econometrics Journal*, 18(2), pp. C1-C21.
- KLEIN, J. L. (1997). *Statistical Visions in Time A History of Time Series Analysis 1662-1938*. Cambridge (United Kingdom): Cambridge University Press.
- LAM, C. y YAO, Q. (2012). Factor modeling for high-dimensional time series: inference for the number of factors. *Annals of Statistics*, 40(2), pp. 694-726.
- LUTKEPOHL, H. (2013). *Introduction to Multiple Time Series Analysis*. Springer Science & Business Media.
- MELO, L. F., NIETO, F. H., POSADA, C. E., BETANCOURT, Y. R. y BARÓN, J. D. (2001). Un índice coincidente para la actividad económica de Colombia. *ENSAYOS Sobre Política Económica*, 19(40), pp. 46-88.
- NIETO, F. H., PEÑA, D. y SABOYÁ, D. (2016). Common Seasonality in Multivariate Time Series. *Statistica Sinica*, 26(4), pp. 1389-1410.
- PEÑA, D. (2005). *Análisis de Series Temporales*. Madrid: Alianza Editorial.
- PEÑA, D. y BOX, G. E. P. (1987). Identifying a simplifying structure in time series. *Journal of the American Statistical Association*, 82(399), pp. 836-843.
- PEÑA, D. y PONCELA, P. (2006). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference*, 136(4), pp. 1237-1257.
- PEÑA, D., PONCELA, P. y RUIZ, E. (2021). *Nuevos métodos de predicción económica con datos masivos*. Madrid: Funcas.
- PEÑA, D. y TSAY, R. S. (2021). *Statistical Learning for Big Dependent Data*. Wiley.

- PONCELA, P., RUIZ, E. y MIRANDA, K. (2020). Factor extraction using Kalman filter and smoothing: This is not just another survey. *Working paper 2020-05, Statistics and econometrics, Universidad Carlos III de Madrid*. Madrid.
- REINSEL, G. C. (1997). *Elements of Multivariate Time Series Analysis*. Springer.
- STOCK, J. H. y WATSON, M. W. (2016), Dynamic factor models: A brief retrospective. En: E. HILLEBRAND y S. J. KOOPMAN (eds.), *Dynamic Factor Models* (Advances in Econometrics, Vol. 35). Emerald Group Publishing Limited.
- STOCK, J. H. y WATSON, M. W. (2017). Twenty years of time series econometrics in ten pictures. *Journal of Economic Perspectives*, 31(2), pp. 59-86.
- TSAY, R. S. (2014). *Multivariate Time Series Analysis*. New Jersey: John Wiley & Sons.