

Ajuste estacional de series de tiempo económicas en México

Seasonal Adjustment of Mexican Economic Time Series

Víctor Manuel Guerrero Guzmán,* Jesús López-Pérez** y Francisco de Jesús Corona Villavicencio***

La intención de este trabajo es brindar al usuario de series de tiempo económicas ajustadas por estacionalidad las herramientas que le permitan usarlas de forma adecuada, para poder tomar decisiones mejor informadas. Por ello, se describe la metodología, se enfatizan algunos aspectos de carácter técnico y se brindan algunas recomendaciones que podrían mejorar la calidad del ajuste. Asimismo, se presenta un ejemplo ilustrativo del ajuste estacional de una serie relevante para México. Los resultados aportan elementos al usuario para elegir el paquete *X-13ARIMA-SEATS* como el mejor en la actualidad para realizar este procedimiento. Además, se resalta la importancia de comprender mejor las herramientas que dan soporte a los procesos considerados automáticos para interpretar en forma apropiada los resultados que se obtienen.

Palabras clave: ajuste directo e indirecto; calendario mexicano; días hábiles; observaciones anómalas; Semana Santa; *X-13ARIMA-SEATS*.

The aim of this work is to provide the user of seasonally adjusted economic time series with the tools that allow her/him to interpret those series adequately in order to make better-informed decisions. To that end, we describe the seasonal adjustment methodology, emphasize some technical aspects and make some recommendations that may help improving the quality of the adjustment. We also present an illustrative example of a relevant series for Mexico. The results are useful to understand that the *X-13 package* is the best one we can use nowadays to deseasonalize a time series. Furthermore, we address the need of acquiring a deep knowledge of the tools underlying the so-called automatic processes, so that their output can be better understood.

Key words: direct and indirect adjustment; mexican calendar; trading days; outliers; Easter; *X-13ARIMA-SEATS*.

Recibido: 22 de enero de 2018.

Aceptado: 21 de mayo de 2018.

* Instituto Tecnológico Autónomo de México (ITAM), guerrero@itam.mx

** Instituto Nacional de Estadística y Geografía (INEGI), jesus.lopezp@inegi.org.mx

*** INEGI, franciscoj.corona@inegi.org.mx

Nota: los autores agradecen los comentarios y sugerencias de cuatro dictaminadores anónimos que contribuyeron a mejorar la presentación de este documento; este trabajo se pudo llevar a cabo gracias al apoyo de Enrique Ordaz, Gerardo Leyva y Yuriko Yabuta, funcionarios del INEGI; de igual forma, Juan Martínez y Laura Montoya facilitaron la realización del proyecto al poner al alcance de los autores algunos documentos y archivos de datos que les fueron solicitados. La participación de Víctor M. Guerrero en este proyecto fue posible gracias a un periodo sabático otorgado por el Instituto Tecnológico Autónomo de México (ITAM) y a la Cátedra de Análisis de Series de Tiempo y Pronóstico en Econometría, asignada por la Asociación Mexicana de Cultura, AC.



Introducción

El documento tiene como objetivo presentar algunos aspectos que se consideran relevantes en relación con el ajuste estacional de series de tiempo económicas, también conocido como *desestacionalización*, que se realiza en México y, en particular, en el Instituto Nacional de Estadística y Geografía (INEGI) de manera regular y sistemática desde la década de los 90, en concordancia con lo que se hace en las agencias oficiales de estadística (AOE) a nivel mundial.

Las series de tiempo económicas se desestacionalizan porque es más fácil entender el comportamiento de las variables sin el efecto que ocasionan los fenómenos estacionales, los cuales generan movimientos de naturaleza no económica que influyen sobre el comportamiento de la serie. La información que aquéllas proveen debe ser útil para realizar análisis económicos sin la influencia de efectos estacionales que tienden a ser regulares y, por lo mismo, pueden anticiparse y cancelarse en forma relativamente sencilla. De hecho, los analistas de información macroeconómica ya están acostumbrados a recibirla de manera desestacionalizada, sin embargo, también los especialistas pueden beneficiarse de un mejor entendimiento de la metodología de desestacionalización.

Los avances de los métodos computacionales han hecho tan ágiles los procesos que llegan a ser considerados como *cajas negras*, aun por algunos usuarios expertos. En México, es común que la prensa y algunos analistas malinterpreten las cifras desestacionalizadas, lo cual puede originar escepticismo, por ejemplo, acerca de la discrepancia entre el ajuste directo de una serie agregada y el indirecto que se obtiene al ajustar por separado a las componentes que forman dicha serie, o suspicacia sobre el impacto de los días hábiles y su interacción con los feriados del país. Por ello, es conveniente hacer un recuento del surgimiento de esta metodología en México, de los métodos estadísticos que respaldan el ajuste estacional y sus alcances, así como de la interpretación de los resultados de éste, además de otros temas específicos, como es la elección del método (directo o indirecto).

La desestacionalización de series de tiempo empezó a aplicarse de manera masiva en el *Census Bureau* de Estados Unidos de América (EE.UU.) en la década de los 50 con el *Census Method II*, variante *X-11*; posteriormente, se popularizó a nivel mundial el uso del paquete llamado *X-11*, cuyo enfoque sigue vigente en la actualidad; Ladiray y Quenneville (2000-2001) presentaron la cronología de dicho enfoque y en el país fue en 1983 cuando el Banco de México (BM) puso en marcha un proyecto formal y sistemático, según se relata en documentos de trabajo publicados años después (ver Guerrero, 1983a, 1983b, 1990 y 1992). En la actualidad, la desestacionalización oficial de series en México se realiza de manera consensuada entre el INEGI, el BM y las secretarías de Hacienda y Crédito Público y de Economía con el paquete *X-13ARIMA-SEATS* (denotado *X-13* de aquí en adelante), por lo que un grupo de funcionarios de estas instituciones se reúnen de forma regular para establecer acuerdos sobre la aplicación más adecuada de la metodología.

Lo que resta de este documento está organizado como sigue: la segunda sección introduce la notación y presenta el estado del arte del ajuste estacional; en la siguiente se hace énfasis en los ajustes previos a la desestacionalización; la cuarta muestra un resumen de los pasos a seguir para realizar un ajuste estacional; la quinta sección detalla la aplicación de los métodos descritos para el caso del indicador global de la actividad económica (IGAE); y en la última se elaboran algunas conclusiones y recomendaciones que subrayan aspectos de particular importancia acerca de los resultados obtenidos.

Estado del arte del ajuste estacional

El *Census Bureau* utiliza como definición de estacionalidad la componente de la serie que se asocia con las "...fluctuaciones periódicas que ocurren cada año, aproximadamente en las mismas fechas y con la misma intensidad, y lo más importante es que pueden medirse y eliminarse de la serie de tiempo en consideración..." (Guerrero, 1990). Además, la estacionalidad es causada por fuerzas no-económicas que quienes toman decisiones no pueden controlar; de hecho, Granger (1978) señala que hay al menos cuatro causas de las fluctuaciones estacionales: 1) el calendario mismo, 2) las instituciones que fijan fechas para realizar ciertas actividades, 3) el clima y 4) las expectativas de fluctuaciones estacionales.

Con el fin de realizar un ajuste estacional, se considera que la serie de tiempo que será sometida al proceso, $\{Y_t\}$, es mensual o trimestral y está compuesta por los siguientes componentes principales: la tendencia-ciclo $\{T_t\}$, que incluye la tendencia de largo plazo y los movimientos cíclicos; la estacionalidad, $\{E_t\}$, que se supone es fundamentalmente estable aunque puede mostrar ligera movilidad con el paso del tiempo; y la irregularidad, $\{I_t\}$, que a su vez incluye los efectos de observaciones anómalas.

Esta manera de caracterizarlos concuerda con Bloem *et al.* (2001), donde además se menciona que la componente irregular debe ser considerada como el error aleatorio que aparece en los modelos estadísticos. Al hablar de una serie de tiempo se usa la notación $\{Y_t\}$ para indicar que se trata de un conjunto de variables observadas en los momentos $t = 1, \dots, N$, o sea, $\{Y_1, \dots, Y_N\}$, mientras que Y_t corresponde a una de tales observaciones.

Las formas que por lo general se adoptan para ligar las componentes en cada momento de observación $t = 1, \dots, N$, son la descomposición aditiva:

$$Y_t = T_t + E_t + I_t \quad (1)$$

y la multiplicativa:

$$Y_t = T_t \times E_t \times I_t \quad (2)$$

De forma adicional, el paquete *X-11-ARIMA* (ver Dagum, 1980) incluyó la opción *log-aditiva*:

$$\log(Y_t) = \log(T_t) + \log(E_t) + \log(I_t) \quad (3)$$

lo cual equivale a la descomposición multiplicativa después de aplicar logaritmos en ambos lados de la ecuación (2). La diferencia entre las expresiones (2) y (3) no radica en la forma de expresión, sino en la manera como el programa *X-11-ARIMA* efectúa la estimación de las componentes, mediante promedios aritméticos móviles en (2) —al igual que en (1)— y con promedios geométricos móviles en (3) (ver Guerrero, 1990 para más detalle). Asimismo, en la literatura sobre el tema se han considerado otras opciones híbridas, por ejemplo, del tipo señalado por Ladiray y Quenneville (2000-2001):

$$Y_t = T_t \times (1 + E_t) \times (1 + I_t) \quad (4)$$

o del considerado pseudoaditivo en el paquete *X-12-ARIMA* (Findley *et al.*, 1998):

$$Y_t = T_t \times (E_t + I_t - 1) \quad (5)$$

que es útil en particular para series que contienen valores iguales a cero.

El uso del modelo aditivo implica que la estacionalidad se considera estable e independiente de la componente de tendencia durante todo el periodo de estudio de la serie, mientras que al usar las opciones (2) o (3), el supuesto implícito es que la estacionalidad depende del nivel que tenga la tendencia. Este último supuesto es intuitivamente más razonable que el de independencia entre estacionalidad y tendencia debido al comportamiento que muestra la mayoría de las series económicas, por lo que el modelo multiplicativo tiende a ser usado con mayor frecuencia.

Es importante recalcar que en el proceso de desestacionalización es transcendental seleccionar de manera correcta el tipo de transformación que se realizará a la serie; esto no es trivial, pues depende de supuestos distribucionales que deben ser validados por los datos; para una discusión sobre el tema, ver Guerrero (1993). Ahora, dentro del entorno del *X-13*, dicha transformación depende directamente de la determinación de si se efectuará una desestacionalización de tipo aditivo o multiplicativo, es decir, si no se realizará transformación a la serie o si se usará el logaritmo. Esta discusión no será

abordada a detalle aquí, ya que representa una línea de investigación muy importante a retomar en el futuro y el objetivo del trabajo no trata de esclarecer este punto. No obstante, es importante comentar que dentro de la especificación del modelo de series de tiempo de tipo autorregresivo integrado y de promedios móviles (ARIMA, por sus siglas en inglés) como fase del proceso de desestacionalización, la comparación entre dos modelos iguales cuando la variable dependiente es diferente debe ser tratada con cautela, pues los criterios de información que se usan dependen del supuesto distribucional de los datos y para modelos con la misma forma funcional. De hecho, la comparación de modelos utilizando criterios de información permite analizar el intercambio existente entre disminuir la suma de los residuos al cuadrado y penalizar por agregar más parámetros al modelo, pero no sirve para seleccionar la forma funcional entre modelos similares.

Las técnicas para cancelar el efecto estacional son variadas (tanto por sus enfoques como por su complejidad de cálculo) y van desde la simple comparación de tasas de crecimiento anual (que cancelan los efectos estacionales estables), hasta procedimientos iterativos del tipo del *X-11*, que requieren forzosamente del uso de computadoras, según lo señalan Lefrancois y Mamay (2004). Los enfoques más empleados en las AOE son básicamente de dos tipos: 1) el basado en la estimación del modelo ARIMA, seguido de la estimación de los factores estacionales por extracción de señal; esto es lo que hace el programa *TRAMO-SEATS* (*Time Series Regression with ARIMA Noise, Missing observations and Outliers-Signal Extraction in ARIMA Time Series*) que se usa en varias AOE europeas, en especial en el Banco de España, donde fue desarrollado y 2) el que aplica filtros de promedios móviles de manera recursiva para estimar las componentes de la serie, sin hacer referencia explícita a un modelo formal de series de tiempo (ver Wright, 2013). Este último se usa en las AOE norteamericanas, como el *Census Bureau* y el *Bureau of Labor Statistics* de EE.UU., en *Statistics Canada* y en el INEGI.

En este trabajo se hace referencia a modelos ARIMA como se realiza por tradición en la literatura de series de tiempo, donde se entiende que la serie, posiblemente diferenciada hasta que ésta sea estacionaria, puede ser modelada a través de sus retrasos y, también, por valores pasados de su respectivo término de error. Aunque el programa *X-11-ARIMA* menciona un modelo ARIMA, éste solo se usa para extender los dos extremos de la serie, sin que se utilice para realizar la desestacionalización en sí (ver Ladiray y Quenneville, 2000-2001). De igual forma, el *X-12-ARIMA* introdujo mejoras sustanciales en la etapa de los ajustes previos mediante un modelo llamado *regARIMA*, que se aplica antes de realizar la desestacionalización. El tema de los ajustes previos requiere atención especial, según lo señala Guerrero (1992) y, por ello, será tratado a detalle en la sección siguiente de este documento.

Al tener acceso a dos opciones metodológicas (*X-12-ARIMA* y *TRAMO-SEATS*) se presenta el dilema de tener que decidir la que sea preferible para un cierto tipo de series; esto ha conducido a realizar investigaciones específicas dentro de algunas AOE, como se indica en Tiller *et al.* (2007). Incluso, algunos autores (e. g. Hylleberg, 2006) recomiendan aplicar los dos métodos de ajuste estacional a la misma serie y decidir cuál ajuste es mejor de forma empírica, ya sea con herramientas estadísticas formales o con la mera inspección visual de los resultados obtenidos. Por ello, es útil el programa *Demetra* elaborado en Eurostat (la AOE de la Unión Europea) para realizar el ajuste con ambos métodos, el cual evolucionó hasta el actual *JDemetra+* (ver Grudkowska, 2015) que en febrero del 2015 fue recomendado por Eurostat y el Banco Central Europeo como el paquete a utilizar para realizar ajuste estacional de estadísticas oficiales en la Unión Europea.

Otra ruta para hacer uso eficiente de las dos metodologías es la que se siguió para producir el paquete *X-13* (ver US Census Bureau, 2016), en el cual se busca integrar las ventajas de cada una de ellas y evitar sus posibles desventajas, de manera que constituye una opción unificadora que no requiere decidir cuál de las dos es preferible. Este paquete se encuentra ya en uso en toda América del Norte desde el 2017 y, por ello, el resto de este documento se enfoca en su utilización. De manera adicional, existen otras metodologías para desestacionalizar series de tiempo, como la desestacionalización temporal de series de tiempo por *Loess* (Cleveland *et al.*, 1990).

Ajustes previos

Éstos deben realizarse porque la desestacionalización se basa en la aplicación repetida de promedios móviles, es decir, de filtros lineales que son sensibles a la presencia de efectos ajenos a la estacionalidad. La etapa de ajustes previos considera efectos atribuibles a la estacionalidad estable (EE), al número de días hábiles (TD, *Trading Days*) que cambian de un año a otro para el mismo mes o trimestre, a otros efectos del calendario (OC), como los de año bisiesto y la longitud de los meses, además de las festividades movibles, como la Semana Santa (SS), la cual ocurre algunos años en marzo y otros en abril, de manera que para la descomposición aditiva se obtiene la igualdad:

$$E_t = EE_t + TD_t + OC_t + SS_t \quad (6)$$

y expresiones semejantes se pueden obtener para las descomposiciones (2) a (5).

Para el caso del efecto irregular, éste contiene los efectos cuya causa y aparición son impredecibles y que se consideran como errores aleatorios (IA) con características deseables, por ejemplo, media 0 en el modelo aditivo y media 1 en el multiplicativo. También, contempla el efecto de observaciones anómalas (A), que pueden tener distintas causas, como equivocaciones en el registro de los datos o intervenciones cuyas causas y fechas de aparición de efectos son conocidas. Las formas que pueden tomar estos efectos son: aditivos (AO, *Additive Outlier*), que tienen efecto momentáneo sobre la serie y su presencia es capturada con una función de pulso; de cambio permanente de nivel (LS, *Level Shift*), cuyo efecto se mantiene una vez que aparecen y se representan mediante una función de escalón; de cambio transitorio (TC, *Transitory Change*), que una vez que ocurren muestran un patrón de absorción en la serie y desaparecen a la larga; y de efecto rampa (RP, *Ramp*), que se acumulan en forma de línea recta y producen un cambio permanente después de varios periodos. De esta manera se obtienen las igualdades:

$$I_t = IA_t + A_t = IA_t + (AO_t + LS_t + TC_t + RP_t), \quad (7)$$

que son válidas en específico para el modelo aditivo, aunque expresiones semejantes para los casos de los modelos (2) a (5) pueden obtenerse sin mayor dificultad.

Para estimar los efectos de los casos mencionados (TD, SS, OC, AO, LS, TC y RP), se hace uso del modelo *regARIMA*, incluido en el paquete *X-13*, que se expresa como:

$$Y_t = \mu_t + Z_t \quad (8)$$

donde Y_t es la variable que da origen a la serie por desestacionalizar, tal vez expresada en logaritmos o con alguna transformación potencia del tipo *Box-Cox* (ver Guerrero, 1993). Cabe señalar que el paquete *X-13* recomienda utilizar el criterio de información de Akaike —corregido— (AICC, por sus siglas en inglés) para decidir de manera automática entre utilizar los datos de la variable con o sin transformación logarítmica, aunque esto no es del todo apropiado. De hecho, el criterio AICC indica qué tan adecuada es la distribución normal para representar la distribución verdadera de los datos. Entonces, al utilizar una transformación logarítmica se supone de forma implícita que la variable original tiene distribución *Log-normal* y si no se usa el logaritmo, el supuesto es que la variable original tiene distribución normal. Como la variable original y la variable con transformación logarítmica no pueden tener ambas una distribución normal, calcular el AICC con y sin transformación logarítmica no es sensato. Es necesario que la selección de la transformación de la variable se acompañe de un minucioso análisis gráfico o de la aplicación de otro método, como el que propone Guerrero (1993).

La media de la variable Y_t es:

$$\mu_t = E(Y_t) = \sum_{i=1}^r \beta_i x_{it} \quad (9)$$

la cual está explicada por un modelo de regresión con efectos fijos asociados a $r \geq 1$ variables regresoras x_{1t}, \dots, x_{rt} , del tipo considerado necesario para capturar los efectos de ajustes previos recién mencionados. Por su lado, la variable aleatoria Z_t se supone que tiene media 0 y su comportamiento se representa mediante un modelo ARIMA multiplicativo estacional (ver Guerrero, 2009, para una introducción al uso de estos modelos), o sea:

$$\phi(B) \Phi(B^{12}) \nabla^d \nabla_{12}^D Z_t = \theta(B) \Theta(B^{12}) a_t \quad (10)$$

donde B es el operador de retraso tal que $BZ_t = Z_{t-1}$; $\nabla = 1 - B$ es el operador de diferencia regular; $\nabla_{12} = 1 - B^{12}$ es el operador de diferencia estacional, con periodo estacional 12; $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ es un polinomio autorregresivo regular de orden p ; $\Phi(B) = 1 - \Phi_1 B - \dots - \Phi_p B^{12p}$ es un polinomio autorregresivo estacional de orden P ; $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ es un polinomio de promedios móviles regular de orden q ; $\Theta(B) = 1 - \Theta_1 B - \dots - \Theta_Q B^{12Q}$ es un polinomio de promedios móviles estacional de orden Q ; $\{a_t\}$ son variables aleatorias independientes e idénticamente distribuidas como una normal con media 0 y varianza σ^2 ; d y D son los órdenes de diferencia regular y estacional, respectivamente. Los parámetros β_i , $i = 1, \dots, r$ del modelo de regresión, así como los ϕ , Φ , θ , Θ y σ^2 del modelo ARIMA son estimados de manera simultánea por el método de máxima verosimilitud.

Ajuste por días hábiles

Éste es necesario en particular para series de tiempo mensuales que se obtienen como agregados de flujos diarios, en las que puede haber cierta periodicidad para los diferentes días de la semana, el cual se transmite a la serie agregada mensual; por ejemplo, podría suceder que la actividad económica de un mes que tenga un domingo más o uno menos que lo usual se vea afectada de manera significativa y que algún sector económico lo resienta más que otro. Es importante señalar que en este trabajo se considera el término *días hábiles* como sinónimo de frecuencia de *días calendario*; lo anterior radica en que los días calendario menos los días festivos fijos, *i.e.*, días hábiles, están correlacionados con la composición misma del calendario. Siguiendo a Young (1965):

“...cada vez que la Navidad cae en lunes, diciembre empieza en viernes y contiene cinco viernes, sábados y domingos...”. En consecuencia, para tener en cuenta estos efectos, es relevante saber que hay 22 tipos de meses dependiendo del tipo de días de la semana que tengan (ver cuadro 1).

En el cuadro se observa que existe un patrón en el calendario que es repetitivo a lo largo del tiempo. Esto es lo que se trata de aprovechar al definir seis variables regresoras como diferencias del número de veces que en el mes aparece cada día de la semana, menos el número de domingos (a lo cual se le conoce como contrastes), o sea, se define para cada mes t , $d_{1t}=(\#Lunes)_t, \dots, d_{7t}=(\#Domingos)_t$, con # que denota el número, entonces:

$$x_{1t} = d_{1t} - d_{7t}, \dots, x_{6t} = d_{6t} - d_{7t} \quad (11)$$

mientras que una séptima variable regresora se utiliza para tener en cuenta la longitud del mes, la cual se considera explícitamente en el apartado siguiente, o sea:

$$x_{7t} = (Longitud\ del\ mes)_t \quad (12)$$

Al estimar los parámetros β de la ecuación (9) se estiman de manera indirecta los efectos de cada día de la semana.

Ajuste por año bisiesto y longitud de los meses (otros efectos de calendario)

La longitud de los meses podría considerarse como un efecto estacional estable, ya que no cambia de un año a otro, excepto en febrero cuando es año bisiesto. De esta forma se tiene que $\nabla_{12} x_{7t} = 0$ excepto cuando la diferencia involucra a un febrero de año bisiesto, en cuyo caso se tiene que

Cuadro 1

Clasificación de los meses en función de los días de la semana que contienen

Categoría	Días del mes	Día inicial del mes	Días que ocurren cinco veces en el mes
1	31	Lunes	Lunes, martes y miércoles
2	31	Martes	Martes, miércoles y jueves
...
7	31	Domingo	Domingo, lunes y martes
8	30	Lunes	Lunes y martes
9	30	Martes	Martes y miércoles
...
14	30	Domingo	Domingo y lunes
15	29	Lunes	Lunes
16	29	Martes	Martes
...
21	29	Domingo	Domingo
22	28	Cualquiera	Ninguno

Fuente: elaboración propia con base en Guerrero (1983b) siguiendo a Young (1965).

$\nabla_{12}x_{7t}=-1$ si t corresponde a febrero de año bisiesto y $\nabla_{12}x_{7t}=1$ si $t - 12$ corresponde a febrero de año bisiesto. Debido a que los años bisiestos ocurren cada cuatro años, se define la variable *Leap* (LP) que toma los valores:

$$LP_t = \begin{cases} 0.75 & \text{si } t \text{ es febrero de año bisiesto} \\ -0.25 & \text{si } t \text{ es febrero de año no bisiesto} \\ 0 & \text{si } t \text{ no corresponde a febrero} \end{cases} \quad (13)$$

Lo importante de esta definición es que las variables desestacionalizadas son iguales, es decir, $\nabla_{12}x_{7t}=\nabla_{12}LP_t$, por lo cual, en modelos que hagan uso de la diferencia estacional, el utilizar x_{7t} o LP_t produce los mismos resultados; pero esta última variable es más fácil de construir e interpretar, por lo que se prefiere su uso en la práctica. Una manera equivalente de tener en cuenta el efecto de año bisiesto, en el caso de la descomposición multiplicativa, consiste en ponderar los datos de febrero aplicando el factor $28.25/x_{7t}$ cuando $t = \text{febrero}$, donde 28.25 es la longitud promedio de ese mes. Para una serie de saldos, el uso de las variables anteriores no es apropiado y, en su lugar, se deben utilizar las variables definidas en el manual del paquete *X-13* (ver US Census Bureau, 2016, p. 32).

Ajuste previo por SS (festividad movable)

Para tener en cuenta el efecto de Semana Santa, se usa el hecho de que esta festividad religiosa está fija dentro del calendario lunar y, por ello, no es regular en el calendario gregoriano que comúnmente se utiliza. Se fija de acuerdo con la regla de que el Domingo de Pascua (o de Resurrección) debe ocurrir después del primer plenilunio que sigue al equinoccio de primavera; de esta manera, se sabe que acaece entre el 22 de marzo, en caso de que el 21 fuese sábado de plenilunio (la última vez que esto ocurrió fue en 1818 y la siguiente será en 2285, según Ladiray y Quenneville, 2000-2001) y el 25 de abril, como caso extremo en el que se fija esta fecha para evitar la coincidencia con la pascua judía. Al calcular cuántas veces el Domingo de Pascua ocurre en marzo y cuántas en abril de 1900 al 2099, se encuentra que en 78% de las veces sucede en abril, de forma que lo extraño es que ocurra en marzo y, por eso, el efecto de Semana Santa se mide en referencia a que haya ocurrido en ese mes.

Por otro lado, mientras más tardía sea la aparición del Domingo de Pascua en abril, menor será el efecto de adelanto de actividad económica en marzo, al grado de que el efecto de esta festividad tiende a desaparecer de marzo y a formar parte entonces de la estacionalidad de abril, lo cual fue verificado de forma empírica por Dagum y Cholette (2006, p. 35). Esto hace que el efecto de Semana Santa se mida en función del número de días previos al Domingo de Pascua que pudo haber afectado su ocurrencia; desde luego, no pueden ser más de 25 días debido a la fecha más tardía en que acaece éste. Así se define la variable *Easter* para flujos, dada por la expresión:

$$E(w)_t = (\#Días anteriores a Pascua en el mes t)/w, \quad (14)$$

que será 0 en todos los meses, excepto en abril, marzo y febrero (en este mes solo si $w > 22$). Debe subrayarse que la dinámica de los efectos de Semana Santa se debe decidir en forma empírica y que es difícil estimar adecuadamente tales efectos, porque aparecen solo una vez al año, de manera que hay pocas observaciones involucradas para hacer la estimación.

Ajuste por observaciones anómalas (*outliers*) e intervenciones

En ocasiones se presentan efectos asociados con eventos que afectan al valor de la variable en un mes en particular y cuya causa puede o no ser conocida. Si lo es, se habla de un efecto de intervención y si no, se le considera como una atipicidad en los datos. Para los fines de limpiar la serie de efectos indeseables, que básicamente distorsionan la dinámica de ésta, no importa si se conoce o no la causa de tales efectos. Éstos pueden ser de los siguientes tipos:

$$AO(t_0)_t = \begin{cases} 1 & \text{si } t = t_0 \\ 0 & \text{si } t \neq t_0 \end{cases} \quad (15)$$

que corresponde a una función de pulso que cambia en el mes t_0 en el que aparecen los efectos;

$$LS(t_0)_t = \begin{cases} -1 & \text{si } t < t_0 \\ 0 & \text{si } t \geq t_0 \end{cases} \quad (16)$$

que señala un escalón en el mes t_0 ;

$$TC(t_0, \alpha)_t = \begin{cases} 0 & \text{si } t < t_0 \\ \alpha^{t-t_0} & \text{si } t \geq t_0 \end{cases} \quad (17)$$

que muestra un decaimiento de magnitud $0 < \alpha < 1$, hacia el nivel previo al del mes t_0 ; y:

$$RP(t_0, t_1)_t = \begin{cases} t_0 - t_1 & \text{si } t \leq t_0 \\ t - t_1 & \text{si } t_0 < t < t_1 \\ 0 & \text{si } t \geq t_1 \end{cases} \quad (18)$$

que toma la forma de una rampa entre los meses t_0 y t_1 .

En este contexto, la determinación de qué valores son anómalos se relaciona con los de los residuos provenientes del modelo ARIMA que sobrepasan los umbrales dados en, por ejemplo, US Census Bureau (2016). Después de tener en cuenta los ajustes previos a la serie, se aplica la desestacionalización con la descomposición que se haya decidido emplear; acto seguido, se decide la subrutina computacional que se desea utilizar, ya sea la del *X-11* o la de *SEATS* (o ninguna, si solo se aplican los ajustes previos). En ambos casos, se usan filtros de promedios móviles para efectuar la estimación de las componentes de la serie; la diferencia estriba en que los filtros en el primer caso ya están predefinidos y el usuario decide cuáles emplear, de manera que este enfoque se puede considerar no-paramétrico, mientras que en el segundo los filtros surgen de los modelos para la tendencia-ciclo, estacionalidad e irregularidad, que se deducen del modelo ARIMA.

Al tener estimados los efectos de la estacionalidad, se obtiene la serie estimada $\{\hat{E}_t\}$ y se deduce entonces la serie ajustada por estacionalidad como:

$$AE(Y_t) = Y_t - \hat{E}_t = (\widehat{T}_t + \widehat{I}_t) \quad (19)$$

en el caso de la descomposición aditiva, y como:

$$AE(Y_t) = Y_t / \hat{E}_t = (\widehat{T}_t \times \widehat{I}_t) \quad (20)$$

en la descomposición multiplicativa, con expresiones semejantes para otro tipo de descomposición. De manera similar, si se aplican ajustes previos a la serie, lo que se obtiene es una serie nueva que en la descomposición aditiva será del tipo:

$$AC(Y_t) = Y_t - \widehat{TD}_t - \widehat{OC}_t - \widehat{SS}_t = (T_t + \widehat{EE}_t + I_t), \quad (21)$$

mientras que en la descomposición multiplicativa se obtendrá:

$$AC(Y_t) = Y_t / (\widehat{TD}_t \times \widehat{OC}_t \times \widehat{SS}_t) = (T_t \times \widehat{EE}_t \times I_t), \quad (22)$$

donde el ajuste se refiere a los efectos del calendario (que típicamente incluyen efectos por días hábiles y por Semana Santa).

Ajuste directo vs. ajuste indirecto

En la práctica del ajuste estacional se presenta una situación disyuntiva cuando se consideran conjuntos de series que están ligadas contablemente con otras, que surgen como agregados de las primeras —digamos como sumas o promedios— y entonces puede hacerse la desestacionalización de manera directa o indirecta; es decir, se podría ajustar por estacionalidad la serie agregada de forma independiente de sus componentes o hacer el ajuste de las series que constituyen el agregado individualmente y agregar —por suma o promedio— las series desestacionalizadas para obtener la agregada desestacionalizada de manera indirecta.

Las discrepancias surgen sobre todo debido a que los patrones estacionales de cada una de las series componentes y del agregado suelen ser distintos. Adicionalmente, puede ser que el modelo para la descomposición sea aditivo para unas series y multiplicativo para otras. De esta forma, no se espera que los resultados directos e indirectos sean iguales y, para decidir entre las dos opciones, conviene conocer las características más sobresalientes del ajuste indirecto, que son: 1) las series que componen el agregado se desestacionalizan individualmente, por lo que se pueden ajustar por estacionalidad de la mejor manera posible; 2) las relaciones que ligan a las series componentes con el agregado se mantienen al agregar las series desestacionalizadas; 3) no se tienen en cuenta las características estacionales específicas de la serie agregada y podría suceder que el ajuste indirecto deje efectos estacionales en la serie desestacionalizada; y 4) los factores estacionales no se estiman, sino que se deducen de forma implícita al contrastar la serie original con la desestacionalizada. Las primeras dos favorecen la aplicación del método indirecto respecto al directo; las otras, al directo.

Las características 1 y 3 son las más relevantes desde el punto de vista de la metodología estadística, y sus detalles deben cuidarse para que la desestacionalización sea exitosa. Por su lado, la 2 es primordial desde la perspectiva de interpretación de resultados en el contexto económico que corresponda, por lo que es deseable cumplir con ella. Por último, la 4 se puede considerar como un subproducto del uso del método indirecto, sin que sea de mayor importancia.

Desde luego, mantener la *optimalidad* de la desestacionalización aplicada de forma individual a todas y cada una de las series, y satisfacer, al mismo tiempo, las restricciones contables impuestas por la agregación, no siempre es posible. Los analistas de la situación económica tienden a estar de acuerdo en que la opción indirecta es preferible (e. g. Heath, 2015); esta visión es compartida

por Otranto y Triacca (2002) para evitar las incongruencias entre los agregados de series con y sin ajuste estacional, lo cual facilita el análisis que realizan los usuarios de la información; sin embargo, no hay argumentos que se consideren definitivos para favorecer una u otra opción y la decisión depende de cada caso en estudio.

Por otro lado, en el trabajo de Hood y Findley (2003) se muestran algunos puntos que conviene tener en cuenta al evaluar el desempeño de los dos enfoques; en él se buscan indicios de posibles inadecuaciones del ajuste con series agregadas y se presentan algunas medidas que permiten diagnosticar los resultados obtenidos en casos específicos. Otranto y Triacca (2002) enfatizan el hecho de que solo en caso de que haya una gran discrepancia entre los resultados habría que decidirse de manera formal, pues de otro modo es indistinto el método a usar.

Respecto a la decisión de cuál enfoque utilizar, en la guía elaborada por Eurostat para el Sistema de Estadísticas Europeas (ver Eurostat, 2015, p. 34) se mencionan las siguientes opciones de aplicación: a) lo mejor es usar el ajuste directo por su claridad, en particular cuando las series muestren patrones de estacionalidad similares, mientras que el indirecto sería adecuado en otro caso, aunque se debe verificar que la serie ajustada no contenga estacionalidad residual ni efectos de calendario, y también podría convenir aplicar el ajuste directo hasta un cierto nivel de agregación y el indirecto de ese nivel para arriba; b) es aceptable utilizar el enfoque directo junto con alguna técnica de ajuste de tipo *benchmarking* para hacer que se cumpla la restricción contable que liga a las componentes con el agregado o usar el indirecto cuidando que no queden efectos residuales de estacionalidad o efectos de calendario; y c) se debe evitar la aplicación de cualquiera de los dos enfoques sin una justificación para ello.

En la literatura sobre el tema han aparecido diversas propuestas que buscan mantener un equilibrio entre la optimalidad del proceso de ajuste y la satisfacción de restricciones contables de agregación, y que obligan a las series componentes ajustadas en forma directa a sumar a los valores desestacionalizados de manera directa del agregado. Se pueden citar al respecto los trabajos de Quenneville y Rancourty (2005), de Di Fonzo y Marini (2005) y de Quenneville y Fortier (2012); sin embargo, la manera más adecuada para conciliar los dos criterios conduce a un tema de investigación actual, que es la desestacionalización de series multivariadas en forma conjunta (ver McElroy, 2015 y Findley *et al.*, 1990). En ese contexto se busca una respuesta al problema, que conduce al uso de otros enfoques para el ajuste estacional, como pudiera ser el de modelos estructurales (e. g. Birrell *et al.*, 2011), y mientras no se tenga la solución completa del mismo, sería preferible usar el método directo —lo que implica que no se cumplan las restricciones contables— o el indirecto.

Pasos para realizar el ajuste estacional

Para que sea exitoso, se requiere que la serie original a la que se le aplique tenga estacionalidad, que la estimación del efecto sea confiable y que no haya estacionalidad remanente en la serie ajustada. Si al aplicar la validación del ajuste de una serie se encuentra que alguna de estas condiciones no se cumple, Tiller y Evans (2016) aconsejan no desestacionalizarla porque los resultados no serán válidos. Por estos motivos, se debe revisar que se cumplan las siguientes recomendaciones, expuestas en detalle en el documento de McDonald-Johnson *et al.* (2010):

- Visualizar los datos de forma gráfica para tratar de detectar patrones evidentes de estacionalidad y de posibles anomalías en los datos (series muy cortas, datos faltantes y atípicos).
- Verificar la presencia de estacionalidad en la serie y que ésta sea razonablemente estable con el paso del tiempo; de no ser así, decidir si otros ajustes pudieran ser necesarios para facilitar el estudio de la serie. La estabilidad del efecto estacional es un requisito indispensable para poder identificarla y estimarla de forma adecuada (Wright, 2013).
- Elegir la expresión más adecuada para la variable, de manera que el modelo estadístico subyacente en el ajuste estacional sea apropiado.
- Determinar el tipo de ajustes previos que requiere la serie para incorporarlos como parte de un modelo de regresión lineal, que se aplique antes de la desestacionalización.
- Decidir las opciones que se consideren más adecuadas para los filtros que se usan con el fin de estimar las componentes de la serie, de manera que el ajuste estacional sea exitoso.
- Para el caso de un conjunto de series interrelacionadas, elegir si deberá aplicarse el ajuste indirecto o el directo al agregado de las mismas.
- Verificar que no haya efectos estacionales en la serie ajustada por estacionalidad ni en la irregular.
- Verificar la ausencia de efectos del calendario en las series ajustada e irregular y, de no ser así, modificar el modelo de regresión usado para capturar los efectos correspondientes.
- Validar los supuestos de los modelos de re-gresión y de series de tiempo respectivos.
- Evaluar la estabilidad del ajuste estacional y de las revisiones.

Desde luego, se requiere evaluar cada uno de los puntos enunciados con el mayor detalle y formalidad posible para generar confianza, así como brindar credibilidad a los resultados del ajuste estacional. Por ello, es importante reconocer que existen criterios y pruebas estadísticas formales que sirven para validar, por ejemplo, la construcción de modelos de regresión, como se muestra en el texto de Guerrero (2009), donde se indica de manera detallada cómo llevar a cabo la verificación de supuestos de un modelo ARIMA y presenta, entre otros procedimientos, transformaciones estabilizadoras de la varianza, el uso de las funciones de autocorrelación y autocorrelación parcial para la especificación de modelos ARIMA, pruebas de raíces unitarias, métodos tanto para la estimación de parámetros como para la identificación y estimación de efectos exógenos, además de pruebas para la detección de autocorrelación residual.

Ilustración empírica

En esta sección se muestra un ejemplo utilizando el IGAE con cifras mensuales de enero de 1993 a agosto del 2017 (último dato disponible al momento de hacer este ejercicio, por lo cual se utilizaron 296) y año base = 2013. Primero se presentan los resultados por el método directo; enseguida, los del indirecto y, finalmente, se expone un ejercicio para evaluar la interacción de los efectos que se emplean en la parte de ajustes previos.

Método directo

En primer lugar, a través de la prueba F de estacionalidad estable se corrobora la hipótesis de presencia de estacionalidad, con un p -valor de 0.00, por lo que es procedente hacer el ajuste estacional.

La descomposición empleada es de tipo multiplicativo, así que se usa la transformación logarítmica y se consideran los siguientes efectos del calendario: días hábiles, año bisiesto y Semana Santa. El modelo ARIMA estacional para la parte regular tiene los órdenes (0,1, [1,4,7]) y para la parte estacional, los (0,1,1), donde se incluyen parámetros para capturar los efectos de tres anomalías: AO (oct. de 1995), LS (feb. de 1995) y LS (ene. del 2009). Los resultados de la estimación del modelo *regARIMA* para los ajustes previos, junto con la estructura estocástica del error se muestran en los cuadros 2 y 3 para la parte de regresión y en el 4 para el modelo ARIMA complementario. En ellos se observa que todos los parámetros estimados son estadísticamente significativos, de manera que se podría decir que no hay parámetros que estén de sobra.

La interpretación de los resultados del cuadro 2 es interesante porque, al estar la serie expresada en logaritmos, los parámetros estimados representan cambios relativos en la producción económica global, por ejemplo, de 1.28% cada año bisiesto y de -2.85% por Semana Santa. En las gráficas 1 se presentan los resultados del ajuste por efectos del calendario para la serie en niveles y para sus variaciones porcentuales anuales; se puede apreciar que las series ajustadas tienen menos variabilidad *indeseable*.

Cuadro 2

Estimación del modelo de regresión para la serie del IGAE en logaritmos

Variable	Parámetro estimado	Error estándar	Estadístico <i>t</i>
Año bisiesto (LP)	0.0128	0.00237	5.39***
Días hábiles (TD)			
Lunes	0.0009	0.00075	1.25
Martes	0.0034	0.00077	4.42
Miércoles	0.0022	0.00076	2.90
Jueves	0.0011	0.00078	1.45
Viernes	0.0028	0.00078	3.58
Sábado	-0.0028	0.00079	-3.62
Domingo	-0.0076	0.00078	-9.75
Semana Santa (<i>Easter</i> (4))	-0.0285	0.00153	-18.57***
Observaciones anómalas			
LS(1995.Feb)	-0.0720	0.00707	-10.18***
AO(1995.Oct)	-0.0345	0.00575	-5.99***
LS(2009.Ene)	-0.0354	0.00690	-5.13***

Cuadro 3

Pruebas *Ji-cuadradas* por grupos de regresores

Efecto de regresión	<i>df</i>	<i>Ji-cuadrada</i>	<i>p</i> -valor-
Regresores de días hábiles	6	416.46	0.00
Regresores de días hábiles y año bisiesto	7	442.57	0.00

Resultados de la estimación del modelo ARIMA para el IGAE

Modelo ARIMA (0, 1, [1, 4, 7]) x (0, 1, 1)			
Parámetro	Estimación	Error estándar	Estadístico <i>t</i>
Parte regular			
MA(1)	0.1910	0.05387	3.55***
MA(4)	-0.1808	0.05361	3.37***
MA(7)	0.3022	0.05308	5.69***
Parte estacional (orden 12)			
MA(1)	0.4604	0.05246	8.78***
Varianza del error	0.000081	0.000007	---

*** Significativo al 1 por ciento.

De manera semejante, en las gráficas 2 se presentan los resultados de la desestacionalización; se aprecia reducción en las fluctuaciones periódicas respecto a la serie previamente ajustada por efectos del calendario, sobre todo en la serie en niveles; ahora no se nota un cambio muy notorio en las variaciones anuales de la serie.

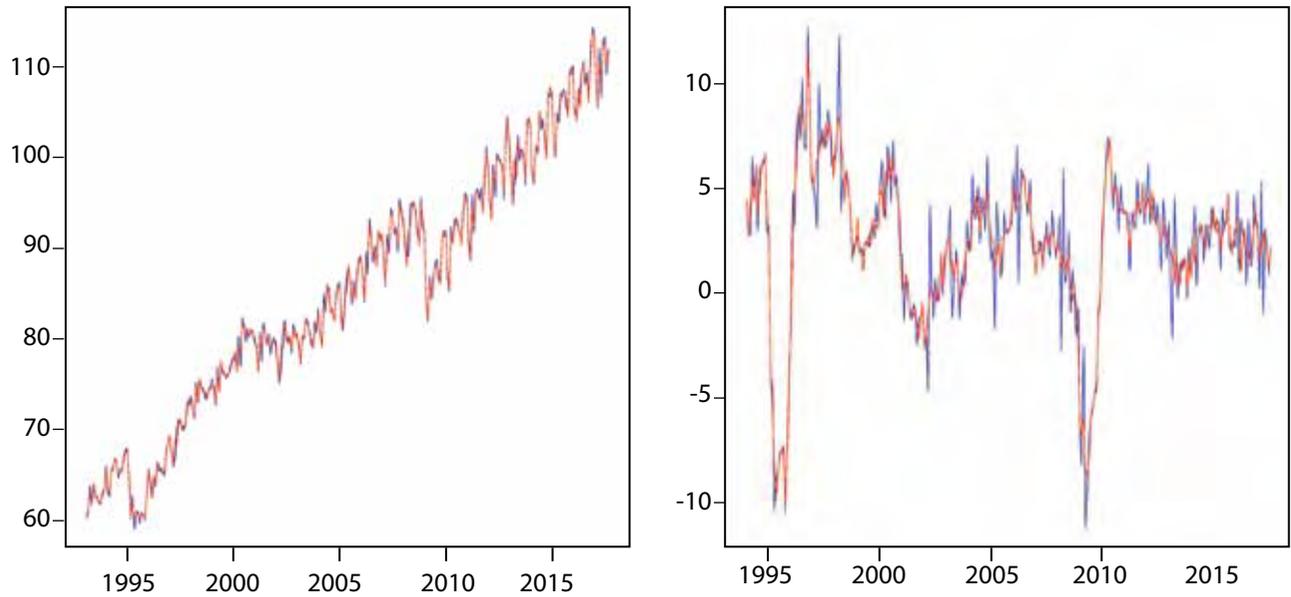
Las gráficas 3 muestran la estimación de los factores estacionales, primero en forma cronológica y después agrupados por mes. En el panel derecho se presentan, además, los factores de tipo IS que miden la razón del efecto irregular al efecto estacional. Es deseable que los factores sean pequeños, así la estimación del efecto estacional es más clara.

En las gráficas 4 se ve el resultado del ajuste estacional y la estimación de la componente de tendencia-ciclo, que difieren debido a la presencia de irregularidad en la primera y que se elimina en la segunda. Esta distinción hace que la serie ajustada por estacionalidad sea útil para análisis de corto plazo, mientras que la otra, para uno de largo plazo.

Una forma conveniente de apreciar la magnitud de cada componente de la serie es como se muestra en las gráficas 5: en las de la serie original, de la ajustada por estacionalidad y de la tendencia-ciclo se observa que mantienen la escala de la variable original; en cambio, los efectos del calendario, los factores estacionales y la componente irregular se expresan en términos relativos debido al uso del modelo multiplicativo. Se puede evaluar la calidad del ajuste mediante el estadístico Q_1 , el cual es aceptable en este caso con un nivel de 0.34. Además, el estadístico QS permite confirmar que ya no hay presencia de estacionalidad; mediante la autocorrelación positiva en el periodo estacional arroja un p -valor de 0.18.

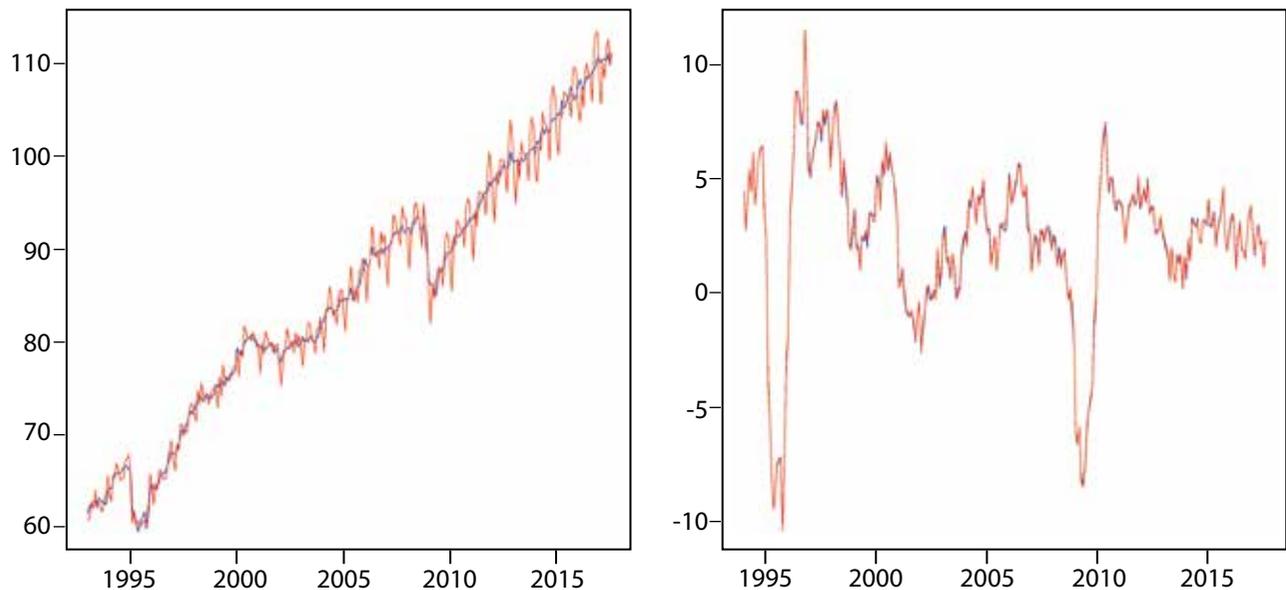
Gráficas 1

Panel izquierdo, serie original IGAE (azul) y serie corregida por efectos del calendario (roja). Panel derecho, variaciones porcentuales de ambas series (mismos colores, respectivamente)



Gráficas 2

Panel izquierdo, serie original IGAE corregida por efectos del calendario (roja) y serie desestacionalizada (azul). Panel derecho, variaciones porcentuales de ambas series (mismos colores, respectivamente)

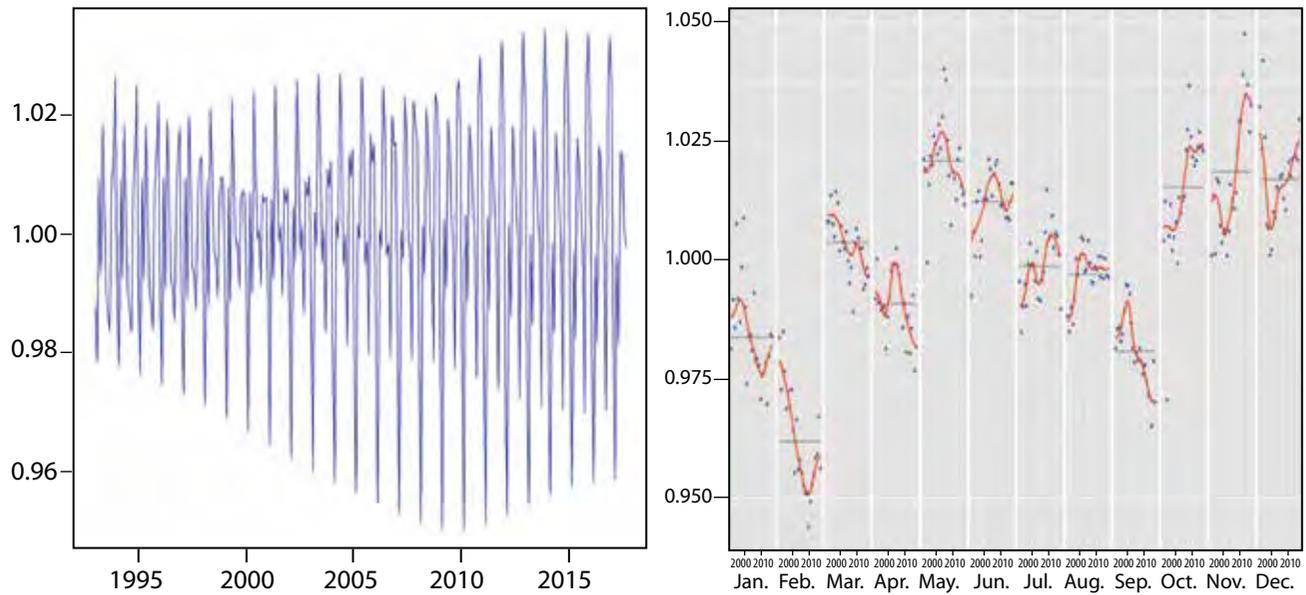


Método indirecto

El ajuste realizado hasta aquí se puede considerar completo y permite realizar análisis con los resultados logrados; no obstante, cuando se considera a la serie como un agregado y se analiza junto con sus componentes lo apropiado es realizar también el ajuste indirecto. Para este caso, se

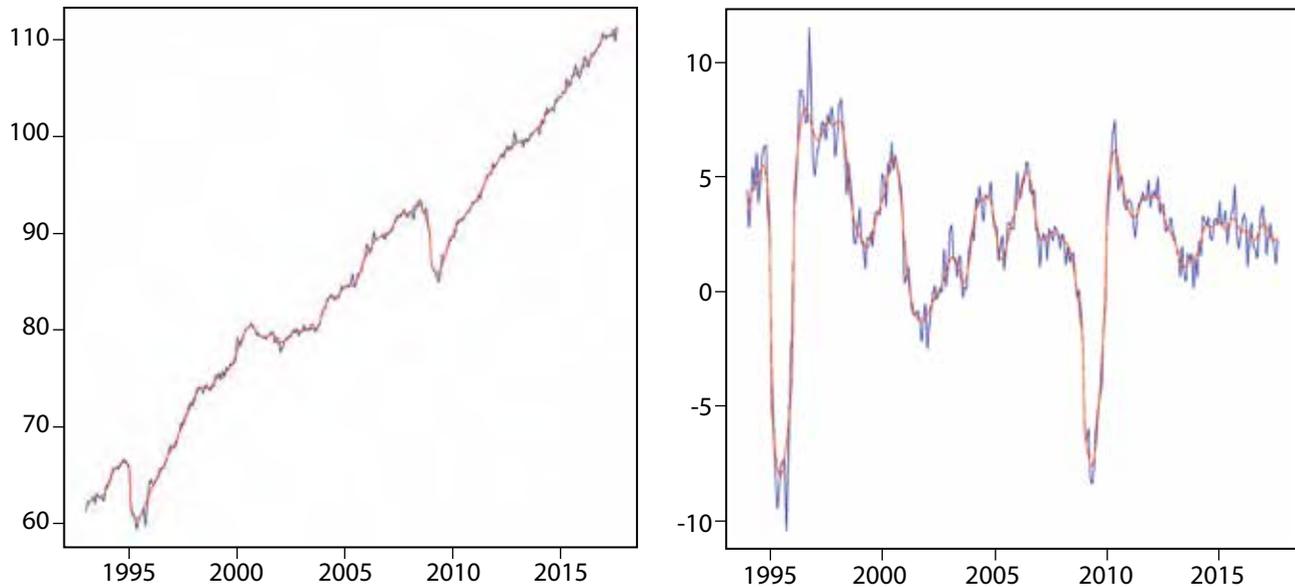
Gráficas 3

Panel izquierdo, serie de efectos estacionales. Panel derecho, factores estacionales y cocientes tipo IS para cada mes



Gráficas 4

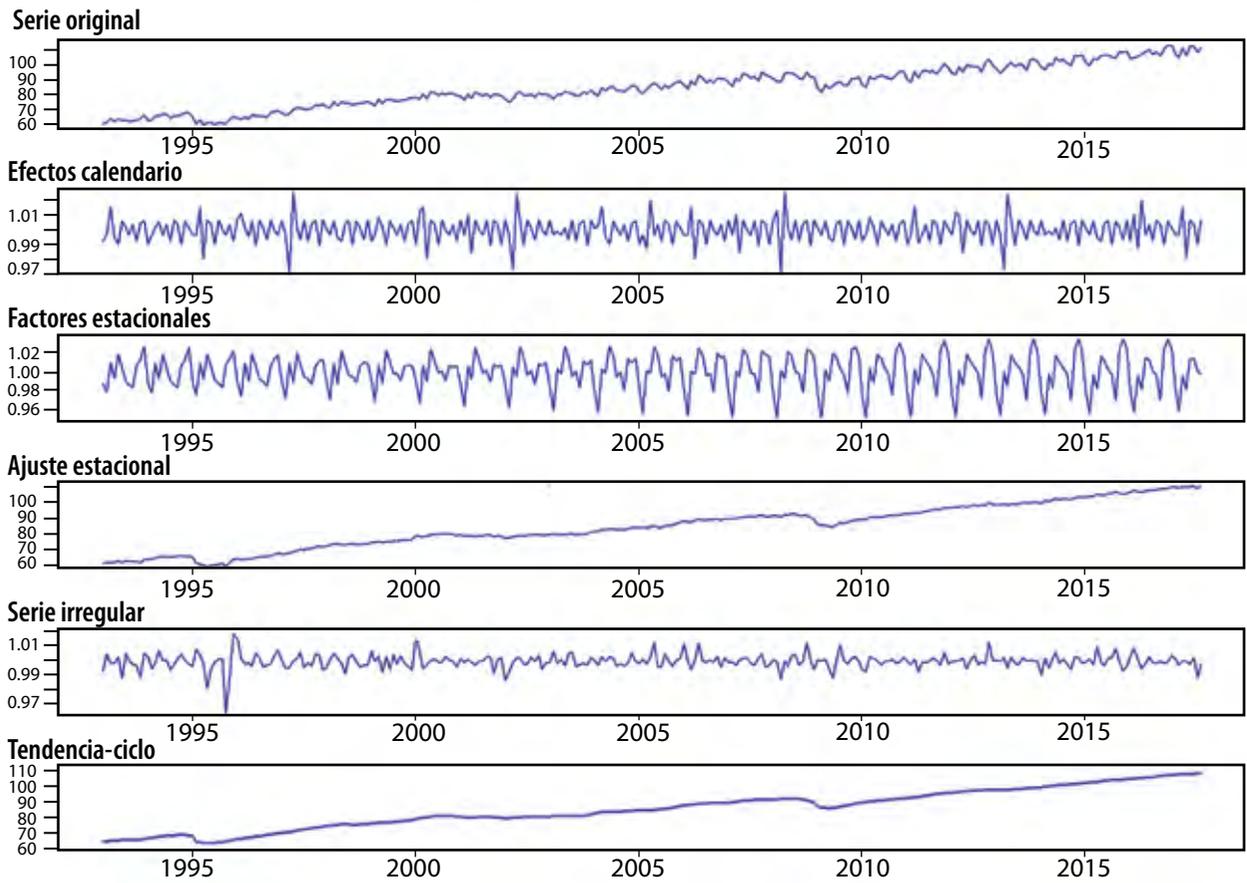
Panel izquierdo, serie del IGAE desestacionalizada (azul) y serie de tendencia-ciclo (roja). Panel derecho, variaciones porcentuales de ambas series (mismos colores, respectivamente)



obtendrá la desestacionalización primero de las series componentes por gran actividad económica del IGAE: primaria (GA1), secundaria (GA2) y terciaria (GA3). El resultado del ajuste estacional de la primera se muestra, junto con los factores estacionales, en las gráficas 6. Debe anotarse que en este caso no hay ajustes previos del calendario porque este tipo de actividad económica no lo requiere.

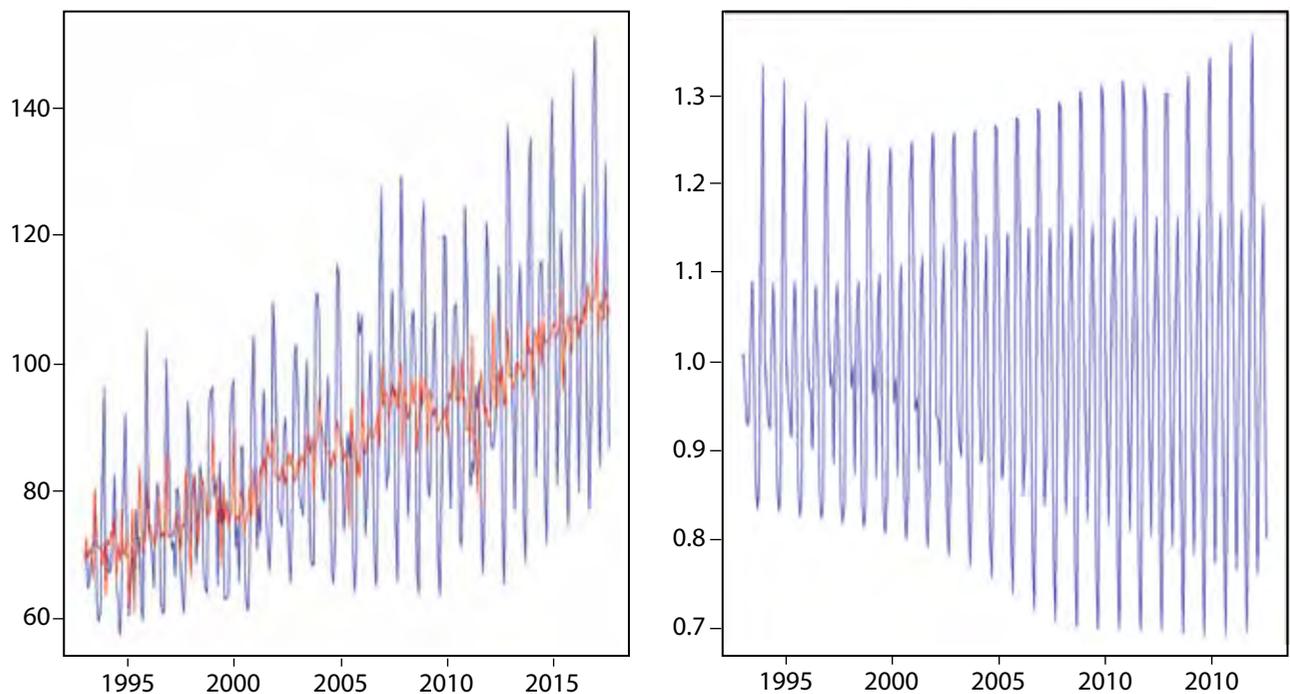
Gráficas 5

Descomposición multiplicativa del IGAE



Gráficas 6

Panel izquierdo, serie GA1 original (azul) y serie desestacionalizada (roja). Panel derecho, serie de efectos estacionales



En las gráficas de las series desestacionalizadas para GA2 y para la GA3 (ver gráficas 7 y 8) sí se realizó el ajuste por efectos del calendario. Se puede observar que los efectos estacionales siguen patrones dinámicos distintos para cada una de las tres GA y eso pudiera usarse como argumento en favor de preferir el método indirecto.

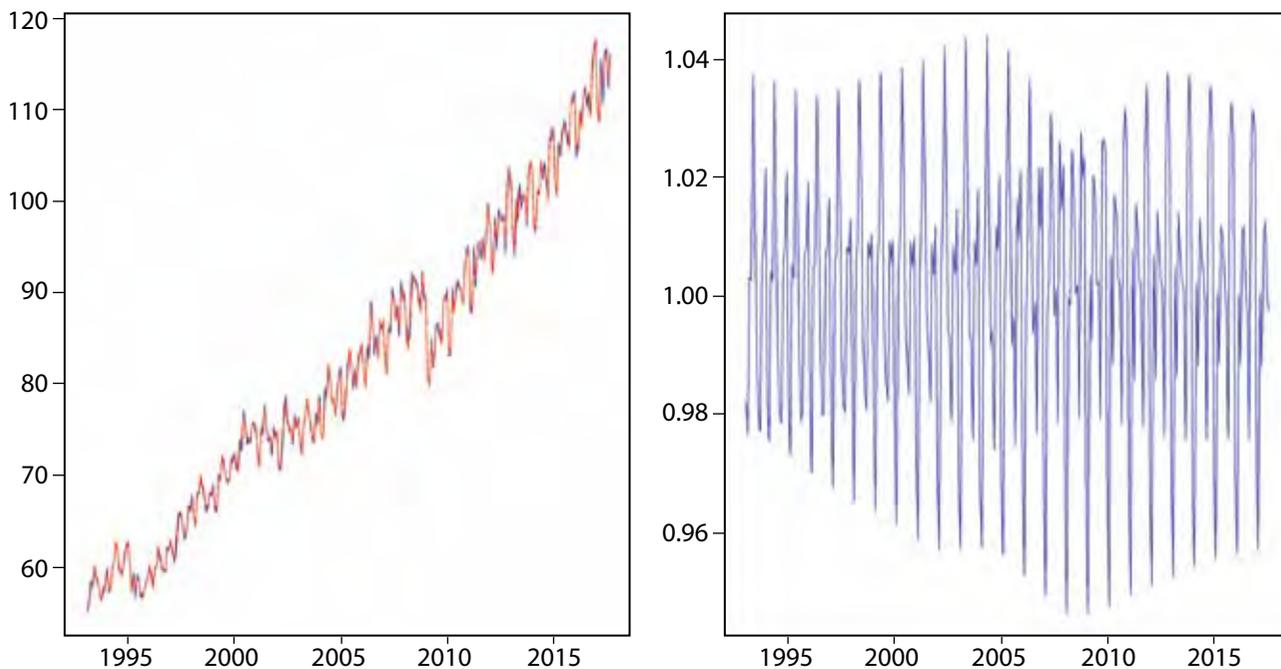
Al agregar las series desestacionalizadas de cada una de las GA se obtiene la serie del IGAE ajustada por estacionalidad de manera indirecta. Nótese que los pesos asociados (de tal forma que se cumpla con la restricción de que una combinación lineal de las GA sea el IGAE) vienen dados por la solución de un sistema de ecuaciones tal que $IGAE_t = \beta_1 GA1_t + \beta_2 GA2_t + \beta_3 GA3_t$. En el panel izquierdo de las gráficas 9 se observa que la serie desestacionalizada en forma indirecta se comporta de manera muy similar a la obtenida con el directo. De igual manera, en el panel derecho se aprecia que las variaciones anuales que se obtienen para ambos métodos son muy parecidas. Por estos motivos, es indistinto usar el directo o el indirecto; sin embargo, dado que las componentes tienen patrones estacionales diferentes, sería mejor usar el indirecto; además, se mantiene así la congruencia con los resultados de la desestacionalización por componentes.

Análisis de las variables explicativas incluidas como ajustes previos

En el 2016 ocurrieron tres efectos del calendario muy cercanos: el 21 de marzo (día de asueto oficial en México) ocurrió el lunes previo al Domingo de Pascua, por lo que en esa semana laboral se suspendieron actividades en tres de los cinco días; además, febrero tuvo 29 días por ser año bisiesto. Por ello, se busca identificar si existió algún efecto de interacción entre estos tres factores que pudieron haber afectado las cifras del IGAE ajustadas por estacionalidad: días hábiles, Semana Santa y año bisiesto.

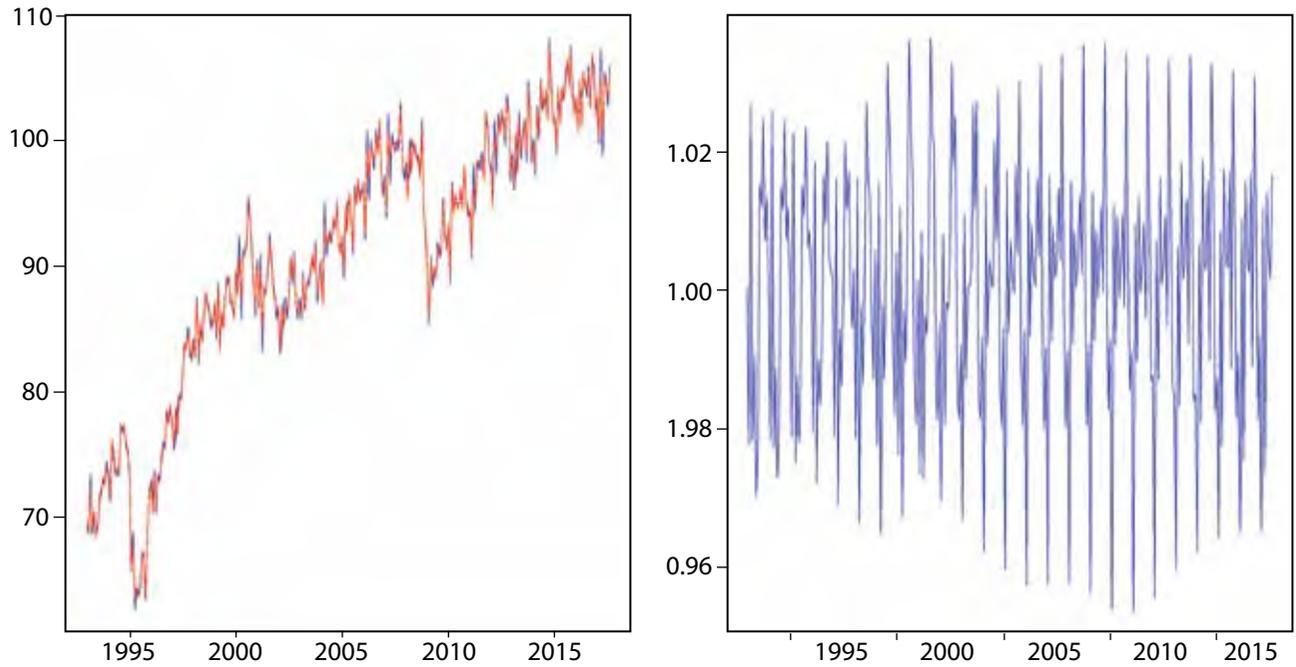
Gráficas 7

Panel izquierdo, serie GA2 original (azul) y serie desestacionalizada (roja).
Panel derecho, serie de efectos estacionales



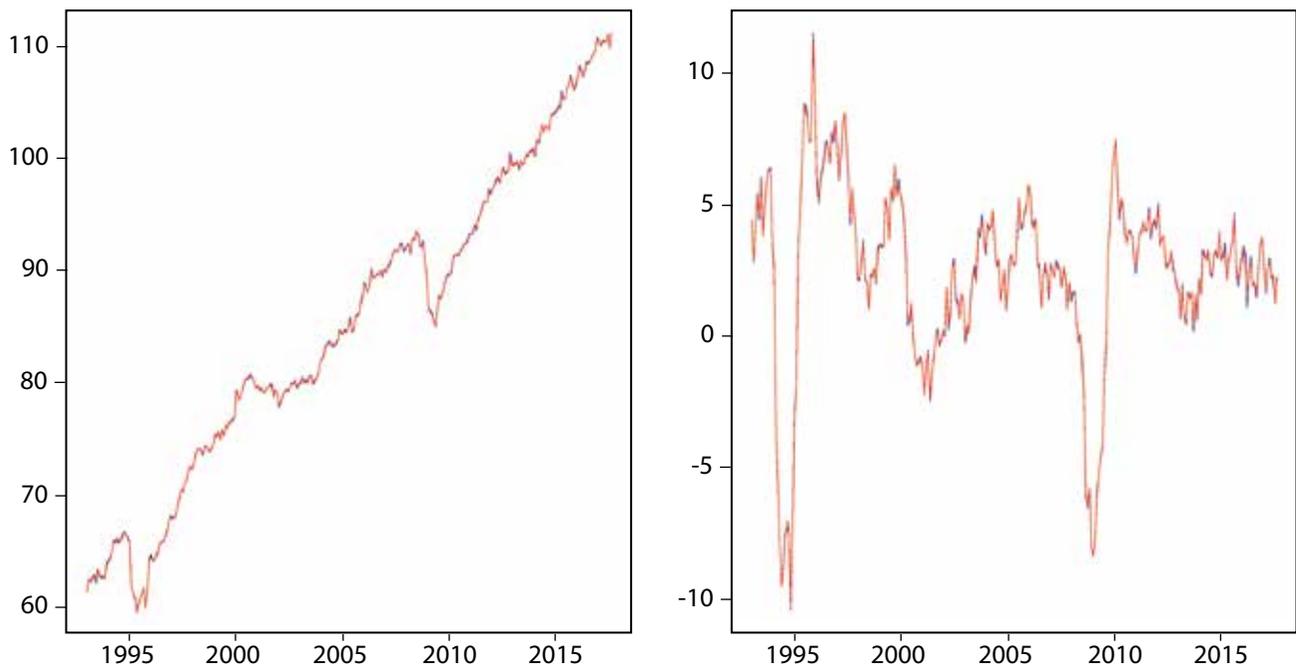
Gráficas 8

**Panel izquierdo, serie GA3 original (azul) y serie desestacionalizada (roja).
Panel derecho, serie de efectos estacionales**



Gráficas 9

**Serie IGAE ajustada por estacionalidad, con método directo (azul) y con indirecto (roja).
Panel derecho, variaciones porcentuales de las series**



Es cierto que, por construcción, la interacción entre días hábiles y Semana Santa no captará efectos específicos de México, pero de resultar significativo nos daría indicios estadísticos de que puede resultar conveniente incorporar un calendario de días hábiles específico para México, dado que los días calendario interactúan con la Semana Santa.

Con un enfoque de diseño de experimentos, se utilizó un diseño factorial que consta de 16 diferentes modelos, siendo éstos todas las combinaciones posibles entre unos con y sin las variables de días hábiles, año bisiesto y Semana Santa, así como la interacción de días hábiles-Semana Santa. Lo anterior, para analizar la variable de respuesta $\nabla_{12} \log(IGAE_{it}) - \nabla_{12} \log(\widehat{IGAE}_{it})$, donde \widehat{IGAE}_{it} es el IGAE ajustado por estacionalidad, $t = 1, \dots, 12$ es el mes e $i = 1, \dots, 16$ es el modelo ajustado y ∇_{12} es el operador en diferencias $(1 - B)^{12}$, donde $Bx_t = Bx_{t-12}$. Se escogió esta variable dado que es de interés observar el efecto introducido al realizar el ajuste estacional respecto a la serie original, aunado a que ésta es estacionaria, lo que hace válidos los resultados obtenidos en los ejercicios de regresión provenientes del análisis experimental. Además, al analizar las diferencias porcentuales, se puede responder a diferentes cuestionamientos, por ejemplo, determinar las causas de por qué es mayor/menor el crecimiento de la serie desestacionalizada con respecto a la original. Para ejemplificar dicha variable, las gráficas 10 muestran los resultados de la diferencia utilizando como modelo ajustado el que se obtuvo en la ilustración empírica.

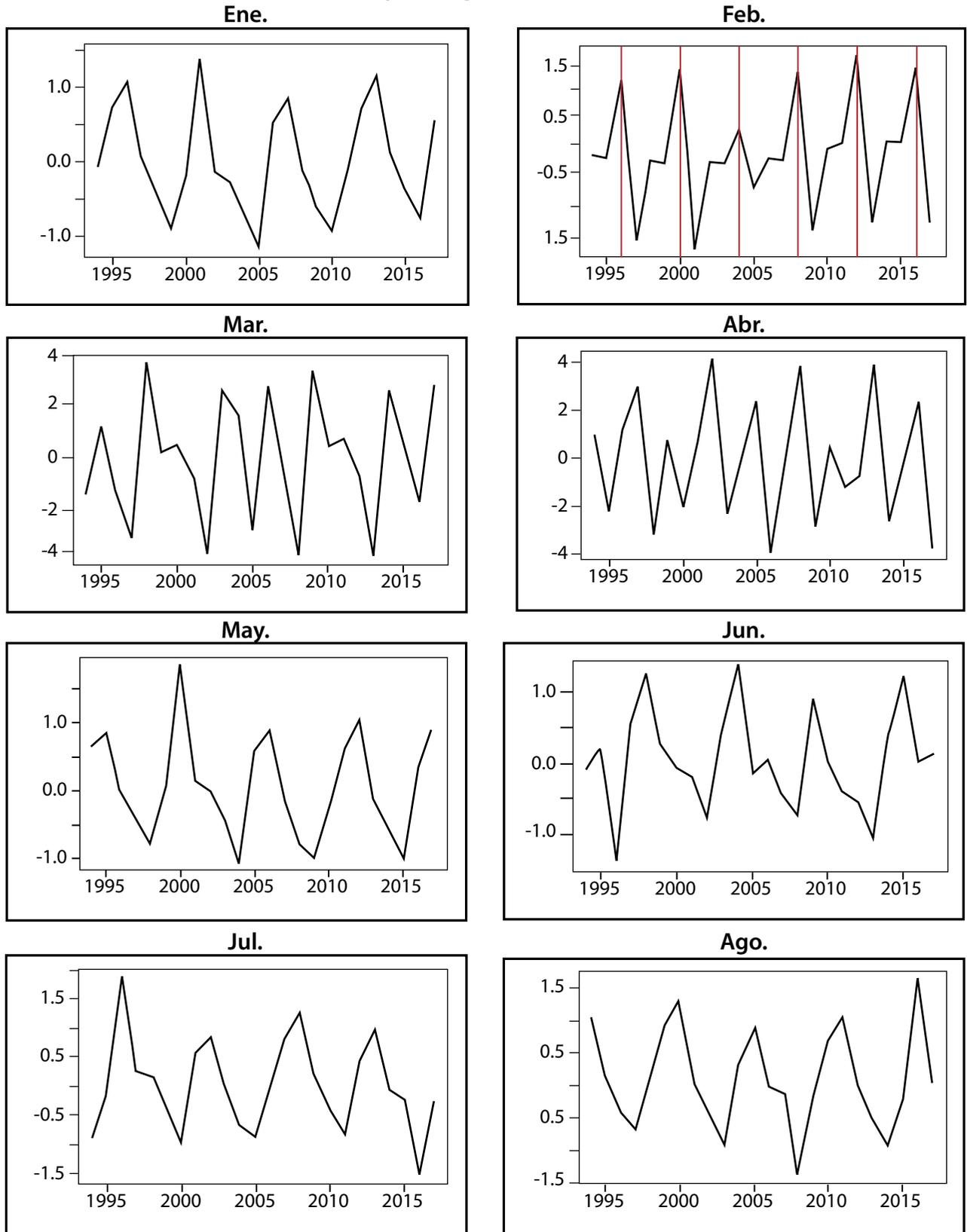
Se puede apreciar de manera visual que el comportamiento es estacionario y, por ejemplo, el crecimiento de la serie original en febrero es mayor en los años bisiestos respecto al crecimiento de la serie desestacionalizada. Para resumir el comportamiento de cada mes y estimar el diseño experimental propuesto, se tomó la mediana de cada mes, esto debido a que al observar las densidades empíricas, las distribuciones no mostraron simetría. De esta forma, el modelo que representa la modelación de todos los efectos es el siguiente:

$$\begin{aligned} \Delta \widehat{IGAE}_t = & \beta_1 + \beta_2 \cdot td_t + \beta_3 \cdot leapyear_t + \beta_4 \cdot easter_t \\ & + \beta_5 \cdot td_t \cdot easter_t + error_t, \end{aligned}$$

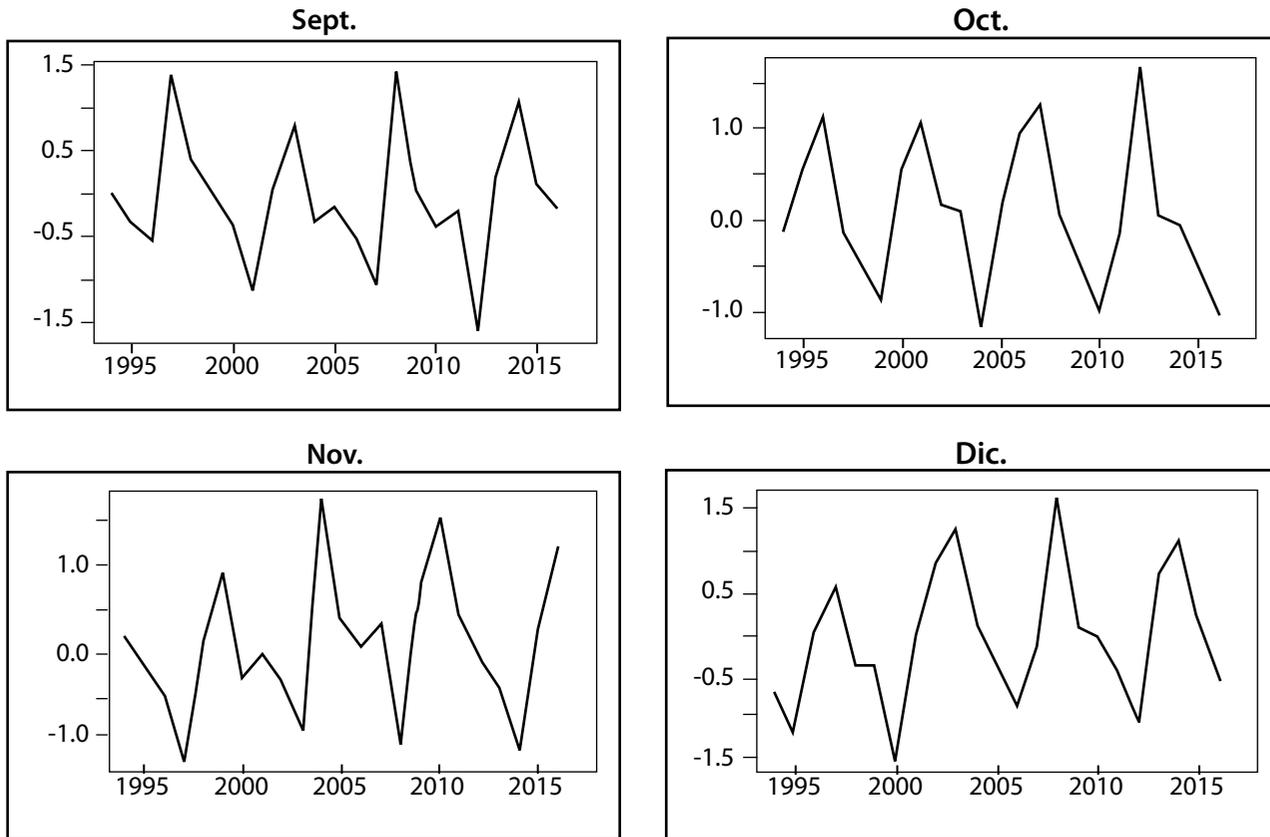
donde $\Delta \widehat{IGAE}_t$ es la mediana de $\nabla_{12} \log(IGAE_{1t}) - \nabla_{12} \log(\widehat{IGAE}_{1t})$ para el modelo 1 (que incluye todos los efectos), β_1 es el intercepto que representa el valor medio de la variable, β_2 , β_3 y β_4 son coeficientes asociados con los efectos principales de días hábiles (td), año bisiesto ($leap year$) y Semana Santa ($easter$), respectivamente. Al considerar que puede existir efecto de interacción td - $easter$ se incluye la variable asociada con el coeficiente β_5 . Nótese que la interacción de días hábiles con año bisiesto puede considerarse, pero ésta no se puede estimar debido a problemas de singularidad en el modelo *regARIMA*. La interacción Semana Santa con año bisiesto es irrelevante y no se incluye. Las gráficas 11 muestran resultados del análisis experimental a través de los *Half-Normal Scores* para cada mes, donde se indican los efectos estadísticamente significativos.

Los meses de interés son febrero, marzo y abril, donde puede haber interacciones. Para febrero, como se esperaba, resultó significativa la variable de año bisiesto. Para marzo, los efectos aislados de días hábiles y Semana Santa resultaron significativos, no así la interacción entre ellos, lo cual podría deberse a que se toman en cuenta los días calendario y no un calendario de días hábiles mexicano que descuenta los festivos (como el lunes más cercano al 21 de marzo). Para abril, los días hábiles y Semana Santa son significativos y también su interacción. Esto implica que los efectos individuales no pueden interpretarse por separado. Al analizar los resultados por completo, se podría considerar un modelo de ajuste estacional que, además, incluya la interacción de días hábiles con Semana Santa. Para ver las diferencias entre éste y el modelo sin interacción, la gráfica 12 muestra los resultados expresados en cambios porcentuales.

Resta de las diferencias logarítmicas del IGAE versus las del modelo para IGAE ajustado por estacionalidad



Resta de las diferencias logarítmicas del IGAE versus las del modelo para IGAE ajustado por estacionalidad

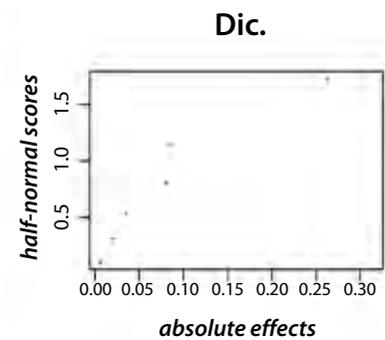
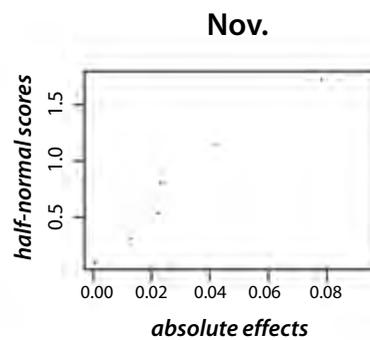
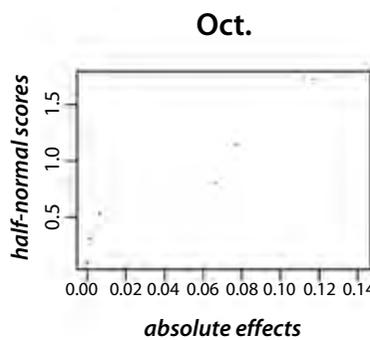
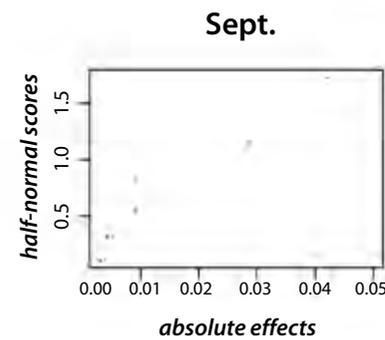
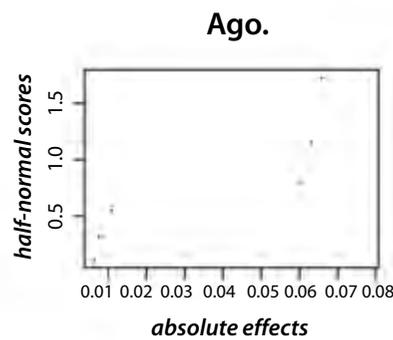
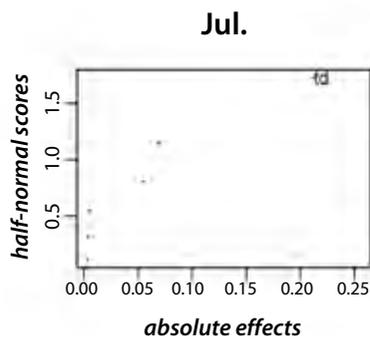
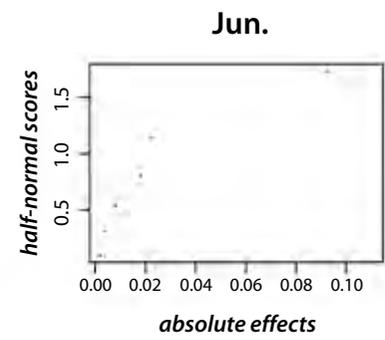
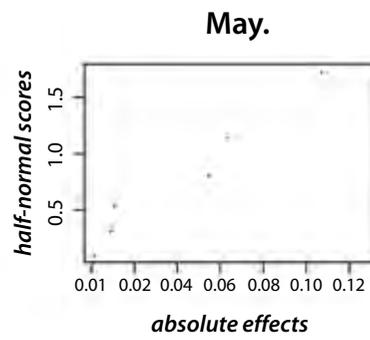
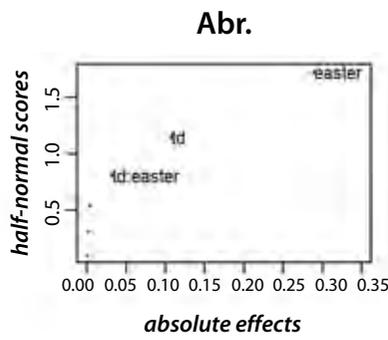
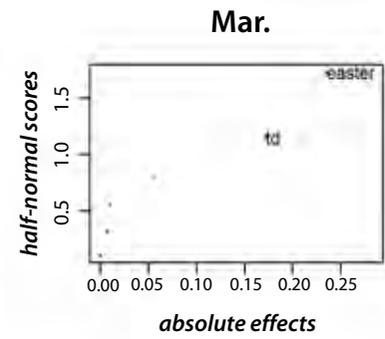
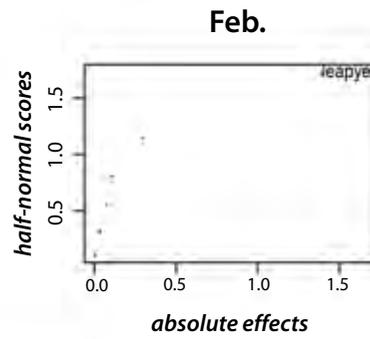
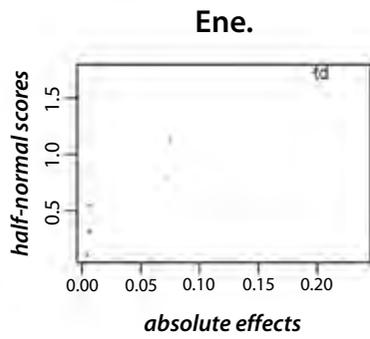


En la gráfica 12 se puede observar que el modelo que incluye la interacción genera una serie *más suave* que el modelo sin interacción. Este resultado puede ser analizado bajo diferentes enfoques: 1) visualmente, donde notamos la presencia de menos *picos* en la serie que incluye la interacción, lo que, por lo general, se busca en el proceso de desestacionalización y 2) estadísticamente, donde el análisis de residuos y estacionalidad residual hace casi indistinguible qué modelo es mejor. Se agregó un tercer criterio más objetivo: realizar una prueba de hipótesis a través de razones de verosimilitud (LRT) para evaluar de forma estadística si las interacciones son en conjunto significativas; con éste se determinó que no se rechaza la hipótesis nula con un p -valor = 0.13, es decir, las interacciones no son necesarias en el modelo. Este resultado no es del todo concluyente, pues visualmente sí se sugiere introducir la interacción, por lo que ésta pudiera ser relevante para algunas series de tiempo económicas mexicanas. Al enfatizar este punto, es evidente que el lunes más cercano al 21 de marzo sí interactúa con la Semana Santa. Por ello, una línea futura de investigación sería observar cómo se da ese proceso de interacción, si a través de los días hábiles, descontando feriados, o bien, modelando los feriados como variables tipo festivas, para todo el calendario mexicano.

Conclusiones

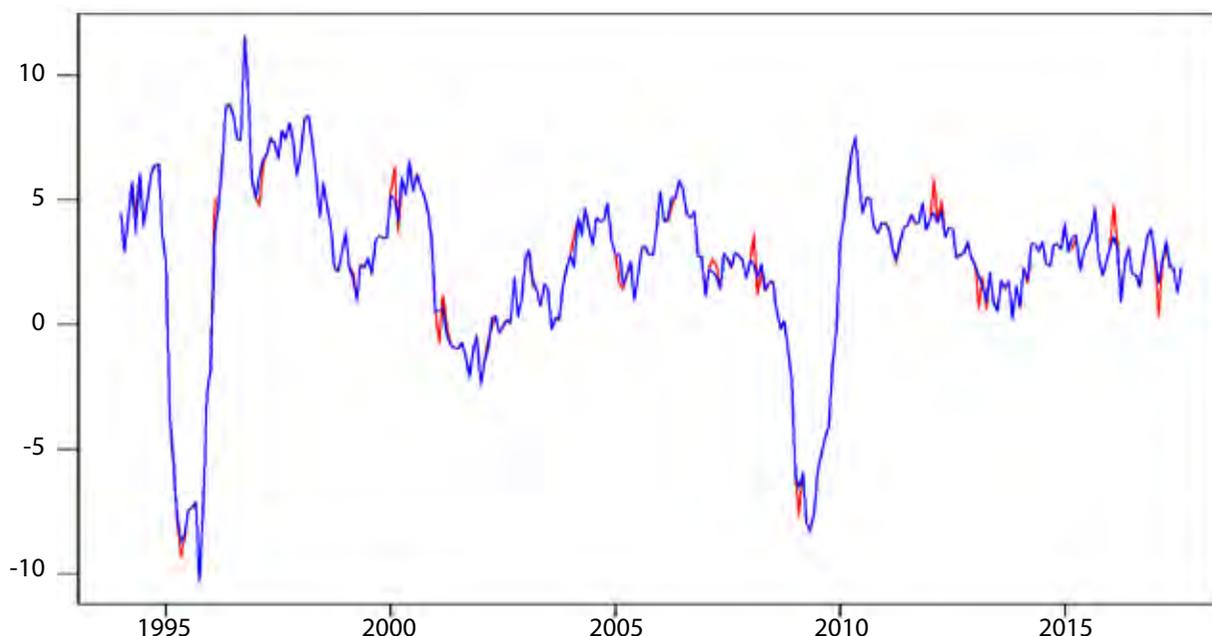
En el presente documento se mostraron diferentes enfoques y herramientas utilizados para realizar el ajuste estacional de series de tiempo económicas en México. Es importante conocer los avances en la literatura especializada en la materia tanto de carácter computacional como de los temas de

Half Normal plot para los meses del año



Gráfica 12

Cambios porcentuales anuales de la serie desestacionalizada que incluye interacción (línea azul) y cambios porcentuales anuales de la serie desestacionalizada que no incluye interacción (línea roja)



investigación sobre asuntos pendientes; sin embargo, se requiere aplicar estos conceptos desde una perspectiva nacional, acorde con el entorno actual. Las recomendaciones aquí presentadas apoyan la decisión de utilizar el procedimiento de ajuste del paquete *X-13* y que el método directo es preferible para series que son resultado de agregación cuando el énfasis está puesto en la parte de teoría estadística, pero la aplicación del indirecto es recomendable cuando el interés es la interpretación práctica —y de carácter contable— de los resultados desestacionalizados.

Desde el punto de vista de la metodología, la práctica estándar de seleccionar la transformación logarítmica para decidir entre un modelo aditivo y uno multiplicativo (o *log*-aditivo) con base en el AICC no tiene buen respaldo de teoría estadística y hace falta considerar otros métodos para elegir de forma adecuada. También, debe subrayarse que entre los pendientes de solución teórica se encuentra el desarrollo de un método para estimar la varianza de los efectos estacionales y de la serie desestacionalizada.

El ejemplo que aquí se presentó, junto con el ejercicio estadístico de diseño de experimentos, contribuye a que el usuario conozca a detalle el proceso de ajuste estacional y no se conforme con los resultados automáticos de los paquetes computacionales. En especial, esta aplicación permitió identificar que sí existen efectos de interacción entre días hábiles y Semana Santa en abril, la cual resultó ser significativa, de modo que al incluirla en el modelo, se obtiene una serie más suave que con el modelo sin interacción. Lo anterior da cabida a considerar relevante la modelación de efectos específicos del calendario mexicano y sus respectivas interacciones. Una primera idea sería construir el calendario de días hábiles, donde se descuenten festivos fijos y movibles, de tal forma que la interacción entre días hábiles y Semana Santa rescate sus particularidades, como la ocurrida en marzo del 2016. De igual manera, los feriados movibles pueden modelarse como variables de días festivos y analizar también su interacción con otras, como los días hábiles y la Semana Santa.

Fuentes

- Birrell, C. L., D. G. Steel y Y-X Lin. "Seasonal Adjustment of an Aggregate Series Using Univariate and Multivariate Basic Structural Models", en: *Journal of Statistical Theory and Practice*. 5 (2). 2011, pp. 179-205.
- Bloem, A. M., R. J. Dippelsman y N. O. Maehle. *Manual de cuentas nacionales trimestrales. Conceptos, fuentes de datos y compilación*. Fondo Monetario Internacional, Washington, DC, 2001.
- Cleveland, R. B., W. S. Cleveland, J. E. McRae e I. Terpenning. "STL: A Seasonal-Trend Decomposition Procedure Based on Loess", en: *Journal of Official Statistics*. 6. 1990, pp. 3-73.
- Dagum, E. B. *The X-11-ARIMA Seasonal Adjustment Method*. Catalogue 12-564E. Statistics Canada, Ottawa, Canada, 1980.
- Dagum, E. B. y P. A. Cholette. "Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series", en: *Lecture Notes in Statistics 186*. Springer-Verlag, New York, 2006.
- Di Fonzo, T. y M. Marini. "Benchmarking Systems of Seasonally Adjusted Time Series", en: *Journal of Business Cycle Measurement and Analysis*. 2 (1). 2005, pp. 89-123.
- Eurostat. *ESS guidelines on seasonal adjustment*. Eurostat, Luxembourg, 2015.
- Findley, D. F., B. C. Monsell, W. R. Bell, M. C. Otto y B-Ch. Chen. "New capabilities and methods of the X-12-ARIMA Seasonal Adjustment Program", en: *Journal of Business and Economic Statistics*. 16 (2). 1998, pp. 127-152.
- Findley, D. F., B. C. Monsell, H. B. Shulman y M. G. Pugh. "Sliding-Spans Diagnostics for Seasonal and Related Adjustments", en: *Journal of the American Statistical Association*. 85. 1990, pp. 345-355.
- Granger, C. W. "Seasonality: causation, interpretation, and implications", en: *Seasonal analysis of economic time series*. National Bureau of Economic Research, Cambridge, MA, 1979, pp. 33-56.
- Grudkowska, S. *JDemetra+ User Guide*. Department of Statistics, Narodowy Bank Polski, Polonia, 2015.
- Guerrero, V. M. "Desestacionalización de series de tiempo: una introducción a la metodología", en: *Documento de Investigación núm. 54*. Subdirección de Investigación Económica, Banco de México, México, 1983a.
- _____ "Desestacionalización de series de tiempo económicas: ajustes previos", en: *Documento de Investigación núm. 55*. Subdirección de Investigación Económica, Banco de México, México, 1983b.
- _____ "Desestacionalización de series de tiempo: una introducción a la metodología", en: *Comercio Exterior*. 40 (11). 1990, pp. 1035-1046.
- _____ "Desestacionalización de series de tiempo económicas: ajustes previos", en: *Comercio Exterior*. 42 (11). 1992, pp. 1042-1053.
- _____ *Análisis estadístico y pronóstico de series de tiempo económicas*. Just in Time Press, México, 2009.
- Heath, J. *Directo o indirecto (DE) blog* consultado en <http://jonathanheath.net/directo-o-indirecto/Arena> Pública el 24 de abril de 2015.
- Hood, C. y D. F. Findley. "Comparing direct and indirect seasonal adjustments of aggregate series", en: M., Manna y R. Peronaci (eds.). *Seasonal Adjustment*. European Central Bank, Frankfurt am Main, Germany, 2003, pp. 9-21.
- Hylleberg, S. "Seasonal Adjustment" (Working Paper No. 2006-4), en: *New Palgrave Dictionary of Economics*. 2nd ed. Department of Economics, School of Economics and Management-University of Aarhus, Denmark, 2006.
- Ladiray, D. y B. Quenneville. "Desestacionalizar con el método X-11", en: *Methodologica*. Núm. 8-9. Université Libre de Bruxelles, Bruxelles, Belgique, 2000-2001.
- Lefrançois, B. y C. Mamay. *Presentation of Seasonally Adjusted Series (Working Paper)*. Short-term Economic Statistics Expert Group (STESeg): Task Force on Data Presentation and Seasonal Adjustment. OECD, Paris, 2004.
- McDonald-Johnson, K. M., B. Monsell, R. Fescina, R. Feldpausch, H. C. C. Harvill y M. Wroblewski. "Census Bureau Guideline. Seasonal Adjustment Diagnostics Checklists", en: *Supporting Document A*. US Census Bureau, Department of Commerce, Washington, DC, 2010.
- McElroy, T. "Multivariate Seasonal Adjustment, Economic Identities, and Seasonal Taxonomy", en: *Research Report Series, Statistics #2015-01*. US Census Bureau, Department of Commerce, Washington, DC, 2015.
- Otranto, E. y U. Triacca. "Measures to evaluate the discrepancy between direct and indirect model-based seasonal adjustment", en: *Journal of Official Statistics*. 18. 2002, pp. 511-530.
- Quenneville, B. y E. Rancourty. "Simple methods to restore the additivity of a system of time series. Statistics Canada", en: *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*. Luxembourg, Eurostat, 2005.
- Quenneville, B. y S. Fortier. "Restoring accounting constraints in time series. Methods and software for a Statistical Agency", en: Bell, W. R., S. H. Holan y T. S. McElroy (eds.). *Economic Time Series: Modelling and Seasonality*. CRC Press, Taylor and Francis Group, Boca Raton, Florida, 2012.

- Tiller, R., D. Chow y S. Scott. *Empirical Evaluation of X-11 and Model-based Seasonal Adjustment Methods (Working Paper)*. Washington, DC, Bureau of Labor Statistics, 2007.
- Tiller, R. B. y T. D. Evans. "Methodology for Seasonally Adjusting National Household Survey Labor Force Series with Revisions for 2016", en: *Current Population Survey (CPS) Technical Documentation*. Bureau of Labor Statistics, Washington, DC, 2016.
- US Census Bureau. *X-13ARIMA-SEATS, Reference Manual. Version 1.1*. Center for Statistical Research and Methodology, Washington, DC, 2016.
- Wright, J. H. "Unseasonal seasonals?", en: *Brookings Papers on Economic Activity* (con discusión). 2013, pp. 65-126.
- Young, A. H. "Estimating Trading-Day variation in monthly economic series", en: *Technical Paper No. 12*. Washington, DC, US Bureau of the Census, 1965.