Notas Metodológicas

Barómetro del Comercio Contenerizado en Colombia

El Barómetro del Comercio Contenerizado es un indicador que mide las desviaciones a corto plazo de la movilización de TEU (Twenty-foot Equivalent Unit) respecto a las tendencias recientes del comercio internacional de mercancías. El indicador está compuesto por los factores de intercambio comercial marítimo relacionados con las exportaciones e importaciones de 11 sectores económicos fundamentales de la economía nacional, este carácter integral lo posiciona como un indicador importante para obtener percepciones sobre la trayectoria actual del comercio internacional de carga contenerizada, identificar puntos de inflexión en los sectores económicos y medir su pulso en la actividad comercial del país.

El proceso metodológico para calcular el barómetro para los factores de intercambio comercial y sectores económicos se ejecuta en tres etapas: **1**. Ajuste estacional y suavizamiento de la serie de tiempo. **2**. Estimación de la tendencia de la serie temporal empleando el Filtro de Hodrick-Prescott. **3**. Cálculo de las desviaciones normalizadas de la serie de tiempo respecto a las tendencias recientes.

1. Ajuste estacional y suavizamiento de una serie de tiempo

Las series de tiempo económicas regularmente exhiben patrones de comportamiento recurrentes con intensidad y frecuencia similar que generan efectos y fluctuaciones de corto plazo que pueden no estar ligados directamente con la evolución natural de la serie temporal, hecho que dificulta la identificación de los movimientos subyacentes de carácter estrictamente económico e impide una evaluación adecuada del comportamiento entre periodos sucesivos de la serie económica.

El ajuste estacional es el proceso de remover los fenómenos estacionales y de calendario de una serie de tiempo con el propósito de proporcionar una comprensión más precisa de las tendencias subyacentes, el ciclo económico y los movimientos a corto plazo de la serie. Uno de principales procedimientos metodológicos para el ajuste estacional comúnmente utilizado por las agencias productoras de datos es el *X-13-ARIMA-SEATS* propuesto por (U.S. Census Bureau, 2017), el cual consiste en una herramienta avanzada que implementa dos métodos alternativos: el *filtro X-11* basado en un proceso iterativo de filtros de medias móviles predefinidos y el *filtro SEATS* basado en técnicas de extracción de señales. Dada la flexibilidad de la metodología *X-13-ARIMA-SEATS*, es el procedimiento recomendado para la generación de datos ajustados estacionalmente (IMF, 2018).

El análisis clásico de series temporales supone que los valores de una variable observada es el resultado de la actuación conjunta de tres componentes no observables: tendencia-ciclo (T_t), estacional (S_t) e irregular (I_t). Por tanto, el modelo general de una serie de tiempo puede expresarse de forma aditiva o multiplicativa de la siguiente manera:

Modelo Aditivo

$$Y_t = T_t + S_t + I_t$$

Modelo Multiplicativo

$$Y_t = T_t \cdot S_t \cdot I_t$$

La lógica general del cálculo de la serie desestacionalizada consiste en remover el componente estrictamente estacional y de calendario de la serie de tiempo original. En el caso del modelo aditivo es equivalente a $Y_t^a = T_t + I_t$. Mientras que en el caso del modelo multiplicativo es igual al producto de sus componentes no observables: $Y_t^a = T_t \cdot I_t$.

La descomposición de la serie de tiempo en los componentes de tendencia-ciclo, estacional e irregular es llevada a cabo a través del módulo de modelado basado en *X-13-ARIMA-SEATS* el cual está diseñado para la construcción de modelos del tipo *regARIMA* con series temporales económicas estacionales y permite la inclusión de varias categorías de variables de regresión predefinidas que incluyen constantes de tendencias, medias generales, efectos estacionales fijos, efectos de días laborales, efectos de días festivos, efectos de outliers, cambios de nivel, efectos de pendiente, entre otros. Los modelos ARIMA fueron propuestos por (Box & Jenkins, 1976) y son frecuentemente implementados para series de tiempo estacionales, estos pueden definirse de forma general de la siguiente manera:

Modelo ARIMA

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D \cdot z_t = \theta(B)\Theta(B^s) \cdot a_t$$

Donde B es un operador de rezagos ($Bz_t = z_{t-1}$), s es el periodo estacional, el operador autorregresivo (AR) regular especificado por $\phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$, el operador autorregresivo estacional $\Phi(B) = (1 - \Phi_1 B - \cdots - \Phi_P B^{Ps})$, el operador de medias móviles (MA) dado por $\theta(B) = (1-\theta_1 B - \cdots - \theta_q B^q)$ y el operador de medias móviles estacionales $\Theta(B^s) = (1-\Theta_1 B - \cdots - \Theta_Q B^{Qs})$. El término de ruido blanco a_t es independiente e idénticamente distribuido con media cero y varianza constante σ^2. El término $(1-B)^d(1-B^s)^D$ implica diferenciación regular de orden d y diferenciación estacional de orden D. Una extensión útil de los modelos ARIMA surge del uso de una regresión lineal para modelar el nivel medio de una variable en el tiempo:

Modelo Regresión Lineal

$$y_t = \sum_i \beta_i x_{it} + z_t$$

El modelo general *regARIMA* definido en el algoritmo *X-13-ARIMA-SEATS* se calcula mediante la combinación de las dos anteriores ecuaciones, cuyo resultado es:

Modelo regArima

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D \underbrace{\left(y_t - \sum_i \beta_i x_{it}\right)}_{z_t} = \theta(B)\Theta(B^s)\cdot a_t$$

El modelo *regARIMA* puede considerarse tanto como una generalización del modelo ARIMA puro al permitir una función de media de regresión especificada por $\sum_i \beta_i x_{it}$. O puede considerarse como una generalización del modelo de regresión al permitir que los errores z_t sigan un modelo ARIMA. La estimación de los parámetros del modelo se realiza usando el *Algoritmo de Mínimos Cuadrados Generalizados Iterativos* (*IGLS*), el cual itera entre la estimación de los parámetros AR y MA obtenidos mediante una regresión de mínimos cuadrados generalizados (*GLS*) y la estimación de los parámetros de regresión β_i computados a través de la estimación de máxima verosimilitud de la serie temporal de errores de regresión $y_t - \sum_i \beta_i x_{it}$ mediante un modelo ARIMA.

El resultado de desarrollar el procedimiento anterior es la serie de tiempo corregida por los fenómenos estacionales y de calendario, la cual es sometida a un proceso de suavización con el objetivo de moderar los ruidos o fluctuaciones aleatorias a corto plazo que puedan dificultar la identificación de patrones significativos, puntos de inflexión y tendencias subyacentes.

El método de suavización empleado en el proceso de construcción del indicador se denomina Loess (*Locally Weighted Regression and Smoothing*) propuesto inicialmente por (Cleveland, 1979), es un modelo no paramétrico basado en la combinación de métodos clásicos como la regresión lineal y no lineal estimada por mínimos cuadrados. Este procedimiento de suavizado ha sido diseñado para ajustar a los datos la siguiente estructura de regresión local:

Regresión Loess

$$y_i = f(x_i) + \epsilon_i$$

Donde $f(\cdot)$ es una función de suavización y el término ϵ_i corresponde a un error aleatorio con media cero y varianza constante. El supuesto de suavidad permite utilizar puntos en un vecindario de $\{(x_i, y_i); i = 1, \ldots, n\}$ que contiene un número determinado de datos para formar la estimación $\hat{y}_i$ a través de polinomios de bajo grado, de modo que no se asume un modelo global fijo sino variable de manera local. Uno de los primeros pasos para estimar la regresión Loess es definir una función de ponderación para una ventana de suavizado $[x_0 - h(x_0),\ x_0 + h(x_0)]$ considerada para la estimación de $f(x_0)$, donde $h(x_0)$ representa el ancho de banda o span. La función de ponderación tradicional usada en Loess adopta la siguiente forma tricúbica:

Tricube Weight Function

$$W(u) = \begin{cases} (1 - |u|^3)^3 & |u| \leq 1 \\ 0 & |u| > 1 \end{cases}$$

Por lo que la secuencia de ponderación puede definirse como $w_i(x_0) = W\left(\frac{x_i - x_0}{h(x)}\right)$ para los k vecinos más cercanos que incluyen un $\alpha\ \times 100\%$ de los datos, donde $\alpha \in (0,1)$. Un segundo paso para estimar la regresión Loess es determinar el grado de la aproximación polinómica, por ejemplo, para el caso de un polinomio cuadrático se puede adoptar una aproximación usando el teorema de Taylor dado por:

Taylor's Theorem

$$f(x) \approx \beta_0 + \beta_1(x - x_0) + \frac{1}{2}\beta_2(x - x_0)^2, \quad x \in [x_0 - h(x_0), x_0 +\ h(x_0)]$$

Luego, para obtener la estimación de la regresión Loess $\hat{f}(x_0)$ se debe encontrar el vector de parámetros $\boldsymbol{\beta} = (\beta_0\ ,\ \beta_1\ ,\ \beta_2\ )'$ que minimice la suma de los errores residuales ponderados:

Problema de Optimización Loess

$$\widehat{\boldsymbol{\beta}} = \min_{\boldsymbol{\beta} \in \mathbb{R}^3} \sum_{i=1}^{n} w_i(x_0) \left[\overbrace{y_i - \underbrace{\left\{ \beta_0 + \beta_1(x - x_0) + \frac{1}{2}\beta_2(x - x_0)^2 \right\}}_{f(x_0)}}^{\epsilon_i} \right]^2$$

Los hiperparámetros del modelo asociados con el ancho de banda y el grado del polinomio, comúnmente son seleccionados a partir de un proceso de validación cruzada de forma tal que se encuentre la combinación de hiperparámetros que minimizan la raíz del error cuadrático medio. En el contexto de series de tiempo la notación se toma como $\{(t_i, y_i); i = 1, \dots, T\}$ donde $x_i = t_i$.

2. Estimación de la tendencia usando el Filtro de Hodrick-Prescott

La metodología descrita en (Hodrick & Prescott, 1997) es elaborada bajo la proposición de revelar patrones interesantes desde el punto de vista de la teoría económica y sus ciclos económicos. La visión tradicional de las fluctuaciones macroeconómicas sustenta que estas variaciones pueden ser explicadas mediante dos componentes: una componente permanente o de tendencia (τ_t) caracterizada por factores de oferta de la economía y una componente cíclica (c_t) caracterizada por factores de demanda de la economía.

Esta visión considera que en el largo plazo lo que determina el movimiento de las series es la oferta agregada, es decir factores como: cambios tecnológicos y demográficos, productividad total de los factores, entorno institucional, sistema tributario y arancelario, entre otros. Mientras que en el corto plazo lo que caracteriza los movimientos de las series macroeconómicas es la demanda agregada, es decir factores como: el gasto de los hogares en bienes y servicios, el gasto en bienes de capital, inversión, tasa de interés, confianza empresarial, condiciones y expectativas económicas, gasto público, tipo de cambio, entre otros. En este marco teórico, las series macroeconómicas corresponden a la suma de la componente permanente y la componente cíclica:

Descomposición Serie Temporal

$$y_t = \underbrace{\tau_t + c_t}_{T_t:\ trend-cycle}$$

con $t = 1, \dots, T$, en el cual la componente permanente o de tendencia puede encontrarse mediante el siguiente problema de optimización:

Filtro de Hodrick-Prescott

$$\min_{\{\tau_t\}_{t=1}^{T}} \left\{ \sum_{t=1}^{T} \left(\underbrace{y_t - \tau_t}_{c_t} \right)^2 + \lambda \sum_{t=2}^{T-1} \left[\underbrace{(\tau_{t+1} - \tau_t)}_{\Delta\tau_{t+1}} - \underbrace{(\tau_t - \tau_{t-1})}_{\Delta\tau_t} \right]^2 \right\}$$

El filtro de Hodrick-Prescott emplea un parámetro de suavizamiento convencional (λ), lo que implica que puede otorgarle un mayor peso a las variaciones recientes que a las más antiguas. Obsérvese que entre menor sea el parámetro, la componente permanente puede fluctuar más, mientras que entre mayor sea más se penalizarán las fluctuaciones de la tendencia. Las perturbaciones grandes o persistentes pueden modificar la tendencia estimada por el Filtro de Hodrick-Prescott en diversos grados, de manera que la tendencia puede ser interpretada como una aproximación a una media móvil ponderada que representa el crecimiento medio de los últimos tiempos.

Una vez calculada la tendencia, podemos determinar las oscilaciones a corto plazo (c_t) que reflejan las características cíclicas del ciclo económico. Este componente cíclico se refiere a fluctuaciones recurrentes causadas por factores económicos, sociales, institucionales u otros fenómenos que afectan la serie temporal, pero no suceden con una intensidad y frecuencia similar a lo largo del tiempo.

Desviaciones a Corto Plazo

$$\mathrm{c_t} = y_t - \tau_t$$

Donde y_t corresponde a la serie de tiempo desestacionalizada-suavizada y τ_t representa la tendencia calculada mediante el Filtro de Hodrick-Prescott.

3. Desviaciones normalizadas de una serie temporal respecto a su tendencia

El objetivo principal del indicador es medir las desviaciones de la movilización de carga contenerizada que surgen en el corto plazo respecto a su tendencia (c_t). Siguiendo el esquema metodológico adoptado por la *Organización Mundial del Comercio* para el cálculo del barómetro sobre el comercio de mercancías planteado en (WTO, 2019)[1], el indicador es normalizado y centrado en 100 para facilitar su interpretación y comparación, ello denota que:

Barómetro del Comercio

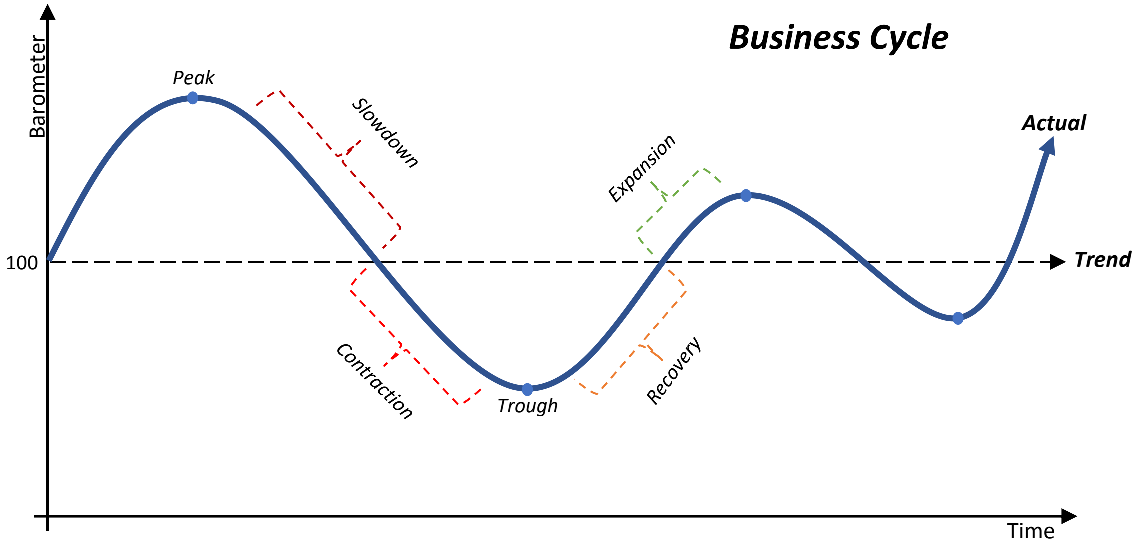
$$Barómetro_t = \left(\frac{c_t - \mu_c}{\sigma_c} + 100\right) \sim N(\mu = 100, \sigma = 1), \quad t = 1, \dots, T.$$

Lo que significa que las desviaciones respecto a la tendencia se expresan en términos de su desviación estándar. Un índice de 100 significa que el comercio se expande acorde a las tendencias recientes, un índice por encima de 100 muestra un crecimiento superior a la tendencia mientras que un índice por debajo de 100 evidencia un crecimiento inferior a la tendencia. La dirección de cambio del indicador

[1] El método empleado para la construcción del barómetro de carga contenerizada en Colombia mantiene fidelidad estadística con la metodología propuesta por la *Organización Mundial del Comercio*.

refleja el impulso de la actividad comercial en comparación con el mes anterior. En consecuencia, el indicador es una guía fundamental para obtener percepciones sobre la trayectoria actual del comercio internacional de carga contenerizada, identificar puntos de inflexión en los sectores económicos y medir su pulso en la actividad comercial del país.

En vista de que la construcción del indicador considera algunos fragmentos de los fundamentos macroeconómicos tradicionales, una guía práctica para una interpretación sencilla del barómetro surge a través de su vínculo con la teoría económica y sus ciclos económicos de la siguiente forma:



Referencias

Box, G. E., & Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control (2nd ed.).* Francisco, CA: Holden-Day.

Cleveland, W. (1979). *Robust Locally Weighted Regression and Smoothing Scatterplots.* Journal of the American Statistical Association, Vol. 74, pp. 829-836.

Hodrick, R. J., & Prescott, E. C. (1997). *Postwar U.S. Business Cycles: An Empirical Investigation.* Journal of Money, Credit and Banking, Vol. 29. Blackwell Publishing.

IMF. (2018). *Quarterly National Accounts Manual 2017.* Washington, DC: International Monetary Fund, Statistics Department.

U.S. Census Bureau. (2017). X-13ARIMA-SEATS Reference Manual. *Center for Statistical Research and Methodology*, Washington, DC.

WTO. (2019). *WTO Goods Trade Barometer Methodology.* World Trade Organization.

Elaborado por

Equipo de Investigaciones Económicas
investigacionyanalisis@sprc.com.co
Octubre 2023