2. NOTA METODOLÓGICA PARA LA ELABORACIÓN DE LAS PROYECCIONES DE EGRESOS

Fundamento Legal

En cumplimiento a lo dispuesto por el Artículo 5, fracción II de la Ley de Disciplina Financiera de las Entidades Federativas y los Municipios, la cual establece que el Proyecto de Presupuesto de Egresos deberá contener una proyección de finanzas públicas, en un horizonte de tiempo de cinco años adicionales al Ejercicio para el que se presupuesta, se presenta la metodología con que fueron realizadas dichas proyecciones.

Base de datos

Para realizar los pronósticos de gasto, se recurrió a un análisis econométrico de series de tiempo, considerando para ello, una base de datos mensual e histórica de las transferencias federales al Estado de Puebla por los conceptos de Ramo General 33 y de Ramo General 28, misma que consta de un lapso de tiempo que abarca desde el año 2002 al mes de julio de 2020. Dicha información fue obtenida en las Estadísticas Oportunas de Finanzas Públicas del Portal de la Secretaría de Hacienda y Crédito Público (SHCP).

Asimismo, se considera también una base de datos de la recaudación por concepto de Impuestos, Derechos, Productos y Aprovechamientos, los cuales pertenecen a los recursos cuya fuente de financiamiento corresponden a los Recursos Fiscales y Otros Ingresos de Libre Disposición, abarcando un periodo que empieza en el mes de enero de 2013 al mes de agosto de 2020. Finalmente, se obtuvo información histórica y mensual del gasto etiquetado diferente al Ramo General 33, misma que corresponde a los Convenios Federales, Fondos Distintos de Aportaciones y Subsidios Federales. Con ello, la información disponible para realizar las proyecciones significó una base de datos con 894 datos observados.

Tratamiento de las Series

De cifras nominales a cifras reales

A fin de modelar adecuadamente las series de tiempo a proyectar, se procedió a deflactar las cifras nominales, con el objetivo de generar una serie de cifras reales. El concepto de deflactar cifras nominales se refiere a eliminar el efecto inflacionario que ocasiona el incremento del nivel general de precios sobre el valor monetario corriente de una cifra. Para ello, se usaron los deflactores implícitos del Producto Interno Bruto (PIB) que se publican en el Banco de Información Económica del INEGI.

$$\boldsymbol{Cifra\ real = \left(\frac{Cifra\ nominal}{Deflactor\ del\ PIB}\right) \times 100}$$

Dado que los deflactores implícitos del PIB se publican con una periodicidad trimestral, se recurrió a deflactar las cifras mensuales con el correspondiente deflactor del trimestre al que pertenecen.

Desestacionalización de series de tiempo

Las series de tiempo son el resultado de tres componentes, la tendencia-ciclo, la estacionalidad y la irregularidad. El primer componente se refiere al comportamiento y los cambios de largo plazo (Islas & Heras, 2012).

Por su parte, la estacionalidad se refiere a las fluctuaciones periódicas de longitud constante y que son causadas por diversos factores, tales como políticas, periodos vacacionales, estación del año, etc. (Islas & Heras, 2012). Gráficamente, una serie de tiempo con componente estacional se identifica como un comportamiento similar y frecuente en cierto periodo de cada año.

Finalmente, la irregularidad es una variable aleatoria cuya ocurrencia carece de explicación alguna y afecta a la serie de tiempo en cuestión.

Dicho lo anterior, el resultado de una serie de tiempo es la suma de sus tres componentes, tal como se indica a continuación:

$$\boldsymbol{O_t = T_t + E_t + I_t}$$

Donde:

$O_t = Serie\ Original\ en\ el\ periodo\ t.$ $T_t = Tendencia - ciclo\ en\ el\ periodo\ t.$ $E_t = Estacionalidad\ del\ periodo\ t.$ $I_t = Irregularidad\ del\ periodo\ t.$

Para desestacionalizar las series asociadas al comportamiento histórico de los Ramos Generales 28 y 33, así como el comportamiento de los Recursos Fiscales y Recursos Federales distintos a Aportaciones, se procedió a emplear la metodología Census X-12, desarrollada por la Oficina del Censo de EE. UU. Dicho método consiste en la aplicación de un modelo Autorregresivo Integrado de Medias Móviles, mejor conocido como modelo ARIMA (p, d, q), el cual considera un proceso de retropolación y extrapolación de la muestra, para finalmente aplicar un suavizamiento de la serie con base en promedios móviles (Guerrero, 1990).

Asimismo, se consideró también la metodología propuesta por Islas & Heras (2012), la cual consiste en aplicar a la serie original un proceso de suavizamiento a base de promedios móviles centrados de 12 meses, obteniendo así, la tendencia-ciclo.

$$\boldsymbol{T_t = \frac{\sum_{s=-6}^{s=5} O_{t\pm s}}{12}}$$

Posteriormente, restando de la serie original la tendencia-ciclo, se obtiene la estacionalidad e irregularidad.

$$\boldsymbol{O_t - T_t = E_t + I_t}$$

Con la finalidad de identificar el patrón estacional que tiene cada mes, el cual es repetitivo en cada año, se procede a promediar las observaciones de la estacionalidad e irregularidad por mes. Así, el patrón estacional del mes j se obtiene como:

$$\boldsymbol{E_j = \frac{\sum_y^Y (E_{j_y} + I_{j_y})}{n}; \quad \forall j = j\ y\ \forall y = y}$$

Donde:

$E_j = Factor\ estacional\ del\ mes\ j.$ $\left(E_{j_y} + I_{j_y}\right) = Estacionalidad\ e\ Irregularidad\ del\ mes\ j\ en\ el\ año\ y.$ $n = Número\ de\ años\ contemplados\ en\ la\ muestra.$

Finalmente, la serie desestacionalizada se obtiene como:

$$\boldsymbol{Serie\ desestacionalizada = O_t - E_j}$$

Modelo Autorregresivo con Rezagos Distribuidos (ARDL)

Modelo Autorregresivo (AR)

Una serie de tiempo se plantea como autorregresiva cuando la variable en cuestión se explica por sí misma (Carter Hill, Griffiths & Lim, 2011). Particularmente, para cualquier y_t, su comportamiento puede explicarse por sus rezagos (valores pasados) y_{t-s}.

$$\boldsymbol{y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-s})}$$

De manera más formal, el modelo autorregresivo planteado se define de la siguiente manera:

$$\boldsymbol{y_t = \alpha + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \cdots + \varepsilon_t}$$

$$\boldsymbol{y_t = \alpha + \sum_{s=1}^{S} \gamma_s y_{t-s} + \varepsilon_t}$$

Donde el término ε_t corresponde al error aleatorio cuyo valor esperado es igual a cero y cuya varianza en el tiempo es constante e igual a σ^2.

Dichos retardos en la variable explican la dinámica del modelo mediante una ponderación de su valor, de tal forma que el efecto de cada valor puede aislarse e interpretarse como el cambio esperado del valor y_t con respecto a un cambio en y_{t-s} (Carter Hill, Griffiths & Lim, 2011).

$$\boldsymbol{\frac{\partial y_t}{\partial y_{t-s}} = \gamma_s}$$

Dichos ponderadores de rezagos son estimados a partir del método de los Mínimos Cuadrados Ordinarios (MCO).

Modelo de Rezagos Distribuidos (DL)

A diferencia del modelo autorregresivo, el modelo con rezagos o retardos distribuidos explica la serie de tiempo asociada a y_t está en función de k variables explicativas y no de sí misma (Carter Hill, Griffiths & Lim, 2011), de tal manera que $y_t = f\left(X_{1_{t-s}}, X_{2_{t-s}}, \dots, X_{k_{t-s}}\right);\ \forall s \geq \mathbf{0}$, o bien:

$$y_t = \alpha + \beta_{1,0}X_{1_t} + \beta_{1,1}X_{1_{t-1}} + \cdots + \beta_{1,s}X_{1_{t-s}} + \beta_{2,0}X_{2_t} + \beta_{2,1}X_{2_{t-1}} \ldots + \beta_{2,s}X_{2_{t-s}} + \cdots + \beta_{k,0}X_{k_t} + \beta_{k,1}X_{k_{t-1}} + \cdots + \beta_{k,s}X_{k_{t-s}} + \cdots \varepsilon_t$$

$$y_t = \alpha + \sum_{s=0}^{s} \beta_{k,s} X_{k_{t-s}} + \varepsilon_t$$

De la misma manera que el modelo autorregresivo, los valores retardados de las variables independientes explican la dinámica del modelo mediante la ponderación de estas, de tal manera que el efecto aislado de cada valor en el tiempo de la variable X_k sobre y_t puede interpretarse como sigue:

$$\frac{\partial y_t}{\partial X_{k_{t-s}}} = \beta_{k,s}$$

Modelo ARDL (p, q)

El modelo autorregresivo con rezagos distribuidos es un modelo integral de los procesos AR y DL vistos anteriormente, por lo que, de manera formal el valor de la serie se define como:

$$y_t = \alpha + \gamma_1 y_{t-1} + \cdots + \gamma_s y_{t-s} + \beta_{k,0}X_{k_t} + \beta_{k,1}X_{k_{t-1}} + \cdots + \beta_{k,s}X_{k_{t-s}} + \varepsilon_t$$

$$y_t = \alpha + \sum_{s=1}^{s} \gamma_s y_{t-s} + \sum_{s=0}^{s} \beta_{k,s} X_{k_{t-s}} + \varepsilon_t$$

Pronósticos en los Modelos ARDL (p, q)

El modelo en cuestión es un proceso muy útil para generar pronósticos; no obstante, se requiere trabajar con valores futuros de las variables explicativas dado que la dinámica del modelo lo exige (Carter Hill, Griffiths & Lim, 2011). Por ello, se simularon los valores futuros de la Recaudación Federal Participable en función del comportamiento del PIB y su crecimiento esperado, de acuerdo con los Criterios Generales de Política Económica para el Ejercicio Fiscal 2021. Estos valores obtenidos, fueron empleados como variable explicativa.

Consideraciones Generales

Toda vez que las series de tiempo suelen tener un comportamiento determinista, se procede a desestacionarizar la serie, a fin de evitar el problema de regresiones espurias (Carter Hill, Griffiths & Lim, 2011). Se dice que una serie es estacionaria si su valor esperado y su varianza es estable en el tiempo. De este modo, una serie determinista debe diferenciarse con su valor rezagado inmediato, obteniendo lo siguiente:

$$\Delta y_t = y_t - y_{t-1}$$

Si una serie debe diferenciarse una vez para llegar a ser estacionaria, se dice que la serie tiene un orden de integración igual a 1. Si esta debe diferenciarse dos veces para ser estacionaria, su orden de integración será igual a 2, y sucesivamente (Carter Hill, Griffiths & Lim, 2011).

Para verificar que la serie de tiempo que se trata es estacionaria, se aplica la Prueba de Raíz Unitaria Dickey-Fuller, en la cual, se prueba que el coeficiente asociado al primer rezago sea menor a 1 en valor absoluto, lo que asegura que el proceso tiene un comportamiento regular y predecible en el tiempo (Carter Hill, Griffiths & Lim, 2011). Formalmente, se demuestra lo siguiente:

$$y_t = \rho y_{t-1} + \varepsilon_t \text{ es estacionario } \leftrightarrow |\rho| < 1$$

Restando y_{t-1} de ambos lados de la ecuación, obtenemos lo siguiente:

$$y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + \varepsilon_t$$
$$\Delta y_t = (\rho - 1) y_{t-1} + \varepsilon_t$$
$$H_0 = (\rho - 1) = 0 \rightarrow \text{El proceso no es estacionario}$$
$$H_1 = (\rho - 1) < 1 \rightarrow \text{El proceso es estacionario}$$

Asimismo, para la selección del número de retardos a incluir en cada modelo, se procedió a tomar el valor mínimo de los siguientes dos criterios a cada modelo de regresión aplicado:

$$\text{Criterio de Akaike } (AIC) = \frac{Ln(SSE)}{T} + \frac{2k}{T}$$

$$\text{Criterio de Schwartz } (SC) = \frac{Ln(SSE)}{T} + \frac{k * Ln(T)}{T}$$

Bibliografía

- Guerrero, V. (1990). *Desestacionalización de series de tiempo económicas: introducción a la metodología*. Comercio Exterior, 1035-1046.
- Islas, J., & Heras, M. (2012). *Descomposición de Series de Tiempo*. Cd. de México: Facultad de Economía, UNAM.
- Carter Hill, R., Griffiths W., & Lim, G. (2011). *Principles of Econometrics*. EE.UU. John Wiley & Sons, Inc.
- Gujarati, D., & Porter, D. (2010). *Econometría*. México, D. F. McGraw-Hill.
- Stata Inc. (2013). *Stata time-series reference manual. Release 13*. Texas. Stata Press.