



**TRIBUTAR ES
PROGRESAR**

Nota Metodológica

Pronósticos Uniecuacionales para los Ingresos Tributarios de Honduras

Jose Carlo Bermúdez

Julio, 2020



www.sar.gob.hn

Pronósticos Uniecuacionales para los Ingresos Tributarios de Honduras

Jose Carlo Bermúdez*

Julio, 2020

Resumen

Este documento presenta un modelo Autoregresivo de Rezagos Distribuidos (ARDL) para pronosticar los ingresos tributarios de Honduras empleando datos de energía y actividad económica. A partir de la metodología de cointegración en niveles se encuentra la presencia de una relación de largo plazo entre las variables analizadas. Posteriormente, la capacidad predictiva de dicho modelo es sometida a prueba y comparada con diez especificaciones diferentes de modelos uniecuacionales. Los resultados sugieren que el modelo obtenido mediante la metodología de cointegración es el que mejor desempeño exhibe dentro de muestra. De hecho, para los cuatro horizontes de tiempo evaluados, las diferentes especificaciones de modelos ARDL parecen superar la capacidad predictiva de los modelos estimados mediante la metodología Box-Jenkins salvo para el horizonte de un trimestre, en el que un modelo ARIMA ofrece mejores proyecciones.

Palabras Clave: ARDL, ingresos tributarios, Honduras.

Clasificación JEL: C52, C53, E37.

*Email: jbermudez@sar.gob.hn. Dirección Nacional de Gestión Estratégica, Servicio de Administración de Rentas (SAR), Honduras. Las opiniones vertidas en el presente documento no representan la opinión institucional del SAR puesto que pertenecen al autor, así como cualquier error u omisión. Este trabajo ha sido publicado en la Revista de Administración Tributaria, véase: <https://biblioteca.ciat.org/opac/book/5743>

1 Introducción

La proyección de los ingresos tributarios es una de las tareas más relevantes al momento de definir los lineamientos para la política presupuestaria de un país. En este sentido, las oficinas de proyecciones de los ministerios de hacienda o bien, las administraciones tributarias usualmente establecen umbrales recaudatorios para períodos no mayores a un año, aunque la programación financiera si suele incluir pronósticos fiscales mayores a dicho horizonte, incluso, los tipos de proyecciones pueden ser de simulación, predicción y para establecimiento de metas (Martín et al., 2012, pp. 11-12). Aunque no existe una metodología específica ampliamente aceptada para realizar las mejores estimaciones de recaudación (puesto que está sujeto a múltiples factores idiosincráticos de la economía y la normativa tributaria de cada país) parece que, *de facto*, predominan el método directo, la extrapolación mecánica, la aplicación de elasticidades, boyanzas y en menor medida, modelos econométricos de series temporales.

En el caso del método directo, las estimaciones suelen realizarse bajo la consideración de información tributaria pormenorizada y con una preponderante aplicación del juicio de experto. La extrapolación mecánica implica usualmente que en el período siguiente la recaudación tendrá un comportamiento similar a lo observado en períodos pasados, esto es, un importante sesgo hacia un patrón determinístico en la recaudación. Las elasticidades¹, que encuentran su genealogía en el trabajo pionero de Groves and Kahn (1952), se refieren a las variaciones porcentuales que experimentan las recaudaciones de impuestos ante movimientos en sus bases gravables y suelen ser estimadas mediante la metodología de cointegración propuesta por Engle and Granger (1987), a su vez dichas elasticidades se incluyen en una ecuación de boyanza también conocida como función de reacción, la cuál permite extrapolar el impuesto deaseado para el período siguiente. De igual manera, las estimaciones econométricas de series temporales permiten pronosticar la recaudación aplicando la metodología de Box and Jenkins (1973), ya sea modelando los impuestos como un proceso Autoregresivo (AR(p)), de Media Móvil (MA(q)) o combinado (ARIMA(p, q)).

El Servicio de Administración de Rentas (SAR), ha empleado los métodos previamente enunciados para diferentes ejercicios de simulación y pronóstico, no obstante, la capacidad predictiva de dichas metodologías no ha sido evaluada con profundidad hasta ahora. En virtud de ello, el presente documento tiene como objetivo el análisis de las estimaciones de ingresos tributarios para Honduras

¹Véase Cardoza (2017) para conocer sobre evidencia empírica de estimaciones de elasticidades y boyanzas de los ingresos tributarios para algunos países miembros del CIAT (Centro Interamericano de Administraciones Tributarias) y Lagravinese et al. (2020) para el caso de países miembro de la OCDE (Organización para la Cooperación y el Desarrollo Económico).

basándose estrictamente en la metodología econométrica de series temporales. Asimismo, se estima un modelo uniecuacional Autorregresivo de Rezagos Distirbuidos (ARDL por sus siglas en inglés) siguiendo la metodología de cointegración a la [Pesaran and Shin \(1999\)](#) y [Pesaran et al. \(2001\)](#) permitiendo identificar la relación de corto y largo plazo entre la recaudación de ingresos tributarios con la oferta y demanda de la economía, al menos en un sentido laxo. A esta línea de modelos [Glenday \(2013\)](#) les denomina como *GDP based models* (modelos basados en el Producto Interno Bruto (PIB)).

Cabe soslayar que el modelo ARDL estimado excluye componentes de legislación tributaria más allá de los que se encuentran implícitos dentro del proceso generador de datos de la recaudación, tampoco incluye variables consideradas como bases gravables, sean de carácter doméstico y/o externo². De esta forma, se ha optado por un modelo parsimonioso y de aplicación para los ingresos tributarios totales (recaudados por el SAR y la administración aduanera³) de forma tal que se pueda ofrecer un marco de análisis robusto comparando diferentes especificaciones econométricas uniecuacionales, como han hecho otros trabajos previamente (véase por ejemplo [Streimikiene et al. \(2018\)](#)). No obstante, la construcción de un modelo estructural más complejo que considere diferentes interrelaciones macroeconómicas y de normativa tributaria es parte de la agenda de investigación futura.

La riqueza del presente estudio se concentra en la posibilidad de contrastar el modelo econométrico más adecuado para ejercicios de pronóstico y simulación de corto plazo del SAR dentro de 11 especificaciones diferentes. Esto no es un ejercicio menor, sobre todo cuando no parece haber un consenso específico en la literatura para la proyección de variables fiscales ya que algunos autores identifican la presencia de un alto sesgo positivo en las proyecciones, el cual tiende a incrementarse en períodos de fuertes expansiones económicas y es aún mayor cuando se revisan pronósticos de mediano plazo ([Frankel, 2011](#)); en cambio, otros encuentran que los modelos de proyección basados en el PIB tienden a ser pesimistas, sobre todo en fases de crecimiento por debajo del promedio ([Brender and Navon, 2010](#)). Lo anterior refuerza la importancia de este trabajo ya que además abre las puertas para que técnicos externos puedan revisar e identificar posibles puntos de mejora a la metodología presentada.

El documento prosigue de la siguiente manera: la sección 2 desarrolla la especificación de los modelos econométricos a emplear; la sección 3 presenta el análisis de los datos; la sección 4 expone con detalle los resultados de las estimaciones, finalmente, la sección 5 concluye.

²[Cerdea et al. \(2019b\)](#) encuentran que incluir variables externas como el tipo de cambio y los precios del petróleo puede ser importante para explicar de los errores de pronóstico en las proyecciones de los ingresos tributarios de países exportadores de materias primas

³Luego de la supresión y liquidación de la Dirección Ejecutiva de Ingresos (DEI) mediante Decreto Ejecutivo Número PCM-083-2015, ambas administraciones fueron separadas.

2 Especificación Econométrica

2.1 Modelo ARDL

Siguiendo a [Pesaran and Shin \(1999\)](#) y [Pesaran and Shin \(2002\)](#) se emplea un modelo ARDL como base para el pronóstico de los ingresos tributarios de Honduras. En este sentido, la metodología de *bounds test* (prueba de límites) desarrollada por [Pesaran et al. \(2001\)](#) permite identificar una relación de cointegración entre diversas variables económicas a través del tiempo sin tomar en cuenta las características del proceso generador de datos, esto es, indistintamente de su orden de integración, lo cual es relevante cuando se realizan estimaciones con métodos de inferencia estándar cuyas propiedades asintóticas son sensibles a la presencia de raíces unitarias. Adicionalmente, dicha técnica es una opción deseable en comparación a la de [Engle and Granger \(1987\)](#) ya que ésta última implica hacer análisis de frecuencia cero para identificar que las variables a incluir dentro del modelo de corrección de errores tengan el mismo orden de integración abriendo paso a la potencial identificación de tendencias espurias que anulan el posterior análisis económico y de pronóstico.

El modelo ARDL(p, q) a estimar es definido en su forma de ajuste parcial:

$$T_t = c_0 + c_1 t + \sum_{i=1}^p \phi_i T_{t-i} + \sum_{i=0}^q \beta_i' \mathbf{x}_{t-i} + u_t \quad (1)$$

Donde $p \geq 1$ y $q \geq 0$; T_t son los ingresos tributarios en niveles, c_0 y c_1 son coeficientes; t es una tendencia lineal del tipo $t = \{1, 2, 3, \dots, n\}$ siendo $n \subseteq \mathbb{N}$ el tamaño de la muestra empleada; \mathbf{x}_{t-i} es un vector de tamaño $k \times 1$ que contiene a las variables que explican la recaudación tributaria, mientras que $u_t \sim \mathcal{N}(0, \sigma^2)$ son los errores del modelo en tanto que se consideran no correlacionados. Hay que mencionar que el vector $\mathbf{x}_{t-i} \equiv \mathbf{x}_t' = \{y_t, en_t\}$ incluye datos de actividad económica (y_t) y de energía eléctrica (en_t) para aproximar la demanda y la oferta de la economía. Modelos para la proyección de ingresos tributarios previos incluyen datos de producción y fuentes de volatilidad externa como el tipo de cambio y los precios del petróleo (véase [Cerda et al. \(2019a\)](#)), sin embargo, para fines de las estimaciones en esta oportunidad si bien se incluye una variable que aproxima la actividad económica, se excluyen variables externas añadiendo el consumo de energía para aproximar los costos de producción de los agentes económicos lo cual es abordado con mayor amplitud en la sección 3.

La metodología de cointegración en niveles de [Pesaran and Shin \(1999\)](#) y [Pesaran and Shin \(2002\)](#) requiere una adecuada identificación a la dinámica del modelo lo cual es posible conociendo el número

de rezagos óptimos (p, q) de la ecuación 1. No obstante, se debe tener en cuenta que el vector \mathbf{x}_{t-i} incluye dos variables explicativas, por lo que la estructura de rezagos es del tipo (p, q, d) . Para tal finalidad se emplea el criterio de información de Akaike (AIC).

Una vez que se conoce la estructura de rezagos óptimos del modelo ARDL, se debe verificar la presencia de cointegración en niveles empleando la prueba de límites de Pesaran et al. (2001). Para ello, conviene reparametrizar la ecuación 1 en su forma de corrección de error, la cual es conocida como modelo condicional de corrección de errores (CEC por sus siglás en inglés), así:

$$\Delta T_t = c_0 + c_1 t - \alpha(T_{t-1} - \theta \mathbf{x}_{t-1}) + \sum_{i=1}^{p-1} \psi_{Ti} \Delta T_{t-i} + \omega' \Delta \mathbf{x}_t + \sum_{i=1}^{q-1} \psi_{xi} \Delta \mathbf{x}_{t-i} + u_t \quad (2)$$

Notar que en la ecuación anterior se incluye el parámetro $\alpha = 1 - \sum_{j=1}^p \phi_j$ el cual proviene de operar la ecuación 1 para la variable T_t y que sirve como el parámetro de corrección de error que captura la velocidad de ajuste del modelo, que dicho en otras palabras, es la velocidad a la que los ingresos tributarios y sus variables explicativas vuelven a su equilibrio de largo plazo o estado estacionario. De igual forma, $\theta = \frac{\sum_{j=1}^q \beta_j}{\alpha}$ es un vector que contiene los parámetros de largo plazo para cada una de las variables explicativas y dado que $\theta < \infty$, los choques de las variables contenidas en \mathbf{x}'_t son transitorios (finitos) sobre los ingresos tributarios.

Es importante que u_t en 2 sean no correlacionados y posean varianza constante (homocedasticidad), caso contrario los estimadores serán ineficientes aunque insesgados y consistentes. Otra condición necesaria previo a verificar la presencia de cointegración en el modelo 2 es que las raíces características $1 - \sum_{j=1}^p \phi_j$ se encuentren dentro del círculo unitario, no obstante, dicha condición no es directamente verificable, por lo que la estabilidad es corroborada mediante la presencia de cointegración empleando una prueba F contra los valores críticos de Pesaran et al. (2001), estos valores implican un contraste de hipótesis en el que si el valor del estadístico F se ubica debajo del valor crítico inferior para $I(0)$, entonces no se verifica la presencia de cointegración, en cambio, si dicho valor F se ubica por encima de los valores críticos del límite superior $I(1)$ se verifica una relación de largo plazo entre T_t y \mathbf{x}'_t .

Luego, la metodología permite proceder con el análisis de inferencia y pronóstico considerando un modelo dinámico que representa una estructura óptima. Dicho esto, la ecuación final en la que se apoyan las estimaciones de proyección de los ingresos tributarios para Honduras viene dada por:

$$T_t = c_0 + c_1 t + \sum_{i=1}^s \gamma_i D_{it} + \sum_{i=1}^p \phi_i T_{t-i} + \sum_{i=0}^d \varphi_i e n_{t-i} + \sum_{i=0}^q \delta_i y_{t-i} + \epsilon_t \quad (3)$$

Dado que la ecuación 3 es estimada con las variables rescaladas en niveles (aplicando logaritmo natural) y con una frecuencia mensual, se incluye D_{it} que es un vector de dummies estacionales donde $s = \{\gamma_1 = 1 \text{ si } t=\text{enero y } \gamma_1 = 0 \text{ otro mes}, \dots, \gamma_{12} = 1 \text{ si } t=\text{diciembre y } \gamma_{12} = 0 \text{ otro mes}\}$. No obstante, al momento de realizar las estimaciones se emplean únicamente $s/2$ variables dummies con la finalidad de evitar caer en la trampa de variables ficticias lo que podría traducirse en multicolinealidad perfecta.

2.2 Modelos Complementarios

Como se ha indicado al comienzo del documento, el ejercicio consiste en comparar la precisión de los pronósticos obtenidos mediante el modelo ARDL respecto a una serie de especificaciones econométricas alternativas, para ello, se apela a estimar cinco modelos autorregresivos diferentes construídos de la siguiente manera:

1. **Modelo AR(p)**. Empleando la metodología de [Box and Jenkins \(1973\)](#) se estima la ecuación:

$$T_t = \phi_0 + \sum_{i=1}^p \phi_i T_{t-p} + v_t \quad (4)$$

2. **Modelo MA(q)**. Empleando la metodología de [Box and Jenkins \(1973\)](#) se estima la ecuación:

$$T_t = \theta_0 + \sum_{i=1}^q \theta_i \epsilon_{t-q} + u_t \quad (5)$$

3. **ARMA(p, q)**. Empleando la metodología de [Box and Jenkins \(1973\)](#) se estima la ecuación:

$$T_t = c_0 + \sum_{i=1}^p \phi_i T_{t-p} + \sum_{i=1}^q \theta_i \epsilon_{t-q} + u_t \quad (6)$$

4. **ARIMAX(p, d, q)**. Tomando la actividad económica y el consumo de energía como variables explicativas del proceso ARIMA(p, q) se estima cada una de las siguientes ecuaciones de manera independiente:

$$T_t = \phi_0 + \phi_1 y_t + \sum_{i=1}^p \phi_i T_{t-p} + \sum_{i=1}^q \theta_i \epsilon_{t-q} + u_t \quad (7)$$

$$T_t = \phi_0 + \phi_1 en_t + \sum_{i=1}^p \phi_i T_{t-p} + \sum_{i=1}^q \theta_i \epsilon_{t-q} + v_t \quad (8)$$

2.3 Evaluación de los Pronósticos

Con la finalidad de verificar cuál es el modelo que ofrece las mejores posibilidades de pronóstico, la literatura de series de tiempo propone una serie de medidas, este trabajo emplea las que se enuncian a continuación:

1. Raíz del Error Cuadrático Medio (RECM):

$$RECM = \sqrt{\frac{1}{n} \sum_{h=1}^n (T_{t+h} - \hat{T}_{t+h})^2}$$

2. Error Absoluto Medio (EAM)

$$EAM = \frac{1}{n} \sum_{h=1}^n |T_{t+h} - \hat{T}_{t+h}|$$

3. Error Porcentual Absoluto Medio (EPAM)

$$EPAM = \frac{1}{n} \frac{\sum_{h=1}^n |T_{t+h} - \hat{T}_{t+h}|}{T_{t+h}}$$

4. Coeficiente de Desigualdad de Theil (THEIL)

$$THEIL = \frac{RECM}{\sqrt{\frac{1}{n} \sum_{h=1}^n \hat{T}_{t+h}^2 + \frac{1}{n} \sum_{h=1}^n T_{t+h}^2}}$$

En este punto se debe tener presente que T_{t+h} es la variable de ingresos tributarios, mientras que $t+h$ denota el horizonte mensual de pronóstico en el que se evalúa cada medida, para el presente análisis se hacen pronósticos para $h = 1, 3, 6, 12$ meses respectivamente. Entre menor sean las medidas de evaluación, mejor es el ajuste del pronóstico en el modelo evaluado.

3 Análisis de Datos

Para la estimación del modelo ARDL se emplean datos de la recaudación tributaria total (T_t) que es la suma de los tributos internos más aduaneros, el producto (y_t) es aproximado por el Índice Mensual de Actividad Económica (IMAE) y el consumo de energía (en_t) es aproximado por las ventas totales de energía medidas en Megavatios por hora (Mwh). Los datos de recaudación han sido tomados de registros del SAR, el IMAE se obtiene del Sistema de Información Macroeconómica y Financiera de la Región (SIMAFIR) publicada por la Secretaría Ejecutiva del Consejo Monetario Centroamericano (SECMCA) y finalmente, los datos de energía han sido tomados de los boletines estadísticos mensuales publicados por la Empresa Nacional de Energía Eléctrica (ENEE)⁴. La datos se organizan en una frecuencia mensual con una muestra que abarca desde enero de 2007 hasta diciembre de 2019 (2007m1-2019m12). Todas las variables han sido reescaladas en logaritmos (véase Anexo 6.1).

El hecho de emplear el IMAE y las ventas de energía como variables explicativas de la recaudación responde a la intención de aproximar la dinámica productiva en general. En este sentido, se espera que el IMAE, al ser un indicador de alta frecuencia, capture la trayectoria de la producción agregada, mientras que las ventas de energía pueden servir como un proxy al comportamiento de la oferta de la economía que incide sobre los costes medios de producción. Con el ánimo de caracterizar la interconexión entre estas variables y la recaudación de impuestos se propone una estrategia sencilla al aplicar un ejercicio de correlaciones móviles o continuas (*rolling correlations*) para una ventana móvil de doce meses calculada como:

$$\hat{\rho}_t^1(n) = \frac{\hat{\sigma}_{Ty,t}(n)}{\hat{\sigma}_{T,t}(n)\hat{\sigma}_{y,t}(n)} \quad (9)$$

$$\hat{\rho}_t^2(n) = \frac{\hat{\sigma}_{Ten,t}(n)}{\hat{\sigma}_{T,t}(n)\hat{\sigma}_{en,t}(n)} \quad (10)$$

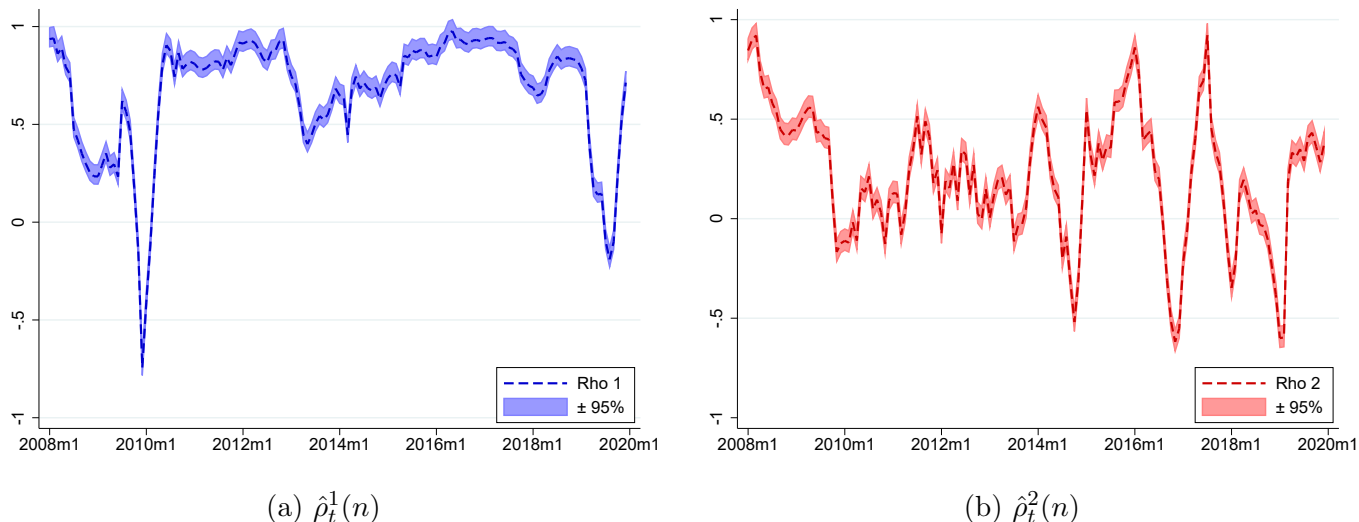
$$t = n, n + 1, \dots, N$$

Donde $\hat{\rho}_t^1(n)$ es la rolling correlation entre la recaudación y el IMAE, mientras que $\hat{\rho}_t^2(n)$ es la rolling correlation entre la recaudación y las ventas de energía. El panel del gráfico 1 muestra las correlaciones obtenidas junto a un intervalo de confianza del 95%. Como primer hecho relevante, los resultados sugieren que las relaciones entre la recaudación con la actividad económica y la venta de energía no han sido constantes a lo largo de la muestra empleada. Específicamente, en el caso del IMAE se

⁴Véase <http://www.enee.hn/index.php/planificacionicono/182-boletines-estadisticos>. Las ventas totales de energía es la suma de las ventas al sector residencial, comercial, industrial, altos consumidores, alumbrado público, gobierno, entes autónomos y sector municipal.

observa una marcada modificación en el patrón de la correlación a lo largo de 2009 como resultado de la recesión económica global, a lo largo de la ventana móvil del año 2019 la correlación exhibió nuevamente un coeficiente menor a cero posiblemente reflejando la desaceleración de la economía nacional como resultado de menores precios en las materias primas y la desaceleración de la presión tributaria. Al margen de los eventos descritos, la correlación se mantiene estadísticamente significativa en la mayoría de la muestra y muy cercana a la unidad, con un promedio de 0.7. En cambio, al revisar la correlación entre la recaudación de impuestos y las ventas de energía se percibe un patrón más irregular, aunque en la mayoría de la muestra la interrelación es estadísticamente significativa; curiosamente, luego de la crisis de 2009 se tornó más común observar ventanas con coeficientes menores a cero, de hecho, el promedio para todo el período asciende a 0.2.

Gráfico 1: Correlaciones Móviles de la Recaudación



FUENTE: estimaciones propias con base a datos del SAR, SIMAFIR y ENEE para una ventana móvil de 12 meses. Las áreas sombreadas representan los intervalos de confianza calculados como $\hat{\rho}_t^i(n) \pm z_\alpha \frac{\sigma}{\sqrt{n}}$, donde $z_\alpha = 1.96$ dada una distribución normal, $\frac{\sigma}{\sqrt{n}}$ es el error estándar de ρ_t^i con $i \in \{1, 2\}$ calculado mediante *bootstrapping* con 400 repeticiones. Así, el valor del error estándar superior es el promedio del intervalo superior obtenido en el remuestreo, mismo caso para el error estándar inferior. Las correlaciones se calculan sobre las variables desestacionalizadas mediante TRAMO-SEATS y en logaritmos.

3.1 Pruebas de Raíz Unitaria

A pesar que la metodología para la estimación del modelo ARDL usando las pruebas de límites no implica que las variables incluídas dentro del modelo tengan el mismo orden de integración como una *conditio sine qua non*, aún es necesario conocer el tipo de tendencia que tienen las mismas debido a que los valores críticos de Pesaran et al. (2001) no son aplicables a variables con un orden de integración mayor a dos (debe cumplirse que $T_t, y_t, en_t \sim I(d); d < 2$). En este sentido, uno de los mayores riesgos

implícitos en la aplicación de toda prueba de raíz unitaria es la identificación de tendencias espurias, sobre todo en la presencia de cambios estructurales o bien, cuando no se dispone de una muestra grande. Como alternativa, este trabajo se apoya en la rutina de [Dolado et al. \(1990\)](#), cuyo algoritmo conlleva la estimación de tres especificaciones diferentes de la prueba de [Dickey and Fuller \(1979\)](#). La tabla 1 muestra los resultados, concluyendo que las tres variables son estacionarias en tendencia ($T_t, y_t, en_t \sim I(0)$), por lo que el modelo ARDL incluye un componente determinístico t que permite remover la tendencia lineal garantizando la estacionariedad de las variables.

Tabla 1: Pruebas de Raíz Unitaria

Variable	Rezagos	Regresor	(1)	(2)	(3)	$I(d)$
T_t	2	T_{t-1}	-0.0576 (-1.79)	-0.0011 (-0.17)	0.0012 (4.07)	I(0)
		ΔT_{t-1}	-0.3069 (-3.70)	-0.3404 (-4.18)	-0.3417 (-4.21)	
		ΔT_{t-2}	-0.1055 (-1.30)	-0.1262 (-1.55)	-0.1268 (-1.57)	
		c_0	0.4648 (1.83)	0.0198 (0.37)		
		t	0.0005 (1.79)			
y_t	3	y_{t-1}	-0.1228 (-2.35)	-0.0012 (-0.16)	0.0009 (4.19)	I(0)
		Δy_{t-1}	-0.4816 (-5.38)	-0.5688 (-6.88)	-0.5703 (-6.93)	
		Δy_{t-2}	-0.1093 (-1.14)	-0.1699 (-1.81)	-0.1708 (-1.83)	
		Δy_{t-3}	0.0839 (1.02)	0.0496 (0.60)	0.0494 (0.60)	
		c_0	0.6241 (2.36)	0.0117 (0.28)		
		t	0.0003 (2.35)			
en_t	1	en_{t-1}	-0.6145 (-6.77)	-0.1035 (-2.51)	0.0002 (0.58)	I(0)
		Δen_{t-1}	-0.0053 (-0.07)	-0.2621 (-3.35)	-0.3126 (-4.06)	
		c_0	7.925 (6.77)	1.352 (2.51)		
		t	0.0012 (6.16)			

FUENTE: estimaciones propias con datos del SAR, SIMAFIR y ENEE. Entre paréntesis se presenta el t estimado. El número de rezagos se obtienen mediante los criterios de información bayesiano (BIC) y de Akaike (AIC). La especificación (1) se refiere a una ecuación que incluye constante (c_0) y tendencia (t); (2) excluye la tendencia y (3) excluye tendencia e intercepto. El valor crítico para (1) es -3.443, para (2) es -2.886 y para (3) es -1.950 para un 95% de confianza.

4 Resultados

4.1 Modelo ARDL

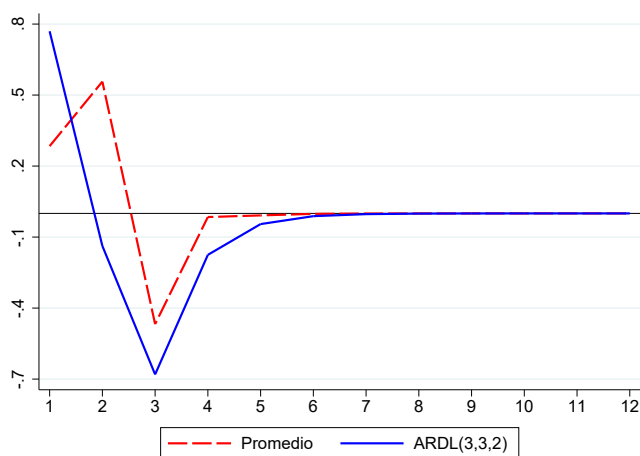
En primera instancia, se estima la ecuación 1 para identificar la estructura ARDL (p, q, d) adecuada. Empleando el criterio de información AIC y luego de 100 iteraciones se concluyó que el mejor modelo posee la especificación del tipo $(3, 3, 2)$, esto permitió estimar inmediatamente el modelo CEC expuesto a través de la ecuación 2. Dicho modelo se obtiene mediante Mínimos Cuadrados Ordinarios (MCO) y sus resultados se presentan en la tabla 2 el cuál resultó poseer errores no correlacionados y homocedásticos (véase tablas 7 y 8 del anexo 6.2), condiciones necesarias para garantizar estimadores eficientes. De esta manera, la prueba de límites de Pesaran et al. (2001) que estima tanto una prueba F como una prueba t para la significancia del coeficiente de corrección de error $\alpha = 1 - \sum_{j=1}^p \phi_j$ sugiere la presencia de una relación de cointegración en niveles entre los ingresos tributarios, el IMAE y las ventas de energía en Mwh puesto que los valores absolutos estimados para la prueba F y para la prueba t exceden los valores críticos de los límites para el umbral $I(1)$ (véase tabla 9 del anexo 6.2). Esta evidencia confirma una relación estable de largo plazo entre las variables bajo estudio.

La estimación para el coeficiente de corrección de error posee el signo esperado y resultó ser estadísticamente significativo, por lo que es posible afirmar que el modelo muestra una valocidad de ajuste del 69% luego de que los ingresos tributarios experimenten algún desvío de su valor de estado estacionario. La magnitud en la velocidad de ajuste es elevada de manera coherente para un modelo que es estimado con una frecuencia mensual. Por otra parte, los coeficientes de largo plazo (θ) del modelo de corrección de errores indican que la energía posee una mayor magnitud en el efecto de equilibrio sobre la recaudación respecto al presentado por el IMAE. Esto puede sugerir que los choques de oferta en la economía tienen una mayor incidencia sobre la trayectoria de largo plazo de la recaudación respecto al efecto que poseen los choques de demanda que por construcción suelen ser de muy corto plazo, es por ello que el coeficiente de largo plazo del IMAE es cercano a cero y no rechaza la hipótesis nula de ser estadísticamente igual a cero (no significativo). No obstante, dicha reflexión deberá ser probada mediante una metodología econométrica que escapa del alcance de la presente investigación.

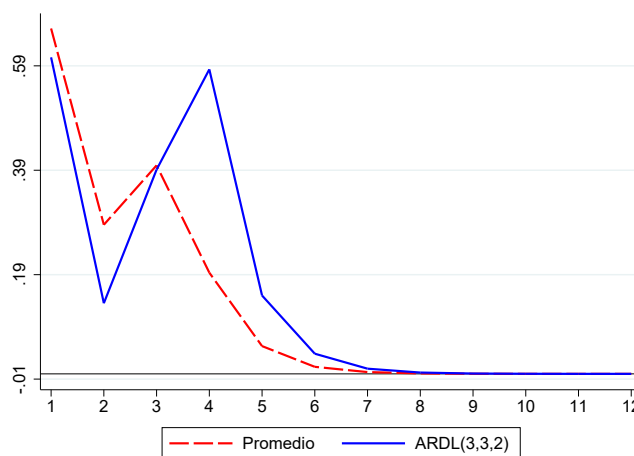
En contraste, los coeficientes de corto plazo resultaron ser estadísticamente significativos y con el signo esperado en tanto que el incremento de las ventas de energía incide de manera negativa en la recaudación, mientras que el IMAE lo hace de forma positiva. Sin embargo, al estimar el efecto multi-

plicador⁵ para ambas variables exógenas (véase gráfico 2), se verifica un particularidad relevante puesto que el efecto del IMAE es positivo solamente de manera contemporánea (en $t = 1$) para luego tornarse menor a cero en el tercer período hasta comenzar a converger al cabo de un semestre. En cambio, las ventas de energía exhiben un efecto multiplicador positivo sobre la recaudación de corto plazo hasta converger a cero luego de siete meses. Como ejercicio de robustez, se incluye también el efecto multiplicador promedio arrojado por las seis especificaciones ARDL (p, d, q) haciendo combinaciones para el número de rezagos óptimos ($p, d \in \{1, 2, 3\}$; $q \in \{1, 2\}$).

Gráfico 2: Efectos Multiplicadores de Corto Plazo



(a) Multiplicador del IMAE



(b) Multiplicador de la Energía

FUENTE: estimaciones propias con base a la ecuación 3. La línea roja corresponde a la media aritmética del efecto multiplicador calculado a partir de seis especificaciones diferentes para cada k -ésimo rezago, así: ARDL (1, 1, 1), ARDL (1, 1, 2), ARDL (2, 2, 1), ARDL (2, 2, 2), ARDL (3, 3, 1) y ARDL (3, 3, 2).

4.2 Análisis de Pronóstico

Esta sección detalla la evaluación de los pronósticos estimados por los 11 modelos empleados para diferentes horizontes de tiempo (h), específicamente, 1, 3, 6 y 12 meses. El primer pronóstico $h = 1$ es en realidad un pronóstico estático dentro de muestra, usualmente conocido en la literatura como *one step ahead forecast* (proyección de un período hacia adelante), las 3 proyecciones restantes corresponden a pronósticos dinámicos fuera de muestra. Para comparar la capacidad predictiva de cada modelo se emplean los estadígrafos enunciados en la sección 2.3, donde el menor de estos indica el modelo con mejor ajuste.

⁵Véase anexo 6.3 para mayor detalle sobre como se calcula el multiplicador de cada variable sobre los ingresos tributarios.

Tabla 2: Estimaciones del Modelo ARDL (3,3,2)

Ecuación 2		Ecuación 3	
T_{t-1}	-0.68884 (0.1576)*	T_{t-1}	-0.0112 (0.0732)
ΔT_{t-1}	-0.3224 (0.1037)**	T_{t-2}	0.0651 (0.0683)
ΔT_{t-2}	-0.2573 (0.0628)*	T_{t-3}	0.2573 (0.0628)
en_{t-1}	1.6134 (0.3700)*	en_t	0.6062 (0.3183)***
Δen_t	0.6063 (0.3183)***	en_{t-1}	0.1422 (0.2746)
Δen_{t-1}	-0.8649 (0.3055)**	en_{t-2}	0.3823 (0.2493)
Δen_{t-2}	-0.4826 (0.2133)***	en_{t-3}	0.4826 (0.2133)
y_{t-1}	-0.0305 (0.5519)	y_t	0.7696 (0.4924)
Δy_t	0.7696 (0.4925)	y_{t-1}	-0.1289 (0.5621)
Δy_{t-1}	.6713 (0.3219)***	y_{t-2}	-0.6713 (.3219)***
$D_{1,t}$	0.0835 (0.0779)		0.0836 (0.0778)
$D_{4,t}$	0.7056 (0.0619)*		0.7056 (0.0619)*
$D_{6,t}$	0.3224 (0.0505)*		0.3224 (0.0505)*
$D_{8,t}$	-0.0669 (0.0429)		-0.0668 (0.0429)
$D_{10,t}$	-0.0653 (0.0389)***		-0.0653 (0.0388)***
$D_{12,t}$	0.3189 (0.0477)*		0.3189 (0.0476)*
t	0.0027 (0.0017)		0.0027 (0.0017)
c_0	-15.324 (4.8822)**		-15.324 (4.8822)**

FUENTE: estimaciones propias. Entre paréntesis se presentan los errores estándar robustos mediante el método de Newey-West.
 (*) Representa la significancia; *1%, **5%, ***10%.

Al revisar el ajuste de los diferentes modelos para $h = 1$ (véase tabla 3) se observa que el modelo que experimenta mejor ajuste dentro de muestra para los ingresos tributarios es el ARDL (3,3,2) que precisamente ha sido estimado mediante la metodología de cointegración en niveles, ya que este

modelo parece replicar adecuadamente el proceso generador de datos de los ingresos tributarios (véase anexo 6.4). No obstante, al comparar las diferentes medidas de evaluación se obtuvo que los pronósticos que mejor se desempeñan para un horizonte de 3 meses (véase tabla 4) es un ARIMA (3,3). De igual manera, para un horizonte de un semestre, los modelos que exhiben mejor desempeño son el ARIMA (3,3) y el ARDL (1,1,1). Finalmente, par pronósticos de un horizonte aún mayor (un año) tanto el ARDL (3,3,1) como el ARDL (3,3,2) son los que mejor capacidad predictiva mostraron.

Con base a lo anterior, se pueden identificar algunas generalidades; en primer lugar, los modelos ARDL (p, d, q) estimados con la metodología de cointegración aparentan ser los de mejor desempeño predictivo en comparación a modelos autorregresivos estándar estimados con la metodología Box-Jenkins, salvo para un horizonte de tiempo de un trimestre. En segundo lugar, a pesar que la metodología de cointegración en niveles es la que parece capturar más adecuadamente el proceso generador de datos de la recaudación dentro de los 11 modelos analizados, ninguno de ellos parece predecir adecuadamente los vencimientos del primer pago a cuenta en junio. Esto invita a pensar una agenda de investigación sobre pronósticos que aborde metodologías macroeconómicas más estilizadas para capturar este tipo de particularidades o bien, explorar diferentes modelos de pronóstico basados en micro-simulaciones para impuestos con vencimientos de frecuencia menor, como es el caso del ya mencionado pago a cuenta.

Tabla 3: Precisión de Pronósticos para $h = 1$

Modelo	RECM	EAM	EPAM	Theil
AR(3)	0.3915	0.3244	0.0353	0.9921
MA(8)	0.3379	0.2721	0.0321	0.8269
ARMA(3, 3)	0.2678	0.2036	0.0238	0.6628
ARMAX (2, 1)	0.2478	0.1963	0.0229	0.6255
ARMAX (1, 1)	0.3074	0.2514	0.0295	0.7593
ARDL (1, 1, 1)	0.1401	0.1052	0.0123	0.3479
ARDL (1, 1, 2)	0.1341	0.1037	0.0122	0.3333
ARDL (2, 2, 1)	0.1326	0.1038	0.0122	0.3282
ARDL (2, 2, 2)	0.1306	0.1019	0.0119	0.3241
ARDL (3, 3, 1)	0.1222	0.0920	0.0107	0.2995
ARDL (3, 3, 2)	0.1206	0.0912	0.0106	0.2960

Fuente: estimaciones propias.

Tabla 4: Precisión de Pronósticos para $h = 3$

Modelo	RECM	EAM	EPAM	Theil
AR(3)	0.0216	0.0027	0.0002	0.0505
MA(8)	0.0561	0.0066	0.0007	0.1265
ARIMA(3, 3)	0.0181	0.0023	0.0002	0.0419
ARIMAX (2, 1, 1)	0.0292	0.0034	0.0004	0.0677
ARIMAX (1, 1, 1)	0.0370	0.0045	0.0005	0.0867
ARDL (1, 1, 1)	0.0200	0.0024	0.0003	0.0464
ARDL (1, 1, 2)	0.0227	0.0027	0.0027	0.0530
ARDL (2, 2, 1)	0.0249	0.0031	0.0003	0.0576
ARDL (2, 2, 2)	0.0260	0.0032	0.0003	0.0604
ARDL (3, 3, 1)	0.0197	0.0026	0.0002	0.0454
ARDL (3, 3, 2)	0.0201	0.0027	0.0003	0.0466

Fuente: estimaciones propias.

Tabla 5: Precisión de Pronósticos para $h = 6$

Modelo	RECM	EAM	EPAM	Theil
AR(3)	0.0527	0.0094	0.0010	0.1227
MA(8)	0.0734	0.0124	0.0125	0.1685
ARIMA(3, 3)	0.0321	0.0056	0.0006	0.0745
ARIMAX (2, 1, 1)	0.0469	0.0078	0.0009	0.1090
ARIMAX (1, 1, 1)	0.0478	0.0086	0.0009	0.1117
ARDL (1, 1, 1)	0.0312	0.0057	0.0006	0.0724
ARDL (1, 1, 2)	0.0339	0.0061	0.0007	0.0792
ARDL (2, 2, 1)	0.0351	0.0065	0.0007	0.0814
ARDL (2, 2, 2)	0.0359	0.0066	0.0007	0.0834
ARDL (3, 3, 1)	0.0329	0.0061	0.0007	0.0759
ARDL (3, 3, 2)	0.0331	0.0061	0.0007	0.0766

Fuente: estimaciones propias.

Tabla 6: Precisión de Pronósticos para $h = 12$

Modelo	RECM	EAM	EPAM	Theil
AR(3)	0.1093	0.0254	0.0028	0.2565
MA(8)	0.1389	0.0329	0.0035	0.3254
ARIMA(3, 3)	0.0564	0.0101	0.0011	0.1333
ARIMAX (2, 1, 1)	0.0726	0.0182	0.0020	0.1696
ARDL (1, 1, 1)	0.0344	0.0083	0.0009	0.0799
ARDL (1, 1, 2)	0.0371	0.0088	0.0009	0.0866
ARDL (2, 2, 1)	0.0380	0.0086	0.0009	0.0879
ARDL (2, 2, 2)	0.0391	0.0089	0.0009	0.0907
ARDL (3, 3, 1)	0.0339	0.0073	0.0008	0.0784
ARDL (3, 3, 2)	0.0344	0.0076	0.0008	0.0798

Fuente: estimaciones propias.

5 Conclusiones

El presente documento tiene como objetivo explorar y contrastar diferentes metodologías econométricas para el pronóstico de los ingresos tributarios de Honduras en varios horizontes de tiempo (1, 3, 6 y 12 meses). Como punta de lanza se emplea la metodología de cointegración en niveles para modelos dinámicos uniecuacionales desarrollada por Pesaran and Shin (1999) y Pesaran et al. (2001). El modelo empírico emplea datos de actividad económica y consumo de energía como vector explicativo de la recaudación observada para una muestra de 2007m1 a 2019m12. El ejercicio confirma la presencia de una relación de largo plazo entre estas variables con una velocidad de ajuste cercana al 70%, en línea con una frecuencia de datos mensual. De igual manera, el modelo permite derivar los efectos multiplicadores de las variables explicativas con lo que se verifica la presencia de efectos transitorios sobre la recaudación. Asimismo, el modelo de corrección de errores especificado parece sugerir que las variaciones en las ventas de energía inciden en mayor magnitud sobre los valores de estado estacionario de la recaudación comparado a la incidencia que tiene la actividad económica, no obstante, mayor evidencia empírica es necesaria para analizar rigurosamente este hecho.

Una vez que se ha revisado el modelo ARDL “óptimo” se contrasta su capacidad de pronóstico contra 10 modelos diferentes, de los cuáles cinco se estiman siguiendo la metodología de Box-Jenkins. Apoyándose en cuatro medidas de evaluación (raíz del error cuadrático medio, error absoluto medio, error porcentual absoluto medio y coeficiente de desigualdad de Theil) se concluye que el modelo ARDL (3,3,2) obtenido mediante la metodología de cointegración es el que mejor desempeño exhibe dentro de muestra (*one step ahead forecast*). De hecho, para los cuatro horizontes de tiempo evaluados, las diferentes especificaciones de modelos ARDL parecen superar la capacidad predictiva de los modelos estimados mediante la metodología Box-Jenkins salvo para el horizonte de un trimestre, en el que un modelo ARIMA (3,3) ofrece mejores proyecciones. Finalmente, se observa de manera general que ninguna de las metodologías de series de tiempo estimadas parece capturar adecuadamente los vencimientos de los pagos a cuenta, sobre todo durante el mes de junio, lo que invita a explorar más detenidamente otras metodologías de mayor complejidad o en el mejor de los casos, explotar las herramientas de microsimulación para dicha finalidad.

Referencias

- Box, G. E. P. and Jenkins, G. M. (1973). Some comments on a paper by chatfield and prothero and on a review by kendall. *Journal of the Royal Statistical Society. Series A (General)*, 136(3):337–352.
- Brender, A. and Navon, G. (2010). Predicting government tax revenues and analyzing forecast uncertainty. *Israel Economic Review*, 7(2):81–111.
- Cardoza, M. (2017). Boyanza y elasticidad de los ingresos tributarios en américa latina y el caribe. Documentos de trabajo, Centro Interamericano de Administraciones Tributarias (CIAT).
- Cerda, R., Gamboni, C., Beyzaga, C., and Gallardo, I. (2019a). Modelos de series de tiempo para la proyección de ingresos fiscales en chile. Estudios de Finanzas Públicas 2019/06, Ministerio de Hacienda, Chile.
- Cerda, R., Vilena, M., Luttini, E., and Beyzaga, C. (2019b). Análisis de las proyecciones de ingresos fiscales en chile. Estudios de Finanzas Públicas 2019/04, Ministerio de Hacienda, Chile.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431.
- Dolado, J. J., Jenkinson, T., and Sosvilla-Rivero, S. (1990). Cointegration and unit roots. *Journal of Economic Surveys*, 4(3):249–273.
- Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–276.
- Frankel, J. (2011). Over-optimism in forecasts by official budget agencies and its implications. *Oxford Review of Economic Policy*, 27(4):536–562.
- Glenday, G. (2013). *Revenue Forecasting*, pages 435–452. Palgrave Macmillan UK, London.
- Gómez, V. and Maravall, A. (1996). Programs seats and tramo: Instructions for the user. Technical report, Bank of Spain.
- Groves, H. M. and Kahn, C. H. (1952). The stability of state and local tax yields. *The American Economic Review*, 42(1):87–102.

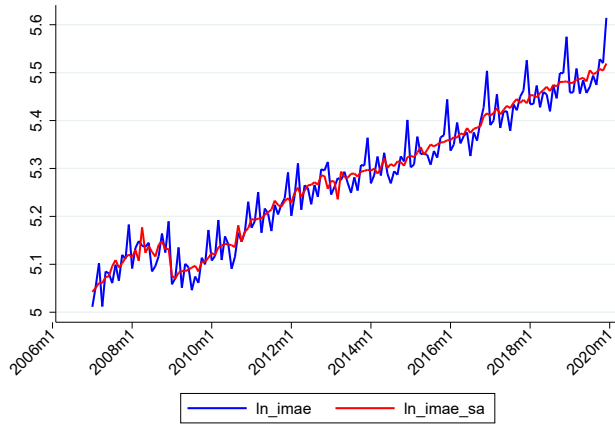
- Lagravinese, R., Liberati, P., and Sacchi, A. (2020). Tax buoyancy in oecd countries: New empirical evidence. *Journal of Macroeconomics*, 63:103189.
- Martín, F., cabello, J., and González, I. (2012). Modelo de proyección de las cuentas fiscales y análisis de la sostenibilidad fiscal. Technical report, Comisión Económica para América Latina y el Caribe (CEPAL).
- Pesaran, M. H. and Shin, Y. (1999). *An Autoregressive Distributed-Lag Modelling Approach to Cointegration Analysis*, page 371–413. Econometric Society Monographs. Cambridge University Press.
- Pesaran, M. H. and Shin, Y. (2002). Long-run structural modelling. *Econometric Reviews*, 21(1):49–87.
- Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3):289–326.
- Streimikiene, D., Ahmed, R. R., Vveinhardt, J., Ghauri, S. P., and Zahid, S. (2018). Forecasting tax revenues using time series techniques – a case of pakistan. *Economic Research-Ekonomiska Istraživanja*, 31(1):722–754.

6 Anexos

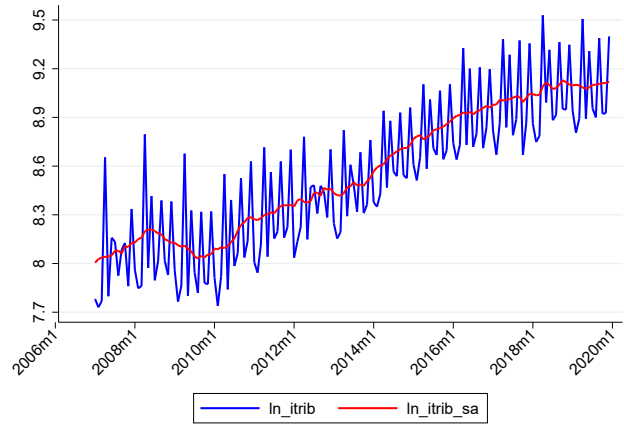
6.1 Descripción de los Datos

Este apartado expone el comportamiento gráfico de las tres variables empleadas para las estimaciones. En línea azul se expresan los datos en niveles y en línea roja, los datos desestacionalizados mediante la metodología de TRAMO-SEATS desarrollada por [Gómez and Maravall \(1996\)](#).

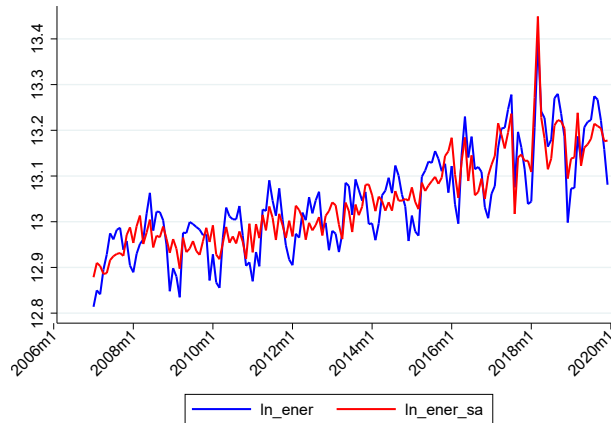
Gráfico 3: Datos Empleados en Logaritmos 2007m1-2019m12



(a) IMAE



(b) Ingresos Tributarios



(c) Energía en Mwh

Fuente: elaboración propia con base a datos del SAR, SIMAFIR y ENEE respectivamente.

6.2 Pruebas al Modelo de Corrección de Errores

Tabla 7: Prueba de Autocorrelación Breusch-Godfrey

Rezagos(p)	Estadístico F	Grados de Libertad	prob > F
1	2.351	(1, 134)	0.1275

Fuente: estimaciones propias con base a ecuación 2.

Tabla 8: Prueba Heterocedasticidad Breusch-Pagan

χ^2	1.19
Prob > χ^2	0.2747

Fuente: estimaciones propias con base a ecuación 2.

Tabla 9: Prueba de Límites de Pesaran et al. (2001)

Estadístico	Estimado	10%		5%		1%		Valor p	
		I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
F	12.259	3.129	4.147	3.773	4.890	5.213	6.521	0.000	0.000
t	-4.587	-2.531	-3.169	-2.840	-3.501	-3.440	-4.133	0.000	0.003

Fuente: estimaciones propias con base a ecuación 2.

6.3 Cálculo de Multiplicadores

Partiendo de la ecuación 3 y excluyendo los parámetros c_0 , c_1 y γ_i por simplicidad, se tiene la especificación ARDL siguiente:

$$T_t = \hat{\phi}_1 T_t + \hat{\phi}_2 T_{t-1} + \hat{\phi}_3 T_{t-2} + \hat{\phi}_4 T_{t-3} + \hat{\varphi}_0 en_t + \hat{\varphi}_1 en_{t-1} + \hat{\varphi}_2 en_{t-2} + \hat{\varphi}_3 en_{t-3} + \hat{\delta}_0 y_t + \hat{\delta}_1 y_{t-1} + \hat{\delta}_2 y_{t-2} + \hat{\epsilon}$$

El efecto multiplicador o marginal de una variable exógena sobre una variable explicativa es definido en manera simple como:

$$\frac{\partial T_t}{\partial \mathbf{x}_t'} = \sum_{j=1}^{d,q} \beta_j$$

De esta forma, el efecto multiplicador de las ventas de energía (en_t) sobre la recaudación en el corto plazo se construye como:

$$\beta_0 = \hat{\varphi}_0$$

$$\beta_1 = \hat{\varphi}_1 + \beta_0 \hat{\phi}_1$$

$$\beta_2 = \hat{\varphi}_2 + \beta_1 \hat{\phi}_2$$

$$\beta_3 = \hat{\varphi}_3 + \beta_2 \hat{\phi}_3$$

$$\beta_j = \beta_{j-1} \hat{\phi}_3 \text{ con } j > 1$$

Análogamente, el efecto multiplicador de corto plazo del IMAE (y_t) sobre la recaudación en el corto plazo se contruye como:

$$\beta_0 = \hat{\delta}_0$$

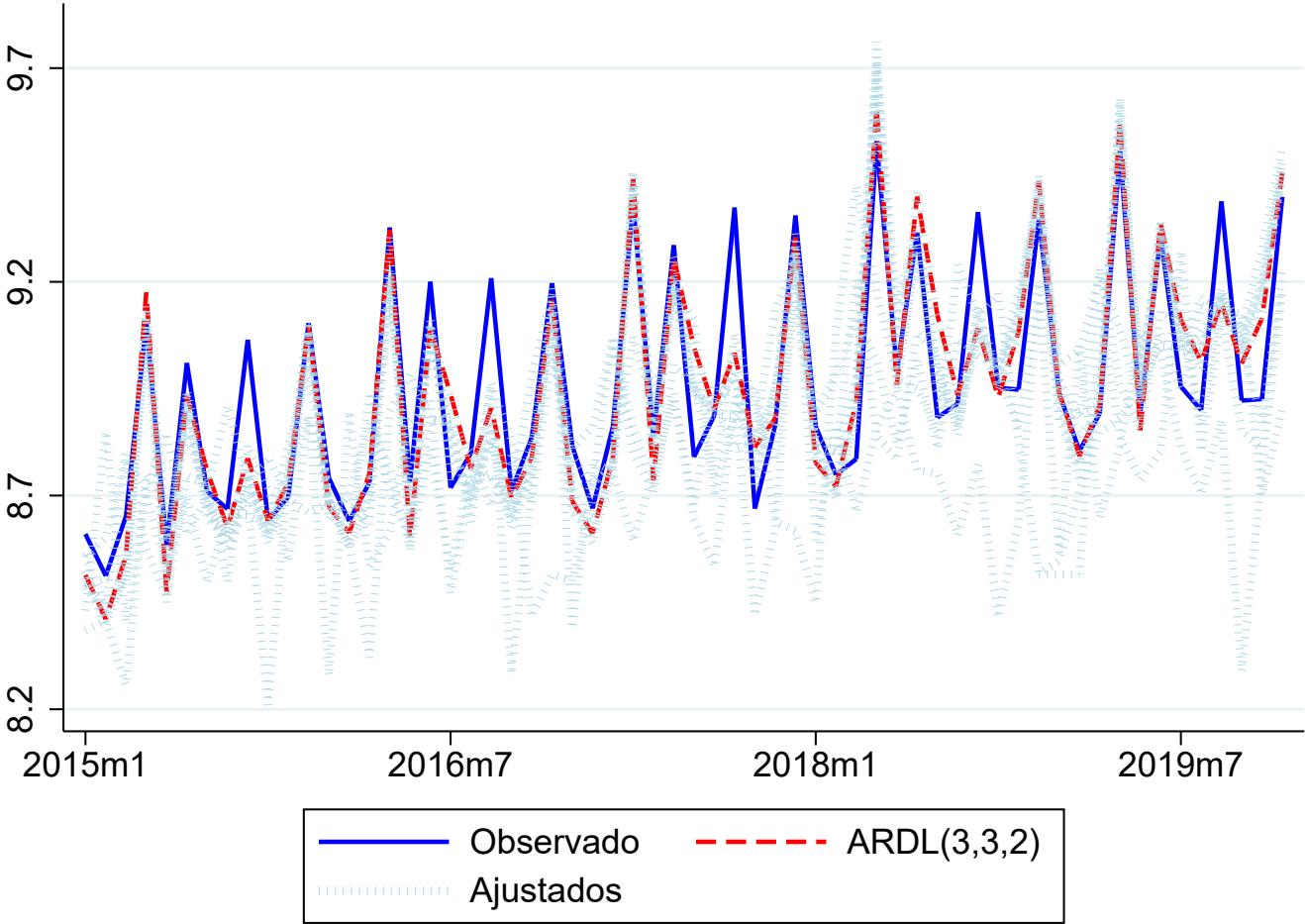
$$\beta_1 = \hat{\delta}_1 + \beta_0 \hat{\phi}_1$$

$$\beta_2 = \hat{\delta}_2 + \beta_1 \hat{\phi}_2$$

$$\beta_j = \beta_{j-1} \hat{\phi}_3 \text{ con } j > 1$$

6.4 Ajuste de Modelos dentro de Muestra

Gráfico 4: Ajuste de Modelos dentro de Muestra $h = 1$



Fuente: estimaciones propias