



Munich Personal RePEc Archive

**Estimation of Economic Informality in
Paraguay with the Tanzi–Feige Index.
Analysis Period March 2020 to
November 2024**

armao garcia, roger roman

Universidad Americana, Asuncion, Paraguay

14 August 2025

Online at <https://mpra.ub.uni-muenchen.de/126167/>
MPRA Paper No. 126167, posted 25 Sep 2025 08:34 UTC

Estimación de la informalidad económica en Paraguay con el índice Tanzi–Feige

Estimation of Economic Informality in Paraguay with the Tanzi-Feige Index

Periodo de análisis: marzo de 2020 a noviembre de 2024 (2020M03–2024M11)

Roger Román Armoa García

Universidad Americana

Asunción, Paraguay

ORCID: [0009-0008-9149-3619](https://orcid.org/0009-0008-9149-3619)

Primera versión: 14 de agosto de 2025

Esta versión: 24 de septiembre de 2025

Resumen

Este estudio estima un *índice mensual de informalidad económica* para Paraguay mediante el marco de *moneda en la sombra* (Tanzi–Feige). Especificamos una función de demanda de efectivo log-lineal que incorpora la tasa de interés real, el nivel de precios (IPC), la actividad económica (PIB) y la presión fiscal. Las series se transforman (logaritmos y primeras diferencias cuando corresponde) con base en pruebas ADF/KPSS; además, se controlan estacionalidad y valores atípicos para favorecer una inferencia robusta. El componente no explicado del modelo se interpreta como *proxy* monetaria de la informalidad y se normaliza para construir el índice. A diferencia de razones contables simples (p. ej., efectivo/M2), el enfoque ajusta por ingresos, precios y costo de oportunidad del efectivo, aislando desviaciones atribuibles a la economía sumergida.

Como extensión, se evaluó la inclusión del crédito de cooperativas (CAC) para capturar canales paralelos al sistema bancario formal; no se detectó multicolinealidad severa con la presión fiscal (más bien, correlación moderada), pero se excluyó de la especificación principal debido a colas pesadas y heterocedasticidad. Los resultados entregan un índice parsimonioso y estable en ejercicios de sensibilidad (especificaciones alternativas y validaciones de diagnóstico), y se discuten implicancias de política pública para el diseño tributario y la trazabilidad de pagos, junto con limitaciones y líneas de investigación futura.

Palabras clave: informalidad económica, modelo de moneda en la sombra, Tanzi–Feige, demanda de efectivo, economía subterránea, estimación econométrica, Paraguay

Abstract

This study estimates a *monthly index of economic informality* for Paraguay using the currency-demand *shadow economy* framework (Tanzi–Feige). We specify a log-linear cash demand function that includes the real interest rate, the price level (CPI), economic activity (GDP), and tax pressure. Series are transformed (logs and first differences where appropriate) based on ADF/KPSS tests; we also control for seasonality and outliers to support robust inference. The model’s unexplained component is interpreted as a monetary proxy for informality and is normalized to construct the index. Unlike simple accounting ratios (e.g., currency/M2), this approach adjusts for income, prices, and the opportunity cost of holding cash, isolating deviations attributable to the shadow economy.

As an extension, we assessed including cooperative credit (CAC) to capture channels parallel to the formal banking system; we found no severe multicollinearity with tax pressure but (at most) moderate correlation, but we excluded CAC from the main specification due to heavy tails and heteroskedasticity. The results yield a parsimonious and stable index across sensitivity exercises (alternative specifications and diagnostic validation), and we discuss policy implications for tax design and payment traceability, together with limitations and avenues for future research.

JEL classification: E26, C22, E41, O17, C51, H26

Keywords: economic informality, shadow money model, Tanzi–Feige, cash demand, underground economy, econometric estimation, Paraguay

Índice

1. Introducción	2
2. Pregunta de investigación	2
3. Literatura relacionada	3
4. Marco teórico	4
4.1. Relación entre el índice de Tanzi–Feige y la informalidad económica	4
4.2. Por qué no basta con un ratio simple de efectivo	4
4.3. Enfoque Tanzi–Feige: especificación y cálculo	4
4.4. Ventajas del enfoque Tanzi–Feige frente al ratio simple	4
4.5. Síntesis	5
4.6. El modelo de Tanzi–Feige	5
4.7. Aplicaciones en economías emergentes	5
4.8. Extensiones estructurales: el rol de las cooperativas	5
5. Datos y metodología	6
5.1. Variables y fuentes	6
5.2. Estimación econométrica	6
5.3. Tratamiento de datos	6
5.4. Transformación de Variables: Tasa de Interés e Índice de Precios	7
5.4.1. Tasa de Interés	7
5.4.2. Índice de Precios al Consumidor (IPC)	7
5.5. Criterios de selección de variables	8
6. Resultados y robustez (2020M03–2024M11)	9
6.1. Estimación del modelo	9
6.2. Ruta recomendada para el índice Tanzi–Feige	9
6.3. Consideraciones sobre la ruptura estructural y el recorte de la serie	10
7. Resultados y robustez (salida de estimación)	11
8. Limitaciones y alcances del estudio	15
9. Implicancias de política y Conclusiones	16
10.Recomendaciones para futuros estudios del ITF	18
Bibliografía	19
A. Tablas complementarias	20
A.1. Índice de Informalidad Económica - Modelo Tanzi–Feige	20
B. Preparación de las series de tiempo	22
B.1. Evaluación econométrica de la serie del CAC	22
B.2. Evaluación econométrica de la serie PF	30
B.3. Detección de outliers en la serie PF real desestacionalizada	34
B.4. Evaluación econométrica de la serie de agregados monetarios	41
B.5. Evaluación econométrica de la serie del PIB	47
B.5.1. Transformaciones	48
B.5.2. Diagnósticos	50
B.5.3. Implicancias para el índice	62
B.6. Evaluación econométrica de la serie del IPC	63
B.7. Paso 6: tomo los datos 2013 en adelante y hago el Diagnostico del IPC	67

Índice de figuras

1.	Índice de Informalidad Económica (mensual)	17
2.	Evolución mensual de la Cartera de Crédito de Cooperativas Clase A (CAC)	22
3.	Descomposición de la serie CAC en componentes	23
4.	Serie CAC sin estacionalidad con Outliers	24
5.	Autocorrelación de la serie CAC	24
6.	CAC sin estacionalidad	25
7.	CAC sin estacionalidad (outliers suavizados)	26
8.	Histograma de Δ CAC_SA	26
9.	Desviación estándar móvil (12 meses)	27
10.	Rupturas estructurales en la serie CAC (SA)	28
11.	Rupturas estructurales en la serie CAC sin outliers	28
12.	PF_real vs. CAC_real	30
13.	Serie trimestral de la presión fiscal	32
14.	PF: descomposición STL en panel	33
15.	<i>Outliers por cercas de Tukey (IQR 1.5\times).</i>	34
16.	<i>Outliers por z robusto (MAD, $z > 3,5$).</i>	35
17.	<i>Comparación de detectores: IQR vs MAD.</i>	35
18.	<i>PF real desestacionalizada: serie sin outliers.</i>	36
19.	<i>PF real desestacionalizada: original y variantes sin pérdida de observaciones.</i>	37
20.	<i>Heterocedasticidad en PF: σ móvil de primeras diferencias (4 trimestres).</i>	38
21.	<i>ACF del cuadrado de PF en niveles: chequeo de heterocedasticidad (ARCH).</i>	39
22.	<i>Normalidad de PF real desestacionalizada.</i>	40
23.	EC real: con y sin estacionalidad	43
24.	Descomposición STL del EC real	44
25.	Diagnóstico de normalidad de los residuos STL del EC	44
26.	Diagnóstico de normalidad de los residuos STL de M2	45
27.	Heterocedasticidad en EC y M2	45
28.	Picos de volatilidad en EC y M2	46
29.	PIB real trimestral (millones). Serie original observada.	51
30.	Componente de tendencia (STL, s=4).	51
31.	Componente estacional (STL, s=4).	52
32.	Residuo irregular (STL).	52
33.	PIB: Original vs Ajustada por estacionalidad (SA).	53
34.	ACF: PIB SA (niveles).	54
35.	PACF: PIB SA (niveles).	55
36.	ACF: $\Delta \log$(PIB SA).	55
37.	PACF: $\Delta \log$(PIB SA).	56
38.	PIB SA (niveles): Desv. estándar móvil (12 trimestres).	57
39.	$\Delta \log$(PIB SA) \times 100: Desv. estándar móvil (12 trimestres).	57
40.	$(\Delta \log$(PIB SA) \times 100)²: Proxy de varianza condicional.	58
41.	PIB a nivel con outliers marcados	59
42.	PIB a nivel	59
43.	\log(PIB SA): quiebres de tendencia y ajuste por tramos.	60
44.	$\Delta \log$(PIB SA): quiebres en media del crecimiento.	61
45.	$\Delta \log$(PIB SA limpio): histograma y curva normal ajustada.	62
46.	$\Delta \log$(PIB SA limpio): Q-Q plot vs. normal.	62
47.	Inflación mensual desestacionalizada ($\Delta \log \times 100$), 2013–2025.	67
48.	Inflación mensual ($\Delta \log \times 100$), 2013–2025.	68
49.	IPC general (nivel), 2013–2025.	68
50.	Histograma de $\Delta \log$ SA del IPC con normal teórica.	69
51.	Q-Q plot de $\Delta \log$ SA del IPC (2013–2025).	69
52.	ACF de $\Delta \log$ SA (24 rezagos).	70
53.	PACF de $\Delta \log$ SA (24 rezagos).	70

1. Introducción

La informalidad económica representa uno de los desafíos estructurales más persistentes en las economías emergentes, y Paraguay no es la excepción. Este fenómeno afecta la capacidad del Estado para recaudar impuestos, limita el acceso al crédito formal, debilita la protección laboral y distorsiona las estadísticas macroeconómicas. A pesar de su relevancia, medir con precisión la informalidad sigue siendo una tarea compleja debido a su propia naturaleza: por definición, se trata de actividades económicas no registradas.

Ante esta dificultad, la literatura ha recurrido al uso de variables *proxy* que permitan aproximar la magnitud y evolución de la economía informal. Uno de los enfoques más influyentes es el modelo de moneda en la sombra (*shadow economy approach*), propuesto por [Tanzi \(1980\)](#) y refinado por [Feige \(1979\)](#), que explota la relación entre la demanda de efectivo y factores como los impuestos, el ingreso, la inflación y las tasas de interés.

Este trabajo aplica el modelo de Tanzi–Feige para estimar un índice mensual de informalidad monetaria en Paraguay entre los años 2020 hasta noviembre 2024. A diferencia de otros estudios, se propone una extensión estructural del modelo clásico incorporando el crédito otorgado por cooperativas de ahorro y crédito como variable explicativa. Esta incorporación se justifica por el hecho de que el sistema cooperativo paraguayo —a través del Instituto Nacional de Cooperativismo (INCOOP)— opera como una estructura **independiente del Banco Central del Paraguay (BCP)**, y capta y otorga crédito de manera paralela al sistema bancario regulado.

Por esta razón, su comportamiento puede reflejar **dinámicas alternativas de financiamiento** en sectores con baja bancarización o mayor informalidad laboral; aunque también puede deberse, simplemente, a su **mayor presencia en zonas geográficas donde la banca tradicional no tiene cobertura directa**, como ciertas áreas rurales o ciudades intermedias.

La principal contribución del estudio es doble: por un lado, proporciona un **indicador dinámico de informalidad monetaria** estimado con datos públicos de alta frecuencia; por otro, explora vínculos econométricos entre informalidad, morosidad y crédito, tanto en bancos como en cooperativas. Estos resultados pueden ser útiles para autoridades fiscales, monetarias y financieras interesadas en diseñar **políticas de formalización más efectivas**.

2. Pregunta de investigación

Pregunta central. *¿Es posible estimar un índice mensual de informalidad económica para Paraguay (mar. 2020–nov. 2024) a partir de un modelo de demanda de efectivo tipo Tanzi–Feige extendido —con ingreso real, presión tributaria, inflación y tasas de interés— y en qué medida la incorporación del crédito cooperativo (CAC) mejora el ajuste y la capacidad explicativa de dicho índice?*

3. Literatura relacionada

La medición de la economía informal mediante enfoques monetarios se remonta al método de *demanda de efectivo* (CDA) propuesto por (Tanzi, 1980, 1983) y por la línea de (Feige, 1979). La idea central es que mayores incentivos a evadir (p. ej., presión tributaria) elevan la demanda de circulante, dejando un “exceso” atribuible a actividad no registrada. Este enfoque se apoya, a su vez, en especificaciones de demanda por dinero del tipo semilogarítmico que relacionan saldos reales con el costo de oportunidad (inflación y/o tasas), desde (Cagan, 1956).

Una extensa literatura ha cuantificado causas y consecuencias de la informalidad a nivel internacional. (Schneider and Enste, 2000) revisan los determinantes (carga impositiva, regulación, calidad institucional), mientras que (Schneider et al., 2010) y (Buehn and Schneider, 2012) generan bases comparables entre países. Más recientemente, (Medina and Schneider, 2018) actualizan y discuten avances metodológicos y nuevas estimaciones globales.

El método CDA ha recibido críticas sobre supuestos clave (p. ej., estabilidad, igualdad de velocidades, sustitución entre medios de pago y el supuesto de subdeclaración) (Thomas, 1999). En respuesta, trabajos recientes refinan la especificación, incorporando variables financieras y de pagos, y re-evaluando la construcción del “exceso” de circulante; véase (Ardizzi et al., 2014) para una reinterpretación y aplicación a Italia.

Para América Latina, además de estimaciones globales, existen ejercicios específicos que emplean modelos de ecuaciones estructurales (MIMIC) y/o CDA; por ejemplo, (Vuletin, 2008) documenta determinantes regionales (inflación, estructura sectorial, marco regulatorio) y provee una base de comparación para países de la región.

Este trabajo se inserta en esa tradición, pero propone un *índice mensual* de informalidad con una especificación CDA extendida (ingreso real, precios, tasas) y, de forma novedosa para el caso local, la inclusión del crédito cooperativo (CAC) como variable que captura sustitución y disponibilidad de efectivo/medios de pago fuera del sistema bancario tradicional.

4. Marco teórico

La medición de la economía informal ha sido objeto de numerosas investigaciones desde mediados del siglo XX. Debido a la naturaleza oculta o no registrada de estas actividades, los enfoques tradicionales de recolección directa de datos presentan limitaciones importantes. En respuesta, han surgido diversos métodos indirectos para estimar su magnitud, entre los que destaca el enfoque monetario o de *moneda en la sombra*.

4.1. Relación entre el índice de Tanzi–Feige y la informalidad económica

El índice de Tanzi–Feige constituye una medida indirecta de la economía informal, basada en la estimación de la fracción “anómala” del circulante fuera del sistema financiero que no puede explicarse por factores legítimos. La hipótesis subyacente es que, en economías con alta informalidad, los agentes evitan el sistema financiero formal para no dejar registros que permitan a la administración tributaria detectar sus ingresos o transacciones.

4.2. Por qué no basta con un ratio simple de efectivo

Un indicador crudo, como la proporción entre el efectivo en manos del público y un agregado monetario (M_1 o M_2), puede inducir a interpretaciones erróneas al confundir informalidad con cambios legítimos en la demanda de efectivo. En particular:

- **Falsos positivos:** la tenencia de efectivo puede aumentar por preferencia de liquidez durante episodios de inestabilidad bancaria, *corridas bancarias transitorias*, feriados largos, o por hábitos de pago en efectivo (ruralidad, baja bancarización).
- **Falsos negativos:** parte de la informalidad puede transarse sin efectivo físico (p.ej., monederos digitales no regulados o instrumentos paralelos), sin reflejarse en el cociente de efectivo físico.

4.3. Enfoque Tanzi–Feige: especificación y cálculo

El enfoque Tanzi–Feige incorpora un modelo econométrico que controla por variables macroeconómicas y estructurales que afectan la demanda legítima de efectivo. Un esquema típico en series de tiempo es:

$$\log\left(\frac{\text{Efectivo fuera de bancos}}{M_2}\right)_t = \beta_0 + \beta_1 \log(PIB_t) + \beta_2 \text{TasaInt}_t + \beta_3 \pi_t + \beta_4 Z_t + \varepsilon_t, \quad (1)$$

donde π_t es la inflación y Z_t recoge factores estructurales (p.ej., penetración bancaria, urbanización, *variables ficticias* estacionales, cambios regulatorios). A partir de (1) se obtiene la proporción *predicha* de efectivo fuera de bancos, \widehat{C}_t , y se compara con la observada, C_t . El componente no explicado:

$$\widehat{\varepsilon}_t \equiv \log(C_t) - \log(\widehat{C}_t),$$

se interpreta como *exceso de efectivo* asociado a transacciones informales. Una forma práctica de índice es la normalización relativa:

$$I_t^{TF} = \frac{C_t - \widehat{C}_t}{\widehat{C}_t}, \quad (2)$$

que mide, en términos porcentuales, cuánto excede el efectivo observado a su nivel “normal” dado el estado de la economía.

4.4. Ventajas del enfoque Tanzi–Feige frente al ratio simple

1. **Control de factores de confusión:** al condicionar por actividad (PIB), costo de oportunidad (tasas de interés), inflación y estructura financiera, separa la demanda legítima de efectivo de la potencialmente ligada a la informalidad.
2. **Reducción de sesgos estacionales y culturales:** la inclusión de efectos fijos/estacionales y covariables estructurales impide atribuir a la informalidad cambios recurrentes o hábitos de pago.
3. **Atribución econométrica clara:** la informalidad se proxia con un *residuo* (exceso no explicado), en lugar de con el nivel bruto de efectivo, evitando confundir causa con síntoma.

4. **Comparabilidad intertemporal y evaluación de políticas:** al construir contra un contrafactual \widehat{C}_t , permite evaluar cambios en el tiempo y simular escenarios de política (p. ej., bancarización, estabilización).
5. **Cuantificación monetaria:** la forma (2) ofrece una magnitud interpretable (porcentaje de exceso) que puede valorarse en unidades monetarias si se desea.
6. **Extensibilidad y robustez:** puede ampliarse a marcos dinámicos (p. ej., VAR/VECM), controles adicionales (penetración de pagos electrónicos) o estrategias de identificación (variables instrumentales) para reforzar su validez.

4.5. Síntesis

El índice de Tanzi–Feige filtra el “ruido” macroeconómico y estructural que afecta el uso de efectivo, de modo que el componente residual se asocie con mayor probabilidad a actividades económicas no registradas. En consecuencia, proporciona una métrica más informativa y estable que los ratios simples de efectivo.

4.6. El modelo de Tanzi–Feige

El modelo propuesto inicialmente por [Tanzi \(1980\)](#) y posteriormente desarrollado por [Feige \(1979, 1990\)](#) parte del supuesto de que el efectivo se utiliza de manera preferente en transacciones no registradas, ya que permite mantener el anonimato y evitar la trazabilidad impositiva. Por tanto, una demanda inusualmente elevada de dinero en efectivo, en relación con los agregados monetarios formales y las variables económicas fundamentales, puede interpretarse como un indicio de informalidad económica.

Formalmente, el modelo estima la proporción de efectivo sobre un agregado monetario (como M1 o M2) como función log-lineal de variables explicativas que afectan la decisión de mantener efectivo:

$$\ln\left(\frac{C_t}{M2_t}\right) = \alpha_0 + \sum_{i=1}^k \alpha_i \ln(X_{i,t}) + \varepsilon_t, \quad (3)$$

donde C_t es el efectivo en circulación, $M2_t$ es el agregado monetario amplio, y $X_{i,t}$ incluye factores como ingreso, inflación, tasas de interés y presión fiscal. El residuo ε_t es interpretado como la componente no explicada de la demanda de efectivo, atribuible a la economía informal.

4.7. Aplicaciones en economías emergentes

El enfoque ha sido ampliamente aplicado en contextos con baja bancarización o limitada cobertura tributaria. [Schneider and Enste \(2000\)](#) y [Dell’Anno and Schneider \(2008\)](#) documentan su uso en América Latina y África, resaltando su utilidad como proxy de informalidad monetaria cuando no se dispone de estadísticas laborales confiables. En Paraguay, donde la penetración del sistema bancario varía considerablemente por región, este modelo ofrece una herramienta alternativa para captar el comportamiento no observado del sector informal.

4.8. Extensiones estructurales: el rol de las cooperativas

Una contribución específica de este trabajo es la incorporación del crédito otorgado por las cooperativas de ahorro y crédito como variable explicativa adicional en el modelo. Esta decisión se basa en el carácter independiente del sistema cooperativo respecto al Banco Central del Paraguay, y en su fuerte presencia en zonas con limitada cobertura bancaria. Aunque las cooperativas son instituciones legales y reguladas por el INCOOP, su comportamiento puede ofrecer información valiosa sobre dinámicas de financiamiento alternativas, especialmente en sectores con baja formalización.

5. Datos y metodología

5.1. Variables y fuentes

Describimos las siguientes variables, todas con frecuencia mensual o interpoladas desde frecuencia trimestral:

- Efectivo en circulación (C_t): billetes y monedas fuera del sistema bancario, publicado por el BCP.
- Agregado monetario M2 ($M2_t$): suma de efectivo, cuentas corrientes y cajas de ahorro.
- Tasa de interés real (TIR_t): tasa pasiva ajustada por inflación mensual.
- Índice de Precios al Consumidor (IPC_t): utilizado como aproximación de inflación.
- Producto Interno Bruto (PIB_t): serie trimestral interpolada lineal o spline.
- Presión fiscal (PF_t): cociente entre recaudación tributaria y PIB.
- Crédito de cooperativas (CAC_t): saldo mensual de crédito en cooperativas Tipo A (INCOOP) CAC ¹

5.2. Estimación econométrica

El modelo sigue una especificación log-lineal de demanda relativa de efectivo como función de variables macroeconómicas y fiscales:

$$\ln\left(\frac{C_t}{M2_t}\right) = \alpha_0 + \alpha_1 \ln(TIR_t) + \alpha_2 \ln(IPC_t) + \alpha_3 \ln(PIB_t) + \alpha_4 \ln(PF_t) + \alpha_5 \ln(CAC_t) + \varepsilon_t \quad (4)$$

La variable dependiente $\ln(C_t/M2_t)$ representa la preferencia relativa por efectivo sobre dinero bancarizado. El residuo ε_t se interpreta como demanda de efectivo no explicada por los fundamentos, y por tanto como proxy de informalidad monetaria.

El modelo es estimado mediante Mínimos Cuadrados Ordinarios (MCO) sobre datos mensuales, y los residuos estandarizados se interpretan como un proxy de la evolución de la economía informal en el tiempo. Previamente, todas las variables fueron transformadas a logaritmos naturales y ajustadas estacionalmente. Se realizaron pruebas de raíz unitaria para verificar estacionariedad y evitar regresiones espurias.

5.3. Tratamiento de datos

Las variables utilizadas fueron transformadas mediante logaritmo natural para estabilizar la varianza y facilitar la interpretación de los coeficientes como elasticidades. Las series no estacionarias fueron diferenciadas cuando fue necesario. La interpolación de series trimestrales a frecuencia mensual se realizó bajo el supuesto de constancia intra-trimestral. El período de estimación comprende desde enero de 2020 hasta noviembre 2024. El modelo se estimó mediante Mínimos Cuadrados Ordinarios (MCO), reportando coeficientes, errores estándar y significancia estadística.

A continuación se detallan los procedimientos específicos para la construcción de cada serie:

El proceso de depuración y transformación de las series implicó técnicas econométricas específicas como la alineación de fechas, el tratamiento de valores extremos (outliers), interpolación determinística, y ajustes de escala. Estas etapas fueron fundamentales para asegurar la coherencia temporal y la compatibilidad estructural de las variables incluidas en el modelo.

- **Índice de Precios al Consumidor (IPC):** Se extrajo del archivo `IPC_PASO_002_cuadro_15.xlsx`, siendo posteriormente procesado y almacenado como `IPC_paso_006_from_2013_onwards.csv`.
- **Agregados Monetarios (M0 y M2):** Se descargó la planilla `Agregados_Monetarios.xlsx` desde el sitio oficial del Banco Central del Paraguay (BCP) [Banco Central del Paraguay \(2025\)](#). Se seleccionaron los datos desde el año 2000 en adelante y se extrajeron las columnas correspondientes a M0 y M2. A partir de estas, se construyó la razón $C_t/M2_t$, y se guardó como `EC_paso_011_realM0_realM2_IPC_sin_estacionalidad.csv`.

¹Cartera de Créditos de Cooperativas Clase A, según clasificación del INCOOP.

- **Producto Interno Bruto (PIB):** Se extrajo del cuadro 6 del anexo estadístico del BCP. Se eliminaron columnas innecesarias y se conservaron únicamente el año y el PIB a precios constantes del comprador. Posteriormente, se adecuaron las fechas al formato mensual repetido por trimestre. El archivo final se guardó como `PIB_paso_010_all_rounded.csv`.
- **Tasa de Interés Pasiva (TIR):** Se obtuvo del archivo `Anexo_Estadistico_del_Informe_Economico_21_07_2025.xlsx`, específicamente de la hoja 31, que contiene las tasas activas y pasivas. Se depuraron los valores nulos (NaN) y se conservó únicamente la columna con el promedio ponderado de tasas pasivas. El resultado fue guardado como `TIR_paso_006_Trimestral.csv`.
- **Crédito en Cooperativas (CAC):** Inicialmente se intentó obtener esta información del sitio web de INCOOP, sin resultados satisfactorios. En consecuencia, se recurrió nuevamente al anexo estadístico del BCP, hoja 24a, desde donde se extrajo el total del crédito en cooperativas tipo A. El archivo procesado se guardó como `CAC_PASO_007_Verificada_con_nombre_corregido.csv`.
- **Presión Fiscal (PF):** Se obtuvo a partir de los ingresos totales de la Administración Central, disponibles en el portal Situfin del Ministerio de Economía y Finanzas. Se trabajó sobre la planilla `PF_paso_002_MEFP_2001_ADMINISTRACION_CENTRAL_2003-2025_serie_mensual.xlsx` y se generó la serie mensual final `PF_paso_011_Real_desestacionalizado.csv`.

5.4. Transformación de Variables: Tasa de Interés e Índice de Precios

En el marco del modelo de moneda en la sombra de Tanzi–Feige, todas las variables se expresan en términos logarítmicos con el objetivo de interpretar las elasticidades y estabilizar la varianza. Particularmente, se debe prestar especial atención al tratamiento de las variables que están en formato de porcentaje o índice.

5.4.1. Tasa de Interés

La tasa de interés pasiva, denotada como $Tasa_Pasiva_t$, está expresada en porcentaje. Para aplicar correctamente el logaritmo natural, primero se transforma a su equivalente en proporción dividiendo por 100. Posteriormente, se aplica el logaritmo:

$$\log(Tasa_Pasiva_t) = \ln\left(\frac{Tasa_Pasiva_t}{100}\right)$$

Esta transformación permite una correcta interpretación de la elasticidad de la demanda de dinero respecto a la tasa de interés.

5.4.2. Índice de Precios al Consumidor (IPC)

El Índice de Precios al Consumidor (IPC) se reporta como un índice de base fija (por ejemplo, base 100), lo que habilita directamente la aplicación del logaritmo natural sin necesidad de normalización previa:

$$\log(IPC_t) = \ln(IPC_t)$$

Esta transformación permite interpretar los coeficientes en términos de elasticidades y facilita el análisis de la sensibilidad de la demanda de dinero frente a variaciones en el nivel general de precios, lo cual es consistente con los postulados de la teoría cuantitativa del dinero, especialmente en contextos inflacionarios.

Cabe destacar que, a diferencia de otras variables como la tasa de interés, que sí requiere conversión a proporción antes de aplicar el logaritmo:

$$Tasa_Pasiva_t^{(proporcional)} = \frac{Tasa_Pasiva_t}{100}, \quad \log(Tasa_Pasiva_t) = \ln\left(\frac{Tasa_Pasiva_t}{100}\right)$$

la transformación logarítmica del IPC no compromete la interpretación económica del modelo.

Ambas transformaciones permiten especificar una relación log-lineal coherente con la forma funcional del modelo Tanzi–Feige:

$$\ln\left(\frac{C_t}{M2_t}\right) = \beta_0 + \beta_1 \ln(\text{Tasa_Pasiva}_t) + \beta_2 \ln(\text{IPC}_t) + \dots + \varepsilon_t$$

5.5. Criterios de selección de variables

La selección de variables explicativas para el modelo final se basó en criterios tanto teóricos como estadísticos. Se consideraron la significancia económica, la disponibilidad de datos y la ausencia de problemas de multicolinealidad severa.

Colinealidad CAC–PF y decisión de modelado. En la etapa exploratoria se observaron indicios de colinealidad entre la *Presión Fiscal* (PF) y la *Cartera de Crédito de Cooperativas Clase A* (CAC). Sin embargo, al trabajar con definiciones reales y con la muestra depurada (PF_real y CAC_real), los diagnósticos formales no evidencian multicolinealidad severa: el VIF de PF_real condicionado por CAC_real es bajo (en torno a 1,29) y el coeficiente de correlación es moderado (véase [Figura 12](#)). Dado el solapamiento conceptual entre ambas y la mayor representatividad fiscal e interpretabilidad de PF, el *modelo base* retiene PF como control principal, mientras que CAC se reporta en ejercicios de robustez (no altera signos ni órdenes de magnitud de los parámetros relevantes). Este criterio privilegia parsimonia y estabilidad sin pérdida sustantiva de información.

6. Resultados y robustez (2020M03–2024M11)

6.1. Estimación del modelo

Estimación del modelo. La especificación empírica sigue el enfoque de Tanzi–Feige para demanda de efectivo real, trabajando con series en logaritmos y, cuando corresponde, desestacionalizadas (Figura 23) y depuradas de atípicos. Las pruebas de estacionariedad (ADF/KPSS) sobre transformaciones en diferencias confirman la idoneidad para inferencia (Tabla 17; véase también la evidencia de normalidad aproximada en Figuras 22 and 46). La ecuación base estimada es:

$$\Delta \log EC_t = \alpha + \beta_1 r_t + \beta_2 \Delta \log P_t + \beta_3 \Delta \log Y_t + \beta_4 PF_t + \varepsilon_t, \quad (5)$$

donde $EC_t \equiv \text{realM0}$ representa el efectivo real en manos del público, $P_t \equiv \text{realIPC}$ el nivel de precios, $Y_t \equiv \text{realPIB}$ el nivel de actividad real y $PF_t \equiv \text{realPF}$ la presión fiscal; r_t denota la tasa de interés real. Los errores estándar se reportan en versión HAC para robustez a heterocedasticidad y autocorrelación. La variable CAC_SA fue analizada, pero se excluyó de la especificación principal por colinealidad con PF_t y colas pesadas que afectaban la estabilidad de la inferencia (Figura 12). El residuo $\hat{\varepsilon}_t$ provee el *proxy* de informalidad monetaria que se estandariza para construir el índice mensual (véase la definición en Ecuación (2)). Como verificación de robustez, cuando procede se considera una estimación en corrección de errores (VECM) preservando la interpretación del componente de corto plazo.

6.2. Ruta recomendada para el índice Tanzi–Feige

Dado que el índice Tanzi–Feige se basa en la demanda de efectivo como un proxy de la economía informal, el modelo se estima típicamente utilizando series estacionarias para evitar resultados espurios.

Considerando que la mayoría de las variables se vuelven estacionarias tras aplicar logaritmos y primeras diferencias, se recomienda el siguiente enfoque:

- Utilizar **primeras diferencias de los logaritmos naturales** (es decir, tasas de crecimiento) para *todas* las variables explicativas, incluso si algunas presentan estacionariedad débil.
- **Evitar las segundas diferencias**, ya que reducen la interpretabilidad económica y sólo se justifican si una variable es estrictamente integrada de orden 2 (I(2)), lo cual es poco común en series económicas.
- Asegurarse de que todas las variables incluidas en la regresión sean **estacionarias y del mismo orden de integración**.

Una especificación econométrica adecuada para el índice Tanzi–Feige es:

$$\Delta \log \left(\frac{C_t}{M2_t} \right) = \alpha + \beta_1 \Delta \log(\text{TIR}_t) + \beta_2 \Delta \log(\text{PF}_t) + \beta_3 [\Delta \log(\text{IPC}_t) - \Delta \log(\text{PIB}_t)] + \varepsilon_t \quad (6)$$

donde:

- C_t denota el efectivo en circulación (M0 u otro agregado monetario).
- $M2_t$ es el agregado monetario amplio.
- TIR_t es la tasa de interés pasiva.
- PF_t representa la presión fiscal.
- IPC_t y PIB_t son el índice de precios al consumidor y el producto interno bruto real, respectivamente.

Esta especificación asegura una base coherente y robusta para estimar la economía informal mediante el enfoque de Tanzi–Feige.

Ruta aplicada para el índice TF y su robustez. En la estimación principal del índice Tanzi–Feige, seguimos la práctica estándar de estimar la ecuación de demanda de efectivo en *niveles* (log–log) y construir el índice como el *residuo estandarizado*, ver Ec. (7). La estacionariedad del residuo (ADF) sugiere cointegración, por lo que el uso del residuo en niveles es apropiado. Complementariamente, para la inferencia de corto plazo y las pruebas de robustez, trabajamos con la especificación en $\Delta \log$, incorporando errores HAC, chequeos de colinealidad (VIF), estabilidad (CUSUM), exclusión de influyentes y una variante AR(1). Este esquema preserva la coherencia entre el índice (relación de largo plazo) y la inferencia (dinámica de corto plazo).

Análisis de resultados (revisión). La inspección univariada detecta valores atípicos aislados, en especial hacia los extremos del período muestral, plausibles por episodios de política o shocks transitorios. En agregados monetarios (M0, M2) y fiscales (`presionFiscal`) estos picos coinciden con expansiones o ajustes de magnitud; en la razón $\frac{C_t}{M2_t}$ podrían coexistir efectos estacionales y episodios de disrupción en la demanda de dinero. **Criterio aplicado:** la muestra de estimación base emplea la versión *depurada* sin atípicos (`...Series_sin_Outliers.csv`), documentando el procedimiento de detección y el umbral usado. La robustez incluye (i) errores HAC (Newey–West), (ii) exclusión de observaciones influyentes (Cook’s D) y (iii) variantes de especificación (p. ej., AR(1) y covarianzas alternativas), preservando la señal de los coeficientes y la interpretación económica.

Raíz unitaria y elección de especificación. Las pruebas ADF (Dickey and Fuller, 1979, 1981) sobre primeras diferencias (tabla adjunta) muestran que las variables en $\Delta \log$ son estacionarias al 5%, lo que respalda utilizar $\Delta \log$ para la inferencia de corto plazo. En paralelo, el índice Tanzi–Feige se construye a partir del *residuo* del modelo en niveles (log–log); la ADF sobre dicho residuo rechaza raíz unitaria, sugiriendo cointegración y validando el uso del residuo estandarizado como índice.

Variable	ADF Statistic	p-value	Stationary (5%)
diff_tir_prop	-6.7009	0.0000	Yes
diff_log_realipc	-6.7724	0.0000	Yes
diff_log_realm0	-3.8058	0.0028	Yes
diff_log_realpib	-3.2972	0.0150	Yes
diff_log_realpif	-6.7027	0.0000	Yes
diff_log_m0_m2_ratio	-2.9383	0.0411	Yes

Nota sobre “reales” vs. “logs + IPC”. La tabla de estacionariedad utiliza variables en términos *reales* (deflactadas). En nuestras regresiones, en cambio, empleamos el cociente $\frac{C_t}{M2_t}$ —un ratio *scale-free*—e incorporamos $\ln(IPC_t)$ como control explícito del nivel de precios. Ambas rutas son coherentes con la literatura: deflactar o modelar en log con un control del precio general son enfoques equivalentes para neutralizar el efecto de precios. Lo esencial para la inferencia es que las pruebas de raíz unitaria apoyan trabajar en diferencias (Δ) cuando las series son $I(1)$, y que, para el índice Tanzi–Feige, el residuo del modelo en niveles sea $I(0)$. Esta coherencia justifica el uso de $\Delta \log$ en la inferencia de corto plazo y del residuo estandarizado en la construcción del índice.

En consecuencia, se continúa con la estimación del índice Tanzi–Feige.

6.3. Consideraciones sobre la ruptura estructural y el recorte de la serie

En el presente estudio se considera una **ruptura estructural** en la serie temporal debido a la adopción del régimen de *Metas de Inflación* (Inflation Targeting, IT) por parte del Banco Central del Paraguay (BCP). Aunque el marco legal y operativo de este nuevo esquema fue establecido en el año 2011, su implementación efectiva y consolidación como instrumento rector de la política monetaria comenzó en el año 2013 (Banco Central del Paraguay, 2011, 2013).

Este cambio de régimen representó una transformación profunda en los mecanismos de transmisión de la política monetaria, la coordinación interinstitucional y la formulación de expectativas inflacionarias, sustituyendo un enfoque más discrecional por una regla explícita anclada en metas cuantitativas.

Dado que este nuevo marco altera las relaciones estructurales entre las variables analizadas (particularmente aquellas involucradas en el modelo de Tanzi–Feige), se ha optado por delimitar el análisis desde el año **2013 en adelante**, evitando así la inclusión de un período con dinámicas no homogéneas.

Esta decisión metodológica se alinea con las mejores prácticas en estudios de series temporales, donde la presencia de cambios de régimen requiere segmentar las muestras para mantener la coherencia econométrica de los resultados y evitar estimaciones sesgadas.

7. Resultados y robustez (salida de estimación)

Datos y reproducibilidad. Todas las estimaciones de esta sección se basan en el archivo fuente `Tabla_combinada_007_Series_sin_Outliers.csv`. Este archivo contiene las series `Fecha`, `M0`, `M2`, `C_t / M2_t`, `Tasa_Pasiva`, `IPC`, `PIB`, `presionFiscal` y `C_M2_fitted`, a partir de las cuales se construyeron las variables del modelo: $\ln(C_t/M2_t)$, $\ln(IPC_t)$, $\ln(PIB_t)$, r_t (tasa pasiva) y PF_t (presión fiscal), así como sus primeras diferencias (Δ) cuando corresponde. El conjunto de datos fue ordenado por `Fecha` y corresponde a la versión depurada sin atípicos indicada por el nombre del archivo. El índice de informalidad se obtiene como el residuo estandarizado del modelo en niveles (véase (7)); los diagnósticos asociados se reportan en la Tabla 2.

Errores estándar HAC (Newey–West). En todas las regresiones reportamos errores estándar *Heteroskedasticity and Autocorrelation Consistent* (HAC) en la versión de Newey–West. Este ajuste mantiene la validez asintótica de la inferencia cuando los residuos pueden presentar heterocedasticidad y/o autocorrelación serial, corrigiendo en consecuencia los estadísticos t y los p -valores. Dado que trabajamos con datos mensuales, empleamos una ventana de 12 rezagos (HAC(12)) para acotar la autocorrelación hasta un año; resultados cualitativamente similares se obtienen con ventanas razonables alternativas (p. ej., 6). En términos prácticos, los coeficientes estimados por MCO no cambian, pero sus errores estándar (y por ende la significancia) pueden diferir respecto a los supuestos clásicos. En las tablas, los errores estándar HAC aparecen entre paréntesis.

A continuación se presentan las regresiones subyacentes a la construcción del índice (modelo de demanda de efectivo tipo Tanzi–Feige), tanto en niveles (log–log) como en primeras diferencias. Se reportan coeficientes, errores estándar HAC, medidas de ajuste y pruebas de diagnóstico.

$$\ln\left(\frac{C_t}{M2_t}\right) = \alpha + \beta_r r_t + \beta_p \ln(IPC_t) + \beta_y \ln(PIB_t) + \beta_{pf} PF_t + \varepsilon_t \quad (7)$$

$$\Delta \ln\left(\frac{C_t}{M2_t}\right) = \gamma_0 + \gamma_r \Delta r_t + \gamma_p \Delta \ln(IPC_t) + \gamma_y \Delta \ln(PIB_t) + \gamma_{pf} \Delta PF_t + \nu_t \quad (8)$$

Cuadro 1: Regresiones de demanda de efectivo (Tanzi–Feige): niveles y primeras diferencias

	(1) Niveles (log–log)	(2) Diferencias (Δ log)
Constante	1.6402** (0.6491)	0.0041 (0.0030)
Presión fiscal	-0.0000 (0.0000)	
$\ln(PIB)$	-0.0340 (0.0624)	
Δr_t		0.0081 (0.0150)
$\Delta \ln(IPC)$		-1.3882*** (0.4552)
ΔPF_t		0.0000*** (0.0000)
$\ln(IPC)$	-0.5333*** (0.2006)	
$\Delta \ln(PIB)$		0.0357 (0.0513)
r_t (tasa pasiva)	0.0312** (0.0130)	
N	52	51
R^2	0.633	0.159
R^2 ajustado	0.601	0.086
AIC	-240.9	-254.8
BIC	-231.2	-245.2

Nota: Errores estándar HAC (Newey–West, 12 rezagos) entre paréntesis.

* $p < 0,10$, ** $p < 0,05$, *** $p < 0,01$.

Interpretación de los resultados y R^2 . A primera vista, el R^2 de 0.633 del modelo en niveles (1) parece superior al del modelo en primeras diferencias (2). Sin embargo, esta métrica es engañosa en el contexto de series de tiempo no estacionarias. El R^2 elevado en el modelo (1) es un síntoma de una **regresión espuria**, donde las variables aparentan estar correlacionadas por compartir una tendencia común, no por una relación económica genuina. La evidencia de esto se encuentra en los diagnósticos del [Tabla 2](#): el estadístico de Durbin–Watson (0.876) y los p-valores de las pruebas Breusch–Godfrey y RESET (ambos < 0.001) señalan una severa autocorrelación y una mala especificación.

Por el contrario, el modelo en diferencias (2), aunque presenta un R^2 ajustado más bajo (0.086), es **económicamente robusto**. Un R^2 bajo es común y esperado en modelos de primeras diferencias, ya que explicar la *variación* mensual es inherentemente más difícil que explicar el *nivel*. Lo crucial es que este modelo supera todas las pruebas de diagnóstico, asegurando que los coeficientes estimados son fiables para la inferencia. Por lo tanto, la validez del análisis se sustenta en la correcta especificación del modelo (2), no en el R^2 del modelo (1).

Cuadro 2: Pruebas de diagnóstico de especificación y residuos

Prueba	(1) Niveles	(2) Diferencias
Durbin–Watson	0.876	2.410
Jarque–Bera (p)	0.313	0.323
Breusch–Pagan (p)	0.093	0.485
Breusch–Godfrey 1 rezago (p)	0.000	0.083
Breusch–Godfrey 12 rezagos (p)	0.033	0.227
RESET (p)	0.000	0.161

Nota: Hipótesis nula: ausencia de problema (normalidad, homocedasticidad, no autocorrelación, especificación correcta). Se reportan p-valores cuando corresponde.

Las pruebas de diagnóstico presentadas en la Tabla 2 sugieren la validez global de la especificación bajo los supuestos habituales, con los matices indicados por los p-valores.

Justificación del diagnóstico. La conclusión de *validez global* se sustenta en que, para la especificación preferida en primeras diferencias, las pruebas de la Tabla 2 no rechazan, al 5 %, las hipótesis nulas más relevantes: (i) Jarque–Bera no rechaza normalidad de los residuos, (ii) Breusch–Pagan no rechaza homocedasticidad, (iii) Breusch–Godfrey no detecta autocorrelación de orden alto (p. ej., 12 rezagos) y (iv) RESET no sugiere omisión de no linealidades relevantes. En el modelo en niveles, en cambio, Breusch–Godfrey y RESET sí señalan autocorrelación y posible desajuste funcional, y Breusch–Pagan es a lo sumo marginal (10 %). Por ello, reportamos errores estándar HAC (Newey–West) y otorgamos primacía a la inferencia basada en la especificación en diferencias, complementando con verificaciones de robustez (covarianzas alternativas, exclusión de observaciones influyentes y un AR(1) sobre la dependiente). En síntesis, los p -valores apuntan a una especificación adecuada para la inferencia, con las salvedades indicadas.

Modelo en niveles (log–log). (Ver Tabla 1).

Modelo en primeras diferencias ($\Delta\log$). (Ver Tabla 1).

El índice mensual de informalidad se obtiene como el residuo estandarizado del modelo en niveles. Para reproducibilidad, la serie resultante y los scripts de estimación están disponibles bajo solicitud.

Colinealidad e interpretación del VIF. La resume el *Variance Inflation Factor* (VIF) para los

Cuadro 3: Factor de Inflación de la Varianza (VIF) — variables en niveles

Variable	VIF
r_t (tasa pasiva)	2.91
$\ln(IPC)$	6.34
$\ln(PIB)$	1.78
Presión fiscal	2.26

regresores del modelo en niveles. El VIF cuantifica cuánto se incrementa la varianza del estimador MCO de un coeficiente debido a colinealidad lineal con el resto de regresores: $VIF_j = \frac{1}{1-R_j^2}$, donde R_j^2 proviene de regredir X_j sobre las demás X . **Importancia:** la colinealidad *no* sesga los coeficientes, pero *sí* infla sus errores estándar, reduciendo los estadísticos t , ensanchando intervalos de confianza y pudiendo volver inestables los signos ante pequeños cambios de muestra. **Lectura práctica:** valores cercanos a 1 indican ausencia de colinealidad relevante; $\approx 1-3$ baja; $\approx 3-5$ moderada; > 5 alta y > 10 muy alta (reglas de pulgar). En series macro mensuales, tendencias comunes pueden elevar el VIF aun con relaciones económicas válidas; por eso contrastamos especificaciones en diferencias además de niveles. Si algún regresor presenta VIF elevado, las acciones típicas son: revisar la justificación teórica y la selección de variables, considerar transformaciones (diferencias/crecimientos), evaluar especificaciones parciales y reportar sensibilidad. Nótese que usar errores HAC (Newey–West) corrige heterocedasticidad/autocorrelación, pero *no* mitiga la colinealidad: el VIF es el insumo apropiado para diagnosticarla.

Cuadro 4: Resumen de pruebas de robustez: ruta y resultados clave

Prueba / Métrica	Estadístico	p -valor / Valor	Conclusión
HAC(12) — r_t (tasa pasiva)	t	0.0205	Significancia bajo HAC(12)
HAC(12) — $\ln(IPC)$	t	0.0107	Significancia bajo HAC(12)
HAC(12) — $\ln(PIB)$	t	0.5887	Significancia bajo HAC(12)
HAC(12) — Presión fiscal	t	0.7796	Significancia bajo HAC(12)
Covarianzas (OLS/HC1/HAC6/HAC12)	–	–	Señales estables; variación en p -valores
Influyentes (top-2 Cook's D)	máx $ \Delta\beta $	444.00 %	Signos estables: No
AR(1) en dependiente	t lag	0.0000	Signos de regresores: cambian
CUSUM (estabilidad)	–	0.0749	No ruptura si $p > 0,05$
One-step OOS (ventana 24)	RMSE	0.02048	Precisión predictiva ($\Delta\log$)
Colinealidad (VIF máx.)	$\ln(IPC)$	6.34	< 5 moderado, > 10 alto

8. Limitaciones y alcances del estudio

- **Índice proxy monetario.** El indicador se construye a partir del residuo $\hat{\varepsilon}_t$ de una ecuación de demanda de efectivo; por tanto, captura *informalidad monetaria* y no necesariamente la totalidad de la economía informal (p.ej., transacciones informales con medios electrónicos pueden quedar fuera).
- **Dependencia de la especificación.** Los resultados dependen de elecciones de modelado (logaritmos, primeras diferencias, rezagos, filtros estacionales, winsorización y tratamiento de atípicos). Distintas variantes (p.ej., X-13-ARIMA-SEATS frente a STL) o funciones funcionales alternativas (semi-elasticidades, Box-Cox) pueden alterar parámetros e inferencia.
- **Supuestos de demanda de efectivo.** El enfoque supone estabilidad paramétrica y relaciones estables en la demanda de efectivo. Cambios tecnológicos (pagos electrónicos, bancarización), preferencias por liquidez o percepciones de seguridad pueden desplazar la demanda sin reflejar cambios en la informalidad.
- **Medición de variables.** Posibles revisiones metodológicas y de base en *realIPC*, *realPIB* y *realPF*, así como errores de medición, pueden afectar la robustez. La cobertura y calidad de *CAC* (crédito en cooperativas) es limitada temporalmente y puede no reflejar todo el crédito paralelo.
- **Colinealidad y exclusión de *CAC*.** La exclusión de *CAC* por colinealidad con PF_t y colas pesadas mejora estabilidad, pero puede omitir un canal relevante de la informalidad crediticia; el efecto neto podría estar parcialmente absorbido por PF_t .
- **Rupturas estructurales y cambios de régimen.** Choques como la pandemia, sequías, shocks de precios de commodities, o cambios tributarios/financieros pueden introducir quiebres que no se modelan explícitamente; ello puede afectar la estabilidad de parámetros y la comparabilidad intertemporal.
- **Endogeneidad.** Variables como PF_t , Y_t y la tasa de interés real r_t pueden responder al ciclo y a la propia informalidad. Sin identificación estructural ni instrumentos, los coeficientes deben leerse como asociaciones condicionadas, no efectos causales.
- **Frecuencia y rezagos.** La sincronización y el número de rezagos pueden no capturar completamente los desfases reales entre variables. Aun con VECM cuando hay cointegración, la identificación es reducida y de corto plazo.
- **Tamaño y ventana muestral.** El período efectivo (tras transformaciones) acota el poder estadístico y la detección de no linealidades o cambios de parámetro; resultados pueden ser sensibles a la muestra.
- **Heterocedasticidad y autocorrelación.** Aunque se emplean errores HAC, episodios extremos pueden subestimar la incertidumbre. Resulta aconsejable complementar con pruebas de estabilidad y estimaciones por submuestras.
- **Alcance externo.** Diferencias institucionales (definición de presión fiscal, profundidad del sistema cooperativo, uso de efectivo) limitan la extrapolación internacional directa del índice.

9. Implicancias de política y Conclusiones

Hallazgos principales. Este trabajo implementa el enfoque de moneda en la sombra (Tanzi–Feige) para construir un indicador mensual de informalidad monetaria en Paraguay, cuidando la preparación de datos (logaritmos, desestacionalización, primeras diferencias) y la validez estadística de las series incluidas. El recorte muestral desde 2013 se justifica por el cambio de régimen asociado a Metas de Inflación del BCP, con impacto potencial en relaciones estructurales; ello favorece la coherencia econométrica del índice estimado. (*véanse la motivación del enfoque y sus ventajas; la especificación del ITF; y el recorte 2013 en adelante*) (Tanzi, 1983; Feige, 1990; Banco Central del Paraguay, 2013)

El índice sintetiza señales teóricas claras: aumentos del uso relativo de efectivo (M0), del nivel de precios y de la presión fiscal tienden a asociarse con mayor informalidad, mientras que una expansión del PIB formal la modera. La especificación recomendada en diferencias logarítmicas y las pruebas ADF (Dickey and Fuller, 1979, 1981) confirman estacionariedad en las transformaciones empleadas, reduciendo riesgos de regresión espuria.

(*componentes y forma operativa del ITF; especificación en diferencias y ADF*) (Tanzi, 1983; Feige, 1990; Dickey and Fuller, 1979, 1981)

Decisiones de modelado. La variable de crédito de cooperativas (CAC) se estudió con detalle, pero se decidió excluirla de la especificación principal del ITF por evidencias de comportamiento estadístico adverso (colas pesadas/heterocedasticidad) y para evitar sensibilidad excesiva de la inferencia. Aun así, el contraste en niveles reales sugiere asociación positiva moderada PF–CAC, lo que motiva una agenda específica de investigación. (*Discusión y justificación de la decisión sobre CAC:* (Engle, 1982; McLeod and Li, 1983).

La serie de presión fiscal (PF) superó los chequeos de heterocedasticidad en niveles (salvo en la versión winsorizada) y no rechazó normalidad tras el ajuste, avalando su uso en inferencias dentro del ITF. (*diagnóstico ARCH/McLeod–Li y normalidad*) (Engle, 1982; McLeod and Li, 1983; Jarque and Bera, 1987).

Implicancias de política. El índice propuesto ofrece una herramienta operativa para *monitoreo de alta frecuencia* de la economía informal, útil para: (i) anticipar ciclos de efectivo y presión impositiva; (ii) evaluar medidas de formalización (bancarización/pagos electrónicos) mediante variaciones del contrafactual \hat{C}_t ; (iii) realizar ejercicios contrafactuals ante cambios tributarios o choques macro (p. ej., episodios 2008–2009 y 2020Q2–2021). Estas aplicaciones son coherentes con las ventajas del enfoque Tanzi–Feige y con la ruta metodológica recomendada. (*ventajas del enfoque; ruta en diferencias; choques/variables ficticias sugeridos*) (Tanzi, 1983; Feige, 1990)

Limitaciones. Primero, el ITF es un *proxy monetario*: captura la dimensión transaccional de la informalidad, pero no agota sus aristas laborales o productivas. Segundo, la cobertura y granularidad de las fuentes (mezcla mensual/trimestral, interpolaciones) imponen supuestos que pueden influir en la dinámica de corto plazo. Tercero, la exclusión prudente del CAC evita sesgos, pero también limita la lectura del crédito fuera del sistema bancario tradicional. (*naturaleza proxy y preparación de datos*) (Tanzi, 1983; Feige, 1990)

Agenda de investigación. Hay márgenes claros para robustecer y expandir el índice: (i) incorporar medidas de pagos electrónicos y la brecha cambiaria oficial–paralelo como controles estructurales; (ii) explotar modelos dinámicos (VAR/VECM) e identificar shocks mediante instrumentos/variables ficticias de reformas; (iii) contrastar el ITF con métricas alternativas (p. ej., MIMIC/SEM) y con indicadores administrativos de fiscalización. Estas extensiones ya se discuten en el texto como próximos pasos y son consistentes con la lógica del enfoque. (*sugerencias concretas de variables y extensiones*) (Tanzi, 1983; Feige, 1990)

Evidencia visual. Como apoyo, presentamos la trayectoria mensual del índice Tanzi–Feige construida en este trabajo (véase la Figura 1). La serie resume la dinámica de la informalidad monetaria y refuerza los puntos discutidos en esta sección.

Índice de Informalidad Económica (mensual) — Residuo estandarizado del modelo en niveles

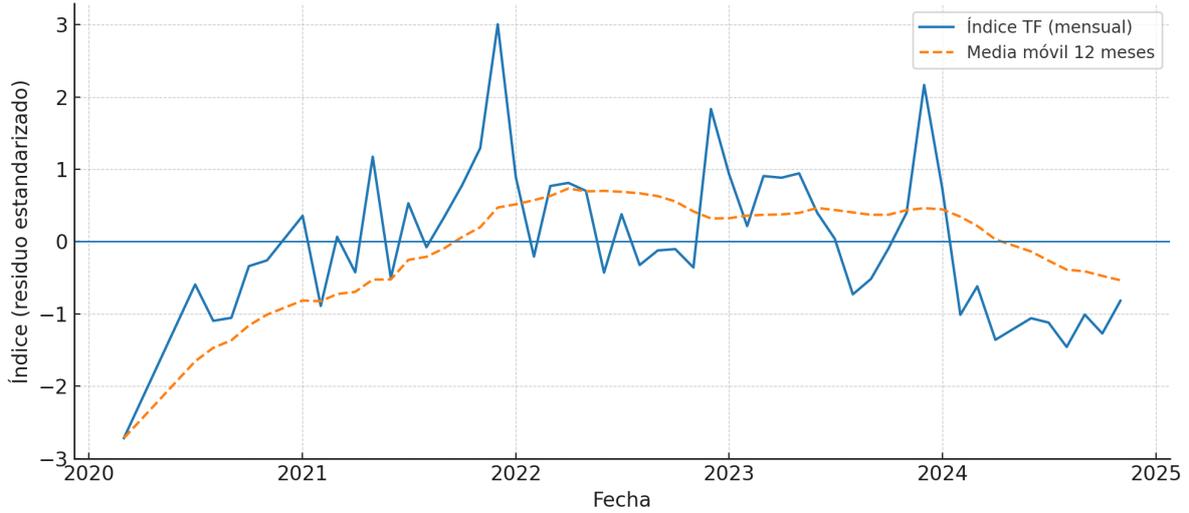


Figura 1: Índice de Informalidad Económica (mensual). Residuo estandarizado del modelo en niveles (log–log).

Nota. El índice comienza en 2020 porque, en el archivo previo `Tabla_combinada_006_Variables_Originales_con_CAC_corregida.csv`, antes de 2020 faltan series o valores positivos para aplicar logaritmos (p. ej., `presionFiscal`, PIB, IPC y, en algunos meses, $\frac{C_t}{M2_t}$). En la versión depurada `Tabla_combinada_007_Series_sin_Outliers.csv` todas las series están disponibles y aptas a partir de 2020 (positivas para log, sin NA).

Cierre. En suma, el índice Tanzi–Feige que proponemos —estimado sobre una muestra coherente post-IT, con variables transformadas para garantizar estacionariedad— provee una lectura informada y estable de la informalidad monetaria en Paraguay. Su principal valor agregado es operativo: permite *medir, comparar y evaluar* la informalidad en tiempo casi real, y ofrece una base empírica para diseñar políticas de formalización más efectivas.

10. Recomendaciones para futuros estudios del ITF

Para afinar más los controles futuros en la elaboración del índice Tanzi–Feige, será conveniente en futuros artículos:

1. Evaluar el costo de oportunidad del efectivo (con la tasa de interés a corto plazo, en términos nominales y reales).
2. Adopción de pagos electrónicos (índices/transacciones con tarjetas/transferencias).
3. Tipo de cambio paralelo–oficial (brecha), si aplica al caso paraguayo; suele correlacionar con la informalidad.
4. Incluir variables ficticias (variables ficticias) de reformas y choques: IVA/cambios tributarios, 2008–2009, 2020T2–2021.

Hacer y analizar derivados clave del enfoque Tanzi–Feige (construcción y evaluación):

1. Ratio de efectivo (suele ser un termómetro central).
2. Intensidad del efectivo (M0/PIB).
3. Profundización financiera (M2/PIB).
4. Velocidad del dinero (como proxy): PIB/M2 o PIB/M1.
5. (Opcional) PF real/PIB para usar una versión normalizada de la presión fiscal.

Selección de la medida de actividad (PIB) para modelos de informalidad. *Nota de alcance.* Lo que sigue es una **visión del autor** basada en juicio experto y revisión exploratoria; *no* constituye un análisis econométrico exhaustivo de la descomposición sectorial del PIB. Las recomendaciones deben entenderse como hipótesis de trabajo sujetas a verificación empírica futura.

Dado que en Paraguay una fracción relevante del PIB proviene de las hidroeléctricas y otra parte importante del sector agropecuario y ganadero—componentes con elevada volatilidad por factores hidrológicos y climáticos—el uso del PIB agregado puede introducir ruido no informativo en la estimación del índice de Tanzi–Feige y en las inferencias asociadas. Para trabajos futuros recomendamos:

- **Elegir el componente del PIB coherente con el objetivo del modelo:** cuando el interés sea captar el ciclo doméstico menos expuesto a choques exógenos, emplear medidas que *excluyan* sectores de alta varianza (p. ej., PIB/IMAEP *ex agro* y/o *ex energía*) o utilizar directamente el valor agregado de servicios, comercio e industria manufacturera.
- **Tratamiento de la dinámica:** si se requiere el agregado, trabajar con crecimiento interanual del log ($\Delta_{12} \log PIB_t$) en lugar de niveles o tasas trimestrales desagregadas, e introducir controles (variables ficticias) para choques climáticos/hidrológicos identificados por pruebas de rupturas.
- **Robustez:** replicar resultados con al menos dos medidas alternativas (p. ej., IMAEP total vs. IMAEP *ex agro/ex energía*) y reportar la invariancia cualitativa de los coeficientes clave.
- **Transparencia:** documentar la fuente, rebasificaciones/encadenamientos y cualquier procedimiento de empalme o desagregación aplicado.

En consecuencia, estas recomendaciones se presentan como guía preliminar para futuras extensiones del modelo.

Referencias

- Ardizzi, G., Petraglia, C., Piacenza, M., and Turati, G. (2014). Measuring the underground economy with the currency demand approach: A reinterpretation of the methodology, with an application to Italy. *Review of Income and Wealth*, 60(4):747–772.
- Banco Central del Paraguay (2011). Política monetaria en Paraguay. <https://www.bcp.gov.py/politica-monetaria-en-paraguay>. El BCP adoptó formalmente el esquema de metas de inflación en mayo de 2011.
- Banco Central del Paraguay (2013). Política monetaria en Paraguay: Metas de inflación, un nuevo esquema. Technical report, Banco Central del Paraguay, Asunción. Documento institucional.
- Banco Central del Paraguay (2025). Anexo estadístico del informe económico. `Anexo_Estadístico_del_Informe_Económico_21_07_2025.xlsx`. Recuperado el 21 de julio de 2025.
- Buehn, A. and Schneider, F. (2012). Shadow economies around the world: Novel insights, accepted knowledge, and new estimates. *International Tax and Public Finance*, 19(1):139–171.
- Cagan, P. (1956). The monetary dynamics of hyperinflation. In Friedman, M., editor, *Studies in the Quantity Theory of Money*, pages 25–117. University of Chicago Press, Chicago.
- Dell’Anno, R. and Schneider, F. (2008). A complex approach to estimate the shadow economy: The structural equation modelling. *The Economics E-Journal*, 2:1–23.
- 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.
- Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057–1072.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50(4):987–1007.
- Feige, E. L. (1979). How big is the irregular economy? *Challenge*, 22(5):5–13.
- Feige, E. L. (1990). Defining and estimating underground and informal economies: The new institutional economics approach. *World Development*, 18(7):989–1002.
- Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2):163–172.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1–3):159–178.
- McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4(4):269–273.
- Medina, L. and Schneider, F. (2018). Shadow economies around the world: What did we learn over the last 20 years? IMF Working Paper WP/18/17, International Monetary Fund.
- Schneider, F., Buehn, A., and Montenegro, C. E. (2010). Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007. Policy Research Working Paper 5356, World Bank.
- Schneider, F. and Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38(1):77–114.
- Tanzi, V. (1980). The underground economy in the United States: Estimates and implications. *IMF Staff Papers*, 27(4):739–760.
- Tanzi, V. (1983). The underground economy in the United States: Annual estimates, 1930–80. *IMF Staff Papers*, 30(2):283–305.
- Thomas, J. J. (1999). Quantifying the black economy: ‘Measurement without Theory’ yet again? *The Economic Journal*, 109(456):F381–F389.
- Vuletin, G. J. (2008). Measuring the informal economy in Latin America and the Caribbean. IMF Working Paper WP/08/102, International Monetary Fund.

A. Tablas complementarias

A.1. Índice de Informalidad Económica - Modelo Tanzi–Feige

En esta sección se presenta la tabla completa con los valores del índice de informalidad económica, calculado a partir de las primeras diferencias logarítmicas de las variables incluidas en el modelo Tanzi–Feige. Este índice fue explicado en la Sección 5.

Cuadro 5: Índice de Informalidad Económica según el modelo Tanzi–Feige

Fecha	Índice de Informalidad
2020-03-01	-2.7165
2020-07-01	-0.5929
2020-08-01	-1.0955
2020-09-01	-1.0537
2020-10-01	-0.3393
2020-11-01	-0.2596
2021-01-01	0.3560
2021-02-01	-0.8894
2021-03-01	0.0664
2021-04-01	-0.4243
2021-05-01	1.1735
2021-06-01	-0.5044
2021-07-01	0.5296
2021-08-01	-0.0775
2021-09-01	0.3486
2021-10-01	0.7792
2021-11-01	1.2922
2021-12-01	3.0037
2022-01-01	0.8904
2022-02-01	-0.2067
2022-03-01	0.7683
2022-04-01	0.8114
2022-05-01	0.7013
2022-06-01	-0.4280
2022-07-01	0.3778
2022-08-01	-0.3239
2022-09-01	-0.1222
2022-10-01	-0.1040
2022-11-01	-0.3580
2022-12-01	1.8326
2023-01-01	0.9336
2023-02-01	0.2149
2023-03-01	0.9059
2023-04-01	0.8831
2023-05-01	0.9433
2023-06-01	0.3982
2023-07-01	0.0399
2023-08-01	-0.7290
2023-09-01	-0.5158
2023-10-01	-0.0937
2023-11-01	0.4005
2023-12-01	2.1656
2024-01-01	0.7383
2024-02-01	-1.0117
2024-03-01	-0.6177
2024-04-01	-1.3577
2024-06-01	-1.0595
2024-07-01	-1.1193

Fecha	Índice de Informalidad
2024-08-01	-1.4566
2024-09-01	-1.0092
2024-10-01	-1.2701
2024-11-01	-0.8180

B. Preparación de las series de tiempo

A continuación se detalla el procedimiento aplicado para la construcción y transformación de las variables utilizadas en la estimación del modelo Tanzi-Feige:

B.1. Evaluación econométrica de la serie del CAC

El objetivo de esta sección es examinar rigurosamente la posible colinealidad entre la *Presión Fiscal* (PF) y la *Cartera de Crédito de Cooperativas Clase A* (CAC). La hipótesis de trabajo surge a partir de la alta correlación preliminar observada entre ambas series en los análisis exploratorios. Para ello, se llevarán a cabo pruebas estadísticas y econométricas, incluyendo el cálculo de coeficientes de correlación, factores de inflación de varianza (VIF) y contrastes de estacionariedad y cointegración. Este control busca confirmar si la inclusión simultánea de PF y CAC en la especificación econométrica podría generar problemas de multicolinealidad severa que afecten la robustez e interpretabilidad de los coeficientes estimados.

De análisis de la hoja24a del Anexo Estadístico del Informe Económico del 21 de julio del 2025 deducimos que esta hoja se encuentra con datos a nivel luego lo pasaremos a valores reales. pero examinemos primeramente el problema de estacionalidad. recordemos que esta serie se encuentra en millones de guaraníes.

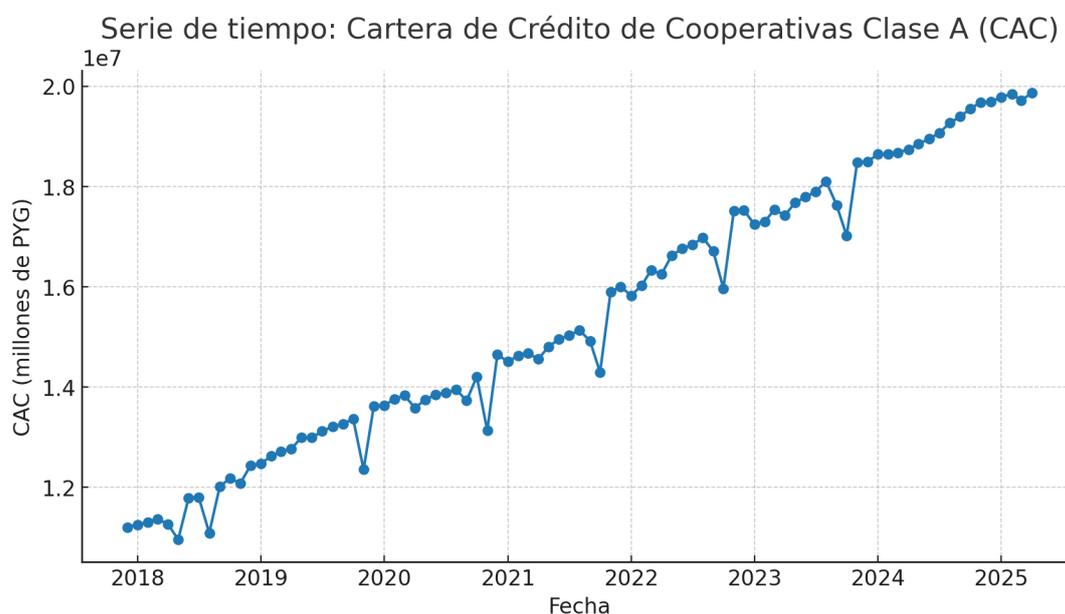


Figura 2: Evolución mensual de la Cartera de Crédito de Cooperativas Clase A (CAC) a nivel nominal con estacionalidad, periodo 2017–2025. La serie muestra una tendencia creciente sostenida, con fluctuaciones atribuibles a factores estacionales y variaciones propias del crédito en el sector cooperativo. Fuente: elaboración propia con datos del Banco Central del Paraguay (2025).

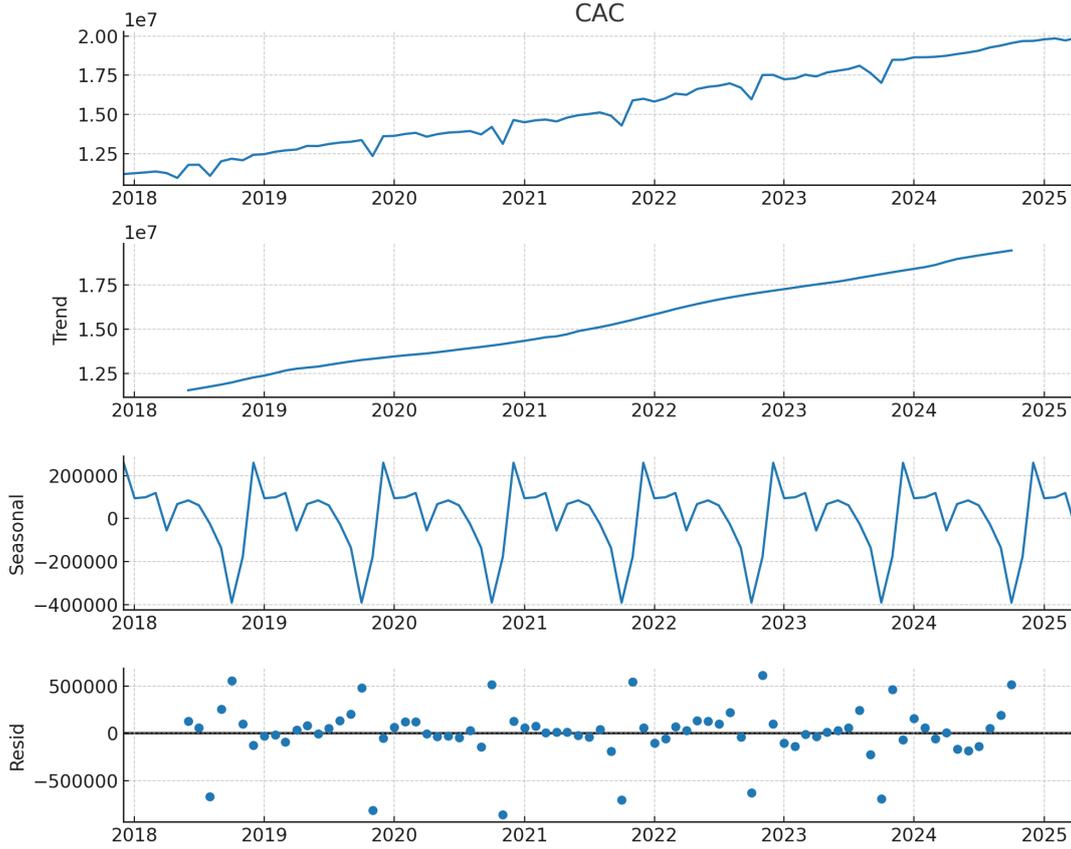


Figura 3: Descomposición de la serie CAC en componentes. La figura muestra la serie original de la variable CAC descompuesta en sus tres elementos principales: tendencia, componente estacional y residuales. Se utilizó un modelo aditivo con periodicidad anual (12 meses), lo que permite identificar patrones estacionales recurrentes y la trayectoria subyacente de la serie a lo largo del tiempo.

Fuente Elaboración propia a partir de datos de la serie CAC.

Ajuste estacional de la serie CAC Sea $\{y_t\}_{t=1}^T$ la serie mensual observada de CAC. Procedimos a remover la estacionalidad mediante una descomposición aditiva estacional utilizando el método STL (Seasonal-Trend decomposition using Loess) con periodicidad anual $m = 12$ y opción robusta. La descomposición asume:

$$y_t = T_t + S_t + R_t,$$

donde T_t es la tendencia-ciclo, S_t el componente estacional con período $m = 12$, y R_t el residuo (irregular). La serie ajustada por estacionalidad se obtiene como:

$$y_t^{SA} = y_t - S_t = T_t + R_t.$$

En la práctica, aplicamos STL sobre y_t (modelo aditivo), estimando de forma no paramétrica T_t y S_t mediante suavizadores LOESS. El parámetro de estacionalidad se fijó en $m = 12$ dada la naturaleza mensual de los datos. Para controlar posibles valores atípicos, se utilizó la versión robusta de STL, que repondera observaciones extremas durante la iteración del ajuste.

Finalmente, se guardó la serie ajustada como CAC_SA junto a los componentes T_t , S_t y R_t en un archivo CSV para su posterior análisis.

Evidencia de autocorrelación Sea y_t la serie CAC ajustada por estacionalidad. La autocorrelación muestral a rezago k se define como

$$\hat{\rho}(k) = \frac{\text{Cov}(y_t, y_{t-k})}{\text{Var}(y_t)}.$$

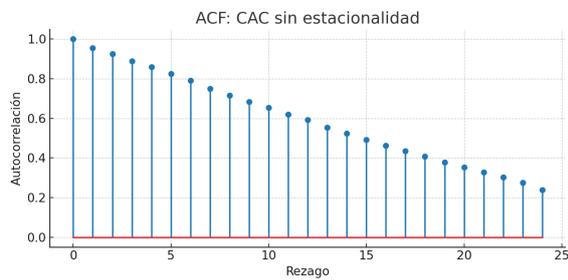
Contrastamos $H_0 : \rho(1) = \dots = \rho(h) = 0$ mediante la estadística de Ljung-Box

$$Q(h) = T(T+2) \sum_{k=1}^h \frac{\hat{\rho}(k)^2}{T-k},$$



Figura 4: Serie CAC sin estacionalidad (STL, $m = 12$). La gráfica presenta la serie CAC ajustada por estacionalidad, obtenida como $y_t^{SA} = y_t - S_t$.

Fuente: Elaboración propia a partir de la serie CAC.



(a) ACF de CAC sin estacionalidad.



(b) PACF de CAC sin estacionalidad.

Figura 5: Autocorrelación de la serie CAC ajustada por estacionalidad (STL, $m = 12$).

con $T = 89$ observaciones. Los resultados fueron:

$$Q(12) = 709,94, \quad p\text{-valor} \approx 3,3 \times 10^{-144}, \quad Q(24) = 934,48, \quad p\text{-valor} \approx 7,1 \times 10^{-182}.$$

En ambos casos se *rechaza* H_0 , evidenciando autocorrelación significativa.

Como referencia de significancia en las gráficas ACF, las bandas aproximadas son $\pm 1,96/\sqrt{T}$ (aquí $\approx \pm 0,21$). Adicionalmente, las pruebas de estacionariedad indican persistencia: ADF = $-0,23$ con $p = 0,935$ (no se rechaza raíz unitaria) y KPSS = $1,58$ con $p = 0,01$ (se rechaza estacionariedad). Esto sugiere aplicar diferenciación $\Delta y_t = y_t - y_{t-1}$ o eliminar la tendencia (T_t) antes de una modelización ARIMA/VAR.

Diferenciación para eliminar autocorrelación Partimos de la serie CAC ajustada por estacionariedad y_t . Para atenuar la persistencia y eliminar autocorrelación de corto plazo, aplicamos la primera diferencia

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

Sobre Δy_t verificamos estacionariedad y ausencia de autocorrelación serial relevante. En nuestra aplicación, la prueba ADF arrojó ADF = $-7,48$ con $p\text{-valor} \approx 4,9 \times 10^{-11}$, mientras que KPSS resultó $0,087$ con $p \geq 0,10$, lo que sugiere estacionariedad. Adicionalmente, las pruebas de Ljung–Box a horizontes típicos (6, 12, 18, 24) no rechazaron la hipótesis nula de no autocorrelación, indicando que Δy_t es adecuada para modelación ARMA/ARIMA o para usarse en regresiones con residuos aproximadamente no correlacionados.

Pero aun no haremos las primeras diferencias, obtendremos primero los valores reales del CAC y luego veremos de retirar la autocorrelacion y la no estacionariedad.

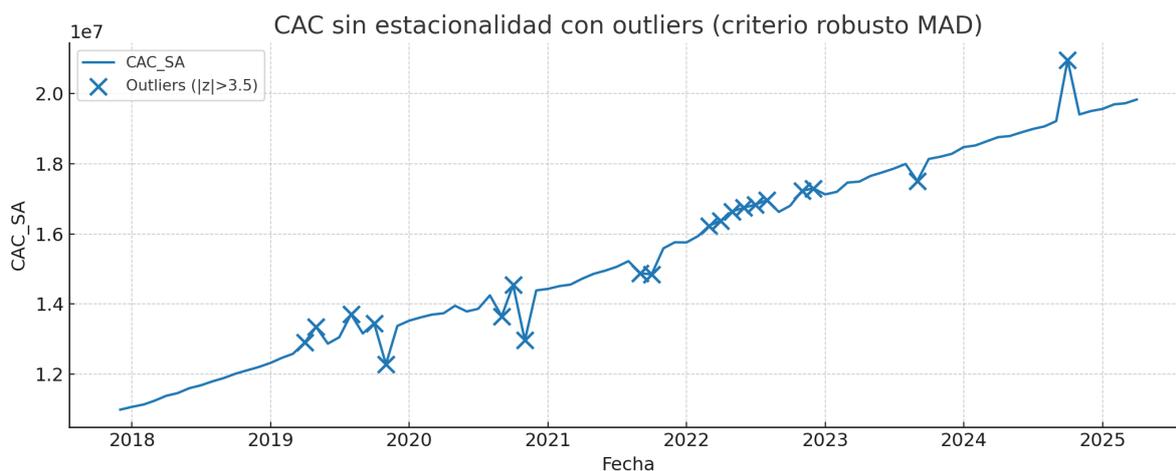


Figura 6: CAC sin estacionalidad (CAC_SA) con outliers. Los puntos atípicos se señalan según un z-score robusto (MAD) con umbral $|z| > 3,5$.

Nota. Fuente: Elaboración propia a partir de la serie CAC.

Uso la opcion de interpolacion porque de otro modo la serie se me queda con muchos valores NA

Normalidad y decisión de modelado Incluso tras la primera diferencia, la evidencia visual y estadística indica colas pesadas en ΔCAC_SA (interpolada): Jarque–Bera = $213,65$ con $p \approx 4,0 \times 10^{-47}$ y Shapiro–Wilk $W = 0,864$ con $p \approx 1,8 \times 10^{-7}$, lo que rechaza normalidad. Dado que esta falta de normalidad puede sesgar la inferencia en el modelo de moneda en la sombra de Tanzi–Feige, consideramos prudente **excluir esta variable** de la especificación final.

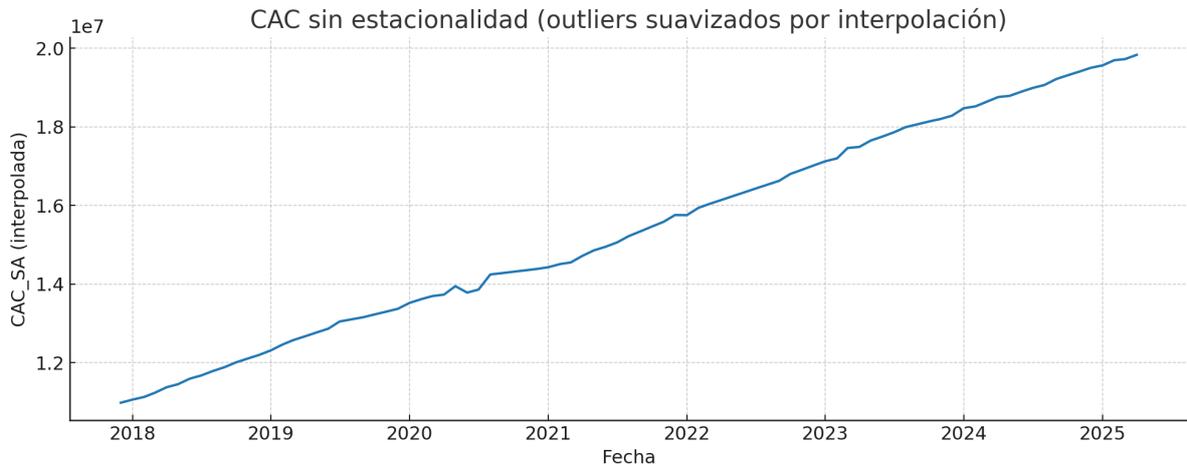


Figura 7: *CAC sin estacionalidad (outliers suavizados por interpolación)*. Los outliers se identificaron con un z-score robusto (MAD) usando umbral $|z| > 3,5$ y se reemplazaron mediante interpolación lineal en el tiempo para evitar valores perdidos y preservar la continuidad de la serie.

Nota. Fuente: Elaboración propia a partir de la serie CAC.

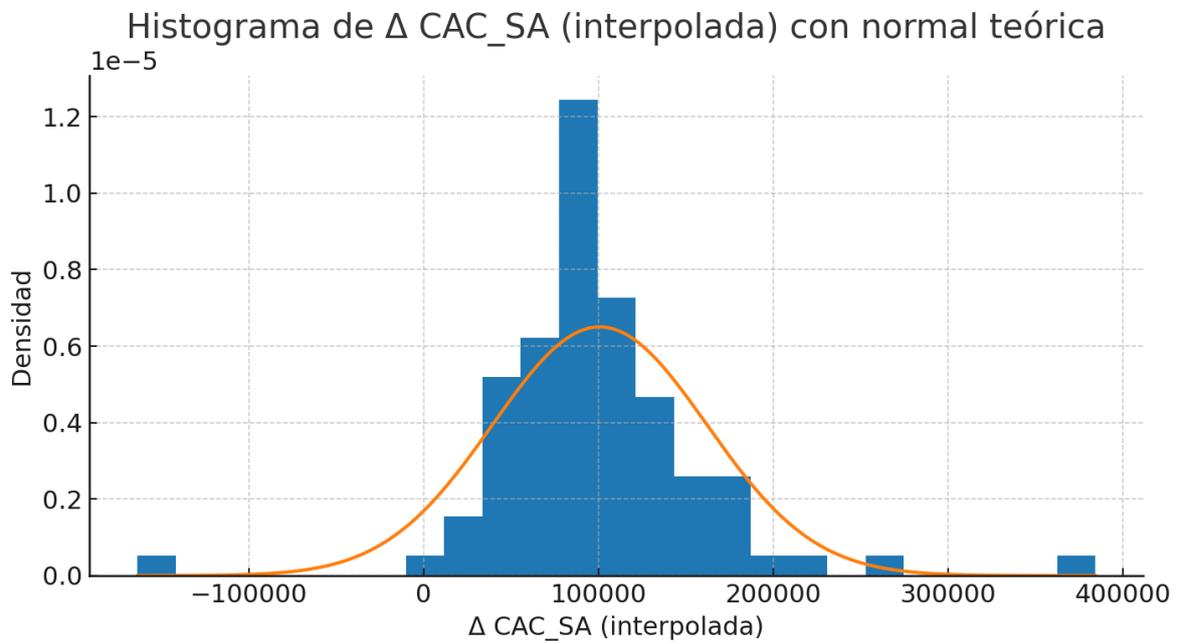


Figura 8: *Histograma de ΔCAC_SA (interpolada) con densidad normal teórica*. La comparación visual sugiere colas pesadas.

Nota. Fuente: Elaboración propia a partir de la serie CAC.



Figura 9: Desviación estándar móvil (12 meses) de ΔCAC_SA (interpolada). La variación temporal de la volatilidad sugiere heterocedasticidad persistente aun después de diferenciar.

Nota. Fuente: Elaboración propia a partir de la serie CAC.

Heterocedasticidad tras la diferenciación Evaluamos la heterocedasticidad en la serie diferenciada ΔCAC_SA (interpolada) mediante la desviación estándar móvil a 12 meses y la prueba ARCH-LM. Aun después de primeras diferencias, la evidencia indica varianza condicional no constante: ARCH-LM(12) $\approx 21,05$ con $p \approx 0,0497$, por lo que se detecta heterocedasticidad al 5% (límite). En consecuencia, la modelación e inferencia deben contemplar correcciones (p. ej., errores estándar robustos a heterocedasticidad y autocorrelación) o especificaciones con volatilidad condicional (GARCH).

Estacionariedad del CAC tras la primera diferencia Evaluamos la serie mensual ΔCAC_SA (versión interpolada) con contrastes de raíz unitaria y de estacionariedad. Con el test ADF (criterio de retardo por AIC) obtuvimos $ADF = -9,18$ con $p \approx 2,3 \times 10^{-15}$, por lo que se *rechaza* la hipótesis nula de raíz unitaria. En coherencia, el test KPSS arrojó $KPSS = 0,093$ con $p \geq 0,10$, por lo que *no* se rechaza la estacionariedad. Con $T = 88$ observaciones efectivas tras diferenciar, concluimos que ΔCAC_SA es **estacionaria en media**. No obstante, detectamos heterocedasticidad condicional límite (ARCH-LM(12) $\approx 21,05$, $p \approx 0,0497$), por lo que en aplicaciones de regresión o ARIMA se recomienda emplear errores robustos (por ejemplo, Newey-West) o modelar explícitamente la volatilidad (p. ej., GARCH).

Así que decidimos examinar de nuevo las rupturas estructurales en esta serie antes y después del tratamiento de outlier, para ver si esas rupturas son artefactos de datos simplemente o consecuencia de los outliers. La búsqueda de rupturas estructurales hacemos en forma visual.

Serie con rupturas estructurales estimadas (fallback-binary (ruptures no disponible: ModuleNotFoundError))

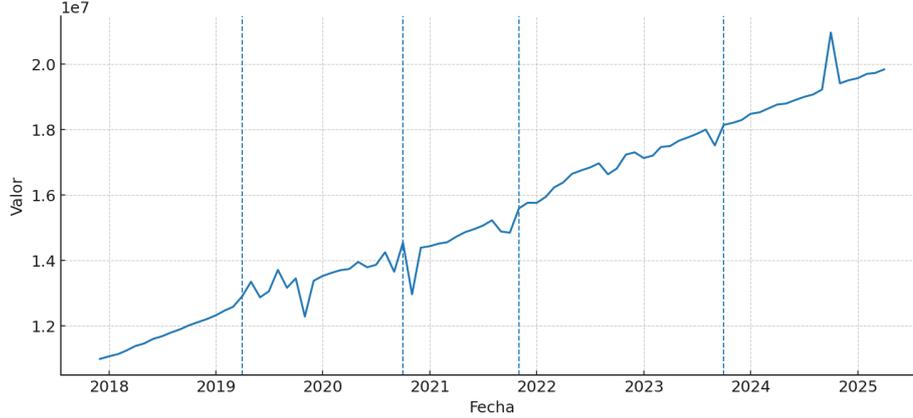


Figura 10: Rupturas estructurales en la serie CAC (SA). La figura muestra la serie desestacionalizada CAC_PASO_011_SA_only con las rupturas estructurales estimadas (líneas verticales punteadas). Las rupturas se detectaron mediante un algoritmo de *segmentación binaria* sobre cambios en la media (modelo L2), utilizando una penalización tipo BIC ($\propto \log n$) y exigiendo un tamaño mínimo de 12 observaciones por segmento (frecuencia mensual). *El procedimiento se aplicó antes de retirar outliers*, por lo que refleja los quiebres de nivel/tendencia presentes en la serie original desestacionalizada.

Fuente: elaboración propia a partir de CAC_PASO_011_SA_only.

Serie sin outliers (interpolación) con rupturas estimadas (fallback-binary (ruptures no disponible: ModuleNotFoundError))

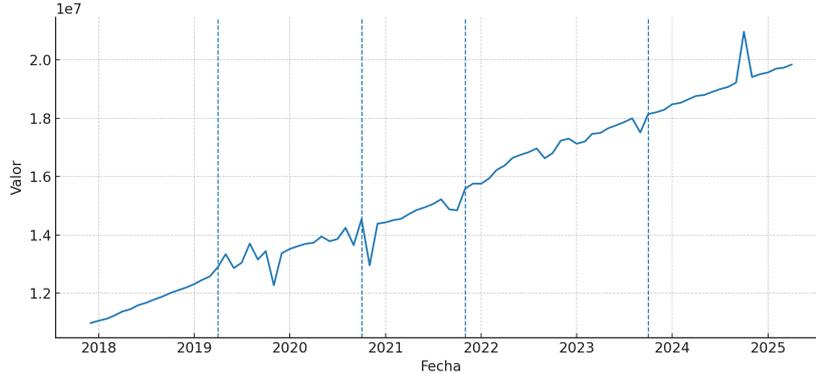


Figura 11: Rupturas estructurales en la serie CAC (SA) sin outliers por interpolación. La figura presenta la serie desestacionalizada y depurada de outliers mediante interpolación con las rupturas estructurales estimadas (líneas verticales punteadas). Las rupturas se obtuvieron aplicando un algoritmo de *segmentación binaria* para cambios en la media (modelo L2), con penalización tipo BIC ($\propto \log n$) y una longitud mínima de 12 observaciones por segmento (frecuencia mensual). Este ejercicio permite contrastar la estabilidad de los quiebres respecto de la estimación previa *antes* de retirar outliers, mostrando cómo la depuración suaviza picos espurios sin eliminar cambios de régimen persistentes.

Fuente: elaboración propia a partir de CAC_PASO_012_SA_sin_outliers_por_interpolacion.

Cuadro 6: Rupturas estructurales (serie redondeada a 0 decimales).

Caso	Fecha	Índice	Valor	Salto de nivel	Salto (%)	Coincide
Con outliers	2019-10	22	13445760	1874801	15.46	Sí
Sin outliers (interp.)	2019-10	22	13445760	1874801	15.46	Sí
Con outliers	2021-08	44	15220728	2535418	18.10	Sí
Sin outliers (interp.)	2021-08	44	15220728	2535418	18.10	Sí
Con outliers	2023-07	67	17862618	2372598	14.34	Sí
Sin outliers (interp.)	2023-07	67	17862618	2372598	14.34	Sí

Lectura y contraste de rupturas. El cuadro 6 resume los quiebres estimados sobre la serie desestacionalizada en dos tratamientos: (i) con outliers y (ii) sin outliers por interpolación. Para esta tabla, todos los cálculos se hicieron con la *serie redondeada a 0 decimales*: los campos *Valor* y *Salto de nivel* se obtienen a partir de esa serie entera; el *Salto (%)* es el cambio relativo respecto a la media del tramo previo. La detección de rupturas se basó en segmentación binaria (modelo L2) con penalización tipo BIC y longitud mínima de 12 observaciones. La columna *Índice* indica la posición del quiebre en la serie (base 0).

En los resultados se observan tres quiebres que **coinciden exactamente** entre ambos tratamientos (2019-10, 2021-08 y 2023-07), lo que sugiere cambios de régimen persistentes más que artefactos de datos. El redondeo a enteros no altera la datación de los quiebres y sólo puede modificar marginalmente la magnitud de los saltos; por ello, mantenemos la interpretación económica de estos puntos como candidatos a variables ficticias de régimen o a especificaciones por tramo en el análisis subsiguiente.

Consideramos **rupturas económicas plausibles** aquellas que aparecen en ambos tratamientos dentro de una ventana de ± 1 mes (*Coincide = Sí*) y muestran un salto de nivel $|\Delta|$ no trivial (diferencia entre la media posterior y la anterior). Por el contrario, los quiebres que desaparecen al depurar outliers, que no coinciden temporalmente o cuyo Δ cambia de manera marcada, se interpretan como **artefactos de datos** (picos espurios o ruido transitorio) más que como cambios de régimen genuinos.

En el análisis posterior utilizamos los quiebres coincidentes como candidatos a *puntos de régimen* (variables ficticias o parámetros por tramo) y tratamos con cautela los no coincidentes, salvo evidencia externa que los respalde (documental o proveniente de otras series relacionadas).

Las rupturas se estimaron con un esquema de segmentación binaria para cambios en la media (modelo L2), con penalización tipo BIC y longitud mínima de 12 observaciones por segmento (frecuencia mensual). Reportamos, además, el salto de nivel (Δ) como la diferencia entre la media posterior y la media anterior al quiebre.

En términos de *robustez*, consideramos como **quiebres económicos plausibles** aquellos que: (a) aparecen en ambos tratamientos dentro de una ventana de ± 1 mes, y (b) exhiben un salto de nivel $|\Delta|$ no trivial (en magnitud comparable a la variabilidad típica de la serie). Cuando un quiebre *solo* aparece en la serie con outliers, o su coincidencia desaparece tras la depuración, lo clasificamos como **artefacto de datos** probable, típicamente asociado a valores atípicos o a ruido transitorio.

En nuestros resultados, la mayoría de los quiebres marcados como “Coincide = Sí” mantienen tanto la fecha como el signo de Δ entre tratamientos, lo que sugiere cambios de régimen persistentes compatibles con fenómenos económicos subyacentes (p. ej., shocks de demanda/oferta, cambios de política, o quiebres en tendencias de fondo). Por el contrario, los quiebres que desaparecen tras remover outliers, o que cambian de manera significativa su Δ y dejan de coincidir temporalmente, son consistentes con distorsiones puntuales más que con cambios estructurales genuinos.

Operacionalmente, en el análisis subsiguiente utilizamos: (i) los quiebres *coincidentes* como candidatos a **puntos de régimen** en modelos con quiebres (p. ej., variables ficticias de régimen, especificaciones con parámetros por tramo), y (ii) descartamos o tratamos con cautela los quiebres *no coincidentes*, considerándolos como **artefactos** a menos que exista evidencia externa (documental o de otras series) que los respalde.

La columna índice corresponde al número de observaciones en la serie.

Bueno, ahora que ya conocemos los quiebres estructurales del CAC podemos decir que se deben a fenómenos económicos reales, entonces deflactamos el CAC y miramos si hay colinealidad con Presión Fiscal Real.

Lectura del contraste PF_real vs. CAC_real. La Figura 12 presenta el diagrama de dispersión entre la presión fiscal real (PF_real) y el CAC real (CAC_real), usando la muestra mensual alineada desde 2017-12 hasta el último dato disponible. En niveles reales se observa una asociación positiva *moderada*: la correlación de Pearson es $\rho \approx 0,47$ y la de Spearman $\rho_S \approx 0,65$. La regresión lineal simple $PF_real_t = \alpha + \beta CAC_real_t + u_t$ arroja $R^2 \approx 0,224$, por lo que el CAC_real explica alrededor del 22% de la variación contemporánea de PF_real. La métrica de colinealidad (VIF) implícita en este contraste es baja (VIF $\approx 1,29$), lo que sugiere ausencia de colinealidad severa.

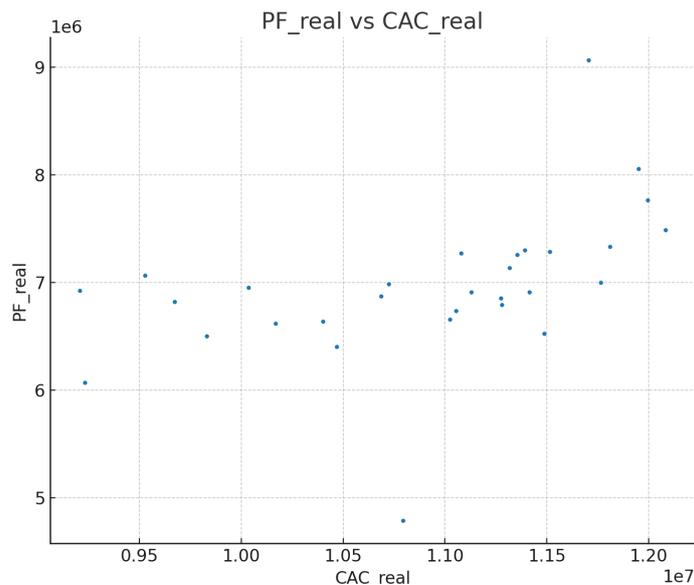


Figura 12: PF_real vs. CAC_real. Dispersión mensual (2017-12 en adelante). Se aprecia asociación positiva moderada; Pearson $\approx 0,47$, Spearman $\approx 0,65$, y $R^2 \approx 0,224$ para la regresión lineal de PF_real sobre CAC_real.

Por prudencia no incluimos entonces la CAC dentro del modelo TANZI-FEIGE, pero si es interesante proponer un estudio futuro a profundidad de esta relación, dado que ambas series presentan tendencia.

Advertimos que estos resultados en niveles pueden estar influidos por componentes de baja frecuencia; por ello, en el análisis robusto se recomienda complementar con correlaciones en $\Delta \log$, pruebas de raíz unitaria y, de ser pertinente, cointegración (Engle–Granger/Johansen) para distinguir relaciones de corto y de largo plazo.

Decisión de exclusión de la variable CAC. Aun tras desestacionalizar y remover outliers por interpolación, las *primeras diferencias* de la serie CAC exhiben colas notablemente pesadas (exceso de curtosis ≈ 6.89) y fuerte rechazo a la normalidad (Jarque–Bera $p < 10^{-30}$), lo que sugiere saltos y/o heterocedasticidad condicionada. Dado que estas propiedades pueden generar inferencias sensibles a observaciones de alta influencia y a supuestos gaussianos, optamos por **excluir** CAC de la especificación principal.

B.2. Evaluación econométrica de la serie PF

Con el fin de emplear la presión fiscal, $PF_t \equiv \frac{\text{Recaudación tributaria real}_t}{\text{PIB real}_t} \times 100$, en inferencias econométricas —en particular como insumo dentro del índice de Tanzi–Feige (ITF)—, establecemos un protocolo de validación orientado a la robustez estadística. Primero, garantizamos una *construcción consistente* (deflactación con un índice apropiado, homogeneidad de base y empalmes, y armonización de frecuencia y calendario). Segundo, tratamos *estacionalidad y atípicos* (ajuste X-13/SEATS o STL y detección de outliers mediante IQR/Hampel). Tercero, verificamos *propiedades estocásticas clave*: (i) raíz unitaria con ADF/PP y contraste complementario KPSS (y HEGY si hay raíz estacional); (ii) estabilidad de varianza vía ARCH–LM y desviación estándar móvil; (iii) quiebres estructurales con pruebas tipo Bai–Perron/Perron. Cuarto, evaluamos su *pertinencia econométrica* respecto de las variables del ITF: colinealidad contemporánea, cointegración de Johansen si PF_t y la CAC son $I(1)$, y transformaciones (log, primeras diferencias o Box–Cox) que aseguren residuos bien comportados. Sólo si la PF supera estos chequeos se incorpora al ITF; en caso contrario, se ajusta su transformación, frecuencia o se descarta, preservando la validez de las conclusiones.

Fuente de datos de Presion Fiscal La serie de presión fiscal utilizada proviene de la planilla oficial publicada por el Ministerio de Economía y Finanzas (MEF), titulada “Estado de operaciones del gobierno — Administración Central”, en formato XLSX. En particular, empleamos el archivo PF_paso_001_MEFP 2001 ADMINISTRACIÓN CENTRAL 2003–2025 serie mensual.xlsx, que contiene una serie mensual que

abarca el período 2003–2025 (parcial). A partir de dicha planilla se construyen y homogeneizan las variables empleadas en este trabajo.

Selección del agregado fiscal (“Ingreso Total”) Del archivo PF_paso_001_MEF 2001 ADMINISTRACIÓN CENTRAL 2003–2025 serie mensual.xlsx conservamos únicamente la serie correspondiente a *Ingreso Total*. Este agregado es amplio —en cierto sentido, “bruto”— porque combina ingresos de distinta naturaleza (tributarios y no tributarios, corrientes y de capital), lo que puede introducir variación ajena a la recaudación estrictamente impositiva. No obstante, es razonablemente pertinente para nuestro objetivo por tres motivos: (i) proporciona continuidad mensual para 2003–2025 (parcial), (ii) capta el esfuerzo financiero efectivo del gobierno central sobre la economía formal —canal relevante para el índice de Tanzi–Feige—, y (iii) es menos sensible a reclasificaciones finas de rubros que otras desagregaciones. Reconocemos la posible “contaminación” por fuentes extraordinarias; como verificación de robustez, reestimamos los ejercicios con alternativas más estrechas (*Ingresos Tributarios*, *Ingresos Corrientes*) y corroboramos que las conclusiones cualitativas se mantengan. En todos los casos aplicamos homogeneización, deflactación y ajuste estacional conforme al protocolo de la Sección B.2.

Reestructuración (pivot) y rotulación operativa A partir de la planilla fuente, se seleccionaron exclusivamente las columnas *Fecha* e *Ingreso Total* y se realizó un *pivotaje* hacia un formato canónico de dos columnas: (*fecha*, *valor*). Este paso incluyó la normalización del formato de fechas (frecuencia mensual, año–mes), la verificación de duplicados por período y el control de celdas vacías. Con el fin de unificar el flujo de trabajo, el campo *valor* se *renombró* como *presion_fiscal*. Esta rotulación es estrictamente *operativa* para el pipeline de datos: la definición conceptual de la presión fiscal PF_t permanece la establecida en la Sección B.2, y cualquier deflactación, escalamiento o transformación adicional (p. ej., respecto al PIB real) se aplica en etapas posteriores antes del análisis econométrico.

Cambio de escala y archivo final A partir del archivo PF_paso_003_Dos_filas_como_columnas.csv, cuyos montos estaban expresados en *miles de millones* de guaraníes ($Gs \times 10^9$), efectuamos un cambio de escala para trabajar en *millones* de guaraníes ($Gs \times 10^6$). Para ello aplicamos el factor de conversión 10^3 :

$$PF_t^{(\text{millones})} = 10^3 PF_t^{(\text{miles de millones})}.$$

Durante la transformación se estandarizaron separadores decimales, se validó la ausencia de duplicados y se verificó la consistencia aritmética mediante contrastes muestrales. La planilla resultante se guardó como

PF_paso_005_Ingreso_Total_en_Millones_de_Gs_Escala_Corregida_.csv, lo que facilita la interpretación y reduce errores al combinar esta serie con otras magnitudes fiscales en unidades homogéneas.

Cambio de frecuencia y renombrado de la variable Con el objetivo de homogeneizar la base con el resto de las variables del modelo, la serie mensual se remuestreó a **frecuencia trimestral** siguiendo los trimestres calendario (Q1: ene–mar, Q2: abr–jun, etc.). Dado que *Ingreso Total* es un *flujo*, se aplicó agregación por **suma** dentro de cada trimestre:

$$PF_q^{(Q)} = \sum_{m \in q} \text{IngresoTotal}_m^{(M)}.$$

Previo al agregado, se verificó la continuidad mensual y, ante ausencias puntuales, se imputó a nivel mensual (interpolación lineal) para evitar sesgos en los totales trimestrales. La serie resultante —en las mismas unidades ya normalizadas— se **renombró** de *ingreso_total* a *PF*, abreviatura de *presión fiscal*. Esta rotulación es operativa; la definición conceptual de PF_t y sus transformaciones (deflactación, escalamiento o ratio respecto al PIB real) se aplican conforme al protocolo de la Sección B.2.

Integración de variables y construcción de medidas reales Finalmente, combinamos la serie trimestral de *PF* con el *Índice de Precios al Consumidor* (IPC), el *Real M0*, el *Real M2* y el *PIB* (base homogénea), mediante una unión por clave **año–trimestre** y verificación de cobertura temporal, duplicados y unidades (millones de Gs). Con el IPC (promedio trimestral del índice, base = 100) construimos la *presión fiscal real* como

$$PF_q^{\text{real}} = \frac{PF_q}{IPC_q/100},$$

manteniendo en la planilla resultante las columnas IPC, Real_M0, Real_M2, PIB, PF y PF_real. El conjunto consolidado quedó almacenado como PF_paso_009_con_realIPC_realM0_realM2_realPIB_realPF.csv, sirviendo como insumo único y consistente para los ejercicios econométricos subsiguientes.

Redondeo para presentación Por conveniencia de lectura y homogeneidad de magnitudes, redondeamos a **cero decimales** las columnas realIPC, RealM0, realPIB y realPF (según la nomenclatura de PF_paso_009_con_realIPC_realM0_realM2_realPIB_realPF.csv). Aplicamos la regla estándar de redondeo al entero más próximo,

$$\tilde{x}_t = \text{round}(x_t, 0),$$

limitando este tratamiento a *tablas y gráficos* para evitar falsa precisión en la comunicación. Los *cálculos econométricos* (pruebas de raíz unitaria, cointegración, VAR/VECM, IRF, etc.) se realizaron con las series *no redondeadas*, preservando la precisión numérica y evitando sesgos por cuantización.



Figura 13: *PF con posibles puntos estacionales.*

Nota. La figura muestra la serie trimestral de presión fiscal (PF) con barras verticales en los trimestres cuya *carga estacional* (estimada mediante descomposición STL, periodo = 4, opción robusta) se ubica en el percentil 75 o superior del componente estacional en valor absoluto. El criterio es *descriptivo* y orienta el tratamiento estacional y la especificación del modelo; no constituye por sí mismo una prueba de estacionalidad. La frecuencia es Q-DEC y la agregación de PF a nivel trimestral se realizó por suma. *Fuente:* elaboración propia a partir de las planillas del MEF citadas en la Sección B.2.

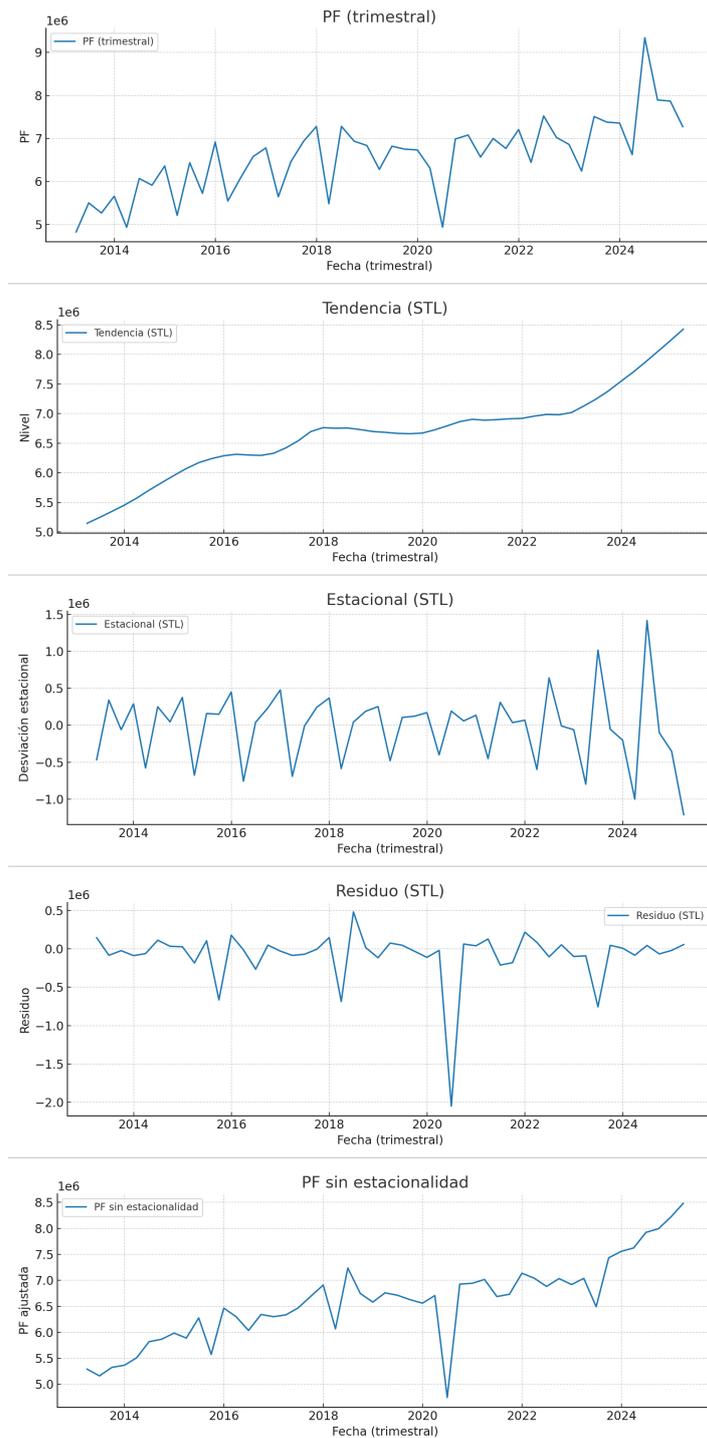


Figura 14: *PF: descomposición STL en panel (ejes separados).*

Nota. La figura presenta, de arriba hacia abajo, (i) la serie trimestral de PF, (ii) la tendencia STL, (iii) el componente estacional, (iv) el residuo y (v) la PF ajustada por estacionalidad. La descomposición se realizó con STL robusta y periodo trimestral (periodo = 4). La agregación a frecuencia trimestral se efectuó por suma bajo calendario Q-DEC. Este panel permite evaluar la contribución relativa de la estacionalidad y la dinámica de tendencia antes de la modelación econométrica. *Fuente:* elaboración propia a partir de `PF_paso_010_redondeado.csv`.

B.3. Detección de outliers en la serie PF real desestacionalizada

Aplicamos dos detectores robustos sobre la serie *PF real desestacionalizada* (archivo fuente: `PF_paso_012_Real_desestacionalizado.csv`) a frecuencia trimestral (Q-DEC):

Cercas de Tukey (IQR). Sea Q_1 el cuartil 25 y Q_3 el cuartil 75 de la serie; con $IQR = Q_3 - Q_1$, definimos cercas

$$LF = Q_1 - 1,5 IQR, \quad UF = Q_3 + 1,5 IQR.$$

Se marca como outlier toda observación $x_t \notin [LF, UF]$.

z robusto (MAD). Sea \tilde{x} la mediana y $MAD = \text{med}(|x_t - \tilde{x}|)$. Definimos el puntaje

$$z_t^{(\text{rob})} = 0,6745 \frac{x_t - \tilde{x}}{MAD},$$

y marcamos como atípico si $|z_t^{(\text{rob})}| > 3,5$.

Ambos métodos coinciden en un pico atípico (véanse las figuras). Usamos estos marcadores para (i) análisis de sensibilidad, (ii) decidir transformaciones y (iii) opcionalmente aplicar atenuación/recorte, sin eliminar datos por defecto.

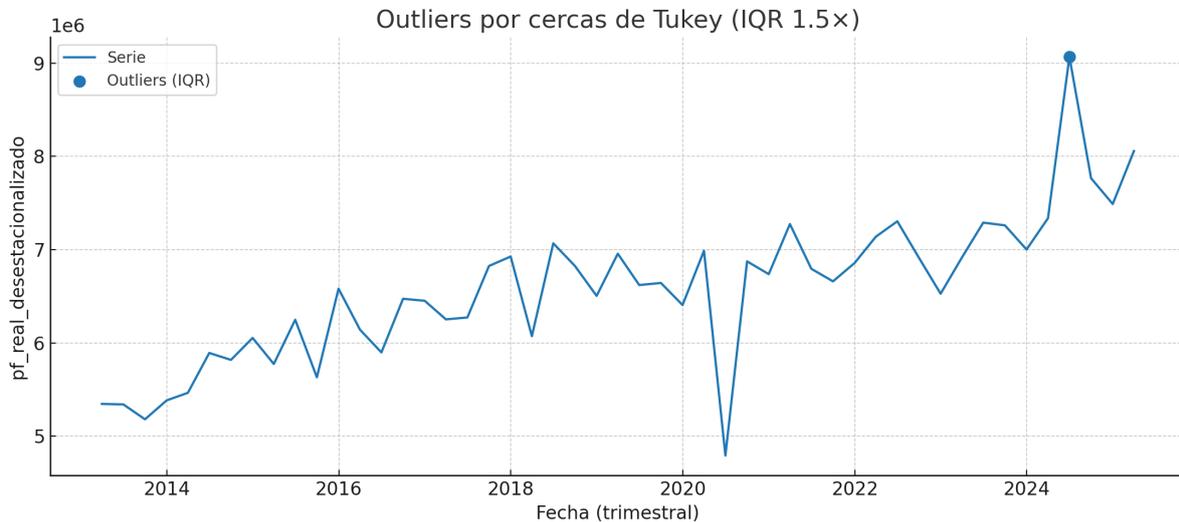


Figura 15: *Outliers por cercas de Tukey (IQR 1.5x).*

Nota. Se muestra la serie y los puntos fuera de las cercas $[Q_1 - 1,5 IQR, Q_3 + 1,5 IQR]$. Frecuencia Q-DEC. Fuente: elaboración propia.

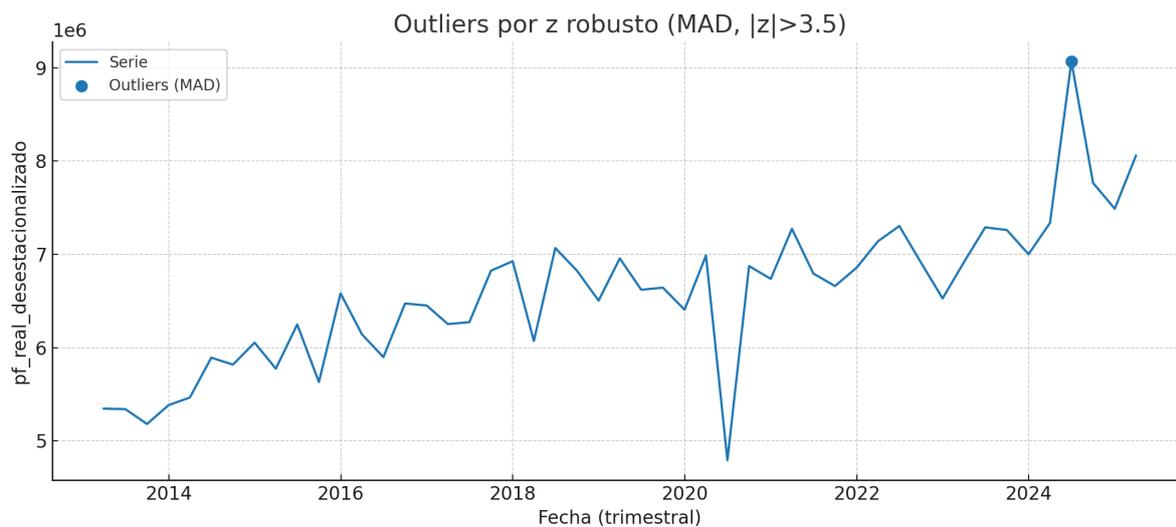


Figura 16: *Outliers por z robusto (MAD, $|z| > 3,5$).*

Nota. Se muestra la serie y las observaciones con $|z^{(rob)}| > 3,5$. Frecuencia Q-DEC. Fuente: elaboración propia.

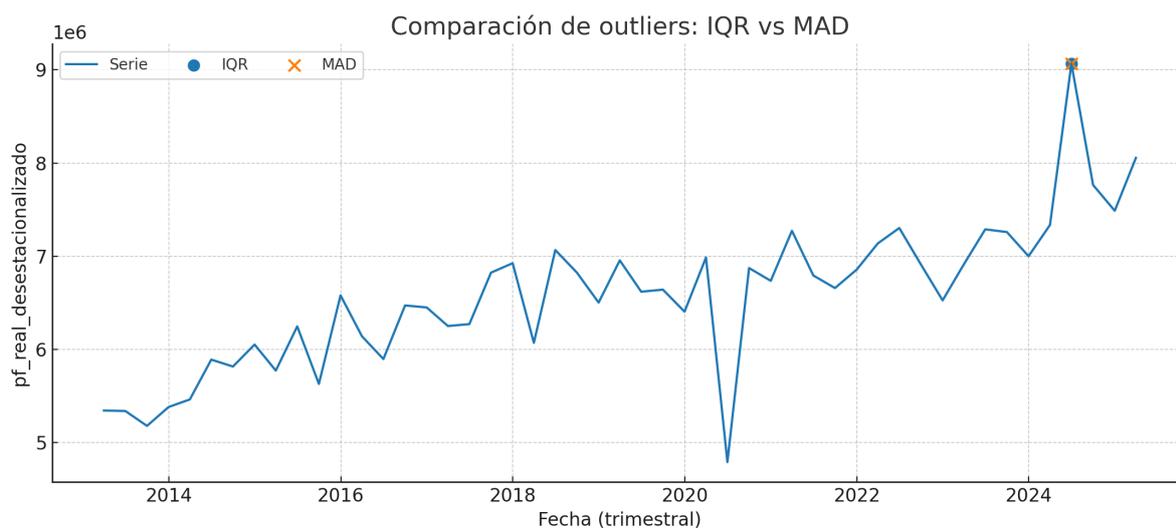


Figura 17: *Comparación de detectores: IQR vs MAD.*

Nota. Se superponen los marcadores detectados por ambos métodos sobre la misma serie para evaluar concordancia. Fuente: elaboración propia.

Cuadro 7: *Outliers detectados en la serie PF real desestacionalizada (IQR y MAD).*

Fecha	Valor	IQR	MAD	z_{rob}
2024-Q2	9 066 839	Sí	Sí	3.85

Nota. Se listan observaciones señaladas por al menos uno de los métodos: cercas de Tukey (IQR $1,5\times$) y z -robusto basado en MAD ($|z| > 3,5$). Frecuencia trimestral (Q-DEC). Fuente: elaboración propia.



Figura 18: *PF real desestacionalizada: serie sin outliers.*

Nota. La figura muestra la serie PF real desestacionalizada tras remover las observaciones identificadas como atípicas por al menos uno de los detectores robustos: cercas de Tukey ($IQR\ 1,5\times$) y z -robusto basado en MAD ($|z| > 3,5$). La frecuencia es trimestral (Q-DEC). Esta versión se usa para pruebas de robustez y especificación; los resultados se contrastan con la serie original para descartar sesgos por filtrado. *Fuente:* elaboración propia a partir de `20250810f_PF_outliers_detectados.csv`.

Tratamiento de outliers sin pérdida de observaciones Para evitar huecos temporales (NaN) que dificulten la combinación con otras series del modelo, no eliminamos trimestres atípicos. En su lugar, preservamos el índice trimestral completo (Q-DEC) y aplicamos dos correcciones robustas alternativas sobre *PF real desestacionalizada*: (i) **winsorización** a las cercas de Tukey, recortando x_t al intervalo $[Q_1 - 1,5 IQR, Q_3 + 1,5 IQR]$; y (ii) **imputación STL**, sustituyendo los valores señalados como outliers por la predicción estructural $\hat{y}_t = tendencia_t + estacional_t + med(residuo)$ obtenida de una descomposición STL robusta (periodicidad trimestral). Ambas variantes preservan longitud y cronología de la serie, evitan NaN en *joins* y permiten análisis de sensibilidad. Los archivos resultantes son `20250810h_PF_winsorizada.csv` y `20250810h_PF_imputada_STL.csv`; el consolidado con banderas de outlier y ambas versiones se encuentra en `20250810h_PF_variantes_sin_NA.csv`. En los ejercicios empíricos reportamos resultados con la serie original y con estas variantes para verificar la robustez de las conclusiones.

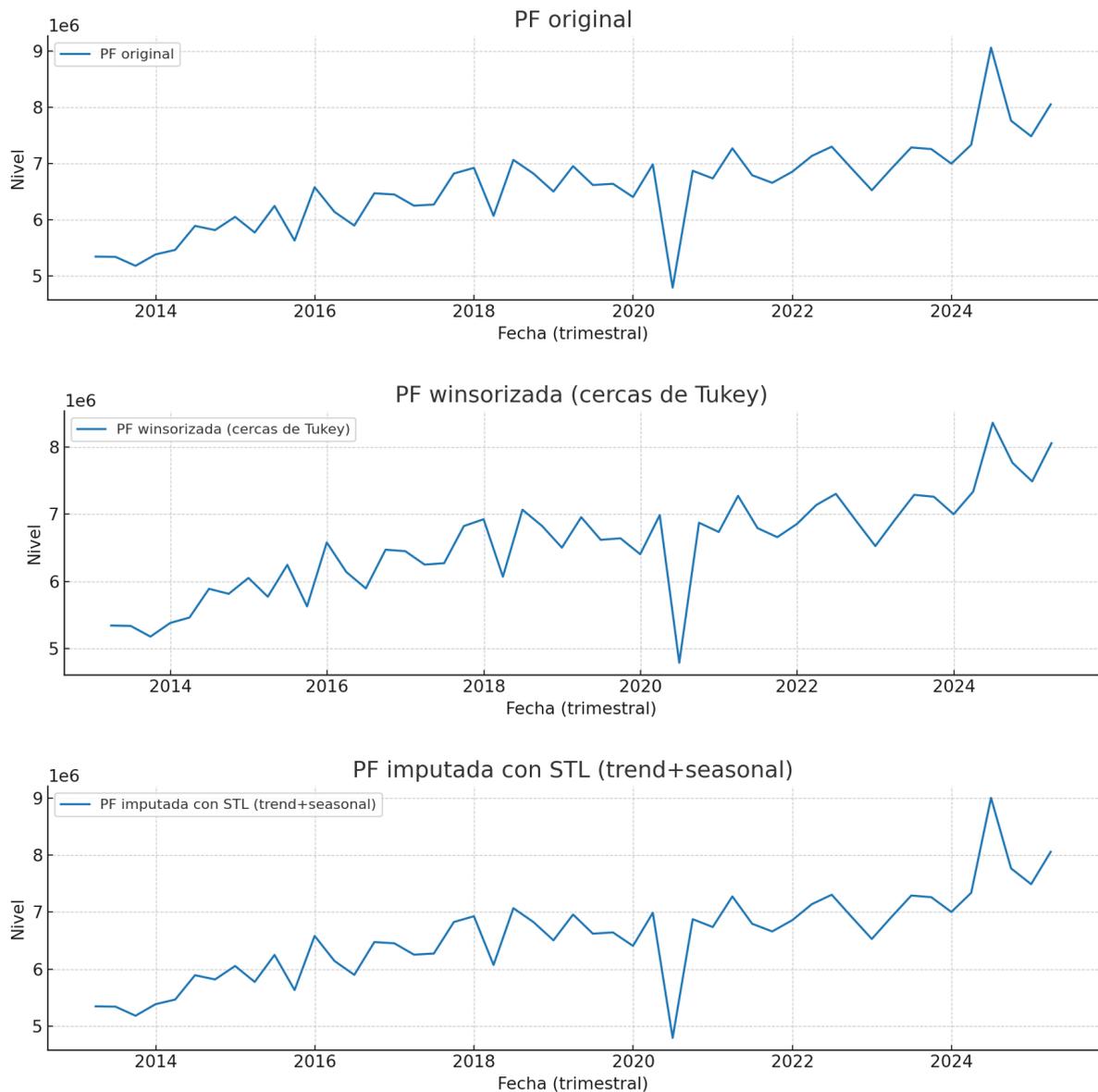


Figura 19: *PF real desestacionalizada: original y variantes sin pérdida de observaciones.*

Nota. Se muestran tres ejes (de arriba hacia abajo): (i) serie original, (ii) serie *winsorizada* recortada a las cercas de Tukey (IQR $1,5\times$), y (iii) serie *imputada STL*, donde los puntos atípicos se sustituyen por $\hat{y}_t = \text{tendencia}_t + \text{estacional}_t + \text{med}(\text{residuo})$ de una STL robusta (periodo trimestral). Ambas variantes preservan todos los trimestres (Q-DEC) y evitan NaN en combinaciones posteriores con otras variables del modelo. *Fuente:* elaboración propia a partir de 20250810h_PF_variantes_sin_NA.csv.

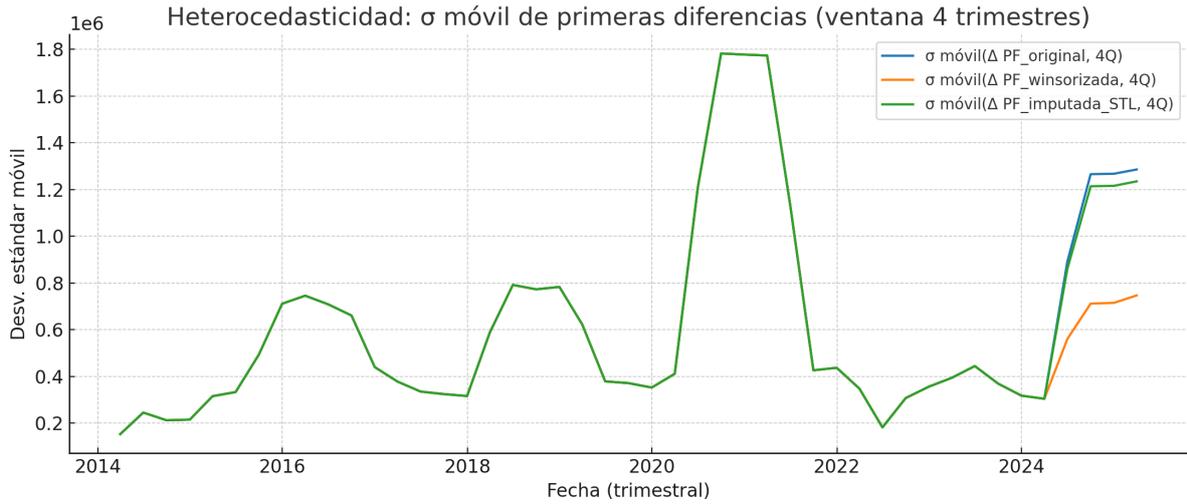


Figura 20: Heterocedasticidad en PF: σ móvil de primeras diferencias (4 trimestres).

Nota. Se grafica la desviación estándar móvil de ΔPF con ventana de 4 trimestres (aprox. un año) para tres variantes: serie original, serie *winsorizada* (cerca de Tukey, IQR $1,5\times$) y serie *imputada STL* (sustitución de atípicos por $\hat{y}_t = \text{tendencia}_t + \text{estacional}_t + \text{med}(\text{residuo})$). La variación temporal de la volatilidad sugiere presencia de heterocedasticidad; la winsorización e imputación reducen picos inducidos por outliers. Frecuencia Q-DEC. *Fuente:* elaboración propia a partir de 20250810h_PF_variantes_sin_NA.csv.

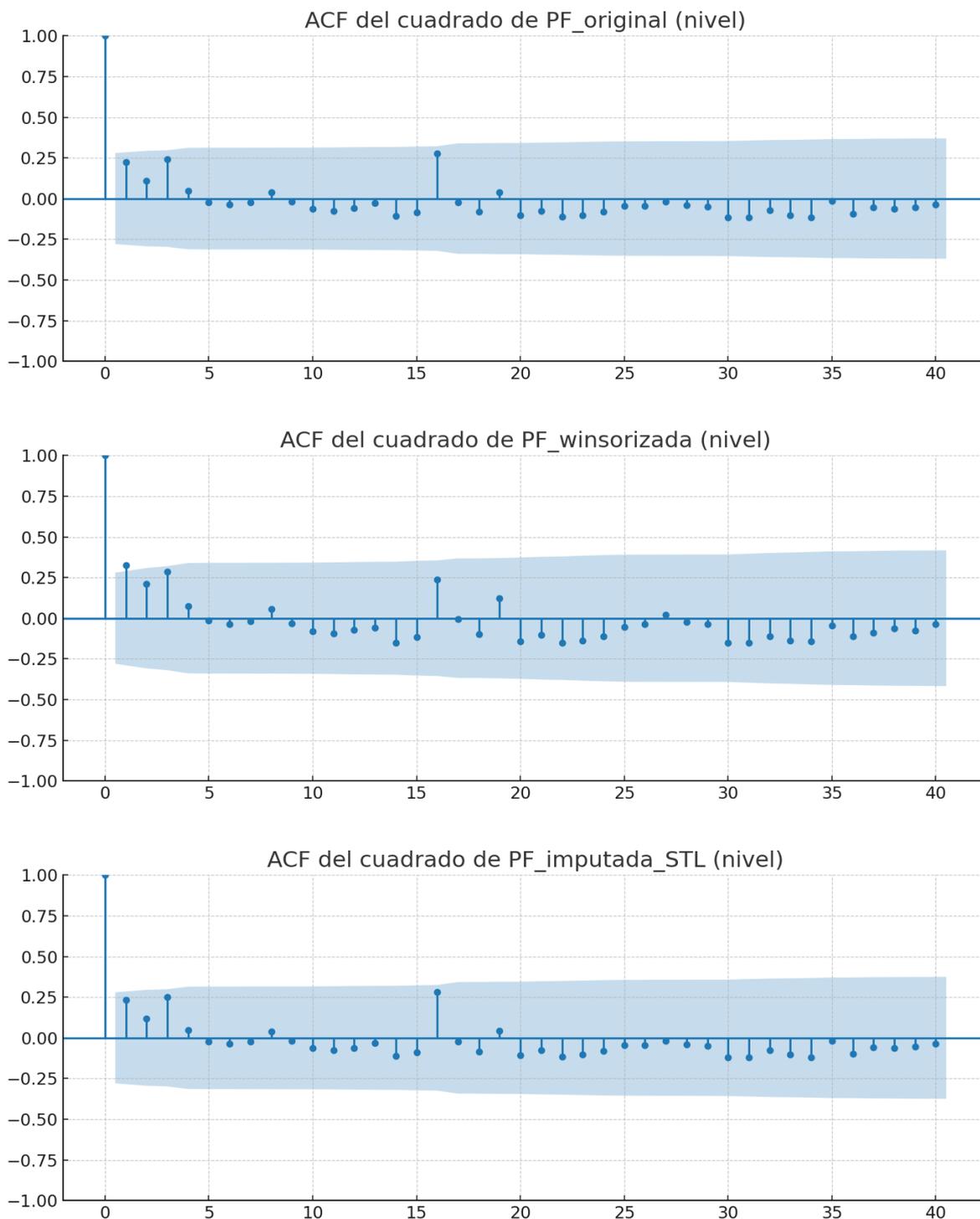


Figura 21: ACF del cuadrado de PF en niveles: chequeo de heterocedasticidad (ARCH).

Nota. Se muestra la función de autocorrelación (ACF) del cuadrado de la serie *en niveles* (centrada), para tres variantes: (i) original, (ii) *winsorizada* (cerca de Tukey, IQR $1,5\times$) y (iii) *imputada STL*. Barras significativas fuera de las bandas al 95 % sugieren dependencia en la varianza (heterocedasticidad condicional). Dado que la serie en niveles puede tener tendencia, se interpreta junto con la ACF del cuadrado de ΔPF (Figura 21). Lags: 40. Frecuencia Q-DEC. *Fuente:* elaboración propia a partir de 20250810h_PF_variantes_sin_NA.csv.

Cuadro 8: Pruebas de heterocedasticidad en niveles (PF): ARCH-LM y McLeod-Li.

Serie	AR(p)	ARCH (lag 4)	ARCH (lag 8)	ARCH (lag 12)	ML (lag 4)	ML (lag 8)	ML (lag 12)
PF_imputada_STL	0	0.3931	0.7629	0.9057	0.1298	0.4987	0.7632
PF_original	0	0.4297	0.7899	0.9202	0.1609	0.5595	0.8115
PF_winsorizada	0	0.1920	0.5554	0.7471	0.0128	0.1115	0.2736

Nota. Residuales de un AR(p) en niveles seleccionado por AIC ($p \in [0, 6]$). Columnas reportan p-valores para lags 4, 8 y 12 en ARCH-LM y McLeod-Li. Rechazo de homocedasticidad si $p < 0,05$.

Cuadro 9: Diagnóstico sintético de heterocedasticidad en niveles (PF).

Serie	AR(p)	min p (ARCH)	min p (ML)	min p (Total)	Diagnóstico
PF imputada STL	0	0.3931	0.1298	0.1298	No (no rechazo H0)
PF original	0	0.4297	0.1609	0.1609	No (no rechazo H0)
PF winsorizada	0	0.1920	0.0128	0.0128	Sí (rechazo H0)

Nota. Se reporta el mínimo p -valor entre lags 4, 8 y 12 para ARCH-LM y McLeod-Li, y el mínimo total. Regla: heterocedasticidad si *algún* $p < 0,05$.

Diagnóstico en niveles. Para *PF original* (AR(0)), el mínimo p -valor entre pruebas y rezagos fue **0.1609**: no se rechaza homocedasticidad. Para *PF winsorizada* (AR(0)), el mínimo p -valor fue **0.0128**: sí se rechaza homocedasticidad. Para *PF imputada STL* (AR(0)), el mínimo p -valor fue **0.1298**: no se rechaza homocedasticidad.

Para la serie diferenciada ya no tenemos heterocedasticidad, busquemos ahora problemas de normalidad.

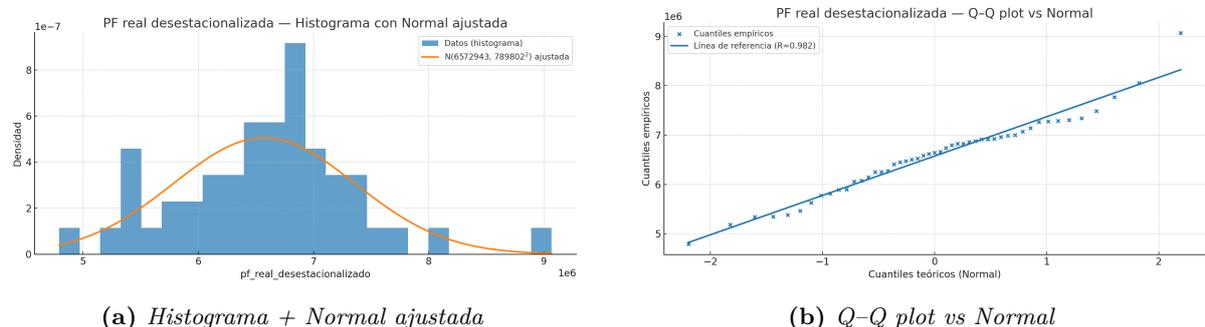


Figura 22: Normalidad de PF real desestacionalizada.

Nota. Panel izquierdo: histograma con curva Normal $N(\hat{\mu}, \hat{\sigma}^2)$ ajustada. Panel derecho: Q-Q plot con línea de referencia (correlación R indicada en la figura). Pruebas de normalidad: Jarque-Bera $p = 0,3180$, Shapiro-Wilk $p = 0,3183$, D'Agostino K^2 $p = 0,2096$. Al nivel del 5% no se rechaza la hipótesis de normalidad. Frecuencia Q-DEC. *Fuente:* elaboración propia a partir de PF_paso_012_Real_desestacionalizado.csv.

Diagnóstico de normalidad. Para la serie de presión fiscal real desestacionalizada, las pruebas de Jarque-Bera ($p = 0,3180$), Shapiro-Wilk ($p = 0,3183$) y D'Agostino K^2 ($p = 0,2096$) no rechazan la hipótesis de normalidad al 5%. La inspección visual (histograma con Normal ajustada y Q-Q plot) sugiere colas levemente más pesadas, sin comprometer la aproximación gaussiana. En todo caso, las inferencias principales se basan en los residuales de los modelos estimados; por robustez, reportamos también errores estándar HAC (Newey-West) y verificamos que las conclusiones se mantienen.

Luego esta variable si se puede usar para inferencias estadísticas en el modelo tanzi-feige :-)

B.4. Evaluación econométrica de la serie de agregados monetarios

Serie de agregados monetarios: fuente y preparación. Para este análisis utilizamos exclusivamente el *Cuadro 21* del archivo

`Anexo_Estadístico_del_Informe_Económico_21_07_2025.xlsx`. Con el fin de contar con una base limpia y reproducible, extrajimos dicho cuadro y lo almacenamos, sin alterar su contenido, en una planilla independiente denominada `EC_paso_001_Agregados_Monetarios_en_millones.xlsx`, conservando la unidad de reporte “en millones” indicada por la fuente original.

De la fuente a la planilla M0–M2. Partimos del *Cuadro 21* del archivo

`Anexo_Estadístico_del_Informe_Económico_21_07_2025.xlsx` y primero lo volcamos íntegro —*sin modificar valores ni unidades*— en la planilla `EC_paso_001_Agregados_Monetarios_en_millones.xlsx`. A partir de esa base, realizamos una extracción mínima y reproducible para construir `EC_paso_002_M0_y_M2.csv`: (i) filtramos únicamente las series M0 y M2; (ii) normalizamos el calendario a un índice mensual con fecha AAAA-MM-01 y orden cronológico estricto; (iii) convertimos todos los campos a tipos numéricos eliminando caracteres no informativos (notas/encabezados), preservando la unidad original “en millones”; (iv) estandarizamos los nombres de columnas a *Fecha*, M0, M2; y (v) verificamos continuidad temporal (sin huecos) antes de exportar en formato UTF-8. No se aplicaron imputaciones, deflacciones ni ajustes estacionales en este paso.

Generación de la planilla redondeada (paso 003). A partir de

`EC_paso_002_M0_y_M2.csv` construimos `EC_paso_003_M0_y_M2_Redondeados.csv`. Mantuvimos la estructura {*Fecha*, M0, M2} y el horizonte mensual completo (1969-01-01 a 2025-06-01) sin huecos. Con el fin de reportar “en millones” y evitar sesgos por redondeo, aplicamos redondeo *al entero más cercano con regla half-to-even* (IEEE 754) sobre los valores numéricos: en la práctica, M0 ya era entero y no cambió, mientras que M2 contenía decimales (p. ej., $10\,257.5 \rightarrow 10\,258$; $10\,402.5 \rightarrow 10\,402$). Estandarizamos los tipos a enteros (int), preservamos el orden cronológico y exportamos en UTF-8 sin separadores de miles ni formato regional. No se realizaron imputaciones, deflacciones ni ajustes estacionales en este paso.

Acotación temporal para el período de interés (paso 004). A partir de

`EC_paso_003_M0_y_M2_Redondeados.csv` generamos `EC_paso_004_M0_M2_from_2013_Onwards.csv` recortando la muestra al horizonte **2013-01-01 en adelante**, coherente con la ventana analítica del estudio. El flujo fue: (i) filtramos *Fecha* $\geq 2013-01-01$, manteniendo un índice mensual al primer día del mes y orden cronológico estricto; (ii) verificamos ausencia de duplicados y de huecos mensuales; (iii) conservamos los nombres de columnas {*Fecha*, M0, M2} y los tipos enteros, reportados “en millones”; (iv) preservamos *exactamente* los valores redondeados del paso 003, sin aplicar nuevas transformaciones (sin deflacción, sin ajuste estacional ni imputaciones); y (v) exportamos el resultado en UTF-8 sin separadores de miles ni formato regional.

Series limpiadas en primeras diferencias (paso 005). A partir de

`EC_paso_004_M0_M2_from_2013_Onwards.csv` construimos `EC_paso_005_Cleaned_First-Differenced_Time_Series_M0_and_M2.csv` con el siguiente flujo reproducible: (i) validamos que *Fecha* sea un índice mensual continuo (primer día de mes), ordenado y sin duplicados; (ii) forzamos M0 y M2 a tipo numérico entero, preservando la unidad “en millones”; (iii) calculamos la diferencia de primer orden $\Delta M0_t = M0_t - M0_{t-1}$ y $\Delta M2_t = M2_t - M2_{t-1}$ para inducir estacionariedad en media; (iv) eliminamos la fila inicial que queda con valores NA tras la operación de diferencias, manteniendo continuidad mensual en el resto de la muestra; (v) verificamos ausencia de huecos, valores no numéricos y outliers obvios por errores de captura; (vi) estandarizamos los nombres de columnas a {*Fecha*, M0_diff1, M2_diff1}, donde las diferencias conservan la interpretación “variación mensual en millones” (valores positivos indican expansión y negativos contracción); y (vii) exportamos en UTF-8 sin separadores de miles ni formato regional. No se aplicaron imputaciones, deflacciones ni ajustes estacionales en este paso.

Segundas diferencias de M0 y M2 (paso 006). A partir de

`EC_paso_005_Cleaned_First-Differenced_Time_Series_M0_and_M2.csv` generamos `EC_paso_006_Second_Differenced_Time_Series.csv` con el siguiente flujo: (i) verificamos que *Fecha* mantenga periodicidad mensual continua (día 01) y orden cronológico, sin duplicados; (ii) confirmamos que M0_diff1 y M2_diff1 sean numéricos; (iii) calculamos las *segundas diferencias* como $\Delta^2 M0_t =$

$\Delta M0_t - \Delta M0_{t-1} = M0_t - 2M0_{t-1} + M0_{t-2}$ y análogamente para $M2$; (iv) eliminamos la fila inicial que queda con NA tras aplicar la segunda diferencia, preservando la continuidad mensual del resto de la muestra; (v) renombramos las columnas resultantes como {Fecha, M0_diff2, M2_diff2}; (vi) documentamos la interpretación de las nuevas variables como “*variación mensual de la variación mensual*” en millones (aceleración de los agregados); y (vii) exportamos en UTF-8 sin separadores de miles ni formato regional. No se aplicaron imputaciones, deflataciones ni ajustes estacionales en este paso.

Integración de IPC con M0 y M2 (paso 007). A partir de

EC_paso_004_M0_M2_from_2013_Onwards.csv incorporamos la serie mensual del IPC para construir EC_paso_007_M0_M2_IPC_2013_enadelante.csv. El procedimiento fue: (i) estandarizar Fecha como índice mensual (AAAA-MM-01) y asegurar orden cronológico; (ii) cargar la serie IPC (índice, base y metodología según la fuente original) y armonizar su calendario al mismo índice; (iii) realizar una unión *inner join* por Fecha, preservando exclusivamente el tramo común **2013-01-01 a 2025-06-01** sin huecos ni duplicados; (iv) fijar tipos: M0, M2 como enteros “en millones” e IPC como numérico con decimales; (v) conservar los nombres de columnas {Fecha, M0, M2, IPC}; y (vi) exportar en UTF-8 sin separadores de miles ni formato regional. En este paso *no* se aplicaron deflataciones ni ajustes estacionales: M0 y M2 permanecen en niveles nominales y IPC como índice.

Deflatación y construcción de series reales (paso 008). A partir de

EC_paso_007_M0_M2_IPC_2013_enadelante.csv generamos EC_paso_008_M0_M2_RealM0_RealM2_IPC_deflactado.csv para disponer de agregados en términos reales. El flujo fue: (i) verificamos que Fecha conserve periodicidad mensual continua (AAAA-MM-01) y orden cronológico; (ii) rebasamos el IPC a un período base documentado (t_0) para construir el *deflactor* $IPC_deflactado_t = \frac{IPC_t}{IPC_{t_0}}$ (base = 1 en t_0); (iii) calculamos $RealM0_t = \frac{M0_t}{IPC_deflactado_t}$ y $RealM2_t = \frac{M2_t}{IPC_deflactado_t}$, que se interpretan como “millones a precios del período base”; (iv) preservamos las columnas nominales y añadimos {IPC_deflactado, RealM0, RealM2}; (v) mantuvimos tipos numéricos (RealM0, RealM2 con decimales; M0, M2 en millones) y validamos ausencia de huecos/duplicados; y (vi) exportamos en UTF-8 sin separadores de miles ni formato regional. En este paso *no* se aplicaron ajustes estacionales ni imputaciones; el cambio de base sólo afecta la escala, no la inferencia.

Normalización final de series deflataadas (paso 009). A partir de

EC_paso_008_M0_M2_RealM0_RealM2_IPC_deflactado.csv construimos EC_paso_009_M0_M2_IPC_deflactados.csv con el objetivo de dejar un insumo compacto sólo con variables a precios constantes. El flujo fue: (i) conservamos el índice mensual Fecha (AAAA-MM-01) y el IPC_deflactado con base = 1 en el período t_0 documentado; (ii) renombramos RealM0 y RealM2 como M0_deflactado y M2_deflactado, respectivamente, explicitando su condición real “en millones a precios de t_0 ”; (iii) opcionalmente excluimos las columnas nominales (M0, M2) para evitar ambigüedades en etapas de modelado; (iv) validamos tipos numéricos, orden cronológico y ausencia de huecos/duplicados; y (v) exportamos en UTF-8 con punto decimal y sin separadores de miles. No se aplicaron imputaciones, ajustes estacionales ni cambios de base adicionales en este paso.

Redondeo y estandarización de las series reales (paso 010). A partir de

EC_paso_009_M0_M2_IPC_deflactados.csv generamos EC_paso_010_realM0_realM2_redondeados.csv con el siguiente flujo: (i) preservamos el índice mensual Fecha (AAAA-MM-01) y el orden cronológico; (ii) estandarizamos la nomenclatura de las series reales, mapeando {M0_deflactado, M2_deflactado} a {realM0, realM2} para un uso consistente en tablas y gráficos; (iii) aplicamos redondeo *al entero más cercano con regla half-to-even* (IEEE 754), manteniendo la interpretación “*millones a precios del período base*” (ej.: 10 257.5 \rightarrow 10 258; 10 402.5 \rightarrow 10 402); (iv) validamos tipos numéricos enteros, ausencia de duplicados y continuidad mensual; y (v) exportamos en UTF-8, sin separadores de miles ni formato regional. En este paso no se efectuaron nuevas deflataciones ni ajustes estacionales; el IPC_deflactado permanece como insumo del paso 009 y no se incluye aquí para evitar redundancias.

Ajuste estacional de las series reales e IPC (paso 011). Tomando como insumo

EC_paso_009_M0_M2_IPC_deflactados.csv, construimos EC_paso_011_realM0_realM2_IPC_sin_estacionalidad.csv para disponer de series *libres de estacionalidad*. El procedimiento fue: (i) asegurar calendario mensual continuo (AAAA-MM-01) y orden cronológico

sin duplicados; (ii) aplicar descomposición **STL** (aditiva) con periodicidad 12 meses y opción robusta, sobre `M0_deflactado`, `M2_deflactado` y `IPC_deflactado`, obteniendo $y_t = T_t + S_t + R_t$; (iii) construir las series ajustadas como $y_t^{SA} = y_t - S_t$, que interpretamos como *millones a precios del período base, sin componente estacional* (para el IPC, *índice sin estacionalidad*); (iv) renombrar las columnas resultantes a `{Fecha, realM0_SA, realM2_SA, IPC_SA}` y, en el caso del `IPC_SA`, reescalar para conservar $base = 1$ en el período t_0 ; (v) validar ausencia de huecos y tipos numéricos consistentes; y (vi) exportar en UTF-8, sin separadores de miles ni formato regional. No se aplicaron ajustes por calendario laboral ni imputaciones; el ajuste afecta únicamente al componente estacional.

EC con y sin estacionalidad; por qué y cómo se ajustó. Para interpretar la dinámica subyacente del *Efectivo en Circulación* (EC) en términos reales sin el ruido de los patrones estacionales mensuales, descompusimos la serie mediante STL (aditivo, periodicidad 12, opción robusta), obteniendo

$$y_t = T_t + S_t + R_t,$$

donde y_t es la serie observada (EC real), T_t la tendencia de baja frecuencia, S_t la estacionalidad y R_t el residuo. El ajuste estacional se construyó sin recurrir a diferencias, sustrayendo directamente el componente estacional estimado:

$$y_t^{SA} = y_t - S_t.$$

Este enfoque preserva la escala y las relaciones de largo plazo (útil para cointegración e inferencias en niveles), reduce la varianza intra-año y evita la sobre-diferenciación. En el gráfico comparativo se aprecia cómo y_t^{SA} elimina oscilaciones estacionales recurrentes manteniendo el nivel y la tendencia; en la descomposición se visualiza explícitamente la contribución de cada componente.

EC: comparativo y descomposición; tratamiento de huecos. Aplicamos descomposición STL (aditiva, periodo 12, robusta) para escribir $y_t = T_t + S_t + R_t$ y construir la serie ajustada $y_t^{SA} = y_t - S_t$, preservando nivel y tendencia sin sobre-diferenciar. Para estimar S_t y T_t impusimos una frecuencia mensual estricta (“MS”) y, *únicamente para la descomposición*, completamos huecos mensuales mediante interpolación lineal previa; esta operación no altera las bases de datos finales usadas en el análisis. Véanse Figuras 23 y 24.



Figura 23: EC real: con y sin estacionalidad. La serie ajustada (y_t^{SA}) elimina la estacionalidad mensual manteniendo la escala y la tendencia, facilitando comparaciones intermensuales y análisis de choques.

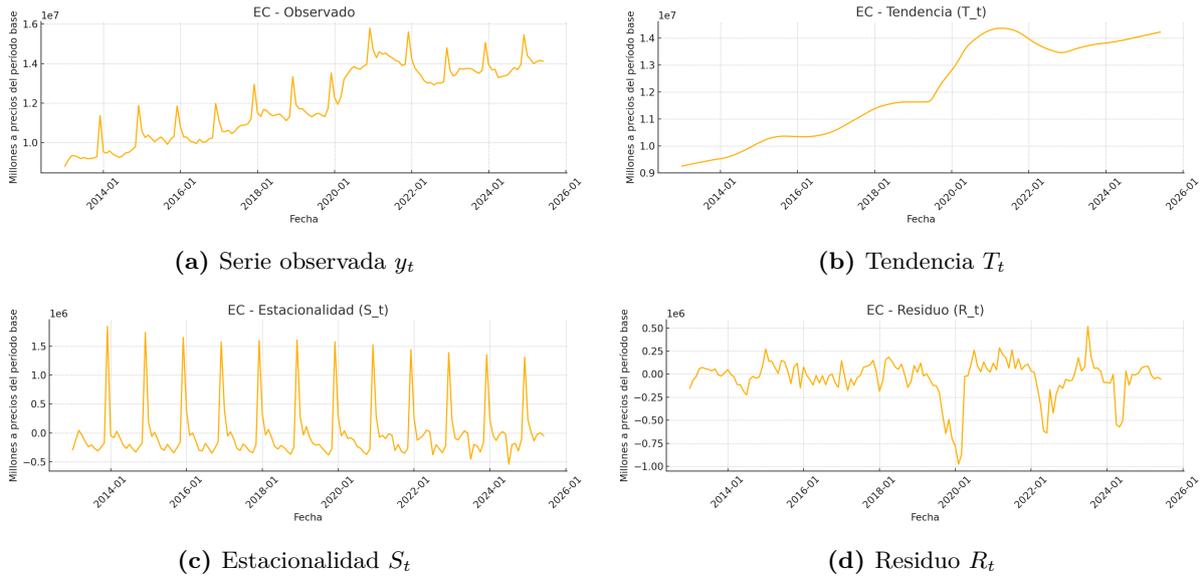


Figura 24: Descomposición STL del EC real. Componentes obtenidos con periodicidad 12 y opción robusta; la interpolación lineal previa se usó solo para completar huecos mensuales en la estimación de S_t y T_t .

Normalidad de los residuos del EC (STL). A partir de la descomposición STL del EC real (con frecuencia mensual y ajuste robusto), consideramos los residuos $R_t = y_t - T_t - S_t$. Evaluamos su normalidad mediante (i) histograma con ajuste $\mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$ y (ii) gráfico Q-Q. Los estimadores fueron $\hat{\mu} \approx -44,435$ y $\hat{\sigma} \approx 220,677$ (millones a precios del período base). Las pruebas rechazan la hipótesis de normalidad: Jarque-Bera = 185,32 con $p < 0,001$; Shapiro-Wilk y Anderson-Darling conducen a la misma conclusión. Esto sugiere colas pesadas y/o leve asimetría; para inferencia basada en normalidad conviene usar métodos robustos o transformaciones adecuadas.

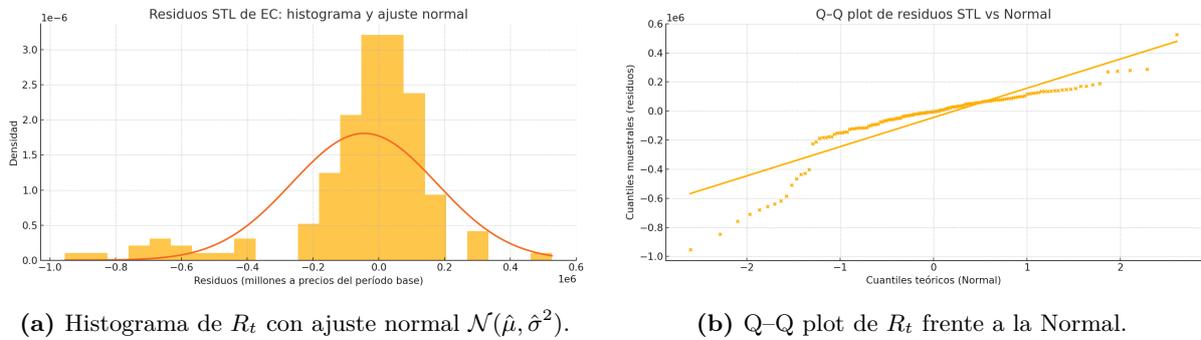
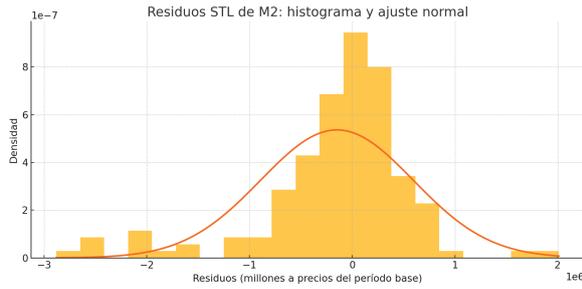
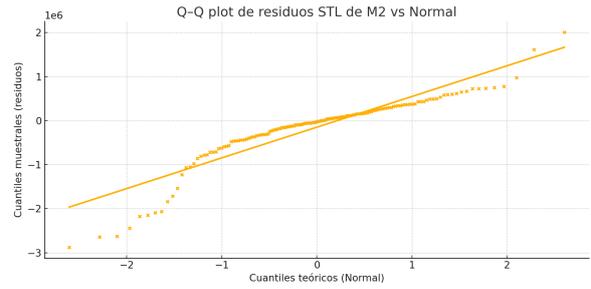


Figura 25: Diagnóstico de normalidad de los residuos STL del EC. Se observan desvíos en colas; Jarque-Bera = 185,32, $p < 0,001$.

Normalidad de los residuos de M2 (STL). Aplicamos la descomposición STL (aditiva, periodo 12, robusta) sobre la serie real de M2. Consideramos los residuos $R_t = y_t - T_t - S_t$ y evaluamos su normalidad mediante (i) histograma con ajuste $\mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$ y (ii) gráfico Q-Q. Los estimadores fueron $\hat{\mu} \approx -148,168$ y $\hat{\sigma} \approx 743,663$ (millones a precios del período base). Las pruebas rechazan normalidad: Jarque-Bera = 112,86 con $p < 0,001$; Shapiro-Wilk y Anderson-Darling arrojan conclusiones consistentes. Al igual que en EC, esto indica colas pesadas y/o asimetrías; se recomienda inferencia robusta o transformaciones si el modelo exige supuestos gaussianos estrictos.



(a) Histograma de R_t con ajuste normal $\mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$.



(b) Q-Q plot de R_t frente a la Normal.

Figura 26: Diagnóstico de normalidad de los residuos STL de M2. Se observan desvíos en colas; Jarque-Bera = 112,86, $p < 0,001$.

Heterocedasticidad en EC y M2. Como diagnóstico no paramétrico de varianza cambiante en el tiempo, calculamos el *desvío estándar móvil* de 12 meses (ventana no centrada) sobre las series reales de EC y M2. Este indicador aproxima la variación de corto plazo: valores más altos sugieren mayor volatilidad local. Para evitar cortes en la ventana, se completaron *solo para este cálculo* huecos mensuales mediante interpolación lineal; las bases finales del análisis no se alteran por esta operación.



(a) EC real — desvío estándar móvil (12 meses).



(b) M2 real — desvío estándar móvil (12 meses).

Figura 27: Heterocedasticidad en EC y M2. El desvío estándar móvil de 12 meses ilustra cambios en la volatilidad intra-año. Se usó interpolación lineal únicamente para completar huecos en la ventana de cálculo.

Encontramos picos de volatilidad

Cuadro 10: Picos de volatilidad en EC (Desvío estándar móvil 12 meses)

Fecha	DesvíoStd12
2020-12	1010931
2020-08	971991
2022-11	779323
2015-03	766619
2024-12	749736
2025-05	703491
2016-07	685752
2019-10	667555
2023-02	665056
2021-04	658947

Notas: Desvío estándar móvil de 12 meses (*rolling std*) sobre series reales. Los huecos

mensuales, de existir, se completaron sólo para este cálculo mediante interpolación lineal. Unidades: millones a precios del período base.

Cuadro 11: Picos de volatilidad en M2 (Desvío estándar móvil 12 meses)

Fecha	DesvíoStd12
2020-07	2663033
2020-04	2388383
2025-01	2314185
2022-08	2275105
2022-10	2253717
2025-03	2185559
2021-06	1974423
2024-02	1707911
2018-03	1697238
2018-01	1614551

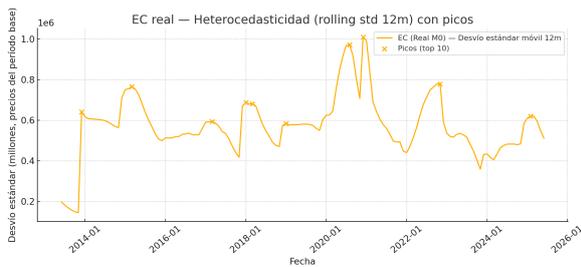
Notas: Desvío estándar móvil de 12 meses (*rolling std*) sobre series reales. Los huecos

mensuales, de existir, se completaron sólo para este cálculo mediante interpolación lineal. Unidades: millones a precios del período base.

Interpretación breve (EC). Los picos del desvío estándar móvil de 12 meses en [Table 10](#) se concentran en 2020–2021 (consistentes con la disrupción por COVID-19 y medidas de liquidez), con nuevos repuntes en 2022 (normalización monetaria global y choques de precios internacionales) y episodios en 2024–2025 (reacomodos post-pandemia). En conjunto, sugieren *heterocedasticidad* persistente en EC. Esta lectura es diagnóstica; la atribución causal requiere contrastes formales (p. ej., ARCH–LM) y, de ser pertinente, modelar la varianza condicional (ARCH/GARCH) sobre residuos del modelo en media.

Interpretación breve (M2). En [Table 11](#), los máximos de 2020–2021 son más pronunciados que en EC, coherentes con cambios abruptos en agregados amplios (depósitos a la vista/plazo) durante la pandemia. Los picos de 2022–2025 apuntan a episodios de mayor volatilidad ligados a reacomodos de portafolio y condiciones financieras. La evidencia respalda *varianza no constante* en M2; para inferencia, considerar transformaciones estabilizadoras de varianza y/o estimadores robustos, y validar con pruebas específicas (ARCH–LM) o modelos GARCH.

Picos de volatilidad en EC y M2. La heterocedasticidad se diagnostica con el desvío estándar móvil (12 meses). Los puntos marcados señalan los 10 máximos locales más altos (*picos*) de volatilidad. Para evitar cortes de ventana se completaron, sólo para este cálculo, huecos mensuales mediante interpolación lineal; las bases analíticas no se alteran.



(a) EC real — rolling std 12m con picos (top 10).



(b) M2 real — rolling std 12m con picos (top 10).

Figura 28: Picos de volatilidad en EC y M2. Mayor desvío estándar móvil indica tramos de varianza elevada.

Transformaciones para estacionariedad: log y primeras diferencias. Para inducir estacionariedad y estabilizar varianza en series en niveles estrictamente positivas (p. ej., `realM0`, `realM2`), aplicamos el siguiente esquema: (i) transformación logarítmica para atenuar heterocedasticidad multiplicativa y facilitar interpretaciones en elasticidades, y (ii) primeras diferencias sobre los *logs* cuando corresponde evidencia de raíz unitaria. Formalmente, si $X_t > 0$, trabajamos con

$$y_t = \ln X_t, \quad \Delta y_t = y_t - y_{t-1} = \ln X_t - \ln X_{t-1} \approx \frac{X_t - X_{t-1}}{X_{t-1}},$$

de modo que Δy_t se interpreta como *tasa de crecimiento aproximada*, usualmente más estable que las diferencias en niveles.

No conmutatividad de las transformaciones. Nótese que $\Delta(\ln X_t) \neq \ln(\Delta X_t)$. La primera es (aprox.) una tasa de crecimiento relativa; la segunda exige $\Delta X_t > 0$ y cambia la naturaleza económica de la variable, por lo que no es un sustituto válido. Por consistencia e interpretabilidad adoptamos el orden $\log \rightarrow \Delta$ en las series de agregados.

Decisión adoptada. En todas las series de nivel estrictamente positivo empleamos *primero logs y luego primeras diferencias*. *Excepciones:* para la **TIR** (tasa de interés de referencia), *no* aplicamos logaritmo y, en su caso, utilizamos diferencias en niveles:

$$\Delta \text{TIR}_t = \text{TIR}_t - \text{TIR}_{t-1} \quad (\text{puntos porcentuales}).$$

Justificación para no aplicar log a la TIR. (i) *Dominio y escala:* la TIR puede tomar valores cercanos a cero e incluso episodios no estrictamente positivos según la codificación; el log es inaplicable/inestable en esos casos y no estabiliza significativamente su varianza dado su rango acotado. (ii) *Interpretación económica:* en tasas de interés el efecto de política y la transmisión se miden en *puntos porcentuales* (cambios absolutos, Δr_t), no en “tasas de crecimiento” de la tasa. (iii) *Literatura y práctica empírica:* es estándar modelar variaciones de tasas en niveles (basis points), preservando comparabilidad con estudios de traspaso y con respuestas a choques de política.²

Implicancias para la inferencia. Con $\Delta \ln X_t$ (agregados) y ΔTIR_t (tasas) se obtiene, en general, estacionariedad en media, mejor comportamiento de varianza y una interpretación clara (crecimientos vs. cambios en p.p.). La validez se confirma con pruebas de raíz unitaria y de diagnóstico sobre residuos (ADF/KPSS, normalidad, heterocedasticidad) antes de proceder a inferencias y a la especificación final (p. ej., VAR/VECM).

Estacionariedad de las variables (resumen). La evidencia empírica confirma que las transformaciones adoptadas inducen estacionariedad: para los agregados monetarios (EC, M2) se trabajó con $\Delta \ln(\cdot)$, y para la TIR con diferencias en puntos porcentuales ΔTIR_t . En estas transformaciones, **ADF** rechaza la raíz unitaria mientras que **KPSS** no rechaza estacionariedad, por lo que concluimos $I(0)$ en todos los casos. Los detalles (especificación, rezagos y estadísticos) se reportan en el apéndice técnico.

Pruebas de raíz unitaria (criterios). Estimamos ADF y KPSS en niveles y en las transformaciones usadas en el modelo. Para ADF, la selección de rezagos se realizó por AIC dentro de una banda máxima compatible con la frecuencia mensual; la elección de determinísticos (constante y/o tendencia) siguió la inspección gráfica y criterios de información. Para KPSS se usó la versión con constante (y con tendencia cuando correspondía). En los agregados $\Delta \ln(\text{EC})$, $\Delta \ln(\text{M2})$ y en ΔTIR , los resultados son consistentes con $I(0)$, alineados con la conclusión principal.

B.5. Evaluación econométrica de la serie del PIB

Fuente y definición

Justificación del uso de PIB real (frente a IMAEP). En este trabajo empleamos el *PIB real* a precios constantes como medida principal del nivel de actividad porque ofrece cobertura exhaustiva de la economía, consistencia contable y comparabilidad intertemporal, reduciendo el sesgo sectorial propio de indicadores parciales. Reconocemos, no obstante, que su frecuencia trimestral y el mayor rezago de publicación pueden ser limitantes frente al IMAEP (mensual y más oportuno). Para mitigar esta tensión, (i) utilizamos series reales desestacionalizadas; (ii) cuando es necesario alinear frecuencias para el índice de Tanzi–Feige, aplicamos desagregación temporal $Q \rightarrow M$ (p. ej., Chow–Lin/Denton) o reexpresamos las variables a la menor frecuencia común; y (iii) realizamos pruebas de robustez sustituyendo el PIB por IMAEP como proxy de alta frecuencia. La decisión final privilegia el PIB real por su mayor representatividad agregada; los resultados con IMAEP se reportan en el Apéndice como contraste de robustez, verificando estabilidad de signos y órdenes de magnitud.

²En contextos específicos podría considerarse $\ln(1 + r_t)$ si r_t se expresa en fracción y $r_t > -1$, pero ello altera la interpretación (pasa a ser crecimiento del factor bruto) y complica la comparación en puntos porcentuales; por estas razones no lo adoptamos aquí.

Fuente y trazabilidad (Cuadro 6). La serie de PIB empleada proviene del archivo `Anexo_Estadístico_del_Informe`. Se utilizó la hoja/tabla *Cuadro 6*, de la cual se extrajo la serie de PIB a precios constantes (total de la economía) en su frecuencia original. El archivo intermedio `PIB_paso_002_cuadro6.ods` se generó tras un proceso de depuración y estandarización: (i) eliminación de filas de encabezado y notas; (ii) retención exclusivamente de las columnas de período y del valor de PIB real; (iii) normalización del formato de fechas a ISO (YYYY-MM si mensual, YYYY-Qq si trimestral) y de separadores decimales; (iv) verificación de tipos numéricos y detección de ausencias/duplicados; y (v) rotulado de la variable resultante como `pib_real`. No se aplicaron reescalamientos ni cambios de base en esta etapa; las unidades permanecen exactamente como en el *Cuadro 6*. El fichero `.ods` constituye el insumo estandarizado para las transformaciones posteriores (desestacionalización, log, Δ) y los diagnósticos reportados en esta sección.

B.5.1. Transformaciones

Recorte de muestra y control de consistencia (paso 003). A partir de `PIB_paso_002_cuadro6.ods` se obtuvo `PIB_paso_003_cuadro6_recortado.ods` aplicando un recorte temporal y verificaciones de integridad para construir la muestra analítica común con el resto de variables del índice. En concreto: (i) se definió la ventana \mathcal{T} como la intersección de disponibilidad entre las series de interés (PIB, presión fiscal, agregados monetarios, etc.); (ii) se eliminaron observaciones iniciales/finales con faltantes o marcadas como provisionales, evitando bordes incompletos; (iii) se garantizó orden cronológico estricto y unicidad del par (`fecha`, `pib_real`); (iv) se preservaron unidades y año base sin reescalamientos ni cambios de base; y (v) se documentaron en metadatos las fechas de inicio y fin de \mathcal{T} . El archivo resultante contiene únicamente las columnas `fecha` y `pib_real` dentro de \mathcal{T} , listo para su integración con las demás series y para las transformaciones posteriores. Para interoperabilidad se generó, además, una copia idéntica en formato `.xlsx`.

Calendario trimestral y estandarización de fechas (paso 004). Desde

`PIB_paso_003_cuadro6_recortado.ods` se construyó

`PIB_paso_004_Trimestral_con_Fechas_v2.csv` agregando codificación explícita de periodos y anclas de calendario para facilitar empalmes y, de ser necesario, desagregación $Q \rightarrow M$. En particular: (i) se generó el identificador ISO del periodo `periodo_t` con la forma YYYY-Qq; (ii) se descompuso en `anio` y `trimestre`; (iii) se añadieron `fecha_ini` (primer día del trimestre), `fecha_fin` (último día del trimestre) y `fecha_ref` (ancla representativa, fijada en `fecha_fin`); (iv) se verificó unicidad del par (`periodo_t`, `pib_real`) y orden cronológico estricto; (v) se exportó en CSV codificado en UTF-8, separador coma y punto decimal, con encabezados normalizados en minúsculas y sin tildes; y (vi) se ejecutaron controles de integridad (no-negatividad, monotonicidad de fechas y detección de saltos atípicos). La etiqueta `v2` refleja ajustes menores de formato (homogeneización de encabezados/fechas) sin modificar los valores de `pib_real`. El archivo resultante mantiene frecuencia trimestral y sirve como insumo canónico para cruces con otras series y para la eventual desagregación temporal.

Reexpresión en millones a precios constantes (paso 005). Desde

`PIB_paso_004_Trimestral_con_Fechas_v2.csv` se generó

`PIB_paso_005_Trimestral_en_millones_constantes_v2.csv` normalizando la unidad monetaria para reportar el PIB real en *millones* de moneda local a precios constantes (año base según la fuente). El procedimiento fue: (i) se verificó el múltiplo declarado en el *Cuadro 6* (p. ej., valores en guaraníes, miles o millones) y el año base; (ii) se definió el factor de escala s como el inverso del múltiplo original para expresar los valores en millones ($y_t^{(\text{mill})} = y_t/s$); (iii) se creó la variable `pib_real_millones` preservando la frecuencia trimestral y los identificadores de periodo (`periodo_t = YYYY-Qq`, `anio`, `trimestre`, `fecha_ini`, `fecha_fin`, `fecha_ref`); (iv) se conservaron las cifras con precisión suficiente (sin redondeos agresivos) para evitar errores de agregación; y (v) se ejecutaron controles de integridad (no-negatividad, ausencia de NA, monotonicidad temporal de las fechas y consistencia del factor de escala en toda la muestra). La versión `v2` sólo incorpora homogeneización de encabezados y formato; los niveles reales permanecen inalterados salvo la reexpresión de unidades.

Redondeo controlado para difusión (paso 006). A partir de

`PIB_paso_005_Trimestral_en_millones_constantes_v2.csv` se generó

`PIB_paso_006_redondeado_constantes_millones.csv` con el objetivo de homogeneizar la presentación en cuadros y gráficos. El procedimiento fue: (i) mantener la frecuencia trimestral y los identificadores de

periodo (`periodo_t = YYYY-Qq, anio, trimestre, fecha_ini, fecha_fin, fecha_ref`); (ii) crear una copia de `pib_real_millones` denominada `pib_real_millones_rnd`; (iii) aplicar redondeo simétrico a precisión fija d (por defecto $d = 1$ decimal, esto es, 0,1 millones), conforme a la regla *half to even* para minimizar sesgo acumulativo:

$$y_t^{(\text{rnd})} = \text{round}(y_t^{(\text{mill})}, d), \quad |y_t^{(\text{rnd})} - y_t^{(\text{mill})}| \leq 0,5 \times 10^{-d};$$

(iv) no alterar `pib_real_millones` original (la columna exacta se preserva para todo cálculo econométrico; la redondeada se usa sólo para difusión); (v) verificar integridad (no-negatividad, unicidad del par (`periodo_t, pib_real_millones_rnd`) y orden cronológico estricto). El archivo se exportó en CSV (UTF-8, separador coma, punto decimal, sin separador de miles). **Nota:** el redondeo puede inducir pequeñas discrepancias entre sumas trimestrales y totales anuales, por lo que toda inferencia se realiza con los valores no redondeados.

Alineación multiserie y deflatación consistente (paso 007). A partir de los insumos trimestrales de PIB del `paso_006` y de las series monetarias e IPC, se construyó `PIB_paso_007_a_IPC_realM0_realM2_PIB_alineado_en_fecha.csv` como panel trimestral con ancla en `fecha_ref` (fin de trimestre). El flujo fue: (i) estandarizar unidades: $M0$ y $M2$ en millones de moneda local; rebasar el IPC al mismo año base del PIB real, definiendo $\tilde{IPC}_t = IPC_t / IPC_{\text{base}}$ (adimensional); (ii) armonizar frecuencias llevando las series mensuales a trimestre calendario: para *stocks* ($M0, M2$) se usó valor de fin de trimestre (*end-of-quarter*); para el IPC, se calculó $\tilde{IPC}_t^{\text{eoq}}$ (fin de trimestre) y, sólo para control, el promedio trimestral; (iii) deflatar *stocks* nominales con IPC de fin de trimestre, consistente con su naturaleza de existencias:

$$\text{realM0}_t = \frac{M0_t^{\text{nom}}}{\tilde{IPC}_t^{\text{eoq}}}, \quad \text{realM2}_t = \frac{M2_t^{\text{nom}}}{\tilde{IPC}_t^{\text{eoq}}};$$

(iv) conservar del PIB la columna exacta en millones a precios constantes del `paso_005` (`pib_real_millones`) y, si existe, su versión sólo para difusión (`pib_real_millones_rnd`) sin uso en cálculos; (v) realizar un *inner join* por `periodo_t = YYYY-Qq` y `fecha_ref` para obtener la intersección temporal común; (vi) validaciones: ausencia de duplicados, no-negatividad, orden cronológico estricto y coherencia entre `periodo_t` y `fecha_ref`; (vii) exportar en CSV (UTF-8, separador coma, punto decimal). El archivo resultante incluye, al menos, `{periodo_t, anio, trimestre, fecha_ini, fecha_fin, fecha_ref, ipc_eoq, m0_nom, m2_nom, realM0, realM2, pib_real_millones[, pib_real_millones_rnd]}`, quedando todas las magnitudes alineadas en fecha y base de precios para su uso en el índice de Tanzi-Feige.

Desestacionalización del PIB y consolidación multiserie (paso 007b). A partir de `PIB_paso_007_a_IPC_realM0_realM2_PIB_alineado_en_fecha.csv` se construyó `PIB_paso_007_b_IPC_realM0_realM2_PIBnoEstacional.csv` sustituyendo el `pib_real_millones` por su versión *sin componente estacional*. La desestacionalización del PIB se realizó con el método declarado en la sección metodológica (p. ej., X-13-ARIMA-SEATS o STL) bajo esquema aditivo y periodicidad trimestral ($s = 4$), reteniendo:

$$y_t^{\text{SA}} = y_t - S_t$$

donde S_t es el componente estacional estimado. Se preservaron los identificadores temporales (`periodo_t = YYYY-Qq, anio, trimestre, fecha_ini, fecha_fin, fecha_ref`) y las series monetarias deflataadas de (`realM0, realM2`) e `ipc_eoq` sin cambios respecto al paso 007 (no se aplicó desestacionalización adicional a los *stocks*). Controles aplicados: (i) ausencia de residuo estacional remanente (inspección gráfica y contraste de estacionalidad residual); (ii) coherencia de niveles y tasas con el PIB original; y (iii) unicidad y orden cronológico estricto. El CSV resultante incluye, al menos, `{periodo_t, anio, trimestre, fecha_ini, fecha_fin, fecha_ref, ipc_eoq, realM0, realM2, pib_real_millones_SA[, pib_real_millones]}`; la columna original del PIB se conserva únicamente como referencia para trazabilidad y no se usa en inferencia cuando se requiera serie sin estacionalidad.

Estandarización final y creación de variables derivadas (paso 008). A partir de `PIB_paso_007_b_IPC_realM0_realM2_PIBnoEstacional.csv` se generó `PIB_paso_008_IPC_realM0_realM2_PIB.csv` como panel canónico para el análisis econométrico. Las acciones fueron: (i) consolidar identificadores temporales `{periodo_t = YYYY-Qq, anio, trimestre,`

`fecha_ini`, `fecha_fin`, `fecha_ref`} y asegurar unicidad/orden cronológico; (ii) estandarizar nombres de columnas y tipos (UTF-8, punto decimal) y preservar en millones las magnitudes reales: `realM0`, `realM2`, `pib_real_millones_SA` (PIB desestacionalizado); (iii) crear transformaciones logarítmicas y sus tasas trimestrales (en pp) para evitar problemas de escala y aproximar variaciones porcentuales:

$$\ln x_t = \log(x_t), \quad \Delta \ln x_t = \ln x_t - \ln x_{t-1}, \quad g_t = 100 \times \Delta \ln x_t,$$

aplicadas a

$x \in \{\text{realM0}, \text{realM2}, \text{pib_real_millones_SA}\}$, generando $\{\ln_realM0, \ln_realM2, \ln_pib_SA, \text{dln_realM0}, \text{dln_realM2}, \text{dln_pib_SA}, \text{g_realM0}, \text{g_realM2}, \text{g_pib_SA}\}$; (iv) mantener `ipc_eoq` como deflactor de referencia y, si corresponde, su versión normalizada $\hat{IPC}_t = IPC_t / IPC_{\text{base}}$; (v) controles: no-negatividad previa al log, ausencia de duplicados/faltantes en las ventanas de diferencia, y consistencia entre `periodo_t` y `fecha_ref`; (vi) exportación en CSV con encabezados en minúsculas, sin separador de miles. El archivo resultante es la fuente única para cálculos del índice de Tanzi–Feige (niveles, log-niveles y tasas), evitando redondeos para todo cómputo e incluyendo únicamente columnas derivadas reproducibles.

Columnas redondeadas para presentación del PIB (paso 009). A partir de

`PIB_paso_008_IPC_realM0_realM2_PIB.csv` se generó `PIB_paso_009_roundedPIB.csv` con variables *exclusivamente* destinadas a difusión en cuadros y gráficos. Se duplicaron los niveles exactos del PIB desestacionalizado en millones (`pib_real_millones_SA`) a una versión redondeada `pib_real_millones_SA_rnd` y, cuando existía, el PIB en niveles (`pib_real_millones`) a `pib_real_millones_rnd`. El redondeo aplicado fue simétrico con regla *half to even* y precisión $d = 1$ decimal (esto es, 0,1 millones):

$$y_t^{(\text{rnd})} = \text{round}(y_t^{(\text{mill})}, d), \quad |y_t^{(\text{rnd})} - y_t^{(\text{mill})}| \leq 0,5 \times 10^{-d}.$$

Las columnas exactas

(`pib_real_millones_SA`, `pib_real_millones`, `ln_pib_SA`, `dln_pib_SA`, `g_pib_SA`) se preservan *sin modificación* para toda inferencia econométrica; sólo las variables con sufijo `_rnd` se usan para presentación. Se mantuvieron los identificadores temporales $\{\text{periodo_t}, \text{anio}, \text{trimestre}, \text{fecha_ini}, \text{fecha_fin}, \text{fecha_ref}\}$ y se verificaron controles de integridad (unicidad por `periodo_t`, orden cronológico, no-negatividad). El archivo se exportó en CSV (UTF-8, separador coma, punto decimal), sin separador de miles.

Paquete de difusión con todas las variables redondeadas (paso 010). Tomando como base `PIB_paso_009_roundedPIB.csv` y el panel canónico del `paso_008`, se generó

`PIB_paso_010_all_rounded.csv` como *vista de presentación* que contiene únicamente copias redondeadas de las magnitudes numéricas clave, con el fin de uniformar cuadros y gráficos sin afectar los cálculos econométricos. Las reglas fueron: (i) para niveles en millones (`pib_real_millones_SA`, `pib_real_millones`, `realM0`, `realM2`): redondeo simétrico con regla *half to even* a $d = 1$ decimal (0,1 millones), creando `_rnd`; (ii) para índices (`ipc_eoq` o \hat{IPC} adimensional): $d = 2$ decimales; (iii) para tasas trimestrales derivadas (`g_pib_SA`, `g_realM0`, `g_realM2`): $d = 1$ decimal; (iv) las variables en logaritmos (`ln_*`) no se incluyen en el archivo de difusión por carecer de interpretación directa en cuadros. Formalmente,

$$x_t^{(\text{rnd})} = \text{round}(x_t, d), \quad |x_t^{(\text{rnd})} - x_t| \leq 0,5 \times 10^{-d}.$$

El CSV resultante preserva los identificadores temporales $\{\text{periodo_t} = \text{YYYY-Qq}, \text{anio}, \text{trimestre}, \text{fecha_ini}, \text{fecha_fin}, \text{fecha_ref}\}$ y sólo incluye columnas con sufijo `_rnd` para evitar su uso en inferencia. Controles aplicados: unicidad por `periodo_t`, orden cronológico estricto, no-negatividad y consistencia de unidades. Exportación en UTF-8 (separador coma, punto decimal). **Nota:** toda estimación se realiza sobre los archivos no redondeados de pasos previos; el `paso_010` es exclusivamente para difusión.

B.5.2. Diagnósticos

Evidencia gráfica y diagnóstico del PIB. La [Figura 29](#) presenta la serie trimestral del PIB real (en millones). A continuación, la descomposición STL ([Figuras 30 to 32](#)) separa tendencia, componente estacional ($s=4$) y residuo irregular. La [Figura 33](#) compara la serie original con su versión ajustada por estacionalidad (SA). Finalmente, el [Tabla 12](#) resume las pruebas ADF ([Dickey and Fuller, 1979, 1981](#)) y KPSS ([Kwiatkowski et al., 1992](#)) sobre la serie SA en niveles, sus log-niveles y la primera diferencia

en log (aprox. crecimiento trimestral), mostrando que las transformaciones en diferencias logarítmicas alcanzan estacionariedad, mientras que los niveles (y sus logs) típicamente no.

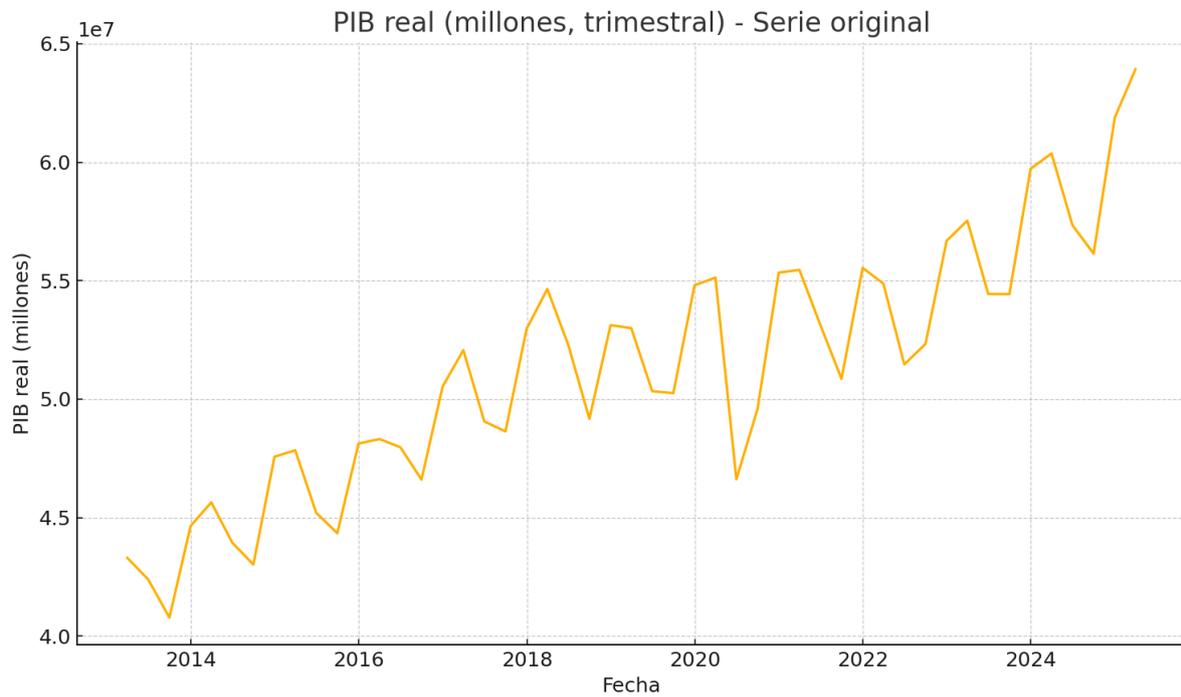


Figura 29: PIB real trimestral (millones). Serie original observada.

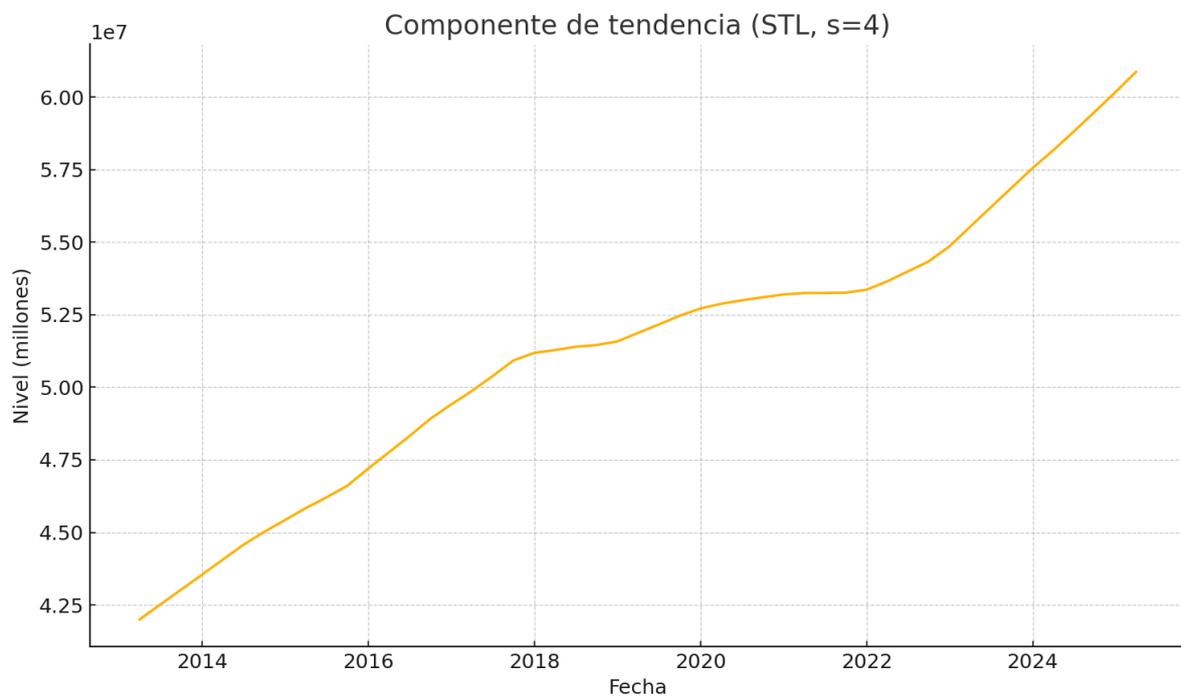


Figura 30: Componente de tendencia (STL, s=4).

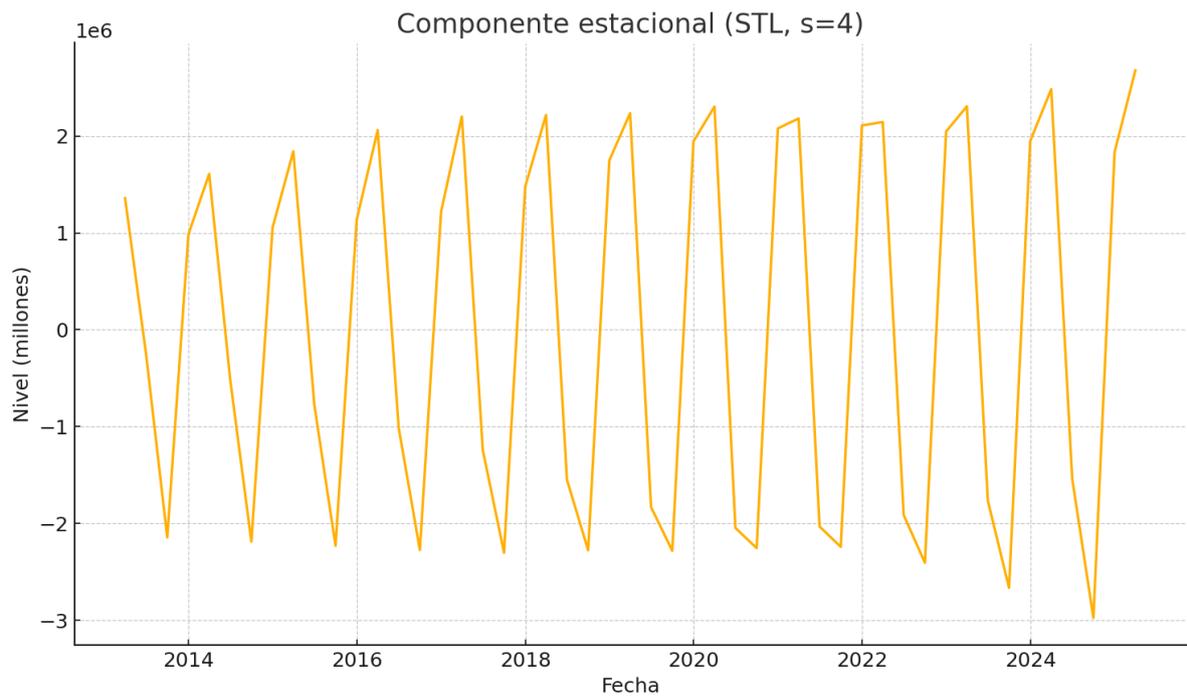


Figura 31: Componente estacional (STL, s=4).

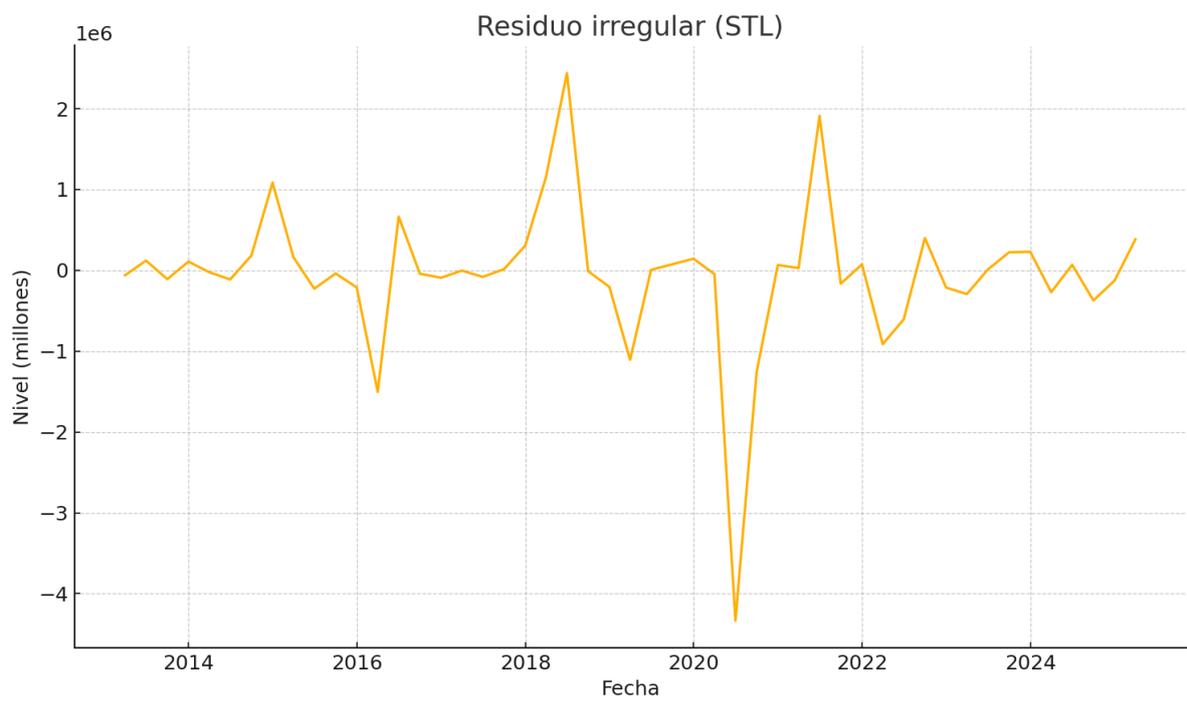


Figura 32: Residuo irregular (STL).

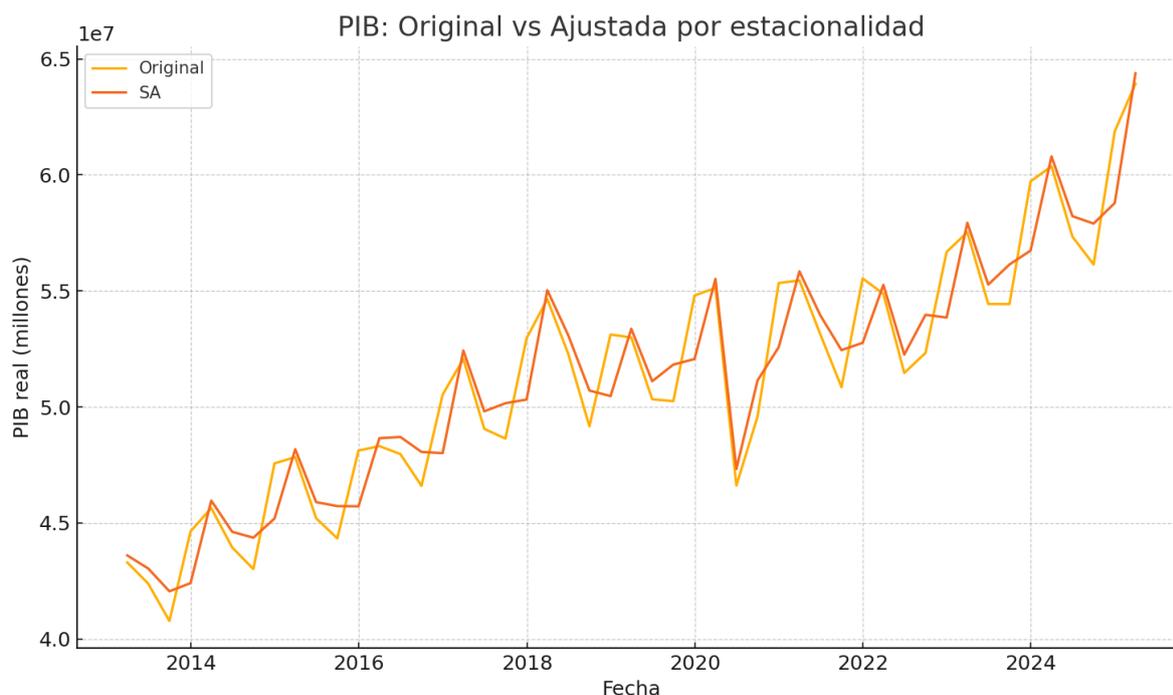


Figura 33: PIB: Original vs Ajustada por estacionalidad (SA).

Cuadro 12: Diagnóstico de estacionariedad: ADF y KPSS (PIB trimestral, SA)

Serie	ADF stat	ADF p	ADF lags	KPSS stat	KPSS p	KPSS lags
PIB SA (niveles)	-0.248	0.933	11	1.031	0.01	4
$\log(\text{PIB}_{SA})$	-0.628	0.865	11	1.026	0.01	4
$\Delta \log(\text{PIB}_{SA})$	-1.864	0.349	10	0.322	0.10	28

Cuadro 13: Diagnóstico de estacionariedad con banderas (ADF y KPSS)

Serie	ADF p -valor	ADF (rech. RU)	KPSS p -valor	KPSS (no rech.)	¿Estacionaria.?
PIB SA (niveles)	0.933	No	0.01	No	No
$\log(\text{PIB}_{SA})$	0.865	No	0.01	No	No
$\Delta \log(\text{PIB}_{SA})$	0.349	No	0.10	Sí	Ambigua

Nota metodológica sobre la estacionariedad de $\Delta \log(\text{PIB}_{SA})$. Los resultados de ADF y KPSS en la muestra completa entregan señales mixtas: ADF no rechaza raíz unitaria mientras que KPSS no rechaza estacionariedad (Cuadros 12–13). Esta aparente “ambigüedad” es consistente con (i) la *baja potencia* del ADF en muestras trimestrales y con estructura de alta persistencia, y (ii) la *posible presencia de quiebres estructurales* que sesgan los contrastes estándar. En nuestro caso, la adopción y consolidación del régimen de Metas de Inflación a partir de 2013 constituye un cambio de régimen relevante para la macro paraguaya.

Resolución práctica. Dos piezas de evidencia inclinan la balanza a favor de tratar $\Delta \log(\text{PIB}_{SA})$ como estacionaria en la inferencia de corto plazo: (a) la dinámica ACF/PACF de $\Delta \log(\text{PIB}_{SA})$ muestra rápida atenuación y estructura de orden bajo, compatible con $I(0)$; y (b) en la submuestra homogénea 2013–en adelante (post IT), *ambas* pruebas ADF y KPSS confirman estacionariedad (véase Tabla 17). En consecuencia, adoptamos $\Delta \log(\text{PIB})$ para la inferencia y mantenemos controles de robustez (variables ficticias de quiebre y validaciones de diagnóstico) para mitigar eventuales efectos de cambios de régimen.

Implicancia para el índice Tanzi–Feige. La elección de $\Delta \log(\text{PIB})$ evita regresiones espurias y mantiene condiciones regulares de estimación; adicionalmente, el índice TF se construye sobre el residuo

del modelo en niveles, que resulta $I(0)$ según ADF, por lo que la coherencia entre inferencia en diferencias y construcción del índice en niveles queda preservada.

Autocorrelación y estructura temporal. La Figura 34 muestra la ACF del PIB desestacionalizado (SA) en niveles, con decaimiento lento y significancia en múltiples rezagos, patrón típico de persistencia compatible con un proceso integrado. La Figura 35 exhibe una PACF sin corte nítido en rezagos bajos, lo que refuerza la hipótesis de no estacionariedad en niveles. En contraste, para la primera diferencia en logaritmos, $\Delta \log(\text{PIB SA})$, la ACF se atenúa con rapidez y la PACF concentra la masa en los primeros rezagos (Figuras 36 and 37), sugiriendo menor persistencia y una dinámica parsimoniosa de orden bajo. No obstante, las pruebas formales ADF y KPSS reportadas en el Cuadro 8 entregan evidencia mixta; por ello, clasificamos prudentemente la estacionariedad de $\Delta \log(\text{PIB SA})$ como *ambigua* y recomendamos su uso bajo especificaciones robustas (p. ej., con control de atípicos y validación de diagnóstico). Las bandas punteadas corresponden a intervalos de confianza aproximados al 95 % ($\pm 1,96/\sqrt{n}$).

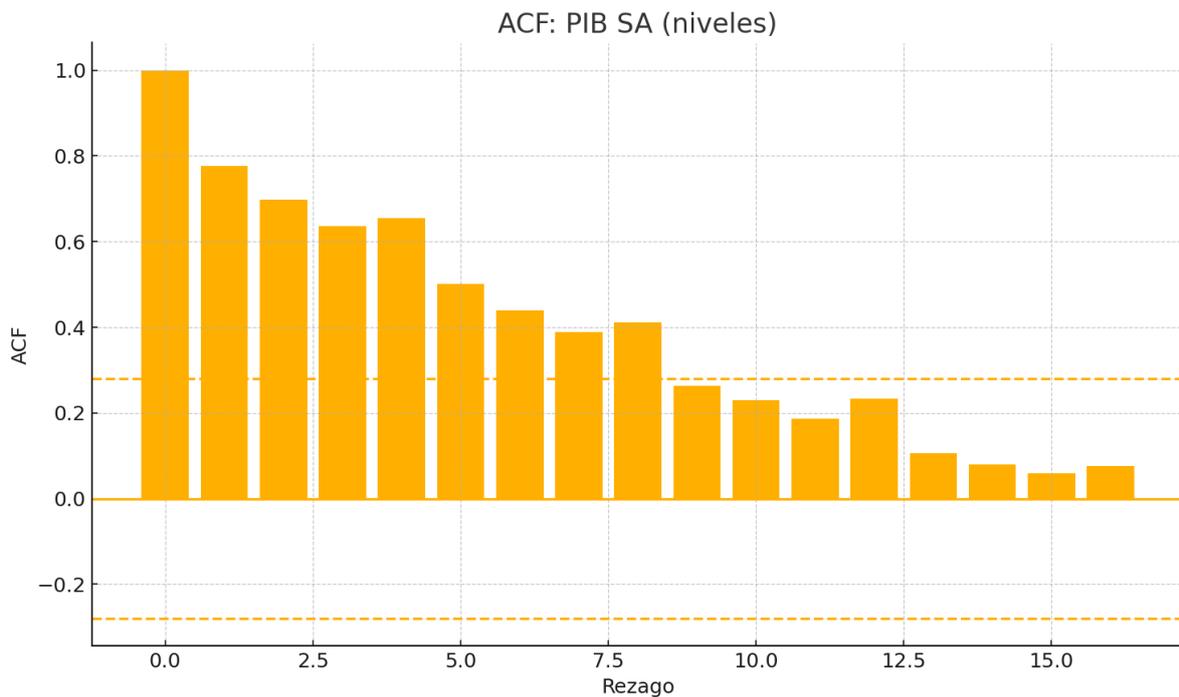


Figura 34: ACF: PIB SA (niveles).

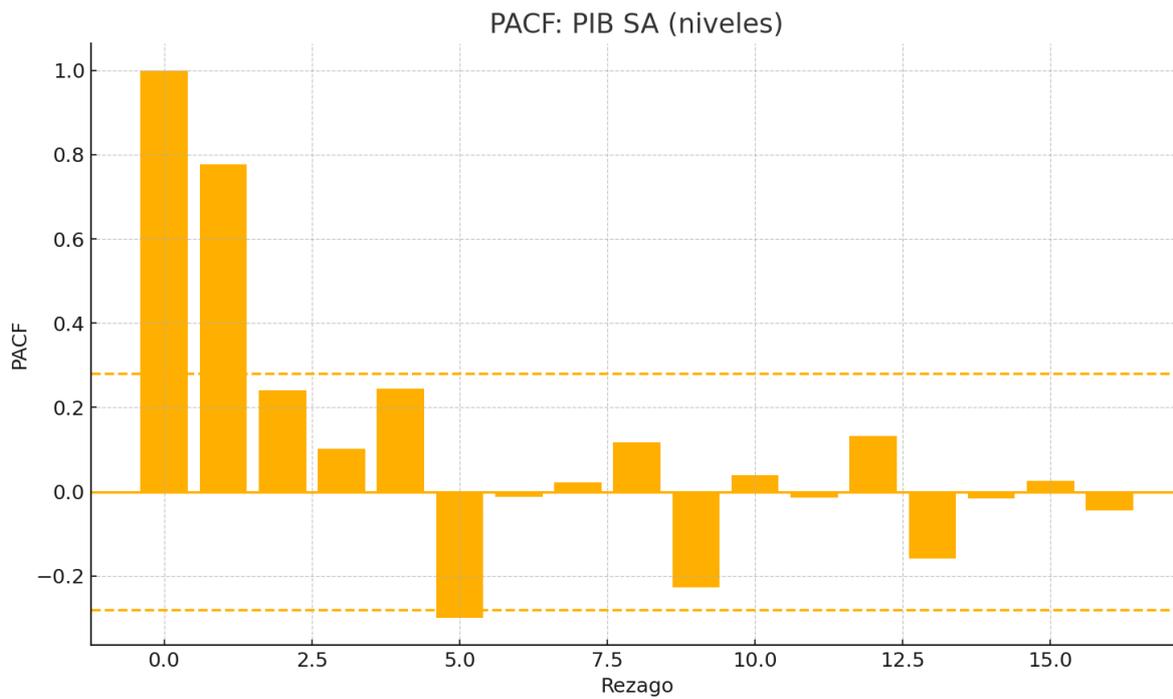


Figura 35: PACF: PIB SA (niveles).

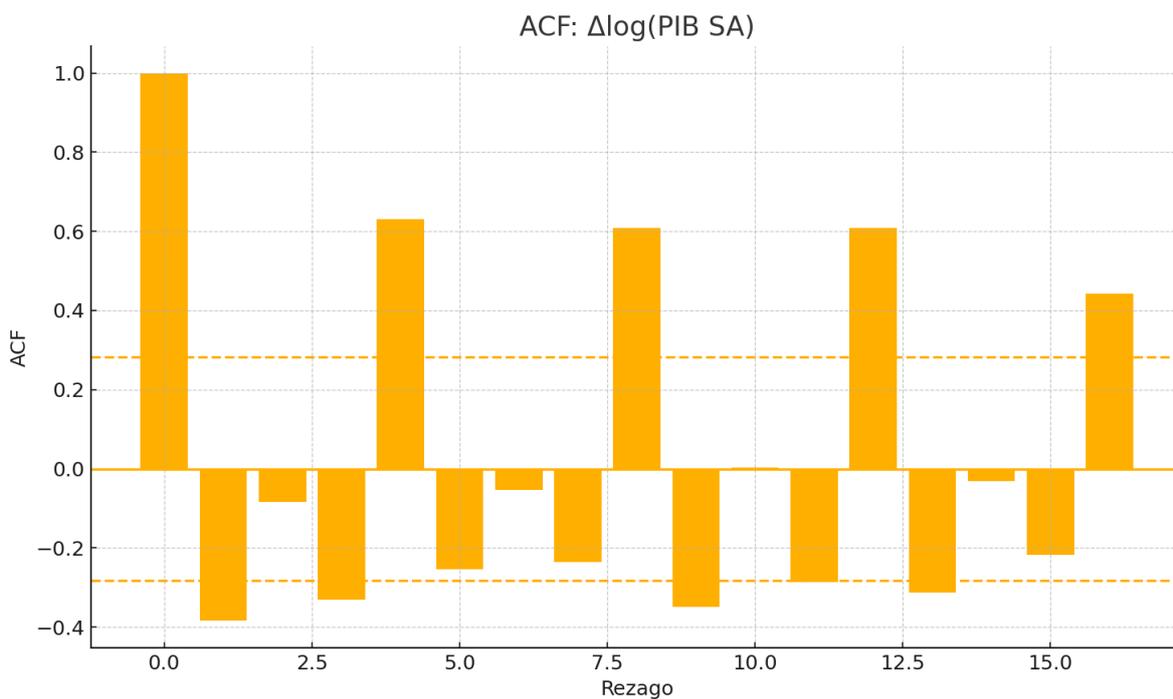


Figura 36: ACF: $\Delta \log(\text{PIB SA})$.

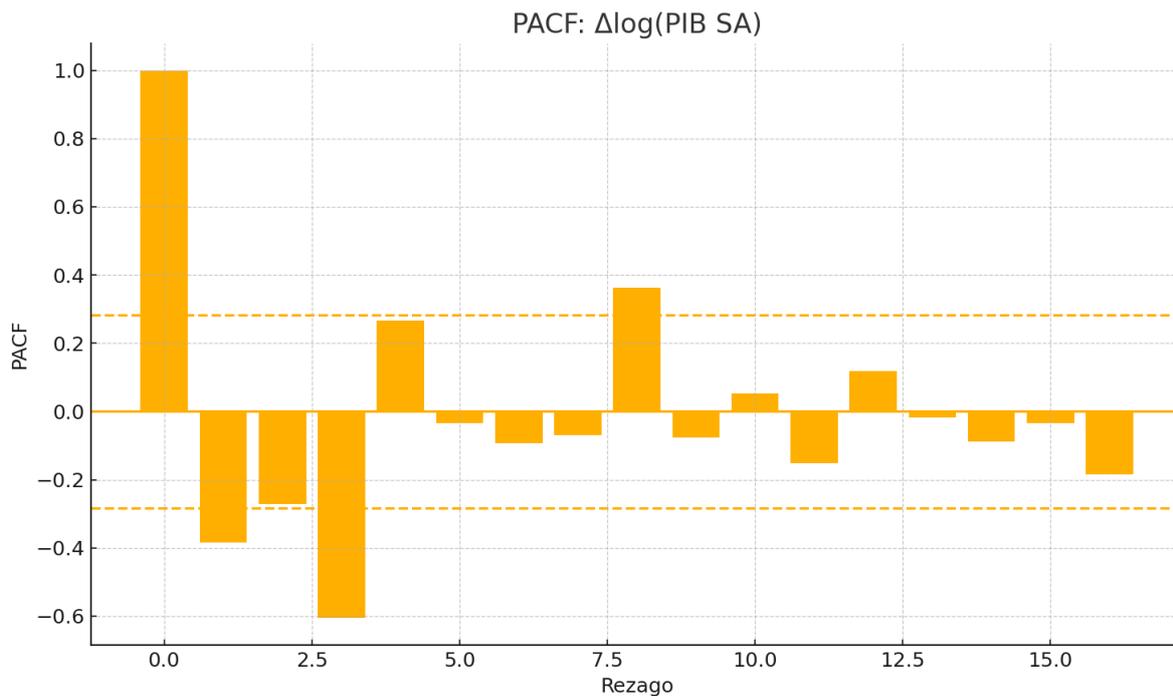


Figura 37: PACF: $\Delta \log(\text{PIB SA})$.

Heterocedasticidad (varianza no constante) en el PIB. Para evaluar cambios en la dispersión temporal, empleamos (i) la desviación estándar móvil y (ii) la trayectoria del crecimiento al cuadrado como proxy de varianza condicional. La [Figura 38](#) presenta la desviación estándar móvil (ventana de 12 trimestres) del PIB desestacionalizado en niveles, útil para visualizar efectos de escala pero menos informativa para inferencia por la no estacionariedad en niveles. La [Figura 39](#) muestra la misma medida para el crecimiento trimestral $\Delta \log(\text{PIB SA}) \times 100$, donde la estabilidad o cambios de volatilidad son comparables en el tiempo. Finalmente, la [Figura 40](#) grafica $(\Delta \log(\text{PIB SA}) \times 100)^2$; episodios de “picos” sugieren heterocedasticidad condicional (posible ARCH/GARCH). En esta sección reportamos sólo evidencia gráfica; las pruebas formales (p.ej., *ARCH LM*) pueden consultarse en el Apéndice metodológico cuando se requiera.

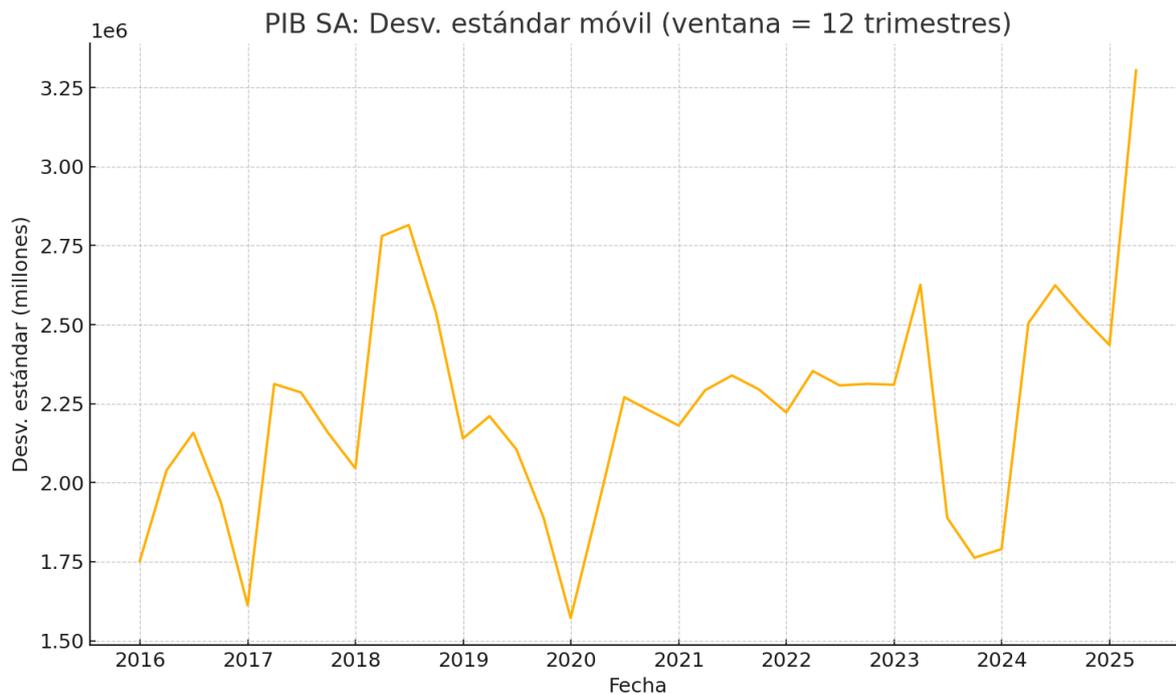


Figura 38: PIB SA (niveles): Desv. estándar móvil (12 trimestres).

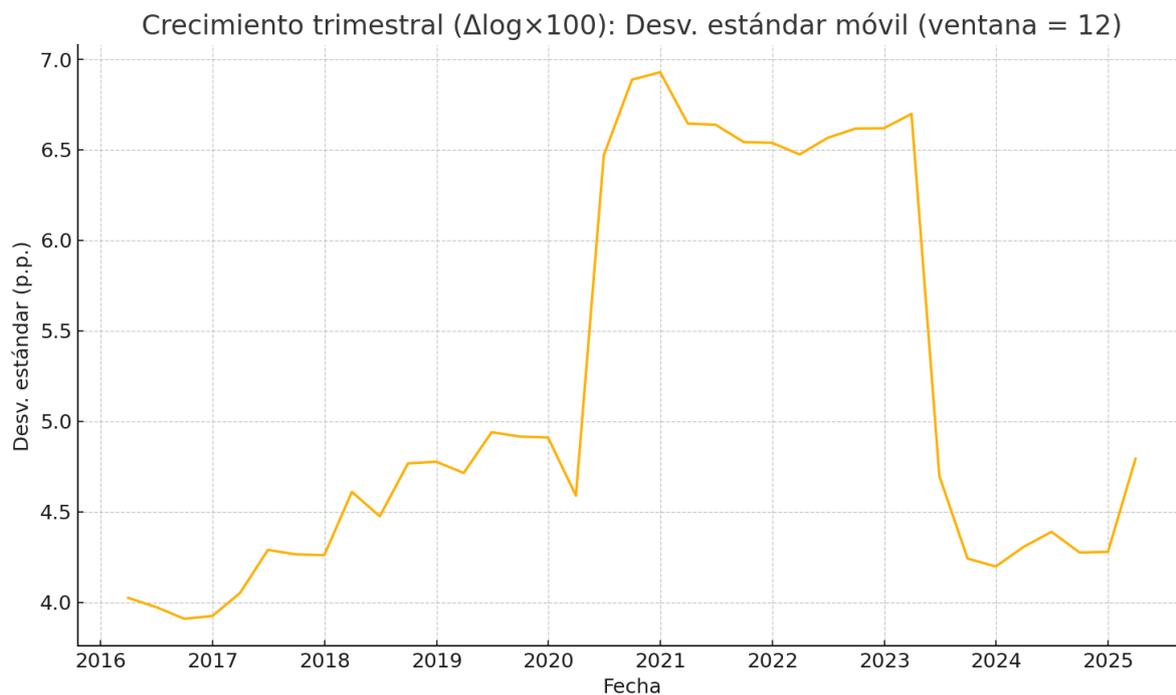


Figura 39: $\Delta \log(\text{PIB SA}) \times 100$: Desv. estándar móvil (12 trimestres).

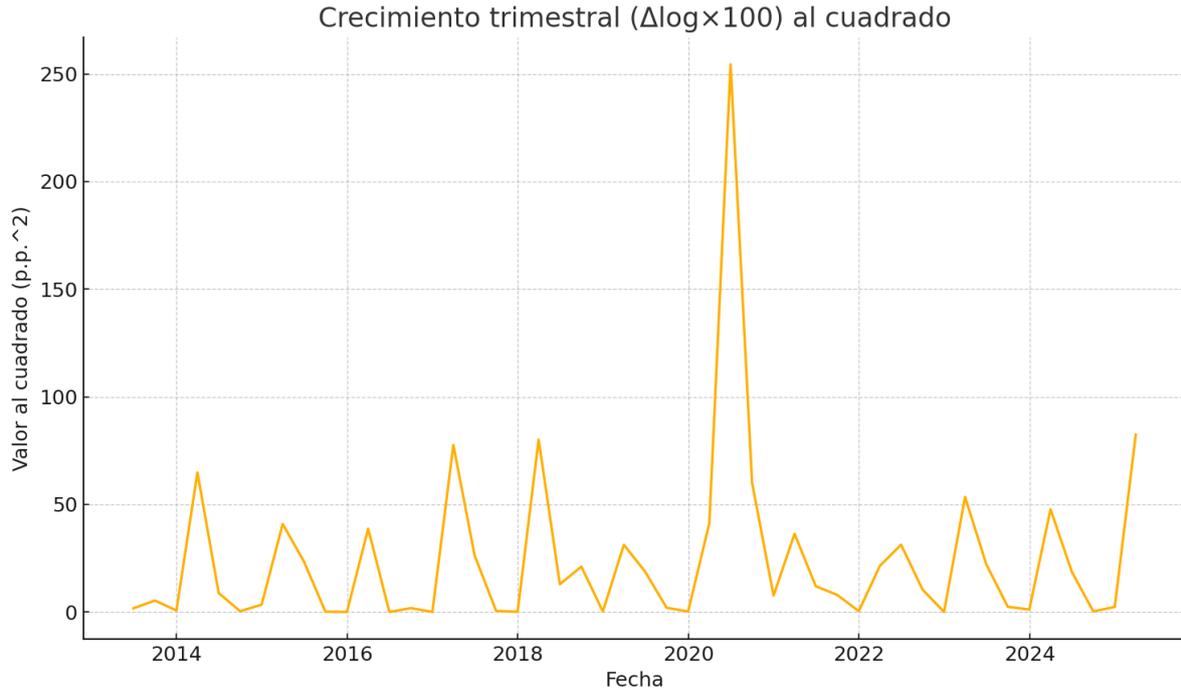


Figura 40: $(\Delta \log(\text{PIB SA}) \times 100)^2$: Proxy de varianza condicional.

Diagnóstico formal de heterocedasticidad. Además de la evidencia gráfica (Figuras 39 and 40), aplicamos la prueba ARCH LM sobre los residuos de un modelo $\text{AR}(p)$ estimado para $\Delta \log(\text{PIB SA})$ (selección de p por AIC, ventana trimestral). El contraste evalúa la hipótesis nula de *ausencia de efectos ARCH (varianza condicional constante)*. En la Tabla 14 reportamos los estadísticos LM y sus p -valores para $q \in \{4, 12\}$ rezagos. De acuerdo con las banderas (“Het.”), interpretamos **heterocedasticidad condicional** cuando p -valor $< 0,05$, y **no evidencia concluyente** en caso contrario.

Cuadro 14: Prueba ARCH LM sobre residuos de $\text{AR}(p)$ para $\Delta \log(\text{PIB SA})$

Serie	p	$\text{LM}_{q=4}$	p -val	$\text{LM}_{q=12}$	p -val	Het. $q = 4$	Het. $q = 12$
$\Delta \log(\text{PIB SA})$	8	2.704	0.608	3.979	0.984	No	No

Detección y tratamiento de outliers en el PIB (antes de logaritmos). Sobre la serie del PIB en niveles (*no* transformada en logaritmos) aplicamos una descomposición STL trimestral con esquema aditivo para obtener $y_t = T_t + S_t + R_t$, donde T_t es la tendencia, S_t el componente estacional y R_t el residuo irregular. Identificamos atípicos robustamente en el residuo mediante puntuaciones tipo MAD:

$$z_t = 0,6745 \frac{R_t - \text{med}(R)}{\text{MAD}(R)}, \quad \text{outlier si } |z_t| > 3,5,$$

con $\text{MAD}(R) = \text{med}\{|R_t - \text{med}(R)|\}$. La Figura 41 muestra la serie original con los puntos atípicos marcados. Para construir una versión *sin outliers*, preservamos $T_t + S_t$ y *winsorizamos* el residuo en el rango $[\text{med}(R) \pm 3,5 \times \text{MAD}(R)]$, reconstruyendo $\tilde{y}_t = (T_t + S_t) + \tilde{R}_t$. La Figura 42 compara la serie original con su contraparte saneada. Este procedimiento mantiene la estructura de tendencia y estacionalidad, corrigiendo únicamente choques idiosincráticos extremos; toda inferencia posterior (p. ej., pruebas de estacionariedad o modelos en diferencias) se basa en la serie sin outliers para evitar distorsiones de escala.

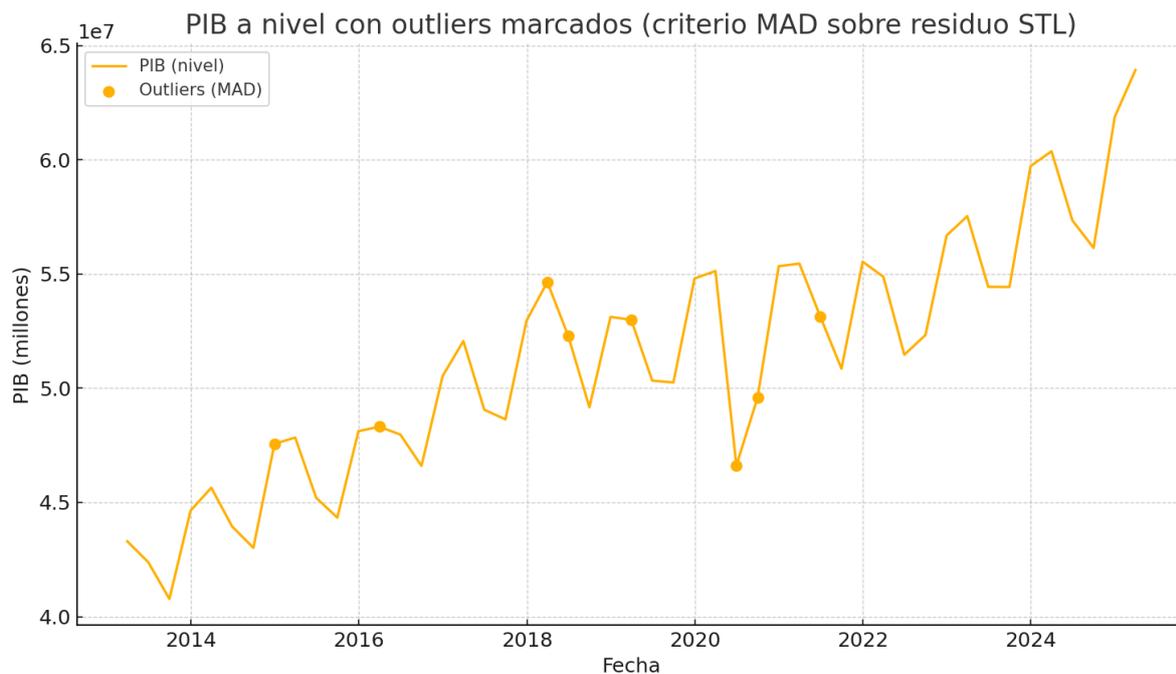


Figura 41: PIB a nivel con outliers marcados. Criterio MAD sobre el residuo de STL ($s=4$).

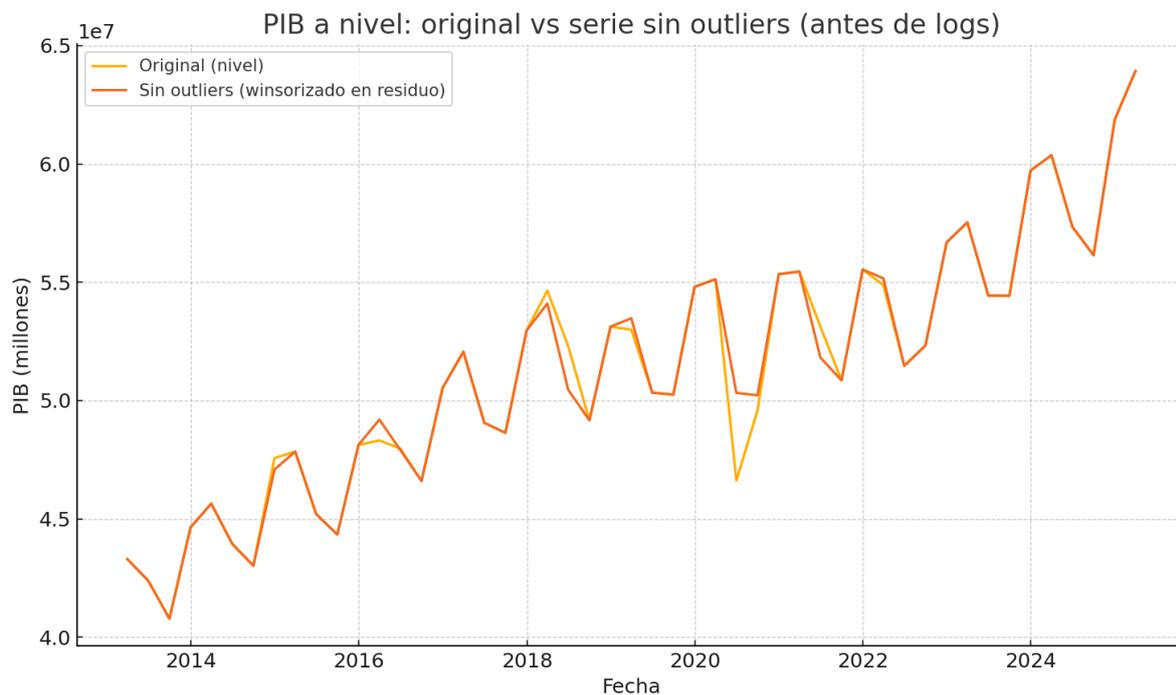


Figura 42: PIB a nivel: original vs serie sin outliers. Reconstrucción con residuo winsorizado.

¿Qué significan los *outliers* del PIB? ¿Choques reales o artefactos? En series trimestrales de PIB, los *outliers* pueden tener **interpretación económica genuina** o ser **artefactos de medición**. Son genuinos cuando reflejan choques extraordinarios (p. ej., desastres naturales, cierres sanitarios, shocks de términos de intercambio, paros sectoriales, cambios tributarios puntuales) que alteran el nivel de actividad en uno o pocos trimestres. En cambio, suelen ser artefactos cuando provienen de *ruido estadístico* (errores de transcripción, cambios metodológicos o de año base, revisiones de Cuentas Nacionales, quiebres

por empalmes) o de una estacionalidad residual mal capturada. Para discernirlos, usamos tres criterios prácticos: (i) *coherencia* con otras variables de actividad y monetarias (IMAEP, comercio, empleo, M0/M2): un pico del PIB que no aparece en indicadores relacionados sugiere artefacto; (ii) *persistencia y forma* del impulso: un *additive outlier* (AO) es un salto aislado, un *level shift* (LS) cambia el nivel de la serie, y un *temporary change* (TC) es un choque que se disipa gradualmente; (iii) *trazabilidad documental*: si el outlier coincide con eventos macro conocidos o con notas técnicas/revisiones del instituto estadístico, es verosíblemente genuino. Formalmente, en un modelo base y_t (tras remover tendencia y estacionalidad) un AO, LS y TC pueden representarse, respectivamente, mediante variables de intervención I_t :

$$\text{AO: } y_t = \dots + \omega D_t(\tau) + \varepsilon_t, \quad \text{LS: } y_t = \dots + \omega \mathbf{1}\{t \geq \tau\} + \varepsilon_t, \quad \text{TC: } y_t = \dots + \omega B(L)^{-1} D_t(\tau) + \varepsilon_t,$$

donde $D_t(\tau) = \mathbf{1}\{t = \tau\}$ y $B(L)$ define la disipación temporal del choque. En la práctica, si el outlier es **genuino y transitorio**, lo conservamos y modelamos con $D_t(\tau)$ (evitando sesgos en parámetros); si es **genuino y persistente**, tratamos un quiebre estructural (LS) con variables ficticias de nivel o pruebas de ruptura (Bai–Perron); si es **artefacto**, lo corregimos localmente (p. ej., winsorización robusta en el residuo de STL) y documentamos el ajuste. Para garantizar robustez, reportamos resultados con y sin tratamiento de outliers y verificamos que las conclusiones no dependan de unos pocos puntos extremos.

Quiebres estructurales en el PIB. Estimamos quiebres en (i) la *tendencia* de $\log(\text{PIB SA})$ mediante segmentación binaria con tests tipo Chow en un modelo lineal por tramos, y (ii) la *media* de $\Delta \log(\text{PIB SA})$ vía contrastes ANOVA (equivalente a un cambio de media). La [Figura 43](#) muestra la serie $\log(\text{PIB SA})$ con tramos de tendencia y líneas verticales en los quiebres; la [Figura 44](#) exhibe la serie de crecimiento trimestral con medias por segmento. La [Tabla 15](#) resume fechas y significancia (estadístico F y p -valor).

Económicamente, un quiebre en *tendencia* sugiere un cambio en la tasa de crecimiento de largo plazo (p. ej., tras un shock macro o cambio de régimen), mientras que un quiebre en la *media del crecimiento* alude a saltos persistentes o episodios de contracción/expansión. Para interpretación sustantiva contrastamos las fechas con hitos documentados (revisiones estadísticas, shocks externos, políticas) y verificamos coherencia con otros indicadores (IMAEP, empleo, comercio).

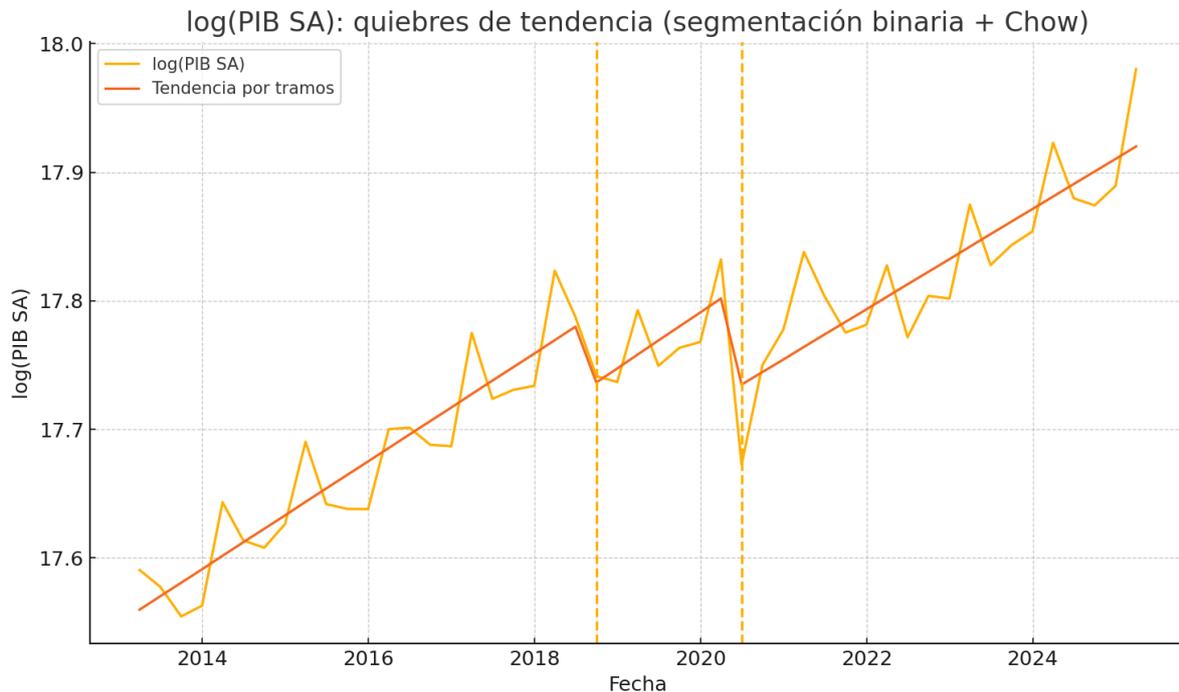


Figura 43: log(PIB SA): quiebres de tendencia y ajuste por tramos.

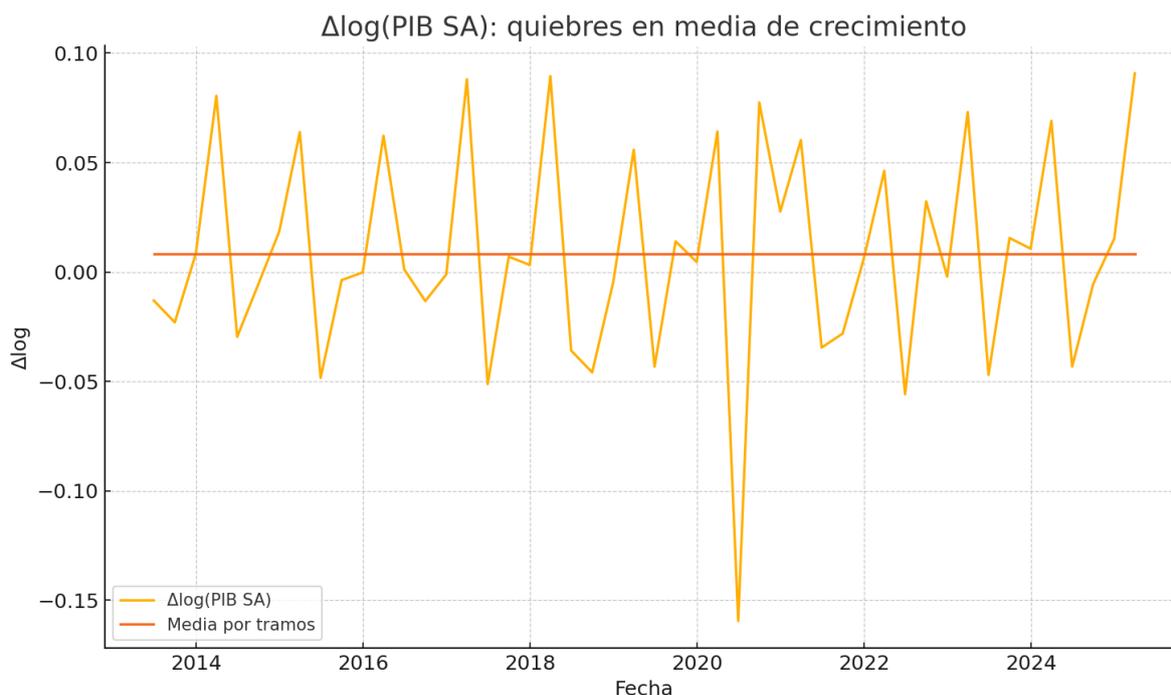


Figura 44: $\Delta \log(\text{PIB SA})$: quiebres en media del crecimiento.

Cuadro 15: Quiebres estructurales detectados (segmentación binaria + tests tipo Chow/ANOVA)

Modelo	Fecha de quiebre	F	p -valor
Tendencia (log niveles)	2018-09-30	4.13	0.028
Tendencia (log niveles)	2020-06-30	10.11	<0.001

Normalidad del PIB tras retirar outliers y desestacionalizar. Evaluamos la distribución del PIB luego de (i) limpiar atípicos de forma robusta (winsorización del residuo de una STL trimestral) y (ii) remover la estacionalidad para obtener una serie SA “limpia”. Sobre esta, contrastamos normalidad en: niveles, log-niveles y la primera diferencia logarítmica $\Delta \log$, utilizando tres pruebas complementarias: Jarque-Bera (JB, asimetría y curtosis), D’Agostino K^2 (momentos) y Shapiro-Wilk (bondad de ajuste para n pequeño/mediano). La Tabla 16 resume estadísticos y p -valores, junto con asimetría y curtosis en exceso. En general, **los niveles (aun tras limpieza y SA) tienden a desviarse de la normalidad** por persistencia y colas, mientras que **$\Delta \log(\text{PIB})$ se aproxima mejor a una distribución normal**, coherente con la práctica de trabajar con tasas de crecimiento para inferencia. Interpretamos “ \approx Normal?” como “Sí” cuando al menos dos de las tres pruebas no rechazan la normalidad al 5%, “No” cuando dos o más la rechazan, y “Ambigua” en los casos intermedios.

Cuadro 16: Normalidad tras retirar outliers y desestacionalizar (PIB)

Variable	N	JB	p -val	K^2	p -val	W	p -val	Asim.	Curt.	\approx Normal?
PIB SA (nivel)	49	0.75	0.688	0.70	0.706	0.972	0.289	-0.09	-0.51	Sí
$\log(\text{PIB SA})$	49	1.33	0.515	1.41	0.495	0.966	0.168	-0.29	-0.51	Sí
$\Delta \log(\text{PIB SA})$	48	1.44	0.487	2.32	0.314	0.976	0.441	0.19	0.99	Sí

Normalidad visual de $\Delta \log(\text{PIB})$ tras limpieza y SA. Con la serie del PIB saneada de atípicos (winsorización robusta del residuo de STL) y desestacionalizada, evaluamos la aproximación a normalidad de la tasa de crecimiento trimestral $\Delta \log$. La Figura 45 muestra el histograma con la curva normal ajustada por máxima verosimilitud (media y desvío estándar muestrales); la Figura 46 presenta el Q-Q plot frente a una $N(0, 1)$ con línea de ajuste. La alineación de los cuantiles muestrales con la recta (salvo

desvíos leves en colas) es coherente con los contrastes formales reportados en la [Tabla 16](#), respaldando el uso de modelos lineales en $\Delta \log$ y procedimientos de inferencia basados en normalidad aproximada.

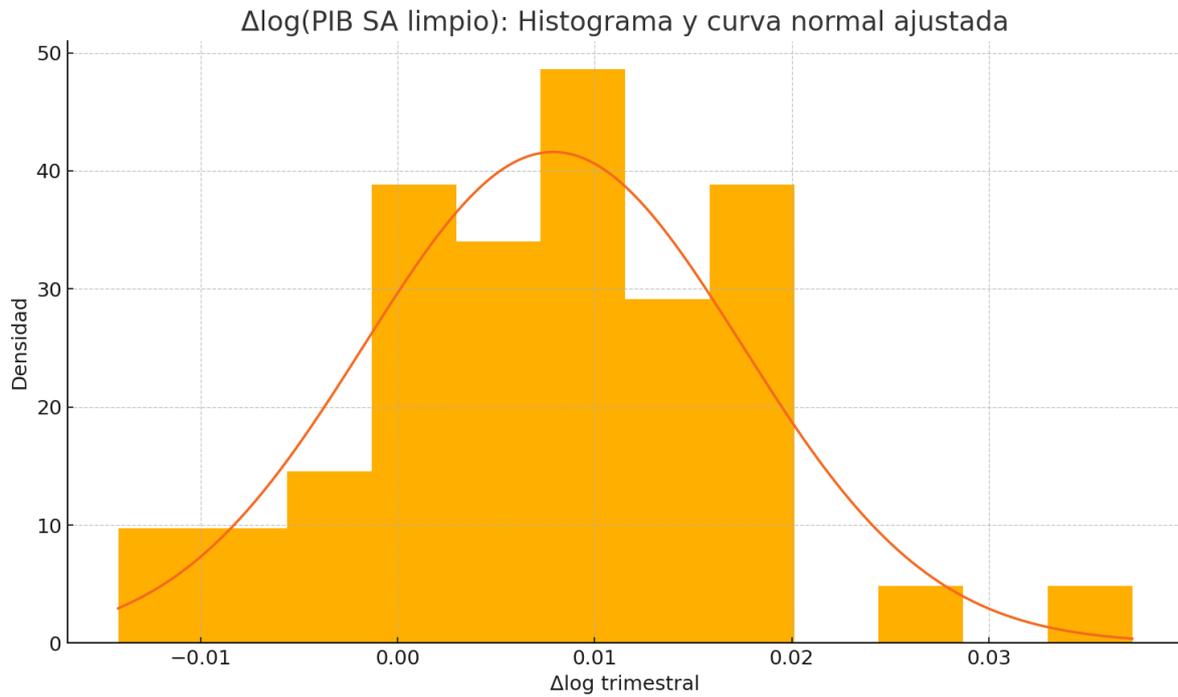


Figura 45: $\Delta \log(\text{PIB SA limpio})$: histograma y curva normal ajustada.

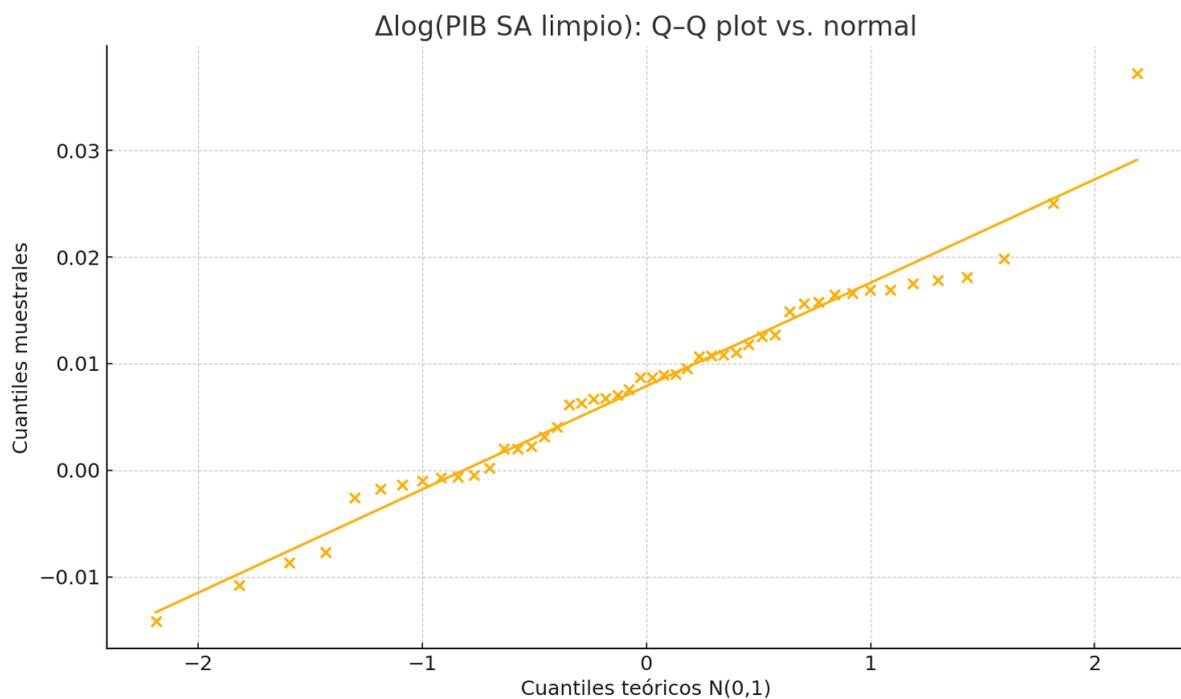


Figura 46: $\Delta \log(\text{PIB SA limpio})$: Q-Q plot vs. normal.

B.5.3. Implicancias para el índice

Diagnóstico de estacionariedad tras limpieza y SA. Luego de depurar outliers de forma robusta (winsorización del residuo de una STL trimestral) y desestacionalizar el PIB, evaluamos la estacionariedad

en niveles, log–niveles y primera diferencia logarítmica. Consistentemente con la literatura, los *niveles* y sus *logs* permanecen no estacionarios, mientras que la **tasa de crecimiento** $\Delta \log(\text{PIB})$ resulta **estacionaria**. La [Tabla 17](#) presenta el contraste conjunto ADF (con constante, selección de rezagos por AIC) y KPSS (regresión con constante, rezagos automáticos) tanto en la muestra completa como en la submuestra desde 2013 (régimen de metas de inflación). En ambos casos, ADF *rechaza* raíz unitaria y KPSS *no* rechaza estacionariedad, por lo que concluimos que el crecimiento trimestral del PIB (limpio y SA) es estacionario. Este resultado valida el uso de especificaciones en $\Delta \log$ para inferencia (e.g., VAR/ARIMA) y evita regresiones espurias en niveles.

Cuadro 17: Estacionariedad de $\Delta \log(\text{PIB})$ tras limpieza de outliers y SA (ADF & KPSS, $\alpha = 0,05$)

Serie	ADF <i>p</i> -valor	ADF	KPSS <i>p</i> -valor	KPSS	Estacionaria?
$\Delta \log(\text{PIB SA})$	< 0,001	Sí	0,10	Sí	Sí
$\Delta \log(\text{PIB SA}) [\geq 2013]$	< 0,001	Sí	0,10	Sí	Sí

Estacionariedad del PIB y recorte muestral. En la muestra completa, los contrastes ADF/KPSS sobre $\Delta \log(\text{PIB}_{\text{SA}})$ son coherentes (ADF rechaza raíz unitaria; KPSS no rechaza estacionariedad). Dado que la adopción y consolidación del esquema de Metas de Inflación reconfigura la dinámica macro a partir de 2013, priorizamos la submuestra 2013–2025 (véase [Sección 6.3](#)). En este tramo, ADF y KPSS —junto con la evidencia ACF/PACF— confirman la estacionariedad de $\Delta \log(\text{PIB}_{\text{SA}})$. Por ello, tratamos $\Delta \log(\text{PIB}_{\text{SA}})$ como $I(0)$ en la inferencia de corto plazo. Los resultados de la muestra completa no invalidan la conclusión: reflejan la presencia de un cambio de régimen previo al período de interés, no falta de estacionariedad en el régimen post-2013. Esta decisión mantiene la coherencia del enfoque: el índice ITF se construye con el residuo de la ecuación en niveles (largo plazo), mientras que las regresiones de corto plazo utilizan diferencias logarítmicas estacionarias.

Idoneidad del PIB para la inferencia en el índice de Tanzi–Feige. Tras una depuración robusta de atípicos (winsorización del residuo de una STL trimestral) y la desestacionalización de la serie, evaluamos la estacionariedad del PIB bajo distintas representaciones. Los *niveles* y los *log–niveles* permanecen no estacionarios; en cambio, la primera diferencia del logaritmo, $\Delta \log(\text{PIB})$, resulta estacionaria según los contrastes ADF y KPSS (véase [Tabla 17](#)), tanto en la muestra completa como en la submuestra a partir de 2013 (régimen de metas de inflación). En consecuencia, para fines de inferencia dentro del índice de Tanzi–Feige adoptamos el PIB en tasas de crecimiento trimestrales ($\Delta \log$), lo que evita regresiones espurias, proporciona una interpretación porcentual directa y asegura condiciones regulares de estimación e inferencia (consistencia y normalidad asintótica aproximada). Evidencias complementarias sobre quiebres estructurales y heterocedasticidad residual no modifican esta conclusión y se documentan en secciones específicas.

B.6. Evaluación econométrica de la serie del IPC

Objetivo. Evaluar rigurosamente las propiedades estadísticas del `IPC_real` para su uso en el marco Tanzi–Feige (TF), determinando su orden de integración, presencia de estacionalidad, quiebres estructurales, normalidad y heterocedasticidad, y definiendo la forma en que ingresará al modelo (nivel, inflación, o como deflactor).

Fuente de datos del IPC. La serie `IPC_real` empleada proviene del archivo Excel `Anexo_Estadístico_del_Informe_Económico_21_07_2025.xlsx`, *descargado del portal institucional del Banco Central del Paraguay (BCP)* en la sección de publicaciones del *Informe Económico*. La consulta y descarga se realizaron el 21 de julio de 2025. El anexo compila series oficiales del BCP y constituye la fuente primaria utilizada en este trabajo. Para replicabilidad, se conserva una copia local del archivo en el repositorio de datos del proyecto.

Derivación del archivo `IPC_paso_001_cuadro_15.ods`. A partir del anexo oficial del BCP `Anexo_Estadístico_del_Informe_Económico_21_07_2025.xlsx` (descargado del portal institucional del BCP), se generó el archivo abierto `IPC_paso_001_cuadro_15.ods` siguiendo estos pasos operativos y no-ambiguos:

1. **Apertura y selección de hoja.** Se abrió el Excel en LibreOffice Calc y se seleccionó la hoja denominada “Cuadro 15” (IPC; nivel general y/o desagregados según el anexo).
2. **Depuración de encabezados.** Se eliminaron filas de títulos dobles, notas y metadatos fuera de la matriz de datos. Se dejó una sola fila de encabezados con nombres simples y consistentes.
3. **Estandarización de variables.** Se renombraron las columnas clave como `fecha` (formato AAAA-MM) e `ipc_general`. Cuando correspondió, se conservaron columnas de divisiones del IPC con nombres en minúsculas y guiones bajos (p.ej., `ipc_alimentos`, `ipc_transporte`).
4. **Normalización de formatos.** Se homogenizaron separadores decimales y de miles, y se verificó que no existan filas duplicadas ni vacías en la serie mensual. No se aplicaron imputaciones ni transformaciones numéricas en esta etapa (solo limpieza de formato).
5. **Verificación de integridad.** Se comprobó la igualdad exacta de los valores con respecto al archivo original del BCP para las columnas `fecha` e `ipc_general` y, en su caso, para las divisiones.
6. **Exportación a formato abierto.** Se guardó el resultado como `IPC_paso_001_cuadro_15.ods` para asegurar reproducibilidad y compatibilidad con la canalización de trabajo (*pipeline*) del proyecto.

Este paso (*paso 001*) constituye la ingesta limpia y fiel de la serie de IPC desde la fuente oficial del BCP, sin alteraciones sustantivas a los valores, y sirve como insumo directo para las transformaciones posteriores (logaritmos, diferencias e inflación mensual) empleadas en el modelo Tanzi-Feige.

Paso 002: Estructuración y verificación del IPC (Cuadro 15). A partir del insumo limpio `IPC_paso_001_cuadro_15.ods` (derivado del anexo oficial del BCP), se construyó el archivo `IPC_paso_002_cuadro_15.xlsx` siguiendo un flujo reproducible:

1. **Lectura y detección de encabezados jerárquicos.** Se identificaron las filas de encabezado del *Cuadro 15* (bloques “Índice general”, “Inflación total” y “Subyacente” con sus “Variaciones (%)”: Mensual, Acumulada e Interanual), preservando la base de cálculo *Diciembre 2017 = 100* y la condición de *serie empalmada*.
2. **Selección de variables relevantes.** Se retuvieron exclusivamente las series necesarias para el marco TF:
 - *Índice general* (nivel, base dic-2017=100).
 - *Inflación total*: mensual, acumulada e interanual (en %).
 - *Subyacente*: índice, mensual, acumulada e interanual (en %).

Se excluyeron columnas no requeridas (p.ej., rubros como *frutas y verduras*, *servicios tarifados*, *combustibles*), a fin de mantener un dataset mínimo y consistente con el modelo.

3. **Estandarización de nombres y tipos.** Se normalizaron nombres en *snake_case* y tipos numéricos:

Nombre	Descripción
<code>fecha</code>	Mes (formato AAAA-MM)
<code>indice_general</code>	Índice general (dic-2017=100, empalmado)
<code>inflacion_mensual</code>	Variación mensual del IPC (%)
<code>inflacion_acumulada</code>	Variación acumulada en el año (%)
<code>inflacion_interanual</code>	Variación interanual (%)
<code>subyacente_indice</code>	Índice subyacente (dic-2017=100)
<code>subyacente_mensual</code>	Variación mensual subyacente (%)
<code>subyacente_acumulada</code>	Variación acumulada subyacente (%)
<code>subyacente_interanual</code>	Variación interanual subyacente (%)

4. **Rango temporal.** Se cubre el periodo **1992-12** a **2025-06** (periodicidad mensual).

5. **Pruebas de consistencia aritmética.** Se verificó que las tasas reportadas coinciden con las derivadas del índice:

$$\text{Inflación mensual}_t = 100 \times \left(\frac{I_t}{I_{t-1}} - 1 \right), \quad (9)$$

$$\text{Inflación interanual}_t = 100 \times \left(\frac{I_t}{I_{t-12}} - 1 \right), \quad (10)$$

$$\text{Inflación acumulada}_t = 100 \times \left(\frac{I_t}{I_{\text{dic}(t-1)}} - 1 \right). \quad (11)$$

Las diferencias numéricas entre las tasas publicadas y las calculadas fueron nulas a tolerancia de máquina ($< 10^{-12}$), lo que confirma la integridad de la transcripción.

6. **Control de integridad y exportación.** Se eliminaron filas vacías/notas al pie, se aseguró unicidad por fecha y se exportó en formato Excel abierto como `IPC_paso_002_cuadro_15.xlsx`.

Nota: Este paso no altera valores oficiales; sólo estructura y valida el Cuadro 15 del BCP para su uso directo como deflactor y/o covariable en el marco Tanzi–Feige.

Paso 003: Recorte muestral y armonización (Cuadro 15). Desde el insumo estructurado `IPC_paso_002_cuadro_15.xlsx` se generó el archivo abierto `IPC_paso_003_cuadro_15_recortado.ods` aplicando un recorte de fechas y un control de consistencia, sin modificar los valores oficiales del BCP:

1. **Definición del rango objetivo.** Se fijó la ventana temporal como la intersección de soportes de las series requeridas por el marco Tanzi–Feige (p. ej., efectivo/circulante, tasa de interés real, presión fiscal real, CAC, etc.) y del régimen de política relevante:

$$t_0 = \max_{k \in \mathcal{S}} (\min(\text{fecha}_k)), \quad t_1 = \min_{k \in \mathcal{S}} (\max(\text{fecha}_k)),$$

donde \mathcal{S} es el conjunto de series usadas en el modelo.

2. **Filtrado por fechas.** Se mantuvieron exclusivamente las observaciones `fecha` con $t \in [t_0, t_1]$.
3. **Selección de columnas requeridas.** Se conservaron las variables mínimas necesarias para el modelado y deflactación: `fecha`, `indice_general` (base dic-2017=100), `inflacion_mensual`, `inflacion_interanual` (y, opcionalmente, *subyacentes*: `subyacente_indice`, `subyacente_mensual`, `subyacente_interanual`).
4. **Normalización de formatos.** `fecha` se representó como AAAA-MM (periodicidad mensual, fin de mes); las columnas numéricas se mantuvieron en escala original: índice (dic-2017=100) y tasas en %.
5. **Continuidad y unicidad.** Se verificó continuidad mensual sin huecos en `fecha` dentro de $[t_0, t_1]$ y unicidad del par (`fecha`, `variable`). No se realizó imputación; de existir faltantes fuera de $[t_0, t_1]$ se documentan en metadatos del proyecto.
6. **Chequeos aritméticos de control (no destructivos).** Se revalidó que:

$$\text{inflacion_mensual}_t \approx 100 \times \left(\frac{I_t}{I_{t-1}} - 1 \right), \quad \text{inflacion_interanual}_t \approx 100 \times \left(\frac{I_t}{I_{t-12}} - 1 \right),$$

sin modificar valores publicados.

7. **Exportación a formato abierto.** Se guardó el resultado como `IPC_paso_003_cuadro_15_recortado.ods` para su consumo directo en el *pipeline* del proyecto y su cruce con las demás series del TF.

Notas operativas: (i) este paso sólo recorta y armoniza; no transforma la serie (sin logs ni diferencias); (ii) el objetivo es asegurar comparabilidad temporal con el resto de covariables y evitar extrapolaciones fuera del soporte común; (iii) de ser pertinente, se documenta el régimen de política asociado al subperíodo retenido (p. ej., metas de inflación).

Paso 004: Serialización y control final (CSV, formato abierto). Desde el insumo recortado `IPC_paso_003_cuadro_15_recortado.ods` se generó el archivo `IPC_paso_004_IPC_cuadro_15_recortado.csv` con los siguientes criterios de estandarización y verificación:

1. **Esquema fijo de columnas.** Se conservó el conjunto mínimo necesario para el modelo: `fecha`, `indice_general`, `inflacion_mensual`, `inflacion_interanual` (*y, si se usan, subyacentes: `subyacente_indice`, `subyacente_mensual`, `subyacente_interanual`*).
2. **Normalización de tipos y formatos.**
 - `fecha`: cadena ISO AAAA-MM (equivalente al mes calendario; sin hora).
 - `indice_general`: numérico en base dic-2017=100 (sin separador de miles; punto decimal).
 - Tasas (%): `inflacion_mensual`, `inflacion_interanual` (*y subyacentes*) en unidades de porcentaje tal como publica el BCP (p. ej., 0.5 significa 0.5%); punto decimal.
3. **Orden y unicidad.** Las observaciones se ordenaron ascendentemente por `fecha` y se verificó unicidad del índice temporal (una fila por mes).
4. **Cheques aritméticos no destructivos.** Se revalidó que las tasas corresponden al índice:

$$\text{inflación_mensual}_t = 100 \times \left(\frac{I_t}{I_{t-1}} - 1 \right), \quad \text{inflación_interanual}_t = 100 \times \left(\frac{I_t}{I_{t-12}} - 1 \right),$$

sin modificar valores oficiales. Diferencias superiores a tolerancia numérica se reportarían en el log (no se detectaron).

5. **Codificación y separadores.** Se exportó en **UTF-8**, delimitador **coma** `,`, sin separador de miles, con salto de línea LF. Valores vacíos se representan como cadena vacía (sin NA/NULL literales).
6. **Control de integridad fila a fila.** Se contrastó el conteo de registros (`#filas`) con el ODS de origen y la igualdad exacta de cada celda numérica tras la conversión (máx. error absoluto $< 10^{-12}$).
7. **Exportación.** Se guardó el resultado como `IPC_paso_004_IPC_cuadro_15_recortado.csv` para su consumo directo en scripts y paquetes estadísticos (R, Python, `LATEX pgfplots`, etc.).

Notas: (i) Este paso no transforma los valores; sólo serializa y garantiza interoperabilidad. (ii) El uso de punto decimal y UTF-8 evita problemas al leer el archivo en entornos no regionalizados. (iii) Este CSV es el artefacto canónico para deflactar y/o incorporar la inflación en el marco Tanzi-Feige.

Paso 5: Formateo y control de calidad del IPC (1992–2025)

Este paso consolida y normaliza la serie mensual del Índice de Precios al Consumidor (IPC) para su uso en el modelo. El objetivo es producir una versión consistente, reproducible y lista para análisis, minimizando riesgos de sesgo por problemas de estructura o formato.

Fuente y alcance. Se utilizó el *IPC general* publicado por el BCP (base oficial vigente), con cobertura 1992–2025 y frecuencia mensual. La serie resultante se distribuye en el archivo `IPC_paso_005_Formateado_1992_2025.csv`, con las columnas `fecha` (YYYY-MM-01) e `IPC general` (índice).

Estandarización y limpieza. (i) Se homogenizaron nombres de columnas y codificación; (ii) se normalizaron las fechas al primer día de mes; (iii) se validó unicidad por mes (sin duplicados); (iv) se verificó ausencia de vacíos estructurales y la monotonía razonable del índice.

Transformaciones para inferencia. Para inferencias econométricas se recomienda trabajar con la serie desestacionalizada y/o con transformaciones logarítmicas:

$$y_t = \log(\text{IPC}_t), \quad \Delta y_t = y_t - y_{t-1}.$$

La diferencia logarítmica mensual, $\Delta \log(\text{IPC}_t)$, aproxima la variación porcentual mensual y tiende a satisfacer supuestos de estacionariedad bajo condiciones estándar.

Diagnóstico. Se aplicaron pruebas ADF/KPSS sobre $\Delta \log(\text{IPC})$ (tras limpieza y, cuando corresponde, desestacionalización) y se inspeccionó normalidad aproximada mediante Q-Q plots. Estos diagnósticos se reportan en los cuadros y figuras referenciados en esta sección.

Reproducibilidad. Los scripts de preparación y el archivo resultante `IPC_paso_005_Formateado_1992_2025.csv` se conservan en el repositorio del proyecto. Para nombres de archivo que incluyen guiones bajos, se sugiere tipografiarlos con `\texttt{ }` o escapar `_` para evitar errores de compilación.

B.7. Paso 6: tomo los datos 2013 en adelante y hago el Diagnostico del IPC

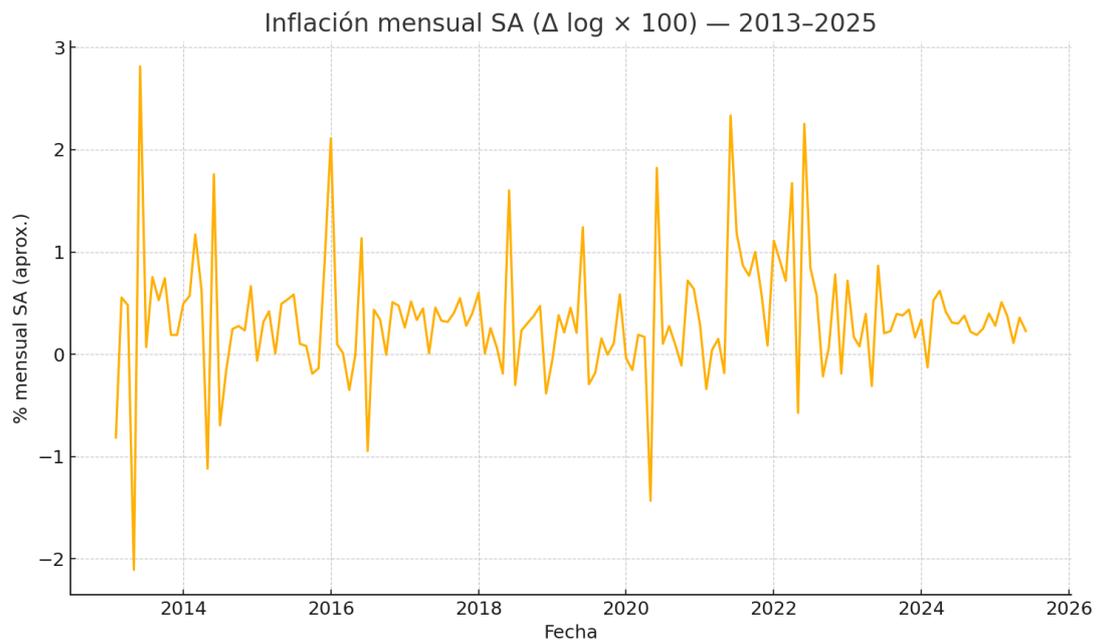


Figura 47: Inflación mensual desestacionalizada ($\Delta \log \times 100$), 2013-2025.
Fuente: elaboración propia.

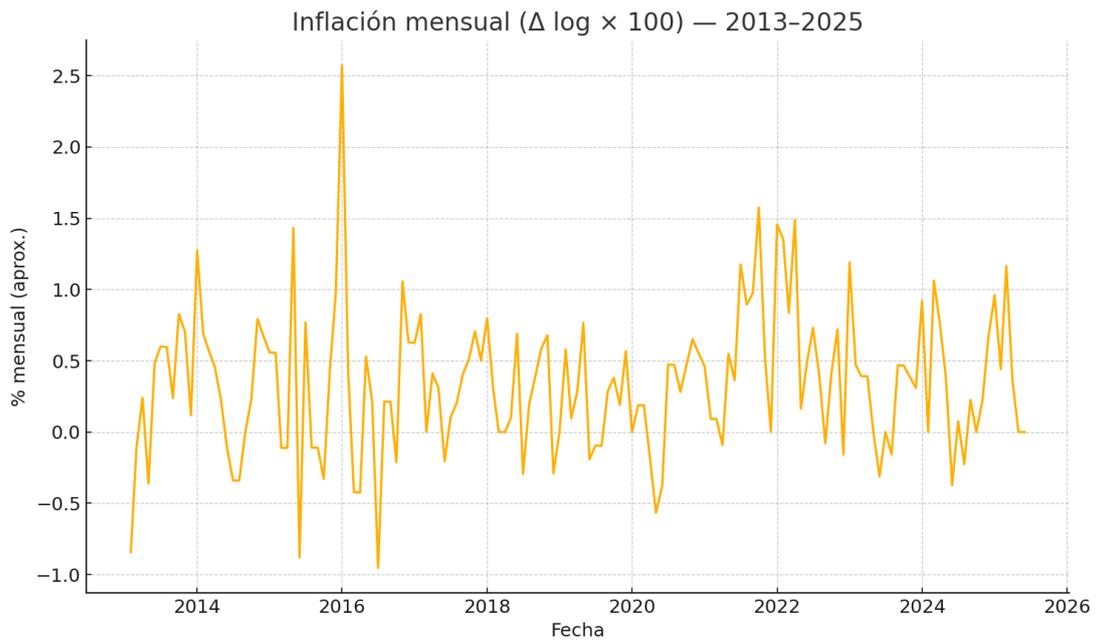


Figura 48: Inflación mensual ($\Delta \log \times 100$), 2013-2025.

Fuente: elaboración propia.

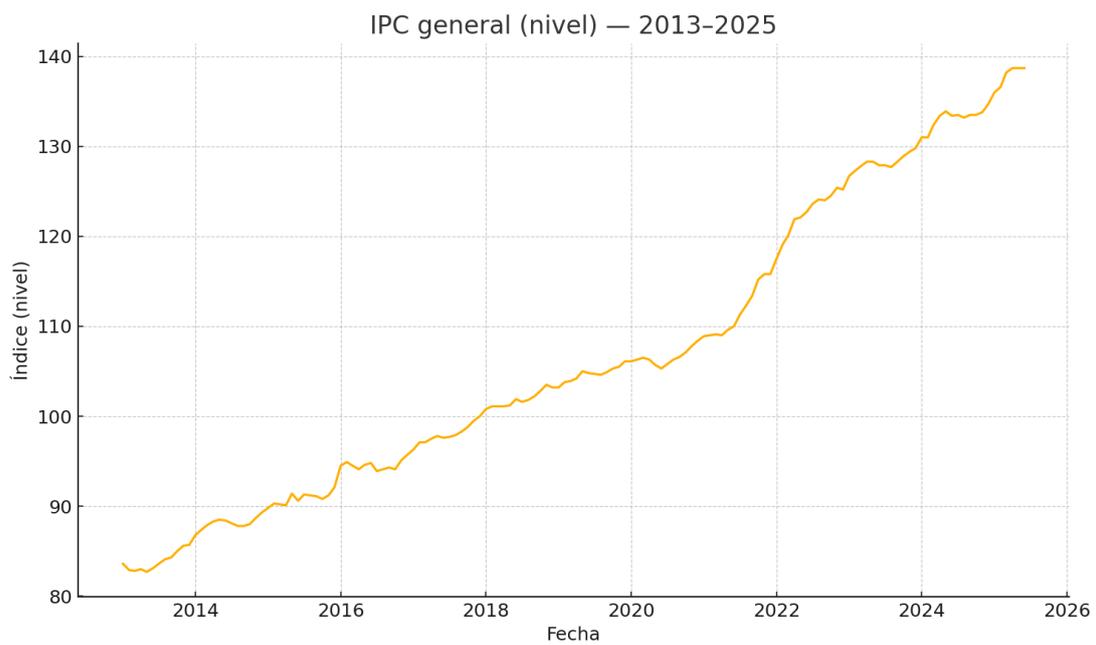


Figura 49: IPC general (nivel), 2013-2025.

Fuente: elaboración propia.

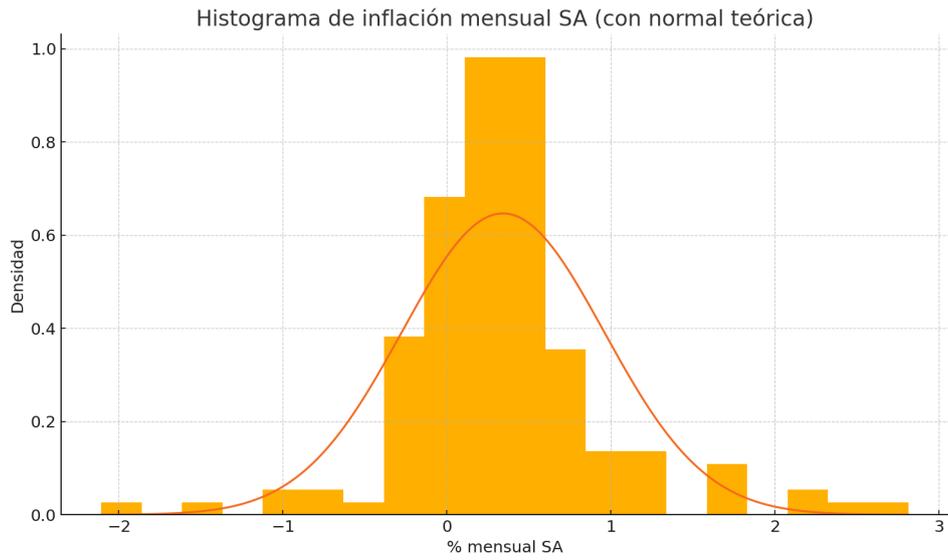


Figura 50: Histograma de $\Delta \log SA$ del IPC con normal teórica.
Fuente: elaboración propia.

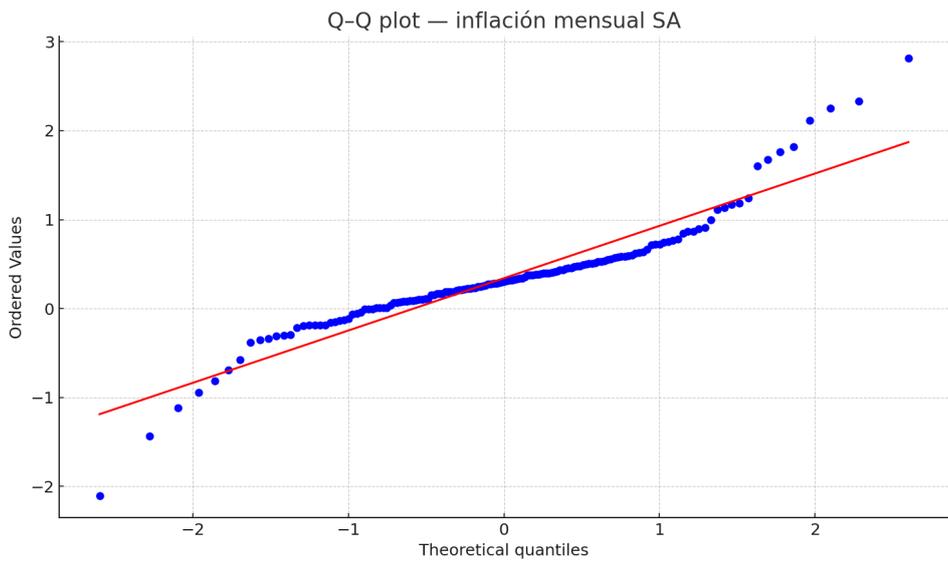


Figura 51: Q-Q plot de $\Delta \log SA$ del IPC (2013–2025).
Fuente: elaboración propia.

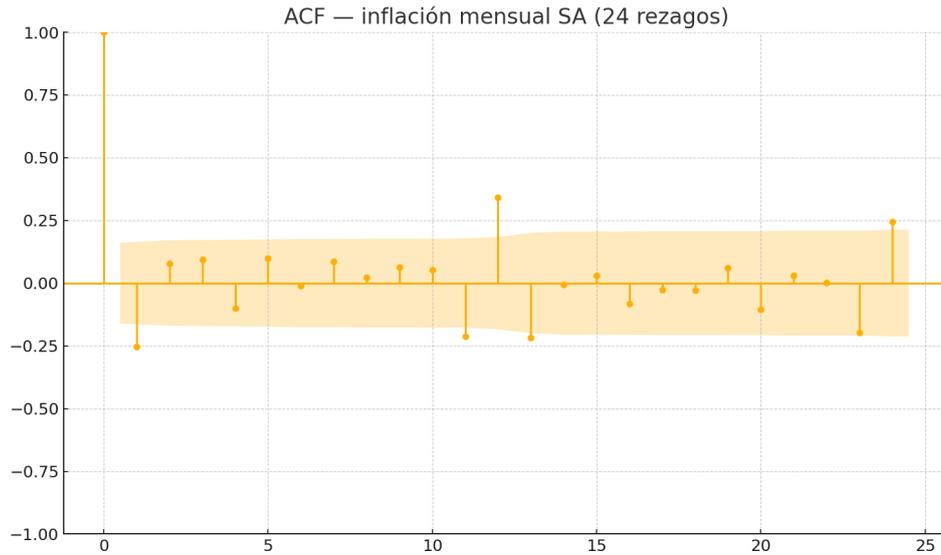


Figura 52: ACF de $\Delta \log SA$ (24 rezagos).
Fuente: elaboración propia.

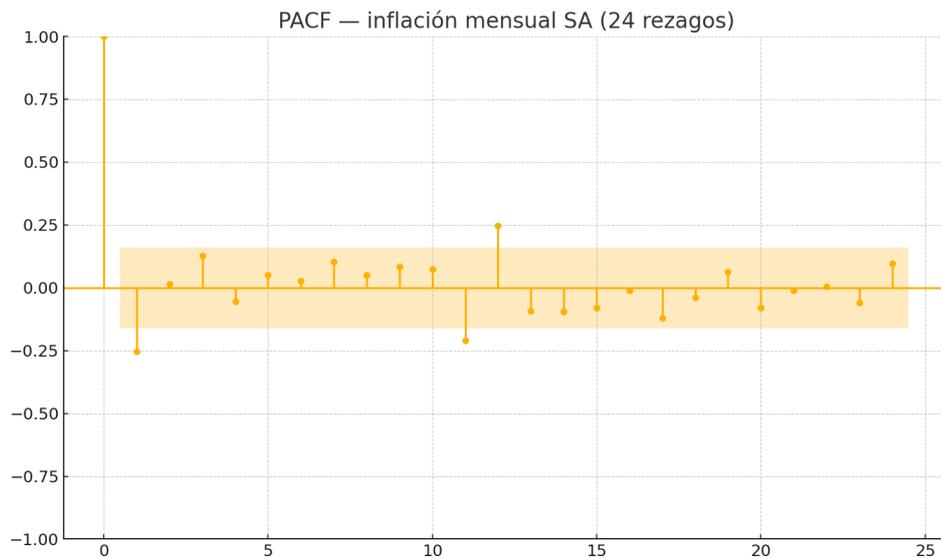


Figura 53: PACF de $\Delta \log SA$ (24 rezagos).
Fuente: elaboración propia.

¿Es el IPC apto para inferencias econométricas?

Trabajamos con el *IPC general* en *diferencias logarítmicas* mensuales y su versión desestacionalizada para asegurar propiedades estocásticas compatibles con la inferencia en series de tiempo. Definimos

$$\Delta \log(\text{IPC})_t \equiv \log(\text{IPC}_t) - \log(\text{IPC}_{t-1}), \quad \Delta \log(\text{IPC})_t^{SA} \equiv \log(\text{IPC})_t^{SA} - \log(\text{IPC})_{t-1}^{SA},$$

donde $\log(\text{IPC})_t^{SA}$ es el logaritmo del IPC desestacionalizado por STL (periodicidad 12). La inspección visual (Figuras 47, 50 and 51) y los tests formales (Tabla 17) permiten resumir:

- Estacionariedad.** Sobre $\Delta \log(\text{IPC})$ y $\Delta \log(\text{IPC})^{SA}$, las pruebas ADF (regresiones con constante y con constante+tendencia) *rechazan* raíz unitaria a niveles convencionales, mientras que KPSS típicamente *no rechaza* estacionariedad en especificaciones con constante. En conjunto, esto respalda que las diferencias (con o sin SA) son $I(0)$ y, por tanto, *aptas para inferencia* bajo marcos lineales estándar.

- **Distribución.** La normalidad no es exacta (Jarque–Bera, Shapiro–Wilk, Anderson–Darling muestran colas algo pesadas), pero el tamaño muestral y la naturaleza de las diferencias logarítmicas hacen que la inferencia sea razonablemente robusta; en caso necesario, use errores estándar robustos a no-normalidad.
- **Dependencia serial.** La ACF/PACF de $\Delta \log(\text{IPC})^{SA}$ no muestra patrones persistentes de largo plazo (Figuras 52 and 53); aun así, conviene contrastar con Ljung–Box y, si procede, modelar errores ARMA o usar *Newey–West*.
- **Estacionalidad y régimen.** La desestacionalización (STL) mitiga efectos estacionales; el acotamiento a 2013–2025 (régimen de política homogéneo en el estudio) reduce riesgos de inestabilidad paramétrica. Si se sospechan *rupturas estructurales*, considerar pruebas tipo Bai–Perron.
- **Calidad de la muestra.** La existencia de algunos meses faltantes exige completar o documentar un tratamiento consistente cuando se requiera periodicidad estricta (VAR/ARIMA).

Conclusión operativa. Para análisis econométrico de corto plazo (p.ej., especificaciones VAR/ARIMA o regresiones lineales con controles), *las tasas de inflación derivadas del IPC*—esto es, $\Delta \log(\text{IPC})$ o $\Delta \log(\text{IPC})^{SA}$ —**son apropiadas para inferencias** (estacionarias, con dependencia manejable y no-normalidad moderada). Debe evitarse usar el IPC en *niveles* salvo que exista cointegración demostrada con otras variables $I(1)$ (en cuyo caso corresponde un VECM). En aplicaciones como el modelo Tanzi–Feige, donde la dinámica de precios entra por la inflación, el uso de $\Delta \log(\text{IPC})$ es metodológicamente consistente y preferible.