

Trabajo Práctico Final

Análisis de Series Temporales

Maestría en Explotación de Datos y Gestión del Conocimiento

Layla Scheli, Franco Lianza, Lucio Scalzo, Ignacio Mujica, Alexis Walker

02 febrero 2022

Índice

1	Introducción	4
2	Marco Teórico	5
2.1	Tests de estacionariedad	5
2.2	Transformaciones para estabilizar la media y la varianza	5
2.3	Vectores autoregresivos	6
3	Análisis de Resultados	7
3.1	Índice de Precios al Consumidor	7
3.2	Índice de Salarios	9
3.3	Base Monetaria	11
4	Vectores Autoregresivos	13
4.1	Modelado	13
4.2	Análisis de diagnóstico	14
4.3	Causalidad e Impulso Respuesta	14
5	Conclusiones	16
6	Referencias bibliográficas	17
7	Apéndices	18
7.1	Modelo <i>VAR</i> con retardo 1 para Índice de Precios al Consumidor y Base Monetaria conjuntamente	18

Abstract

El presente trabajo tiene por objetivo estudiar la relación entre dos variables económicas fundamentales, el IPC (Índice de precios al Consumidor) y el índice de Salarios.

Ambos indicadores son considerados fundamentales para el desarrollo de un país ya que conllevan, en la relación entre ambos, dos efectos antagónicos según se comporten.

Por un lado, cómo evoluciona en nivel de vida de la población. Si los salarios se incrementan en mayor proporción que los bienes y servicios, consiguientemente se mejora el bienestar además del incremento de la producción ocasionando mayor demanda.

Por otro, como se demostrará en el desarrollo del presente. Si los indicadores de incremento de bienes y servicios crecen a mayor ritmo que los salarios, disminuye el bienestar de la masa asalariada.

Como resultado de este proceso inflacionario y de baja demanda interna hay que mencionar que esto perjudica a los sectores productivos internos.

Las causas que originan el fenómeno de inflación no serán tratadas por lo complejo y extenso que es el tema.

1 Introducción

Las series de tiempo son el resultado de observar valores de una variable aleatoria a lo largo del tiempo en intervalos iguales o desiguales. En el análisis de las series temporales se utilizan métodos que permiten interpretarlas y extraer información sobre las relaciones entre los datos de la serie u otras series.

Según Mochón and Beker (2008), “*La inflación es el crecimiento generalizado y continuo de los precios de los bienes y los servicios*” (p. 496), y el índice de precios al consumidor (IPC) representa el costo de la canasta de bienes y servicios consumidos por la economía local.

El fenómeno de la inflación se produce por múltiples causas. Las teorías tradicionales se pueden resumir en: inflación de demanda, que se produce cuando los gastos superan la capacidad de producción de la economía, por causas monetarias, cuando se incrementa la emisión monetaria, la inflación de costos, cuando los precios suben por grupos económicos de presión y la inflación estructural, que es la que se origina debido a los problemas estructurales de las economías en vías de desarrollo.

La percepción individual sugiere que los motivos, en la coyuntura actual, son variados: especulación, emisión desmedida, presión por devaluación, bajas exportaciones, monopolios formadores de precios, elevado gasto público, mercado laboral altamente regulado, etc.

Analizar como evolucionan conjuntamente en el tiempo el Índice de precios al consumidor y el Índice de salarios puede ayudar a clarificar los problemas y soluciones posibles.

Además, esta comparativa muestra cómo se comporta el poder adquisitivo de un empleado. Es posible determinar si se recompone conforme aumentan los precios, acompaña o se abre una brecha que da cuenta de su disminución.

Índice de Precios al Consumidor, Índice de Salarios acumulados y Base Monetaria

Con base Diciembre 2016

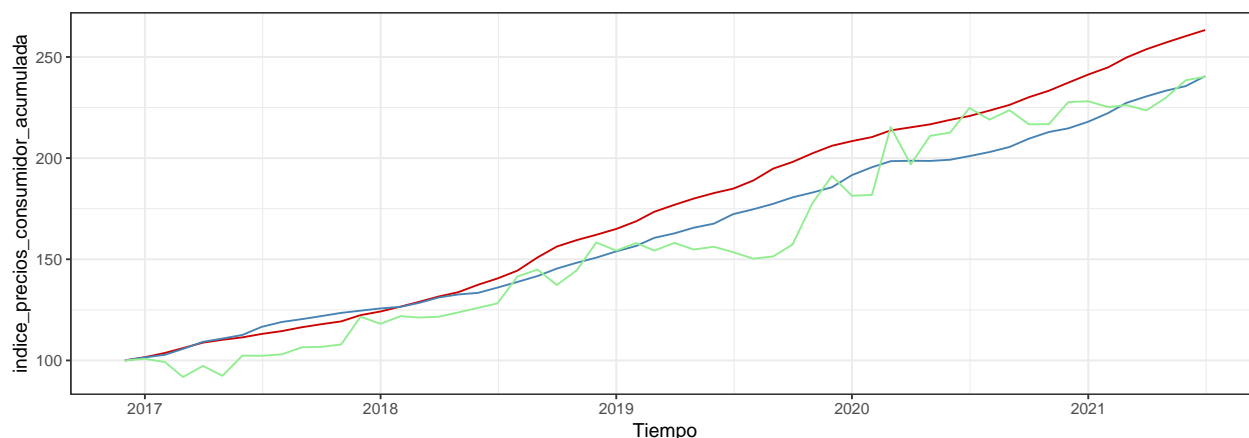


Fig. 1: Inflación versus Salarios

La información fue extraída de la página oficial del Instituto Nacional de Estadística y Censos - INDEC Argentina:

- Índice de precios al consumidor
- Índice de salarios

Se considera la variación mensual acumulada estableciendo a Diciembre 2016 como el 100% y se pretende predecir los cambios de las series para Enero 2022.

2 Marco Teórico

Los modelos autorregresivos y de medias móviles son validos para series de tiempo que son *estacionarias*. Se define que una serie es estacionaria de orden 2 cuando sus segundos momentos son constantes a lo largo del tiempo. Es decir, cuando su media y varianza se mantienen constantes.

Por lo contrario, una serie es *no estacionaria* cuando la tendencia y/o la variabilidad cambian a lo largo del tiempo.

El presente trabajo práctico tiene como objetivo el análisis de la estacionariedad de las series presentadas y su modelado mediante modelos SARIMA.

2.1 Tests de estacionariedad

La estacionariedad puede ser puesta a prueba mediante diferentes tests. Se detallan los dos a utilizar.

2.1.1 Dickey-Fuller Aumentado

En la ecuación de Dickey-Fuller aumentada (1):

$$\nabla Y_t = a_0 + a_1 t + (\phi_1 - 1)Y_{t-1} + \epsilon_t \quad (1)$$

Se pone a prueba el coeficiente que acompaña al rezago del período anterior: $(\phi_1 - 1)$.

El estadístico de este test es:

$$\tau^e = \frac{\hat{\phi}_1 - \phi_1}{\sigma(\hat{\phi}_1)} \quad (2)$$

Se plantea la regla de decisión:

Si $\tau \leq \tau^e < 0 \Rightarrow$ No Rechazar H_0

Si $\tau^e < \tau \Rightarrow$ Rechazar H_0

Como se explica en Peña Sanchez de Rivera (2010), la hipótesis nula establece que la raíz mayor de un $AR(p+1)$ es igual a uno, por lo que el proceso es no estacionario.

2.1.2 Kwiatkowski-Phillips-Schmidt-Shin

Kwiatkowski et al. (1992) definen otra prueba para raíces unitarias en la cual la hipótesis nula es contraria al test ADF, es decir, plantea que la serie temporal observable es estacionaria.

2.2 Transformaciones para estabilizar la media y la varianza

Según lo expresado en Peña Sanchez de Rivera (2010), la identificación de las estructuras no estacionarias consiste en detectar qué transformaciones hay que aplicar para conseguir un proceso con media y varianza constante.

En particular, consiste en determinar si es necesario: transformar la serie para que tenga varianza constante y/o diferenciarla para darle una media constante.

Por un lado, las transformaciones de potencia de Box-Cox son una forma general que permiten convertir el proceso original en un proceso condicionalmente homocedástico.

Por otro, como se puede asumir que la tendencia de una serie en el instante t es muy próxima a la del instante $t-1$, se puede formular una nueva serie de la manera $Y_t = y_t - y_{t-1}$. Este proceso se define como *diferenciación* y permite transformar una serie no estacionaria a estacionaria.

2.3 Vectores autoregresivos

El modelado ARIMA está limitado por su incapacidad para capturar interrelaciones dinámicas entre variables de interés. La generalización directa del modelo ARMA estacionario a la forma multivariante conduce a un modelo vectorial (George Athanasopoulos and Vahid (2012)).

Por variable en el sistema, existe una ecuación, cuyo lado derecho incluye una constante y retrasos de todas las variables presentes.

Para mantenerlo simple, se presenta un modelo VAR de dos variables con retraso igual a 1 ($p = 1$):

$$y_{1,t} = c_1 + \phi_{11,1}y_{1,t-1} + \phi_{12,1}y_{2,t-1} + \epsilon_{1,t} \quad (3)$$

$$y_{2,t} = c_2 + \phi_{21,1}y_{1,t-1} + \phi_{22,1}y_{2,t-1} + \epsilon_{2,t} \quad (4)$$

Donde $\epsilon_{1,t}$ y $\epsilon_{2,t} \sim N(0, \sigma_y)$ y no están autocorrelacionados. $y_{1,t}$ y $y_{2,t}$ son estacionarias. El coeficiente $\phi_{ii,p}$ captura la influencia del *lag* p de la variable y_i sobre si misma. Mientras que $\phi_{ij,p}$ determina el peso del *lag* p de la variable y_j sobre y_i .

3 Análisis de Resultados

3.1 Índice de Precios al Consumidor

3.1.1 Análisis de la serie

En la Figura 2, se observa que la serie de IPC es *no estacionaria* ya que su media varia a lo largo del tiempo y su función de autocorrelación decrece en forma lineal lentamente.

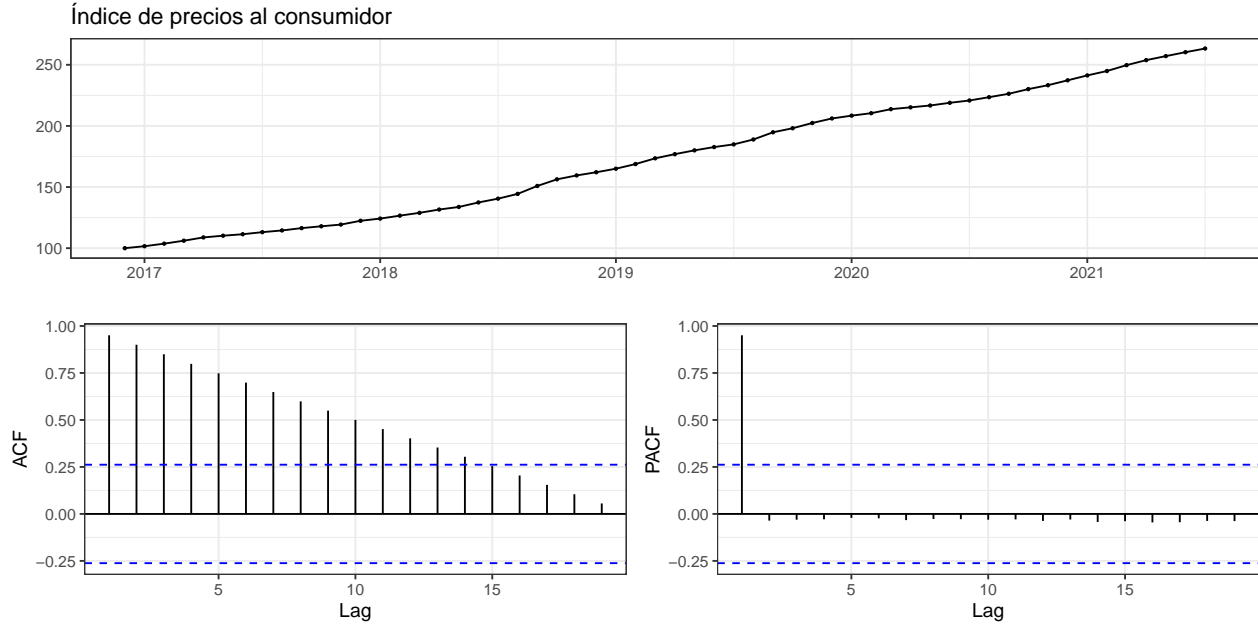


Fig. 2: Índice de precios al consumidor

Es posible someter la evidencia gráfica a distintas pruebas estadísticas para ratificar o rectificar lo observado.

Test de Estacionariedad: ADF

En la Tabla 1 se observa τ^e entre el valor crítico y cero, lo que determina que no hay evidencia para rechazar H_0 (serie no estacionaria).

	Estadístico	1pct	5pct	10pct
tau3	-2.469	-4.04	-3.45	-3.15
phi2	7.057	6.50	4.88	4.16
phi3	3.819	8.73	6.49	5.47

Tabla 1: Prueba de ADF para IPC

Test de Estacionariedad: KPSS

Los resultados de la prueba se encuentran detallados en la Tabla 2.

Estadístico	P-Value	Lag
1.5	0.01	3

Tabla 2: Test KPSS para IPC

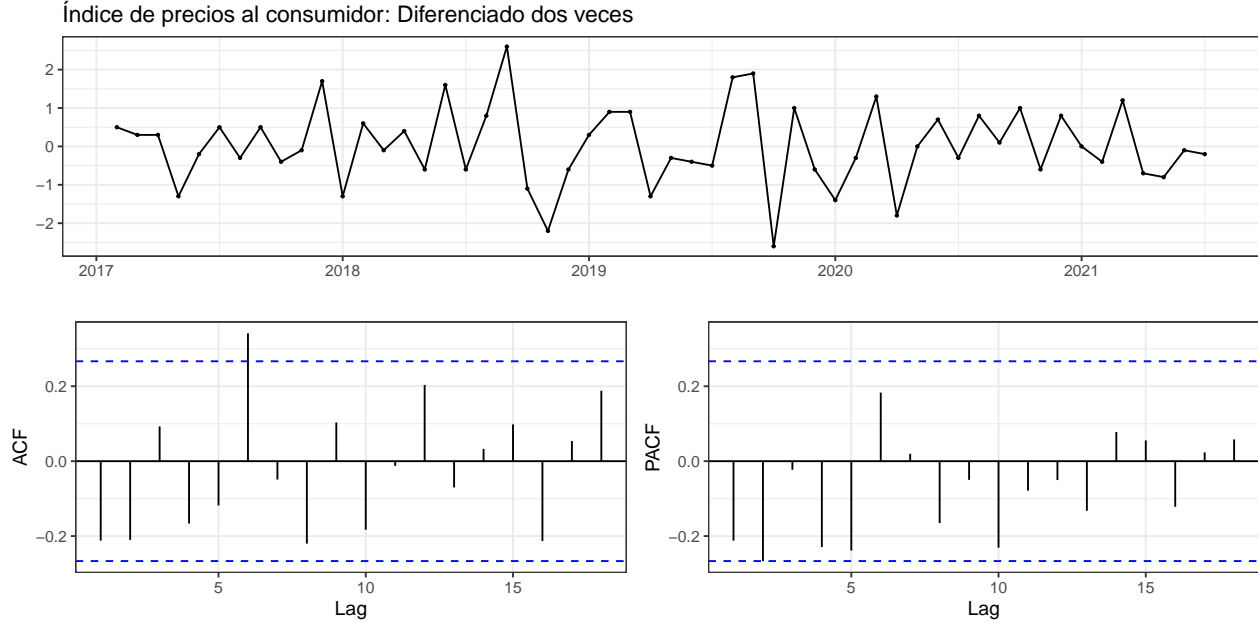
En este caso el p-value es menor al grado de significación 0.05, por lo que se rechaza H_0 .

En conclusión, ambos tests determinan que la serie de IPC es *no estacionaria*.

3.1.2 Diferenciación

Mediante la función `ndiffs` de R, se estima que el numero de diferenciaciones necesarias para convertir a la serie IPC en estacionaria (2 veces).

La Figura 3 muestra el indice de precios al consumidor diferenciado.



Test de Estacionariedad: ADF

Se corrobora que la serie de IPC diferenciada es estacionaria mediante el test Dickey-Fuller aumentado (Tabla 3). El valor de τ^e se encuentra dentro de la zona de rechazo, por lo que se concluye que se rechaza H_0 .

	Estadístico	1pct	5pct	10pct
tau3	-7.139	-4.04	-3.45	-3.15
phi2	16.994	6.50	4.88	4.16
phi3	25.486	8.73	6.49	5.47

Tabla 3: Test de ADF para IPC diferenciada

Test de Estacionariedad: KPSS

El test KPSS coincide con ADF. El p-value es mayor al grado de significación de 0.05, por lo que no se rechaza la hipótesis nula de estacionariedad.

Estadístico	P-Value	Lag
0.055	0.1	3

Tabla 4: Test KPSS para IPC diferenciada

Se confirma entonces que la serie diferenciada es *estacionaria*.

3.1.3 Modelado

En el trabajo práctico anterior, se obtuvo un modelo $ARIMA(0,2,6)$ para modelar la serie IPC (ver summary en Apendice). Si bien la serie no parece presentar algún componente de estacionalidad, se propone un modelo $SARIMA(0,2,0)(0,2,6)[12]$ (ver summary en Apendice) para compararlo con el modelo anterior.

3.1.4 Métricas de performance

Se comparan ambos modelos mediante AIC y BIC en la Tabla 5.

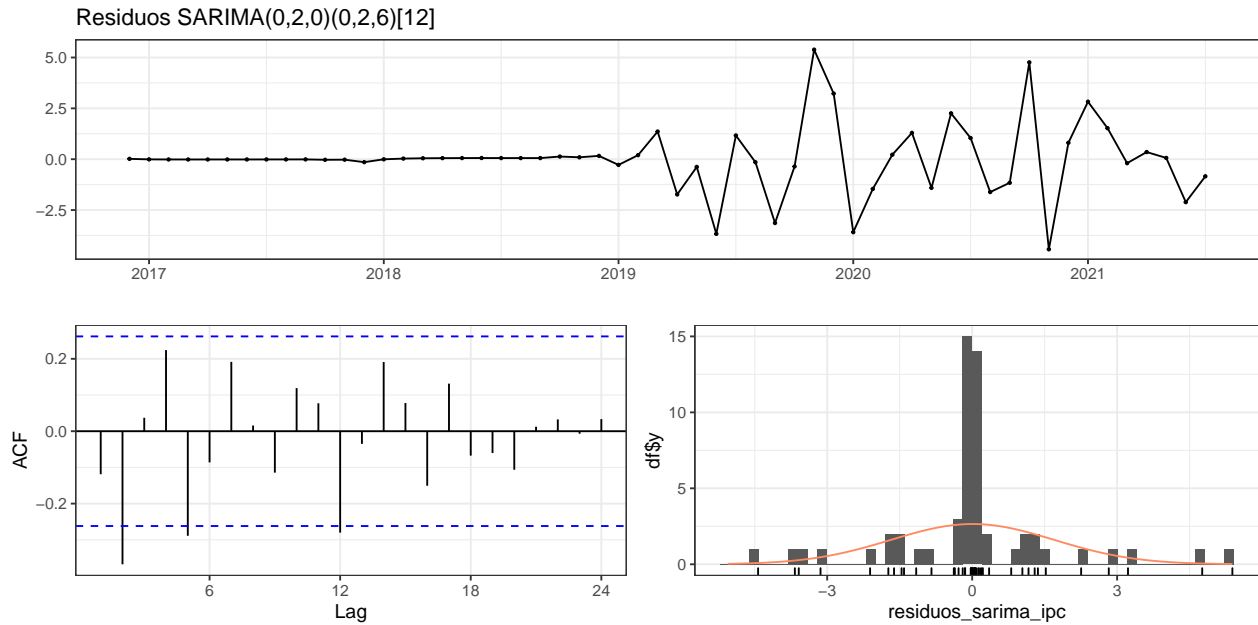
Modelo	AIC	BIC
$ARIMA(0,2,6)$	154.470	158.448
$SARIMA(0,2,0)(0,2,6)[12]$	144.063	148.267

Tabla 5: Metricas modelos propuestos para IPC

Parece que el modelo $SARIMA(0,2,0)(0,2,6)[12]$ es mejor al $ARIMA(0,2,6)$ propuesto en el trabajo anterior. Es necesario realizar un análisis de los residuos del modelo para verificar que presenta un buen ajuste.

3.1.5 Análisis de diagnóstico

Si bien el modelo SARIMA presenta mejores métricas de información que el modelo baseline del trabajo práctico anterior, la nueva propuesta no ajusta de manera correcta los datos. Al graficar los residuos en la Figura 4, se puede observar que éstos presentan autocorrelación y no se distribuyen normalmente.



3.2 Índice de Salarios

La Figura 5 muestra la serie de Índice de Salarios.

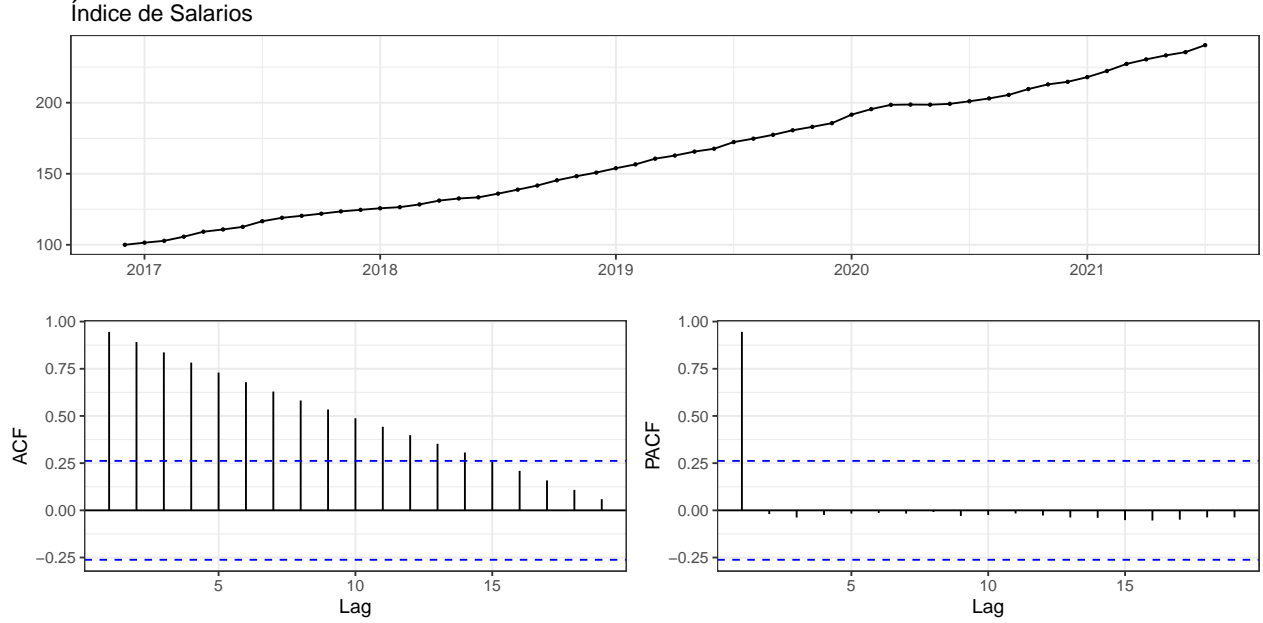


Fig. 5: Índice de Salarios

3.2.1 Análisis de la serie

Test de Estacionariedad: ADF

En la Tabla 6 se observa τ^e entre el valor crítico y cero, lo que determina que no hay evidencia para rechazar H_0 .

	Estadístico	1pct	5pct	10pct
tau3	-1.449	-4.04	-3.45	-3.15
phi2	7.416	6.50	4.88	4.16
phi3	2.343	8.73	6.49	5.47

Tabla 6: Prueba de ADF para Índice de Salarios

Test de Estacionariedad: KPSS

Los resultados de la prueba se encuentran detallados en la Tabla 7.

Estadístico	P-Value	Lag
1.5	0.01	3

Tabla 7: Test KPSS para Índice de Salarios

Ambos tests concluyen que la serie de Índice de Salarios es *no estacionaria*.

3.2.2 Diferenciación

En la Figura 5 se observa que, al igual que la serie de IPC, el Índice de Salarios tampoco es estacionario. Su media varía a lo largo del tiempo y su función de autocorrelación también decrece lentamente en forma lineal.

Mediante la función `ndiffs` de R, se estima que el número de diferenciaciones necesarias para convertir la serie en estacionaria. El resultado obtenido es 1.

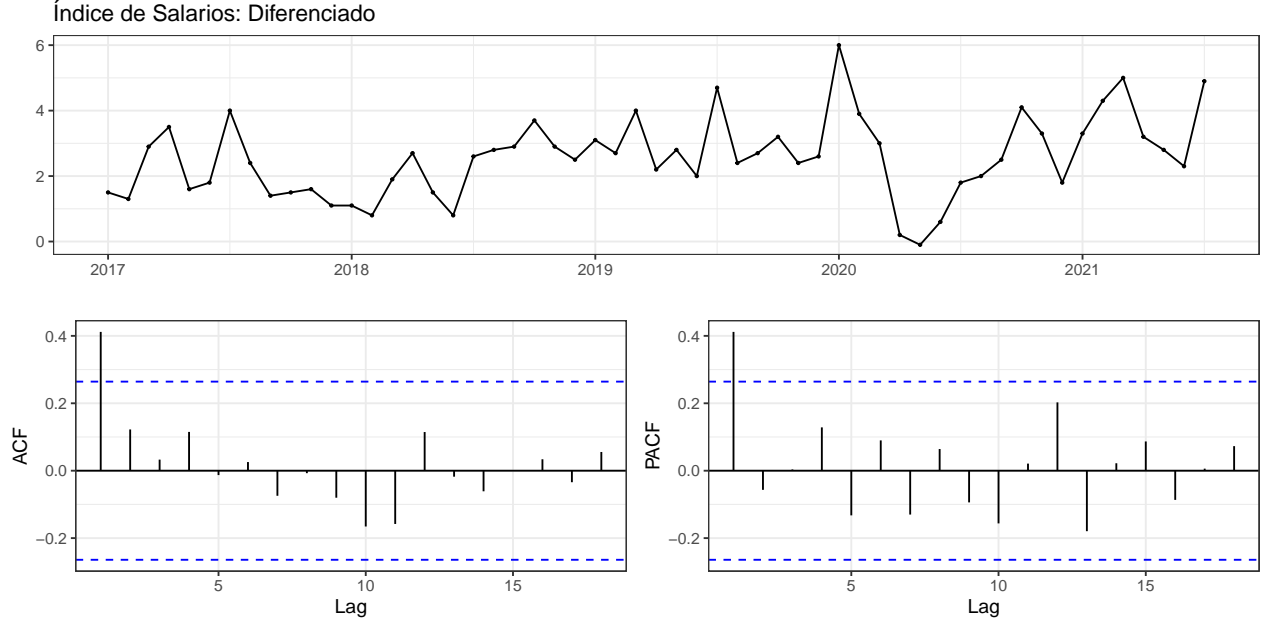


Fig. 6: Índice de Salarios diferenciado

Test de Estacionariedad: ADF

Al aplicar el test ADF a la serie diferenciada, el valor de τ^e se encuentra en la zona de rechazo de H_0 , por lo que se confirma que la serie diferenciada es estacionaria.

	Estadístico	1pct	5pct	10pct
tau3	-4.341	-4.04	-3.45	-3.15
phi2	6.361	6.50	4.88	4.16
phi3	9.435	8.73	6.49	5.47

Tabla 8: Prueba de ADF para Índice de Salarios diferenciado

Test de Estacionariedad: KPSS

El test KPSS sobre la serie diferenciada concuerda con el test ADF, concluyendo que la serie diferenciada es *estacionaria*.

Estadístico	P-Value	Lag
0.367	0.091	3

Tabla 9: Test KPSS para Índice de Salarios diferenciado

3.2.3 Modelado

En este trabajo practico no se propone otro modelo ya que en el anterior el seleccionado fue un $SARIMA(1,0,0)(1,0,0)[12]$ cuyo summary se encuentra en el Apéndice.

3.3 Base Monetaria

3.3.1 Analisis de la serie

En la Figura 7, se observa que la serie de _____ es *no estacionaria* ya que su media varia a lo largo del tiempo y su función de autocorrelación decrece en forma lineal lentamente.

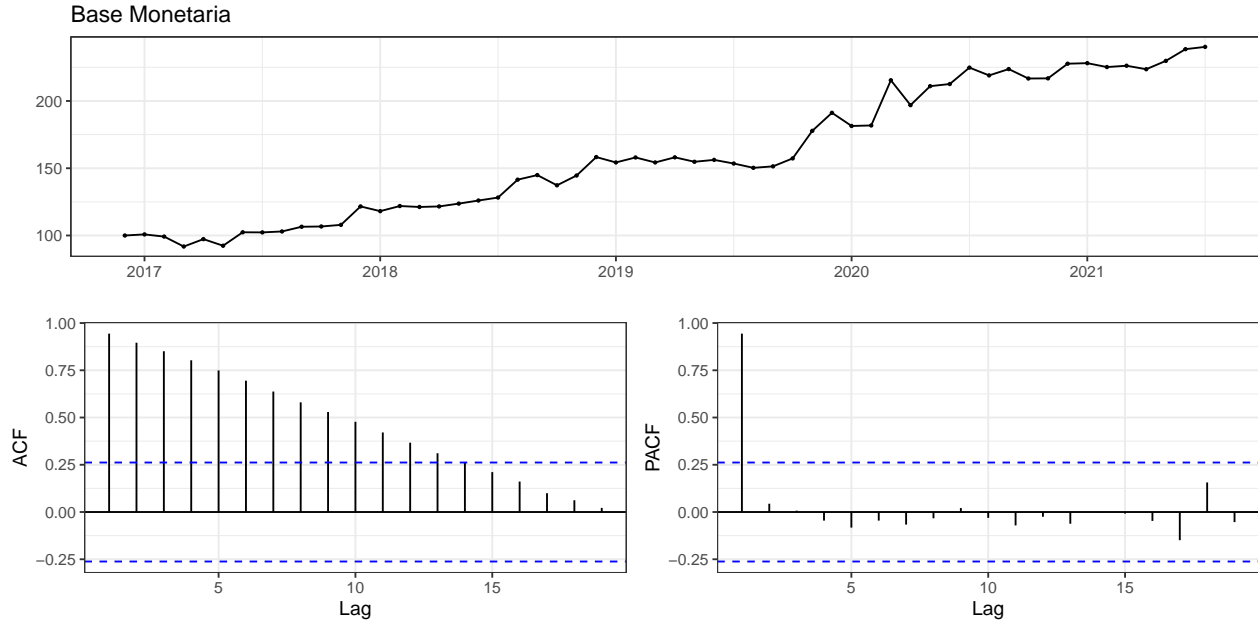


Fig. 7: Índice de precios al consumidor

Es posible someter la evidencia gráfica a distintas pruebas estadísticas para ratificar o rectificar lo observado.

Test de Estacionariedad: ADF

En la Tabla 10 se observa τ^e entre el valor crítico y cero, lo que determina que no hay evidencia para rechazar H_0 (serie no estacionaria).

	Estadístico	1pct	5pct	10pct
tau3	-3.016	-4.04	-3.45	-3.15
phi2	6.808	6.50	4.88	4.16
phi3	4.886	8.73	6.49	5.47

Tabla 10: Prueba de ADF para Base Monetaria

Test de Estacionariedad: KPSS

Los resultados de la prueba se encuentran detallados en la Tabla 11.

Estadístico	P-Value	Lag
1.46	0.01	3

Tabla 11: Test KPSS para Base Monetaria

En este caso el p-value es menor al grado de significación 0.05, por lo que se rechaza H_0 .

En conclusión, ambos tests determinan que la serie de IPC es *no estacionaria*.

3.3.2 Diferenciación

Mediante la función `ndiffs` de R, se estima que el numero de diferenciaciones necesarias para convertir a la serie ____ en estacionaria (1 veces).

La Figura 8 muestra el indice de precios al consumidor diferenciado.

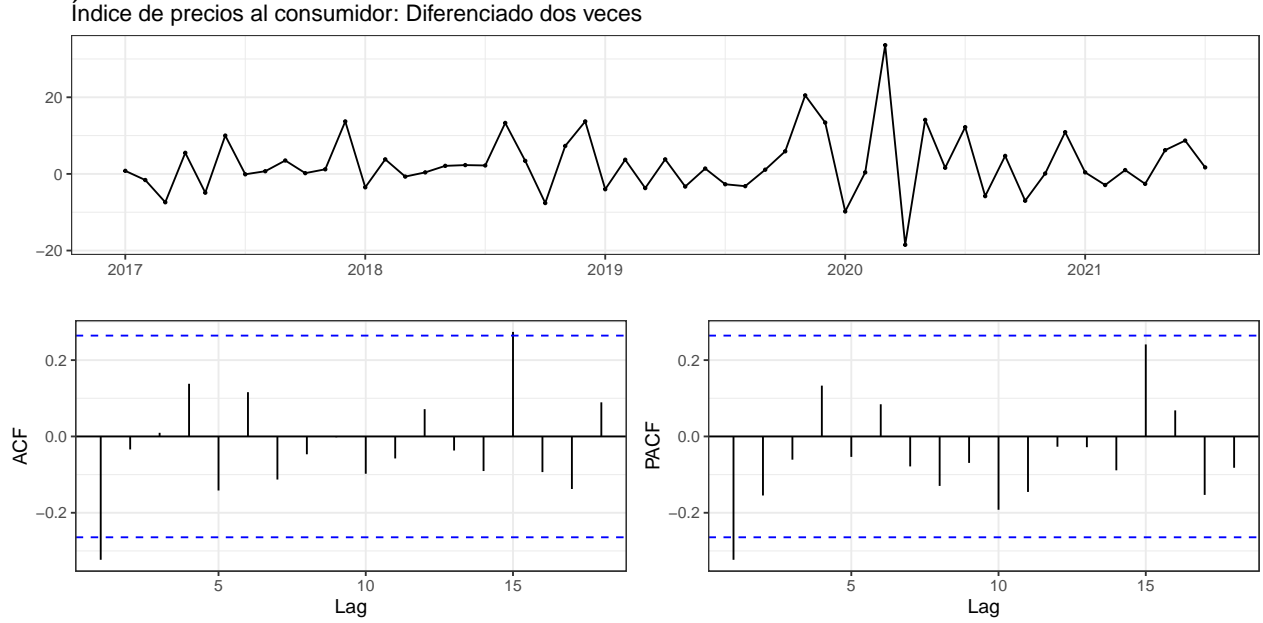


Fig. 8: Base Monetaria diferenciada

Test de Estacionariedad: ADF

Se corrobora que la serie de — diferenciada es estacionaria mediante el test Dickey-Fuller aumentado (Tabla 12). El valor de τ^e se encuentra dentro de la zona de rechazo, por lo que se concluye que se rechaza H_0 .

	Estadístico	1pct	5pct	10pct
tau3	-6.761	-4.04	-3.45	-3.15
phi2	15.245	6.50	4.88	4.16
phi3	22.857	8.73	6.49	5.47

Tabla 12: Test de ADF para Base Monetaria diferenciada

Test de Estacionariedad: KPSS

El test KPSS coincide con ADF. El p-value es mayor al grado de significación de 0.05, por lo que no se rechaza la hipótesis nula de estacionariedad.

Estadístico	P-Value	Lag
0.125	0.1	3

Tabla 13: Test KPSS para Base Monetaria diferenciada

Se confirma entonces que la serie diferenciada es *estacionaria*.

4 Vectores Autoregresivos

blah blah introduccion

4.1 Modelado

Se utilizan distintos criterios de información para decidir el número de retardos a incluir.

AIC(n)	HQ(n)	SC(n)	FPE(n)
1	1	1	1

Tabla 14: Criterios de selección

En la tabla 14, todos sugieren que el uso de $p = 1$ es apropiado para estimar el modelo. El resumen del modelo se incluye en el Apendice).

4.2 Análisis de diagnóstico

Para evaluar el ajuste del modelo se ejecutan algunas pruebas de diagnostico sobre los residuos del modelo.

Estadístico	Parámetro	P Value
57.609	44	0.082

Tabla 15: Residuos: Correlación

Con la prueba de Portmanteau se revisa correlación. Un p-value mayor al 5% indica la ausencia de correlación, no se rechaza la hipótesis nula.

Es posible tener evidencia de homocedasticidad utilizando la prueba Multiplicador de Lagrange, para efectos ARCH.

Estadístico	Parámetro	P Value
114.213	108	0.323

Tabla 16: Residuos: Homocedasticidad

En la Tabla 16 se observa un p-value por encima del valor crítico lo que no brinda evidencia suficiente para rechazar la hipótesis nula de homocedasticidad.

Se utiliza la prueba de Jarque-Bera para comprobar si una muestra de datos tiene la asimetría y la curtosis de una distribución normal.

	Estadístico	Parámetro	P Value
Jarque-Bera	47.209	4	0
Asimetría	18.127	2	0
Curtosis	29.082	2	0

Tabla 17: Residuos: Normalidad

Los resultados resumidos en la Tabla 17 no muestran residuos distribuidos normalmente. Si se analizan los generados por cada ecuación por separado, la prueba de Shapiro-Wilk (p-value igual a 4.4930085×10^{-4}) para el índice de precios señala normalidad, mientras que para la base monetaria no (p-value igual a 0.6791738).

De revisar la Figura 1 se observa un valor atípico para el mes de marzo de 2020, eliminando esta lectura se obtiene un p-value igual a 0.1851587 observando normalidad.

4.3 Causalidad e Impulso Respuesta

En este punto es posible probar la causalidad de Granger, donde la hipótesis nula de que no hay causalidad se descarta en ambas direcciones

Generar funciones de respuesta de impulso para describir la respuesta de la producción a un shock de desempleo

Método	P-Value
Granger causality H0: dif_ts_base_monetaria_star do not Granger-cause dif_ts_indice_precios_consumidor	0.728

Tabla 18: Causalidad de Grange: Base Monetaria

Método	P-Value
Granger causality H0: dif_ts_indice_precios_consumidor do not Granger-cause dif_ts_base_monetaria_star	0.014

Tabla 19: Causalidad de Grange: IPC

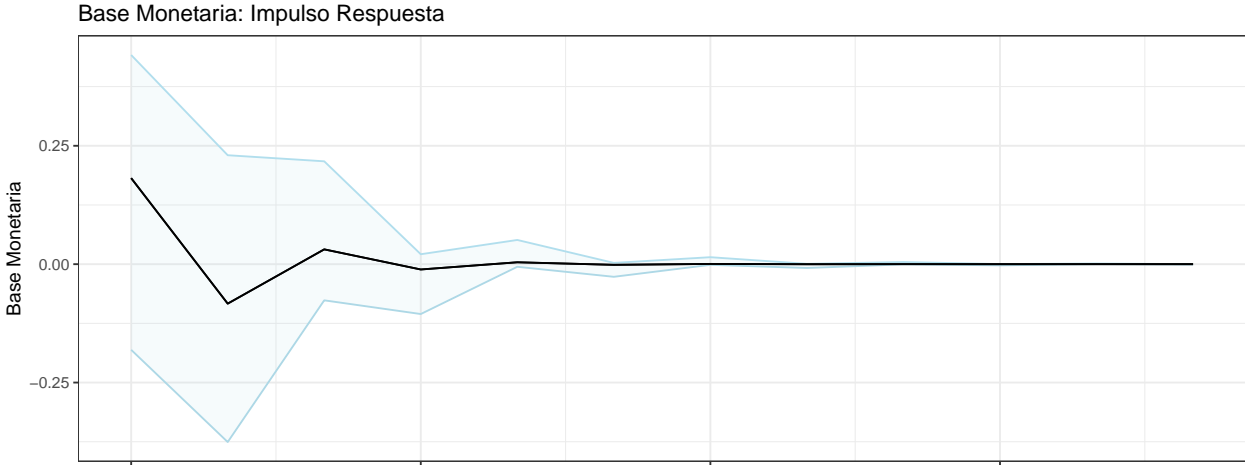


Fig. 9: Base Monetaria: Impulso Respuesta

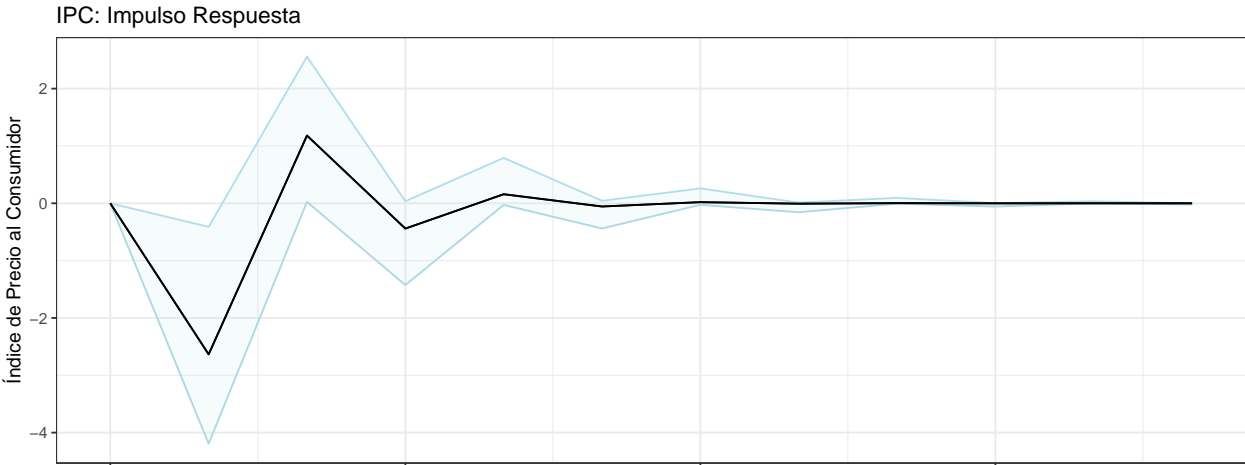


Fig. 10: IPC: Impulso Respuesta

5 Conclusiones

El Índice de Precios al Consumidor no presenta un comportamiento estacional. Esto evidencia una falta de políticas a largo plazo en cuanto a materia económica, constantes y evolutivas año tras año.

La sugerencia de ajuste del Índice de Salarios por la búsqueda iterativa de la función `auto.arima{forecast}` de R, propone un modelo $AR(1)(1)_{12}$ lo que solo evidencia un intento cíclico de recomponer parcialmente la pérdida frente a la inflación.

De no realizarse cambios en el rumbo económico, no se lograrán mejorar los indicadores y el salario real seguirá decreciendo, aumenta la brecha en el poder adquisitivo de la masa asalariada.

6 Referencias bibliográficas

George Athanasopoulos, D. S. Poskitt, and Farshid Vahid. 2012. “Two Canonical VARMA Forms: Scalar Component Models Vis-à-Vis the Echelon Form.” *Conometric Reviews* 31 (1): 60–83.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. “Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” *Journal of Econometrics* 54 (1): 159–78.

Mochón, F., and V. Beker. 2008. *Economía, Principios y Aplicaciones*. McGraw Hill.

Peña Sanchez de Rivera, D. 2010. *Análisis de Series Temporales*. Alianza Editorial S.A.

7 Apéndices

7.1 Modelo *VAR* con retardo 1 para Índice de Precios al Consumidor y Base Monetaria conjuntamente

```
##
## VAR Estimation Results:
## =====
## Endogenous variables: dif_ts_base_monetaria_star, dif_ts_indice_precios_consumidor
## Deterministic variables: const
## Sample size: 53
## Log Likelihood: -255.483
## Roots of the characteristic polynomial:
## 0.3532 0.0952
## Call:
## VAR(y = var_data, p = 1, type = "const", exogen = NULL)
##
##
## Estimation results for equation dif_ts_base_monetaria_star:
## =====
## dif_ts_base_monetaria_star = dif_ts_base_monetaria_star.l1 + dif_ts_indice_precios_consumidor.l1 + const
##
##               Estimate Std. Error t value Pr(>|t|)
## dif_ts_base_monetaria_star.l1    -0.2479    0.1295  -1.915  0.06120 .
## dif_ts_indice_precios_consumidor.l1 -2.5747    1.0270  -2.507  0.01547 *
## const                3.3823    1.0846   3.118  0.00301 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 7.521 on 50 degrees of freedom
## Multiple R-Squared: 0.2063, Adjusted R-squared: 0.1746
## F-statistic: 6.5 on 2 and 50 DF, p-value: 0.003096
##
##
## Estimation results for equation dif_ts_indice_precios_consumidor:
## =====
## dif_ts_indice_precios_consumidor = dif_ts_base_monetaria_star.l1 + dif_ts_indice_precios_consumidor.l1 + c
##
##               Estimate Std. Error t value Pr(>|t|)
## dif_ts_base_monetaria_star.l1    -0.006242    0.017892  -0.349   0.729
## dif_ts_indice_precios_consumidor.l1 -0.200421    0.141938  -1.412   0.164
## const                0.039250    0.149895   0.262   0.795
##
##
## Residual standard error: 1.039 on 50 degrees of freedom
## Multiple R-Squared: 0.04752, Adjusted R-squared: 0.009425
## F-statistic: 1.247 on 2 and 50 DF, p-value: 0.296
##
##
## Covariance matrix of residuals:
##               dif_ts_base_monetaria_star
## dif_ts_base_monetaria_star          56.563
## dif_ts_indice_precios_consumidor      1.367
##               dif_ts_indice_precios_consumidor
## dif_ts_base_monetaria_star           1.367
## dif_ts_indice_precios_consumidor      1.080
```

```
##
## Correlation matrix of residuals:
##               dif_ts_base_monetaria_star
## dif_ts_base_monetaria_star             1.0000
## dif_ts_indice_precios_consumidor        0.1748
##               dif_ts_indice_precios_consumidor
## dif_ts_base_monetaria_star             0.1748
## dif_ts_indice_precios_consumidor        1.0000
```