

¿Cómo se siente el presidente?*

Tomás Pacheco** Abigail Riquelme***

Esta versión: 31 de agosto de 2021

Abstract

Cómo se comunican las personas es un aspecto fundamental en la vida, particularmente en la de un presidente. En este artículo, estudiaremos la comunicación paraverbal del presidente de Argentina, Alberto Fernández, en Twitter desde el momento de su asunción. Para esto utilizamos los textos de sus *tweets* y los clasificamos con un algoritmo de análisis de sentimiento, el cual asigna a cada *tweets* la probabilidad de que este tenga un contenido *positivo* o *negativo*. No se encuentra un modelo que sea estrictamente superador al resto, aunque se pueden diferenciar leves ventajas en el uso de ciertos modelos durante breves períodos de tiempo, como ETS o VAR. Concluimos que no es sencillo pronosticar correctamente este tipo de variables que se caracterizan por tener un alto grado de subjetividad. Dicha dificultad es creciente en la cantidad de pasos adelante según la cual se pronostica.

Palabras clave: Twitter, análisis de sentimiento, Alberto Fernández, pronósticos, COVID-19

* Asumimos la responsabilidad por todos los errores. El código para replicar los pronósticos puede encontrarse en el siguiente [enlace](#). Ante cualquier duda o error, comunicarse con [Abigail Riquelme](mailto:Abigail.Riquelme).

** Estudiante de Licenciatura en Economía en la Universidad de San Andrés, tpacheco@udesa.edu.ar

*** Estudiante de Licenciatura en Economía en la Universidad de San Andrés, ariquelme@udesa.edu.ar

1. Introducción

La comunicación es fundamental en la vida de todas las personas: estudiantes, profesionales, deportistas y políticos, especialmente el presidente. Cuando hablamos de comunicación no solo nos referimos a su contenido, sino a la forma en la que se expresan las ideas. Diversos autores han estudiado el paralenguaje utilizado por distintos presidentes en actos oficiales o campañas ([McCormick y Stuckey, 2013](#)). Sin embargo, se encuentra poca evidencia de lo que expresan los presidentes a través de Twitter.

En este trabajo, analizaremos el paralenguaje del actual presidente argentino, Alberto Fernández, desde un enfoque particular: el análisis de sentimiento (*sentiment analysis*, en inglés). Lo que haremos será construir una serie de tiempo utilizando los *tweets* que ha publicado el presidente desde el inicio de su mandato. A través de este algoritmo, se le asignará un valor entre 0 y 1 a cada publicación en la red social, de forma tal de reflejar la probabilidad de que el texto sea “positivo”. Un valor más cercano a 1 representará un sentimiento mayormente positivo, mientras que un valor más cercano a 0 reflejará un sentimiento negativo. Resulta importante aclarar, que a lo largo de todo este trabajo, nos referiremos al valor que toma este índice con la palabra *sentimiento*.

Nuestro principal interés es identificar la mejor forma de pronosticar cómo serán las emociones del presidente, así como cuáles son las variables económicas, epidemiológicas y climáticas que determinan su comportamiento. También, analizaremos con cuánta anticipación se pueden hacer predicciones de algo tan subjetivo como lo es el sentimiento de Alberto Fernández.

Dentro de los resultados a los que llegamos, podemos destacar que no encontramos grandes ventajas en la utilización de un modelo en específico, aunque pudimos identificar leves beneficios al pronosticar un paso hacia adelante con los modelos ETS y VAR. Además, notamos que resulta muy difícil predecir correctamente esta variable con una anticipación mayor a un día.

El trabajo está dividido en cuatro partes. La primera presenta los datos utilizados. La segunda plantea la metodología y el desarrollo del trabajo. En la tercera parte, se hace referencia a los resultados obtenidos. Finalmente, se presentan las conclusiones del estudio y sus limitaciones.

2. Literatura relacionada

Las redes sociales han tomado un papel muy importante en la sociedad, ya que constituyen un espacio en el que la gente se interrelaciona y expresa opiniones y pensamientos. Es por esto que dicho tipo de plataformas almacenan una gran cantidad de información. El tratamiento y la exploración de estos datos representan, al mismo tiempo, un gran desafío y una gran oportunidad. El análisis de sentimiento surgió como una respuesta a este desafío ya que de alguna manera, permite captar sentimientos positivos o negativos en el texto de cada *tweet*. Esta herramienta ha sido ampliamente utilizada con diferentes fines.

Diversos autores han utilizado este análisis con el objetivo de incorporar información de Twitter a la hora de pronosticar distintos precios. Por ejemplo, [Kordonis, Symeonidis, y Arampatzis \(2016\)](#) encuentran una correlación positiva entre el sentimiento del contenido del texto y el precio de algunas acciones. Además, [Kraaijeveld y De Smedt \(2020\)](#) llegan a la conclusión de que el sentimiento tiene un poder predictivo importante a la hora de pronosticar los retornos de las principales criptomonedas que se encuentran en el mercado. No solo la información es útil para pronosticar retornos, sino que también es relevante para modelizar volatilidades financieras ([Kranefuss y Johnson, 2021](#)). Por último, como muestran [Pai y Liu \(2018\)](#), el contenido de las redes sociales, junto con el valor de las acciones, son esenciales a la hora de hacer pronósticos sobre las ventas mensuales de autos en Estados Unidos.

La incorporación del análisis de sentimiento no solo ha sido relevante en el ámbito financiero, sino que también en el político. Hay una vasta literatura que aplica esta técnica con el objetivo de pronosticar resultados de elecciones presidenciales. Para Indonesia, [Budiharto y Meiliana \(2018\)](#) utilizan los *tweets* de los candidatos a presidente y los *hashtags* relacionados a la elección para hacer los pronósticos. El modelo pronosticó correctamente al ganador de la elección. Lo mismo sucedió para Chile, ya que [Rodríguez et al. \(2018\)](#) predicen correctamente el resultado de la elección presidencial. A su vez, [Liu, Yao, Guo, y Wei \(2021\)](#) desarrollan un modelo predictivo de los resultados de elecciones en Estados Unidos en el que incluyen una variable del sentimiento de *tweets*, cuya tasa de acierto es del 81 %. También relacionado con la política, [Yaquub \(2020\)](#) encuentra una correlación negativa y estadísticamente significativa entre los *tweets* del presidente Donald Trump y la cantidad

de casos de COVID-19 en Estados Unidos.

El análisis de sentimiento, además de generar pronósticos, ha tenido otros usos. Por ejemplo, [Molteni y De Leon \(2016\)](#) lo utilizan para analizar si existe una relación entre la audiencia de las series de televisión en Estados Unidos y el sentimiento de los *tweets* relacionados a dicho programa. De acuerdo a los autores, esta relación es positiva y significativa. Otro ejemplo es el del trabajo de [Beigi, Hu, Maciejewski, y Liu \(2016\)](#), en el que estudian la información contenida en las redes sociales durante desastres naturales y emergencias.

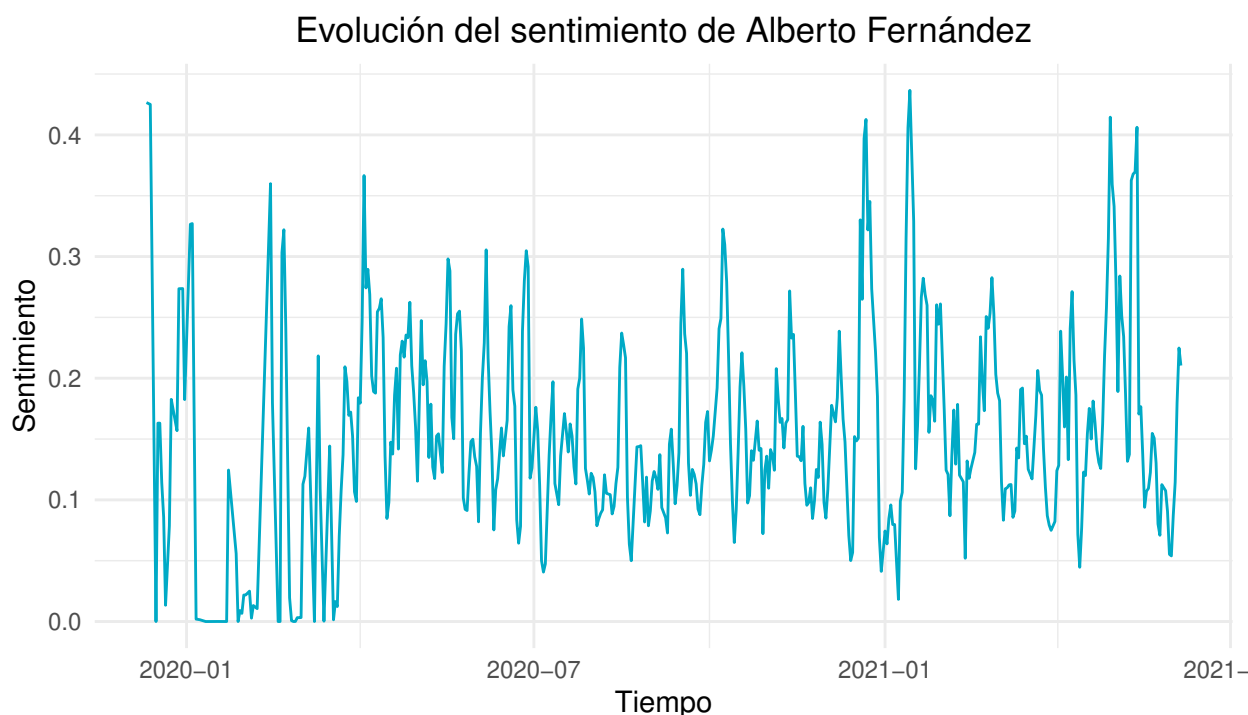
Si bien este tipo de análisis de texto ha sido utilizado en diversas áreas, como ya hemos mencionado, consideramos importante destacar que en todos los trabajos la utilización de esta técnica contribuye a explicar o pronosticar las variables de interés. En términos generales, autores como [Arias, Arratia, y Xuriguera \(2014\)](#) han analizado el diferencial de incluir información de redes sociales, que en muchos casos resulta significativo. A pesar de ser utilizado como componente explicativo, no se identifica una gran literatura que intente explicar el sentimiento de una persona en Twitter. Es por esto que consideramos que este artículo contribuye a analizar la posibilidad de pronosticar contenido con un alto nivel de subjetividad.

3. Datos

3.1. Descripción y fuentes

La principal variable de todo nuestro trabajo, el sentimiento del presidente de la Nación, la construimos a partir de los *tweets* que este publicó en la red social desde su asunción el 10 de diciembre de 2019. El contenido de cada Tweet, como también la cantidad de favoritos y de *retweets*, fue descargado utilizando la API que la misma red social ofrece. A partir de esta información, construimos nuestra serie aplicándole, a cada Tweet, el análisis de sentimiento. Este análisis, realizado con el paquete SENTIMENT-ANALYSIS-SPANISH en Python, lo que hace es analizar el texto y asignar una probabilidad de que este refleje un lenguaje positivo. Por ejemplo: el texto “me divierto mucho con mis amigos” va a tener un puntaje alto, dado que refleja un sentimiento positivo; mientras que la frase “me parece terrible esto que me estás haciendo” va a tener asignada una probabilidad que tiende a cero, es

decir, que revela un sentimiento negativo. El algoritmo utiliza redes neuronales para la predicción y, según su [documentación](#), tiene una tasa de acierto del 88 %. Debido a que hubo días en los que el presidente *twiteó* más de una vez, tomamos el promedio del sentimiento por día. Con el objetivo de ver de manera más clara la tendencia de nuestra serie, aplicamos un suavizado a través de la media móvil tomando cuatro observaciones previas. En la Figura (1) se puede ver nuestra principal serie de interés.



Fuente: elaboración propia

Figura 1: Sentimiento del presidente Alberto Fernández

Para realizar los pronósticos, utilizamos distintas variables explicativas. Las dos primeras, que surgieron del *scraping* de *tweets*, son la cantidad de favoritos y de *retweets* que tiene cada publicación. Al resto de las variables podemos dividirlos en tres grupos: variables económicas, variables epidemiológicas y variables climáticas. Dentro del primer grupo tenemos el tipo de cambio contado con liquidación¹ (pesos por dólares). Estos datos fueron extraídos de [Rava Bursátil](#). La siguiente variable económica que utilizamos fue la tasa de interés mínima de plazo fijo a 30 días, y la serie fue descargada desde el sitio web del Banco Central. De esta misma fuente, utilizamos la serie de

¹Elegimos el tipo de cambio contado con liquidación dado que Argentina tiene un régimen cambiario de tipo de cambio fijo, entonces, el impacto de la situación macroeconómica no se ve reflejada en el tipo de cambio oficial.

reservas del Banco Central y la base monetaria de Argentina.

El siguiente grupo de variables que utilizamos son las relacionadas a la situación epidemiológica por la pandemia de COVID-19. El año 2020 fue un período en el que toda la discusión pública estuvo centrada en torno a la pandemia. Es por esto que suponemos que puede haber cierta influencia de las series que reflejan la situación epidemiológica en Argentina. Elegimos como predictores a la cantidad de casos confirmados de coronavirus y a la cantidad de muertos a raíz de esta enfermedad, extraídos de la página del [Ministerio de Salud de la Nación](#). Como la pandemia fue un suceso mundial, consideramos relevante incluir en nuestros modelos la situación en Argentina relativa al resto del mundo. Por tal motivo, construimos una variable que definimos como el ratio entre la cantidad casos diarios en Argentina y la cantidad de casos en el resto del mundo. Lo mismo hicimos con las muertes por COVID-19. Los datos de casos y muertes a nivel mundial fueron extraídos de [Our World in Data](#).

Diversos autores ([Howarth y Hoffman, 1984](#); [Keller et al., 2005](#)) relacionan el estado de ánimo de las personas con el clima. Dado que con nuestro análisis intentamos predecir el sentimiento de una persona, consideramos relevante incluir este tipo de variables. Particularmente, utilizamos la temperatura máxima y mínima para cada día en la Ciudad de Buenos Aires. Estos datos fueron extraídos de [World Weather Online](#).

3.2. Características

Para todas las variables descritas previamente, contamos con sus valores desde el 11 de diciembre de 2019 hasta el 5 de junio de 2021. Estas series no son continuas debido a que la principal variable (sentimiento del presidente) no tiene valores para todos los días. Esto sucede debido a que el presidente no *twiteó* todos los días dentro del período que consideramos. En total, tenemos 483 observaciones en nuestra base de datos. Para llevar a cabo el análisis, dividimos los datos en dos ventanas. La primera de ellas, la ventana de estimación (*in-sample*), comienza el 11 de diciembre de 2019 y termina el 28 de noviembre de 2020. Los datos restantes (28 de noviembre hasta el 5 de junio de 2021), constituyen nuestra ventana de pronósticos (*out-of-sample*). Esta última representa el 12 % de las observaciones totales.

En la Figura (1) observamos, tal como se especificó anteriormente, nuestra principal variable de interés. En primer lugar, podemos notar que los valores del sentimiento de Alberto Fernández están, durante todo el período analizado, por debajo del 0.5. Es decir, presenta un sentimiento más cercano a cero que a uno, lo cual indica un sentimiento más negativo. En segundo lugar, podemos decir que visualmente se puede identificar una posible tendencia decreciente al inicio del período. Esto parece cambiar levemente entre julio de 2020 y enero de 2021.

Con respecto a las covariables, podemos identificar una tendencia creciente en algún momento del período considerado en el tipo de cambio (Figura 17 en Apéndice), la base monetaria (Figura 18 en Apéndice), la cantidad de vacunas aplicadas (Figura 19 en Apéndice), la cantidad de casos de COVID-19 en Argentina relativos respecto del resto del mundo (Figura 20 en Apéndice), la temperatura mínima (Figura 21 en Apéndice), la temperatura máxima (Figura 22 en Apéndice), cantidad de casos de COVID-19 (Figura 23 en Apéndice) y las reservas del BCRA (Figura 24 en Apéndice). En algunas de estas variables es posible observar, a partir de los gráficos, posibles tendencias decrecientes.

También, resulta útil nombrar que, limitándonos al análisis meramente visual, pareciera que algunas de las covariables presentan estacionalidad; por ejemplo, la temperatura mínima, la temperatura máxima y la cantidad de muertes por COVID-19 en Argentina. Por último, podemos destacar que las variables cantidad de favoritos (Figura 25 en Apéndice) y cantidad de *retweets* (Figura 26 en Apéndice) contienen un valor atípico en, aproximadamente, enero del año 2021.

En general, las conclusiones que se pueden obtener, tanto de la variable de interés como de las covariables, a partir de un análisis visual de los gráficos no son concretas. Por lo tanto, antes de comenzar a entrenar los modelos para luego pronosticar, le hicimos a cada una de nuestras series un *test* de raíz unitaria para evaluar su estacionariedad. En la Tabla (1) se presenta el *test* de Dickey Fuller Aumentado. Cabe mencionar que la longitud de rezago óptima fue seleccionada con el criterio de información de Akaike. Esta información es relevante debido a que, para la estimación de ciertos modelos, como por ejemplo el VAR, necesitamos que las variables sean estacionarias.

Variable	None	Trend	Drift
SentimientoAF	-4,15***	-9,88***	-9,6***
Favoritos	-11,15***	-14,24***	-13,25***
Retweets	-9,22***	-13,36***	-12,21***
Reservas BCRA	-1.09	1.67	-1.28
L_1 (Reservas BCRA)	-11,73***	-12***	-11,76***
Tasa de interés	-1	-4,11***	-2,86*
Base monetaria	0.3	-5,05***	-3,17**
Tipo de cambio	1.76	-2.57	-1.65
L_1 (Tipo de cambio)	-15,96***	-16,19***	-16,18***
Casos Arg.	-2,91***	-7,38***	-4,38***
Muertes Arg.	-2,82***	-5,39***	-4,46***
Vacunas	-3,75***	-6,58***	-4,4***
Casos Arg. Rel.	-2,17**	-5,05***	-3,89***
Muertes Arg. Rel.	-2,3**	-3,43**	-3,43**
Temp. Min.	-1.45	-4,92***	-4,92***
Temp. Max.	-1.49	-4,9***	-4,91***

Nota: *p<0.1; **p<0.05; ***p<0.01

Cuadro 1: Test de raíz unitaria: Dickey-Fuller aumentado

En la Tabla (1) identificamos que solo hay dos variables que no son estacionarias al 10 %: el tipo de cambio y las reservas del Banco Central. En este sentido, comprobamos que la diferencia de estas dos variables sí resulta estacionaria, con lo cual podemos decir que ambas son integradas de grado uno. El resto de las variables consideradas son estacionarias en niveles, es decir, integradas de grado cero. Resulta útil destacar que los resultados de la estacionariedad de las variables consideradas son robustos a la utilización del *test* de raíz unitaria de Phillips-Perron (Tabla 4 en Apéndice).

4. Metodología

Con el objetivo de realizar pronósticos en el período *out-of-sample*, realizamos algunas estimaciones preliminares en la ventana *in-sample*. En particular, comenzamos estimando varios modelos uniecuacionales. Decidimos concentrarnos en los modelos ARIMA, ARIMAX, ETS y ADL. En los modelos ARIMAX y ADL utilizamos las covariables nombradas anteriormente como regresores. También, consideramos relevante estimar modelos multiecuaciones como el VAR y el FAVAR.

Antes de estimar los modelos uniecuacionales, lo que hicimos fue seleccionar aquellos modelos sobre la base de algún criterio establecido. En cuanto al modelo ARIMA, podemos decir que elegimos aquel

que minimiza el criterio de información de Akaike corregido, este es un ARIMA(5,0,1). En otras palabras, seleccionamos un modelo autorregresivo de orden 5 con una media móvil. Los resultados de la estimación de este modelo se encuentran en la Columna (1) de la Tabla (2). Es posible observar que el primer, cuarto y quinto componente autorregresivo son estadísticamente significativos.

Con respecto al modelo ARIMAX, seleccionamos el modelo ARIMA que minimiza el criterio de información de Akaike corregido que, tal como se nombró anteriormente, es un ARIMA (5,0,1). A la estimación de este modelo le incorporamos las covariables como regresores. En la Columna (2) de la Tabla (2) se pueden observar los resultados de esta estimación. Detectamos que, nuevamente, el primer, cuarto y quinto componente autorregresivo son significativos. Luego, con respecto a la significatividad de las variables incorporadas como regresores, solo podemos decir que el tipo de cambio es significativo.

El modelo de suavización exponencial resultante fue un ETS(A,Ad,N). Esto indica que se incluye un error aditivo y una tendencia aditiva *damped*. Los parámetros de suavización estimados se encuentran en la Columna (3) de la Tabla (2).

También, estimamos un modelo ADL con el objetivo de explicar al sentimiento de Alberto Fernández en función de sus propios rezagos y de los rezagos de las demás variables. El algoritmo seleccionó como mejor modelo al ADL(2, 4, 1, 1, 0, 0, 0, 0, 0, 0, 2, 1, 0, 0). Esto significa que se incluyen dos rezagos para la variable que corresponde al sentimiento del presidente y para la temperatura máxima en Buenos Aires, un rezago para la cantidad de *retweets*, las reservas del Banco Central y para la temperatura mínima, y cuatro rezagos para la cantidad de favoritos. Para el resto de las variables no se incluyen rezagos. Los resultados de la estimación de este modelo se puede observar en la Columna (4) de la Tabla (2). A partir de esta Tabla podemos decir que las variables relevantes son el primer rezago de la temperatura mínima y el primer y segundo rezago del sentimiento de Alberto Fernández, las reservas del BCRA, la tasa de interés y la temperatura máxima.

Con el objetivo de validar los modelos nombrados anteriormente, aplicamos el *test* de autocorrelación sobre los residuos de cada uno ellos (Ljung-Box). De esta forma es posible observar si queda o

Cuadro 2: Estimaciones *in-sample*

	<i>Variable dependiente: SentAF</i>					
	<i>ARIMA</i>	<i>ARIMA</i>	<i>ETS</i>	<i>OLS</i>	<i>VAR</i>	<i>FAVAR</i>
	(1)	(2)	(3)	(4)	(5)	(6)
$L_1(\text{SentAF})$	0,503** (0.201)	0.421** (0.197)		0,800*** (0.054)	0,803*** (0.054)	0,799*** (0.052)
$L_2(\text{SentAF})$	0.151 (0.160)	0.191 (0.163)		-0,174*** (0.049)		
$L_3(\text{SentAF})$	-0.003 (0.060)	-0.024 (0.060)				-0,231*** (0.066)
$L_4(\text{SentAF})$	-0,232*** (0.058)	-0,225*** (0.050)			-0,192*** (0.070)	
$L_5(\text{SentAF})$	0,182*** (0.057)	0,153*** (0.045)			0,252*** (0.067)	0,252*** (0.066)
$L_6(\text{SentAF})$					-0,116** (0.051)	
$MA(1)$	0.296 (0.203)	0,369* (0.190)				
ReservasBCRA				0,0001* (0.00004)		
$L_1(\text{ReservasBCRA})$				-0,0001* (0.00004)		
Tasa de Interés		-0.001 (0.001)		-0,002** (0.001)		
Tipo de Cambio		-0,001** (0.0004)		-0.001 (0.0004)		
$L_1(\text{TwRet})$					-0,00001** (0.00000)	
$L_3(\text{CasosArg})$					0,00001* (0.00001)	
$L_2(\text{MaxTemp})$				-0,002** (0.001)	-0,002* (0.0001)	
$L_1(\text{MinTemp})$				0,002* (0.001)		
$L_1(\text{PC2})$						0,013* (0.007)
$L_6(\text{PC4})$						0,021* (0.012)
Alpha			0.9011			
Beta			0.0001			
Phi			0.822			
Constante	0,154*** (0.009)	0,978*** (0.282)		0.371 (0.281)	0.070 (0.061)	0,141*** (0.033)
Tendencia					-0,0002** (0.0001)	-0,0003*** (0.0001)
Observaciones	425	425	425	421	418	417
R ²				0.573	0.633	0.592
Log Likelihood	614.389	618.204				
Ljung-Box <i>p</i> -valor	0.1405	0.6359	0.0000	0.0026	0.0000	0.00236

Nota:

*p<0.1; **p<0.05; ***p<0.01

no dinámica en el modelo. La Tabla (2) muestra que con los modelos ARIMA y al ARIMAX estimados no rechazamos la hipótesis de no autocorrelación de los residuos, es decir no ,queda dinámica en los errores. Por lo tanto, podemos validar estos modelos.

Al aplicar el *test* de Ljung-Box a los residuos de los modelos ETS y ADL, rechazamos la hipótesis nula de no autocorrelación de los errores. Por lo tanto, no podemos validar estos modelos. Sin embargo, decidimos considerar al ETS y ADL para realizar pronósticos debido a que sabemos que al trabajar con datos diarios es muy difícil incorporar toda la dinámica en el modelo. En otras palabras, con datos diarios es difícil no rechazar la hipótesis nula de no autocorrelación de los residuos.

Luego, estimamos dos modelos multiecuacionales: VAR y FAVAR. Para la estimación del VAR, primero seleccionamos el número de rezagos que minimiza el criterio de *Final Prediction Error* de Akaike. Al realizar esto nos queda un VAR(6). En la Columna (5) de la Tabla (2) se presentan los resultados de la estimación de este modelo en el período *in-sample*. Encontramos que los rezagos que influyen a la hora de explicar el sentimiento de Alberto Fernández, es decir, nuestra variable de interés, son el primer, cuarto, quinto y sexto rezago del sentimiento del presidente, el primer rezago de la cantidad de *retweets*, el segundo rezago de la temperatura máxima y el tercer rezago de la cantidad de casos de COVID-19 en Argentina.

Para estimar el modelo FAVAR, aplicamos la técnica de componentes principales a nuestras variables explicativas. Seguidamente, seleccionamos la cantidad de componentes que utilizamos según el valor de los autovalores. Para esta selección, hicimos uso del criterio que se presenta en el trabajo de [Stock y Watson \(2012\)](#). Dichos autores plantean que se eligen los componentes principales cuyos autovalores sean mayores a uno. Esto es porque un autovalor mayor que uno implica que ese componente principal explica más varianza que el promedio. Por lo tanto, elegimos los primeros cuatro componentes principales. Esto se puede ver en la Figura (2), en la que el “codo” se encuentra en el cuarto componente principal.

Posteriormente, lo que hicimos fue estimar un modelo VAR con el sentimiento de Alberto Fernández y los cuatro componentes principales seleccionados. El orden de rezagos seleccionado es aquel que

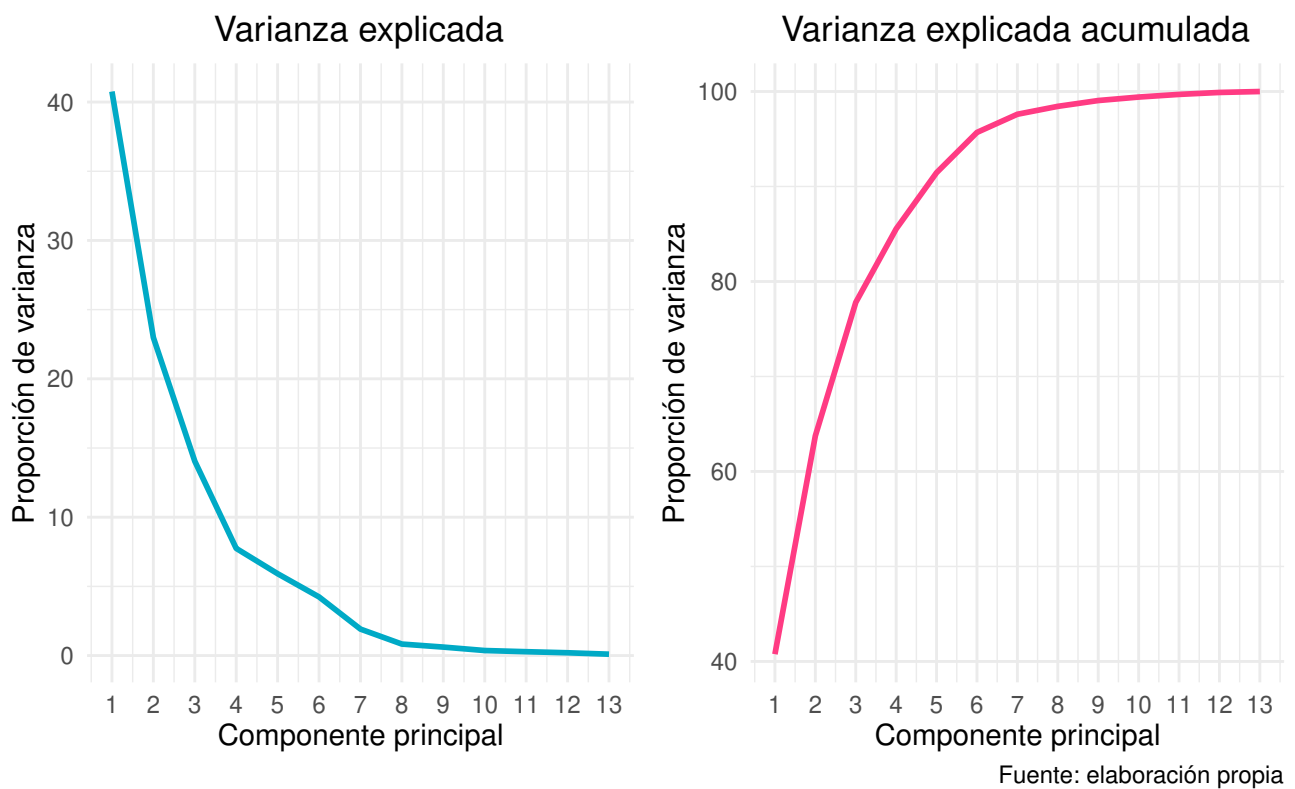


Figura 2: Varianza explicada por los componentes principales

minimiza el criterio de *Final Prediction Error* de Akaike, es decir, un VAR(8). En la última columna de la Tabla (2) se pueden observar los resultados de la estimación de este modelo. Encontramos que son significativos el primer, cuarto y quinto rezago de la variable del sentimiento de Alberto Fernández. También lo son el primer rezago del segundo componente principal y el sexto rezago del cuarto componente principal.

Nuevamente, con el objetivo de validar los modelos, aplicamos el *test* de Portmanteau asintótico a los residuos del VAR y FAVAR estimados. Al 1 % no rechazamos la hipótesis de no autocorrelación de los errores para el modelo FAVAR estimado. Sin embargo, sí rechazamos a todos los niveles comunes de significatividad con el modelo VAR. Si bien no podemos validar mediante el *test* anteriormente nombrado al modelo VAR, decidimos utilizarlo para la realización de pronósticos ya que, tal como se explicó anteriormente, con datos diarios es difícil no rechazar la hipótesis nula de no autocorrelación de los residuos.

Es útil notar que en la estimación del modelo VAR observamos cierta influencia en la dinámica del sentimiento del presidente de algunos rezagos de determinadas covariables. Por lo tanto, considera-

mos relevante utilizar diversas técnicas con el objetivo de poder interpretar en un sentido económico el modelo estimado.

En primer lugar, planteamos las funciones de impulso-respuesta (IRF) para observar de qué forma afectan los *shocks* de los predictores a nuestra variable de interés. En las Figuras (3) y (4) presentamos solo las funciones que revelan un efecto de un *shock* de un desvío estándar en una covariable, en el sentimiento del presidente. No incluimos el resto de los gráficos de las funciones de impulso-respuesta debido a que nuestra variable de interés no presenta reacción alguna antes los *shocks* simulados.

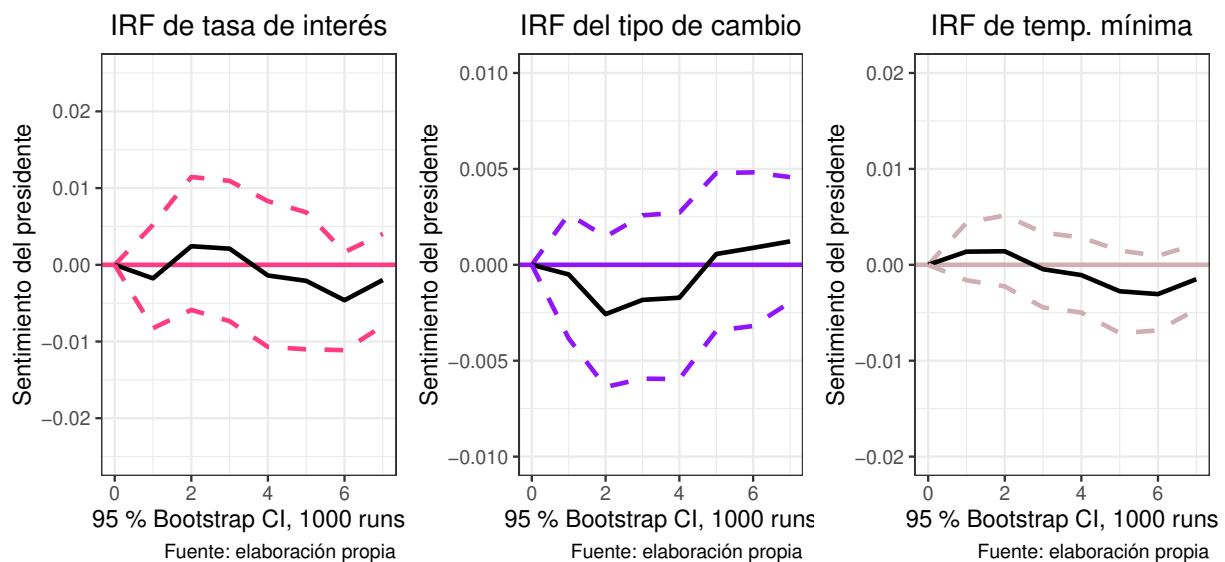


Figura 3: Funciones de impulso-respuesta I

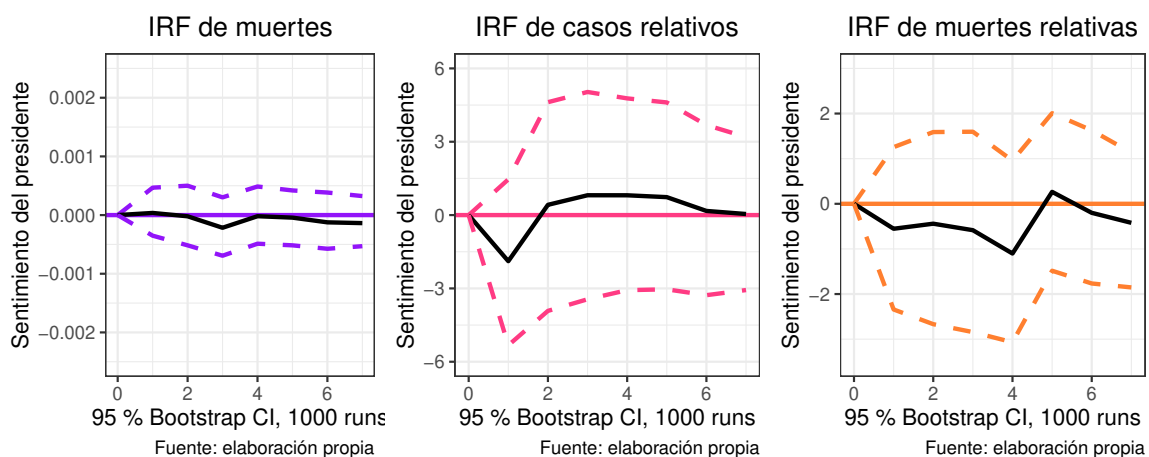


Figura 4: Funciones de impulso-respuesta II

Encontramos que *shocks* en la cantidad de muertes por COVID-19 en Argentina, la cantidad de casos de COVID-19 relativa al resto del mundo, la cantidad de muertes por COVID-19 relativa al resto del mundo, la tasa de interés, el tipo de cambio y la temperatura mínima afectan el sentimiento de Alberto Fernández.

En particular, en la función impulso-respuesta de la tasa de interés (Figura 3) podemos observar que un *shock* en esta variable en cierto momento tiene un efecto negativo en el sentimiento del presidente, con el paso del tiempo este efecto se vuelve positivo (entre el tercer y cuarto día) y, luego, vuelve a caer. Estos resultados tienen sentido debido a que podemos pensar que un aumento en la tasa de interés puede reflejar una política del BCRA que tenga como objetivo mantener una rentabilidad positiva de los ahorristas frente a la suba de precios esperada y por lo tanto, contener la migración de depósitos hacia el dólar. De esta forma, el sentimiento positivo del presidente puede verse afectado de manera negativa por este tipo de *shock* debido a que representa un peor escenario económico.

Con respecto a la función impulso-respuesta del tipo de cambio Figura (3), lo que sucede es que, ante un *shock* en esta variable, el sentimiento del presidente cae hasta aproximadamente el cuarto día y luego comienza a subir. Estos resultados parecen razonables, ya que, que un aumento en el tipo de cambio genera un peor escenario económico para, por ejemplo, los importadores. Por lo tanto, puede suceder que el sentimiento positivo de Alberto Fernández se vea afectado por esta situación.

También, se observa que *shocks* en la temperatura mínima afectan positivamente el sentimiento del presidente. Este efecto comienza a ser negativo, aproximadamente, en el tercer período (Figura 3). Si bien los efectos son pequeños, es interesante destacar que estos resultados son consistentes con los encontrados por (Keller et al., 2005), quienes hayan que las personas tienen un mejor estado de ánimo cuando las temperaturas son más cálidas.

Al observar la función de impulso-respuesta de la cantidad de muertes por COVID-19 (Figura 4) podemos identificar que parece tener un leve efecto negativo en el sentimiento de Alberto Fernández. Este se mantiene durante, aproximadamente, los 7 días considerados en el análisis. Esto puede tener

sentido debido a que un aumento en la cantidad de muertes por COVID-19 refleja una peor situación epidemiológica, lo cual podría impactar en el sentimiento positivo del presidente.

Luego, si examinamos la función de impulso-respuesta de la cantidad de casos de COVID-19 relativo al resto del mundo (Figura 4), notamos que un *shock* en esta variable afecta negativamente al sentimiento del presidente. Este efecto negativo se mantiene hasta, aproximadamente, el período 2. A partir de ese momento, parece aumentar y disminuir de manera leve hasta estabilizarse en los valores iniciales. Nuevamente, esto puede ser razonable ya a que un aumento en la cantidad de casos de COVID-19 respecto del resto del mundo refleja una peor situación epidemiológica de Argentina en relación con los demás países, lo que podría afectar negativamente el sentimiento del presidente.

Por último, si observamos la cantidad de muertes por COVID-19 relativas al resto del mundo (Figura 4) notamos que un *shock* en esta variable afecta negativamente al sentimiento de Alberto Fernández. A partir del cuarto día parece que esta última variable aumenta, pero luego vuelve a bajar. Estos resultados tienen sentido debido al mecanismo explicado para el caso de la cantidad de casos de COVID-19 relativos al resto del mundo.

Para completar el análisis del modelo VAR, decidimos realizar una descomposición de la varianza. Al utilizar esta técnica, es posible observar la proporción de los movimientos de las variables endógenas del sistema que se deben a *shocks* propios o de otras variables. Presentamos el resultado de esta descomposición en la Figura (5).

En la Figura (5) podemos observar que un gran porcentaje de la varianza del sentimiento de Alberto Fernández es explicada por sí misma. Además, notamos que esta proporción va decreciendo levemente con el paso del tiempo y que sucede lo contrario con las demás variables consideradas. Aunque debido a que el porcentaje de la varianza del sentimiento del presidente explicado por cada una de estas variables es muy pequeño, no es posible distinguir cambios si realizamos un análisis visual del gráfico.

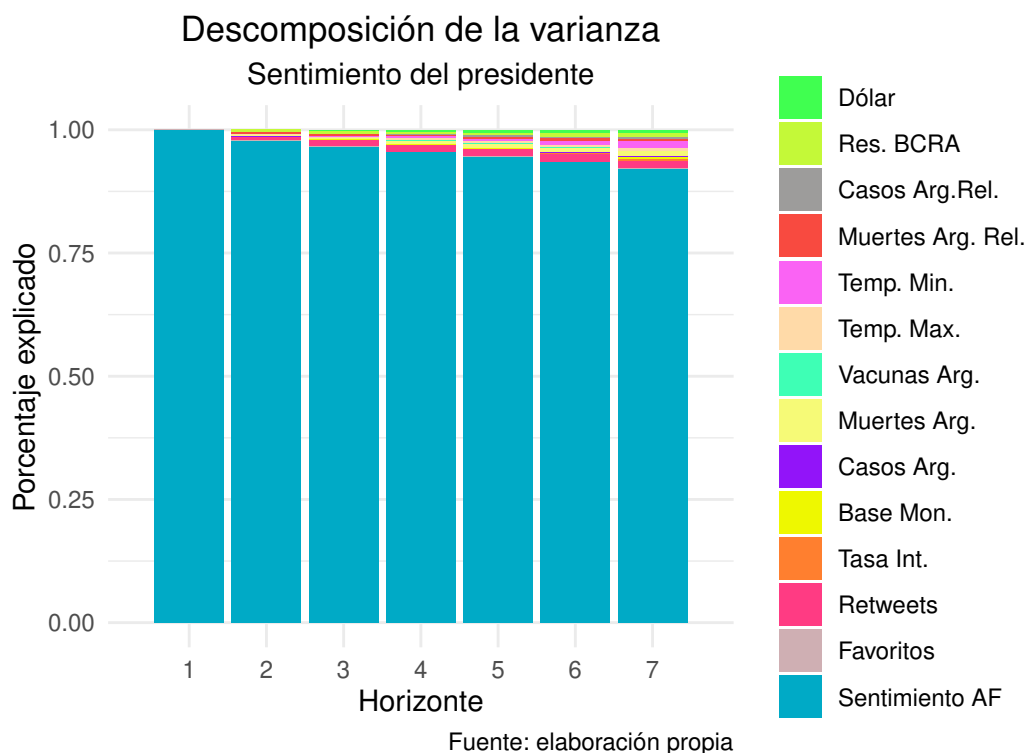


Figura 5: Descomposición de la varianza

5. Resultados

Para realizar los pronósticos del sentimiento de Alberto Fernández, utilizaremos todas las series en niveles, excepto en los casos en los que el modelo requiera una diferenciación de las variables que no son estacionarias (Tipo de cambio y Reservas del BCRA). Los modelos que se utilizaron para pronosticar son el ARIMA, ARIMAX, ADL, ETS, VAR y FAVAR. Además, realizamos los pronósticos con esquemas fijos, recursivos y *rolling* y variamos la cantidad de pasos adelante, a partir de los cuales pronosticamos. En particular, elegimos trabajar con pronósticos uno, dos y siete pasos adelante con el objetivo de evaluar a cuán largo plazo podemos predecir la variable de interés.

También, realizamos pronósticos con los modelos, esquemas y pasos adelante anteriormente nombrados con la técnica de *bagging* (exceptuando el ADL debido a limitaciones computacionales). El objetivo de esto es ver si es posible mejorar nuestros pronósticos remuestrando nuestro período *out-of-sample* utilizando mil muestras *bootstrap* y tomando el promedio².

²En el cuerpo de este trabajo presentaremos solo los gráficos más relevantes. El resto de los pronósticos se encuentran graficados en el [Apéndice Online](#).

En la Figura (6) observamos los pronósticos realizados con los modelos ARIMA, ARIMAX, ETS, ADL, VAR y FAVAR mediante un esquema fijo y un paso hacia adelante. Notamos que, en general, todos los modelos suelen estimar correctamente los cambios en la pendiente de la serie de interés, aunque se distinguen algunas diferencias en cuanto a su precisión. Específicamente, podemos destacar el caso de los pronósticos realizados a partir de los modelos VAR y ETS. El primero de estos modelos sobreestima algunos aumentos y caídas del sentimiento de Alberto Fernández. En cambio, el segundo parece tener un ajuste muy bueno aunque rezagado. En el resto de los modelos se identifica visualmente un desempeño similar. Estos comentarios aplican a los modelos realizados un paso hacia adelante con esquema recursivo y *rolling*, como se puede ver en las Figuras (30) y (31) en el Apéndice.

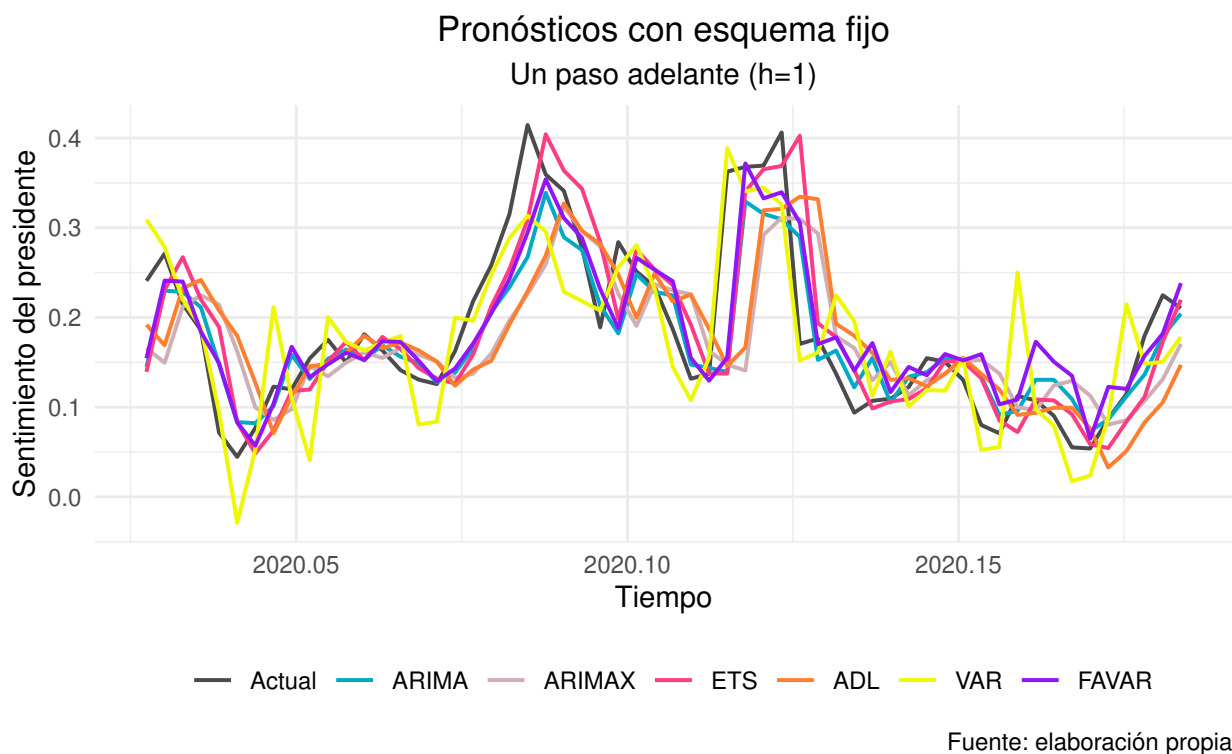


Figura 6: Pronósticos

En la Figura (7) observamos los pronósticos realizados con los modelos ARIMA, ARIMAX, ETS, ADL, VAR y FAVAR mediante un esquema fijo y dos pasos hacia adelante. No podemos identificar ventajas en la utilización de un modelo con respecto a los demás a través de únicamente un análisis visual para la mayoría de los esquemas nombrados anteriormente. Solo los pronósticos realizados a partir del modelo VAR parecen ajustarse en algunos momentos del período *out-of-sample*.

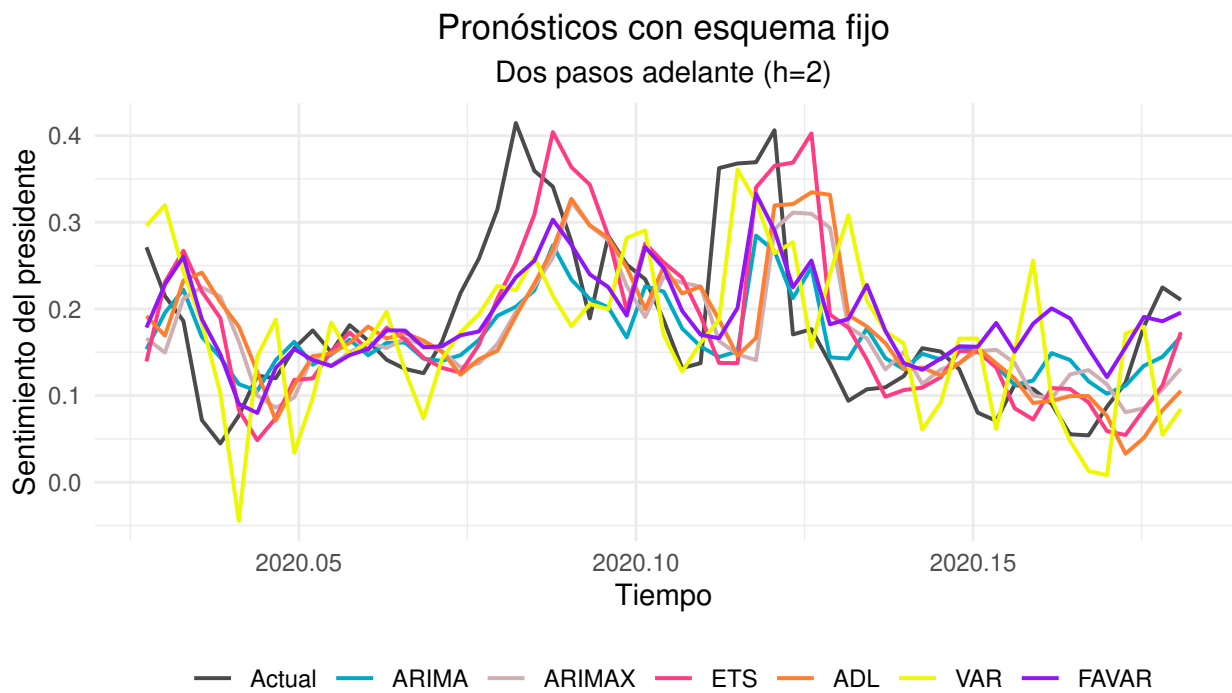


Figura 7: Pronósticos

En la Figura (8) se presentan los pronósticos realizados con los modelos ARIMA, ARIMAX, ETS, ADL, VAR y FAVAR mediante un esquema fijo y siete pasos hacia adelante. A partir de un análisis visual no podemos identificar ventajas en la utilización de un modelo con respecto a los demás para ninguno de los esquemas planteados anteriormente, aunque podemos decir que todos los modelos parecen tener un ajuste no muy bueno. A pesar de esto, podemos notar que, por ejemplo, el modelo ETS pronostica de manera correcta, pero de forma rezagada.

En resumen, al analizar los gráficos de los pronósticos podemos decir que surgen claras desventajas al aumentar la cantidad de pasos adelante de los pronósticos. El hecho de que los pronósticos con dos pasos hacia adelante empeoren puede dar indicios sobre la importancia de la observación inmediatamente anterior a la que se quiere pronosticar (esto se ve en la Tabla (2), ya que el primer rezago es significativo para todos los modelos). La misma conclusión es posible de extender para el caso de siete pasos hacia adelante.

A continuación presentamos la versión *bagged* de la mayoría de los pronósticos analizados anteriormente. Recordemos que la aplicación de esta técnica se realizó con mil replicaciones. En la Figura

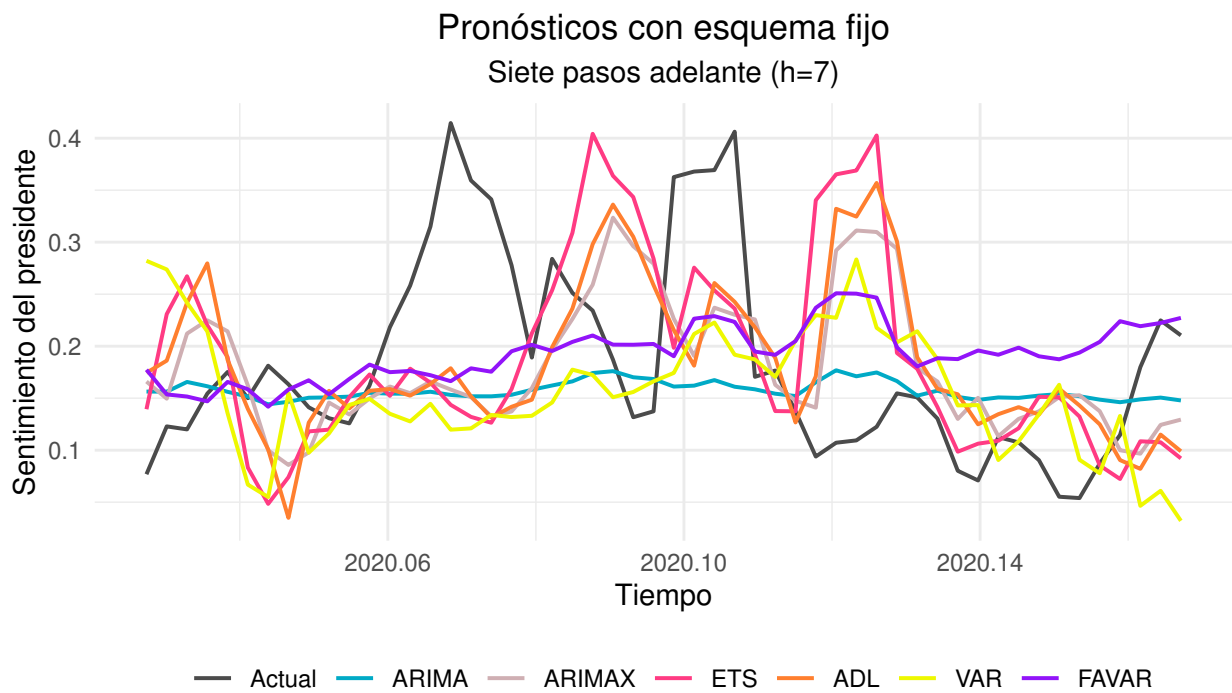


Figura 8: Pronósticos

(9) podemos observar que aquellos pronósticos realizados con un esquema fijo, recursivo y *rolling* un paso hacia adelante. En este caso no se observa un modelo que pronostique correctamente en niveles a lo largo de todo el período *out-of-sample*. Sin embargo, el modelo ETS es el que presenta cierta ventaja relativa al resto de los pronósticos.

Asimismo, podemos analizar la versión *bagged* de los pronósticos realizados con un esquema fijo, recursivo y *rolling* dos pasos hacia adelante. Tal como sucedía para el caso de la versión tradicional de los pronósticos, observamos que al aumentar la cantidad de pasos adelante estos empeoran notablemente. Estas conclusiones se extiende para el caso de los pronósticos realizados siete pasos hacia adelante.

Estos resultados nos muestran la importancia de las observaciones cercanas al valor del sentimiento de Alberto Fernández que se busca pronosticar. Por lo tanto, a priori parece no haber una ventaja por aumentar la cantidad de pasos adelante de los pronósticos.

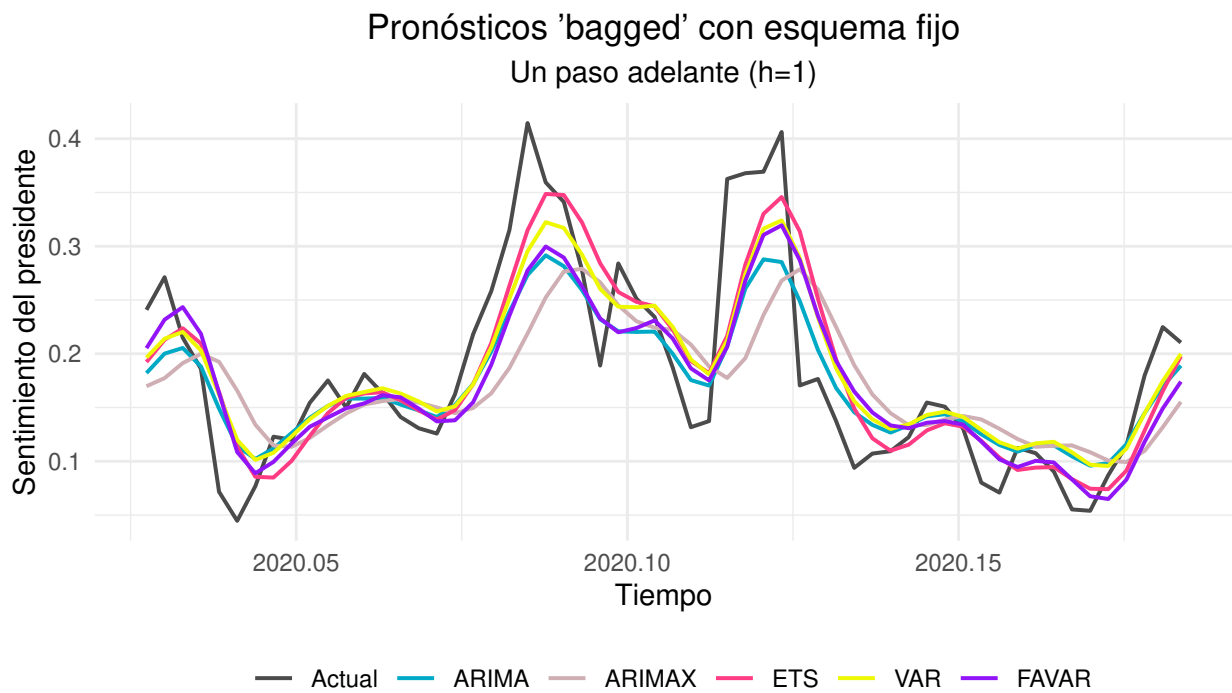


Figura 9: Pronósticos

Resulta pertinente destacar que todas las conclusiones anteriores resultaron solo de una inspección visual de los gráficos de pronósticos. A continuación, evaluaremos el desempeño de los distintos pronósticos realizados. En particular, nos centraremos en tres medidas de desempeño: RMSE, MAE y MAPE. También aplicaremos el *test* de Diebold Mariano a cada pronóstico utilizando como *benchmark* un AR(1).

Para analizar las medidas de *accuracy* dividiremos los pronósticos según la cantidad de pasos adelante que se utilizan. Es útil notar que las medidas de *accuracy* presentes en las Tablas (3), (5), (6) no son comparables entre sí debido a que tienen diferente cantidad de observaciones. Por lo tanto, para poder comparar estas medidas entre todos los modelos, deberíamos eliminar las primeras seis observaciones de los pronósticos que son realizados un paso hacia adelante y las primeras cinco observaciones de los pronósticos que son realizados dos pasos hacia adelante. Los resultados de este ajuste se encuentran en las Tablas (7), (8) y (9).

A continuación, presentamos en la Tabla (3) las medidas de *accuracy* de todos los pronósticos realizados un paso hacia adelante junto con los resultados principales del *test* de Diebold-Mariano. Notemos

que al considerar la raíz del error cuadrático medio (RMSE) podemos decir que el modelo que minimiza esta medida es el ETS *bagged* con esquema *rolling*. Además, 21 de los 32 modelos utilizados generan pronósticos que tienen un valor de dicha medida de desempeño menor al *benchmark*.

Modelo	RMSE	MAE	MAPE	DM	p-val
AR(1)	0.0588	0.0405	25.7162		
ARIMA fijo	0.0541	0.0371	22.9074	-1.8375	0.0713
ARIMAX fijo	0.0787	0.0595	39.896	2.8246	0.0065
ETS fijo	0.0614	0.0427	27.1503	0.609	0.5449
ADL fijo	0.08	0.0612	41.0544	3.2779	0.0018
VAR fijo	0.0509	0.0393	28.1487	-0.7394	0.4627
FAVAR fijo	0.0537	0.0391	27.1415	-1.7819	0.0801
ARIMA recursivo	0.0529	0.0369	23.1069	-2.2371	0.0292
ARIMAX recursivo	0.0779	0.0587	39.5463	3.0967	0.003
ETS recursivo	0.0614	0.0424	26.9193	0.586	0.5602
ADL recursivo	0.0789	0.06	39.8063	3.1851	0.0023
VAR recursivo	0.0509	0.0393	28.1487	-0.7394	0.4627
FAVAR recursivo	0.0523	0.0383	27.4639	-1.6451	0.1055
ARIMA rolling	0.0528	0.0349	21.3249	-2.3198	0.024
ARIMAX rolling	0.076	0.0574	38.8969	2.5375	0.0139
ETS rolling	0.0609	0.0416	26.165	0.4507	0.6539
ADL rolling	0.0771	0.0568	38.4779	3.1532	0.0026
VAR rolling	0.0483	0.0387	27.8408	-1.0005	0.3213
FAVAR rolling	0.052	0.0377	26.6219	-1.8567	0.1036
ARIMA fijo bagged	0.0516	0.0386	24.8783	-1.2374	0.221
ARIMAX fijo bagged	0.0751	0.0578	38.4889	2.6268	0.0111
ETS fijo bagged	0.0491	0.037	24.8465	-1.735	0.0881
VAR fijo bagged	0.0531	0.0415	26.1083	-1.6895	0.0966
FAVAR fijo bagged	0.0506	0.0395	28.4818	-1.1396	0.2592
ARIMA recursivo bagged	0.0506	0.0386	25.2501	-1.4264	0.1592
ARIMAX recursivo bagged	0.0753	0.0586	39.4242	2.7825	0.0073
ETS recursivo bagged	0.0485	0.0364	24.4805	-1.83	0.0725
VAR recursivo bagged	0.052	0.041	26.1257	-1.4439	0.1542
FAVAR recursivo bagged	0.0506	0.0395	28.4818	-1.2808	0.2055
ARIMA rolling bagged	0.049	0.0366	23.5829	-1.6541	0.1036
ARIMAX rolling bagged	0.0748	0.0579	38.7858	2.6021	0.0118
ETS rolling bagged	0.0467	0.0349	23.454	-2.0474	0.0452
VAR rolling bagged	0.0539	0.0424	25.9851	-1.3984	0.1674
FAVAR rolling bagged	0.051	0.0399	28.2342	-0.9734	0.3345

Cuadro 3: Medidas de desempeño - Un paso adelante

Por otro lado, si consideramos el MAE (error medio absoluto), podemos decir que los pronósticos que minimizan esta medida de desempeño son los realizados con el modelo ETS *bagged rolling* y ARIMA *rolling*. En este caso, 18 de los 32 modelos utilizados generan pronósticos que tienen un MAE menor

al *benchmark*. Al tomar en cuenta el MAPE (error medio absoluto porcentual), los pronósticos que minimizan esta medida de desempeño son los realizados con el ARIMA *rolling*. Además, 9 de 32 modelos utilizados tienen un desempeño mejor al del *benchmark*.

Ahora bien, es útil evaluar si las diferencias de desempeño identificadas anteriormente son estadísticamente significativas. Para esto aplicamos el *test* de Diebold-Mariano. En particular, aplicamos dicho *test* a los pronósticos realizados un paso hacia adelante con los modelos estimados a lo largo del trabajo contra el AR(1), es decir, contra el *benchmark*. Los resultados se presentan en la Columna (6) de la Tabla (3). Notemos que debido a la forma en la que realizamos este *test*, los valores positivos del estadístico reflejan un mejor desempeño del *benchmark* y los valores negativos son evidencia de lo contrario.

A partir de los resultados del *test* de Diebold-Mariano, podemos decir que 7 de los 32 modelos considerados tienen un desempeño superior al *benchmark*. Para los modelos ARIMA fijo, FAVAR fijo, ETS *bagged* fijo, VAR fijo *bagged* y ETS recursivo *bagged*, la diferencia resulta estadísticamente significativa al 10 %. En cambio, podemos decir que el ARIMA *rolling*, ARIMA recursivo y ETS *bagged rolling* son superadores con una significatividad estadística del 5 %.

De la misma manera calculamos las medidas de desempeño de los pronósticos realizados dos y siete pasos adelante con los diferentes modelos estimados, cuyos resultados están disponibles en las Tablas (5) y (6) en el Apéndice. En estos casos se observa que el modelo AR(1) tiene un desempeño estadísticamente superior a casi la totalidad de los modelos. Tal como se dijo en el análisis gráfico de los pronósticos, podemos inferir que esto puede deberse a que, para realizar los pronósticos de nuestra variable de interés, son muy importantes las observaciones anteriores al día que se quiere pronosticar.

Es útil recordar que las medidas de *accuracy* evalúan el desempeño medio a lo largo de todo el período *out-of-sample*, pero podría suceder que la *performance* relativa cambie a lo largo del tiempo. Esto se puede testear mediante el Fluctuation Test (Giacomini y Rossi, 2010). Consideramos interesante aplicarle este test a los pronósticos realizados uno, dos y cuatro pasos adelante con los diferentes

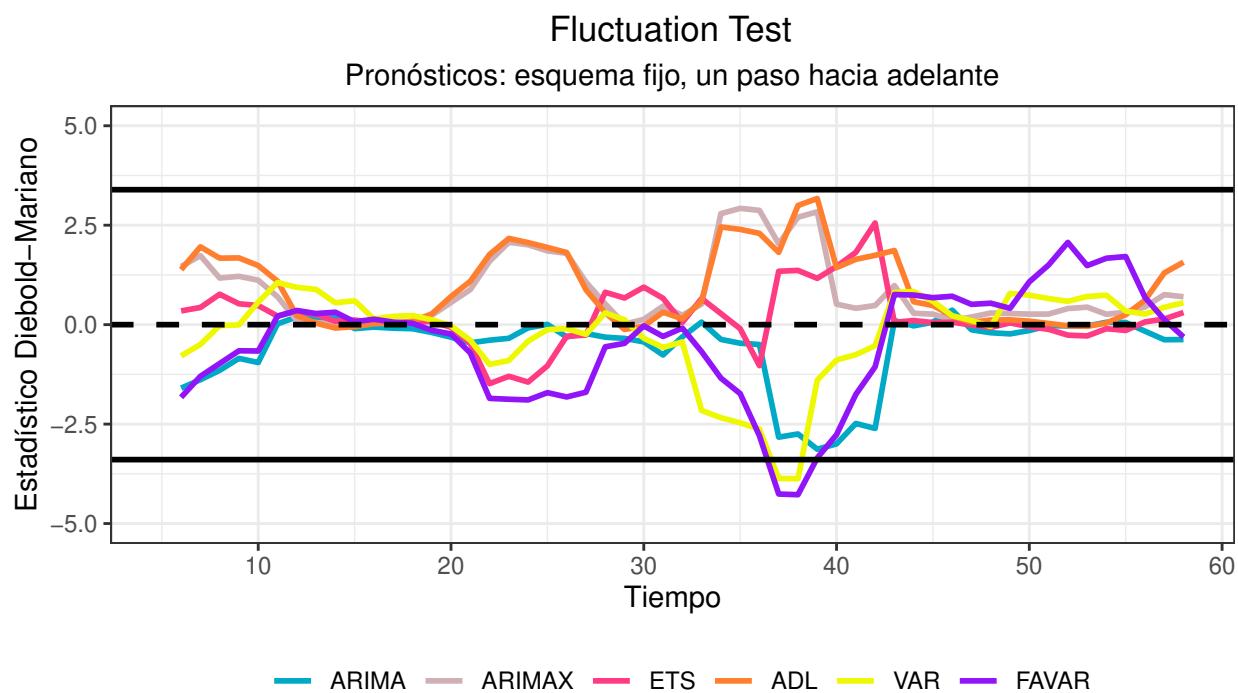
modelos utilizando, nuevamente, como *benchmark* un AR(1). A continuación se presentan algunos gráficos³ que resultan de la aplicación de dicho *test* sobre los pronósticos realizados con modelos fijos uno, dos y siete pasos adelante. Cabe destacar que, dada la forma en la que realizamos este *test*, los valores positivos del estadístico indican un mejor desempeño de los pronósticos que resultan de la utilización del *benchmark*. Presentamos los intervalos de confianza al 95 %.

La Figura (10) muestra la aplicación del test de Giacomini-Rossi sobre los pronósticos realizados con esquema fijo y un paso hacia adelante. En este sentido, encontramos que los modelos VAR y FAVAR resultan superadores al modelo AR(1), aproximadamente, entre las observaciones número 36 y 38 del período *out-of-sample*.

Estas conclusiones son similares al aplicar el *test* de Giacomini-Rossi a los pronósticos realizados con los modelos que cuentan con esquema recursivo y *rolling*. Es importante destacar que para la realización de este *test* elegimos un m igual a 5, es decir, este es el tamaño de la ventana *rolling* a partir de la cual se realiza el *test* de Diebold-Mariano. Por lo tanto, dado que el tamaño de nuestra ventana *out-of-sample* es de 58, el μ que resulta es 0.10, siguiendo lo planteado por [Giacomini y Rossi \(2010\)](#). También, aplicamos el *test* de Giacomini-Rossi sobre los pronósticos *bagged* realizados con esquema fijo (Figura 11). Notamos que para el caso de este tipo de pronósticos ningún modelo resulta superador al AR(1).

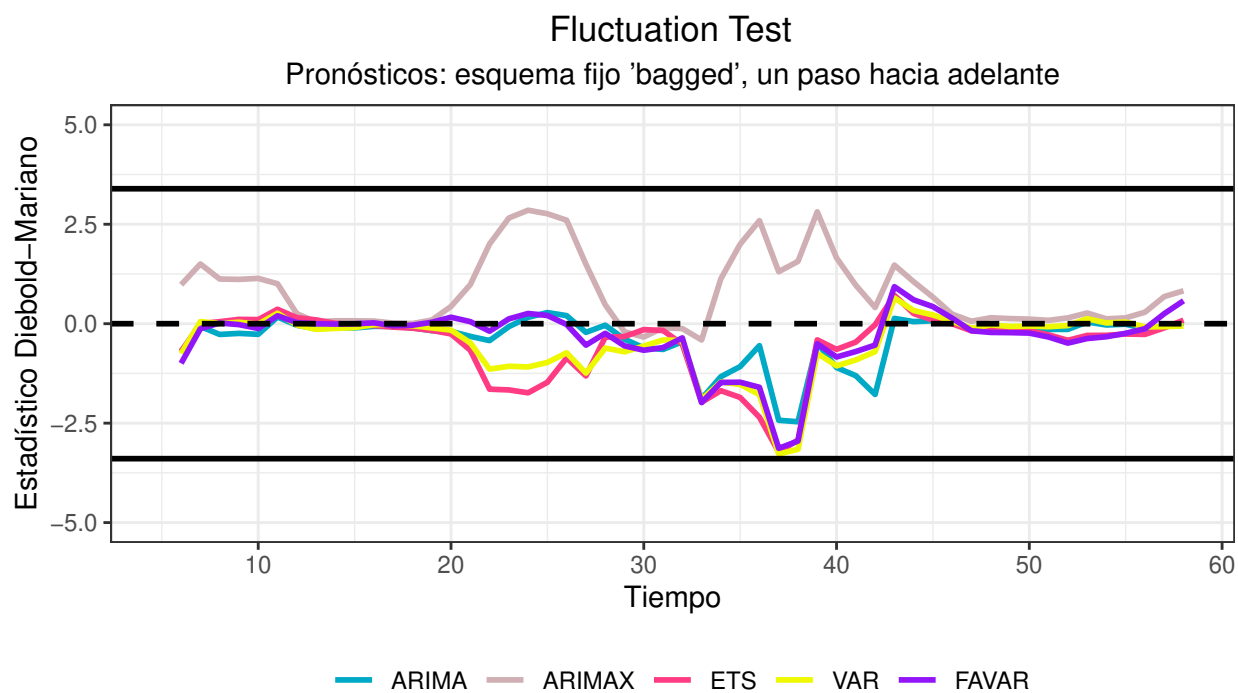
La Figura (12) muestra la aplicación del *test* de Giacomini-Rossi sobre los pronósticos realizados con esquema fijo y dos pasos hacia adelante. A partir de este gráfico, podemos concluir que a lo largo de todo el período *out-of-sample* ningún modelo resulta superador al AR(1). En particular, podemos decir que el modelo AR(1) tiene un desempeño mejor y estadísticamente significativo que los modelos ETS, ARIMAX y ADL entre las observaciones número 30 y 40. Además, realizamos este tipo de análisis con los esquemas esquemas fijos y recursivos dos pasos hacia adelante, así como con la versión *bagged* de cada uno de estos pronósticos (esquema fijo, recursivo y *rolling*). Los gráficos se pueden observar en Figuras (32), (33), (34), (35), (36). Las conclusiones que surgen a partir del análisis de estos gráficos son similares.

³El resto de los gráficos se encuentran en el [Apéndice Online](#).



Fuente: elaboración propia

Figura 10: Test de Giacomini Rossi (esquema fijo, un paso adelante)



Fuente: elaboración propia

Figura 11: Test de Giacomini Rossi (esquema fijo *bagged*, un paso adelante)

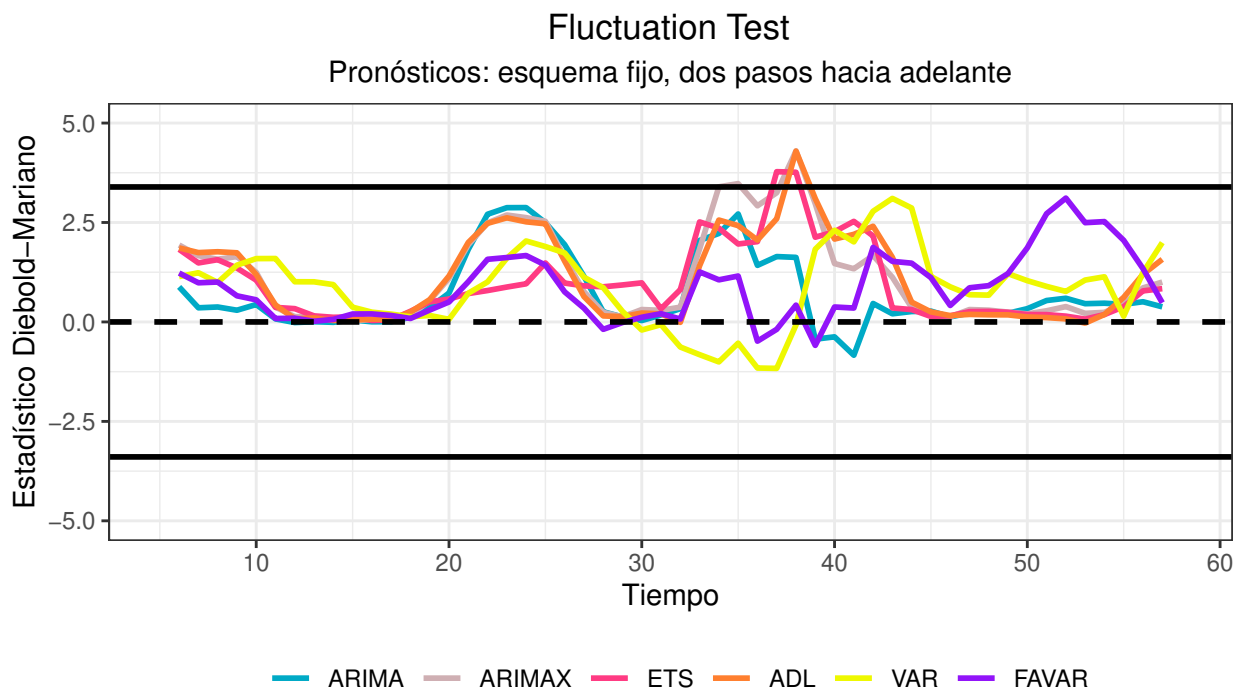


Figura 12: Test de Giacomini Rossi (esquema fijo, dos pasos adelante)

La Figura (13) presenta la aplicación del *test* de Giacomini-Rossi sobre los pronósticos realizados con esquema fijo y siete pasos hacia adelante. A partir de este gráfico, podemos decir que ningún modelo resulta superior al AR(1) a lo largo de todo el período *out-of-sample*. Es útil destacar que el modelo AR(1) tiene un desempeño mejor y estadísticamente significativo al desempeño de todos los modelos considerados (ARIMA, ARIMAX, ETS, ADL, VAR y FAVAR). A la vez, realizamos este tipo de análisis con los pronósticos resultantes de la utilización de esquemas fijos y recursivos siete pasos hacia adelante, así como con la versión *bagged* de cada uno de estos pronósticos (esquema fijo, recursivo y *rolling*). Los gráficos se pueden observar en Figuras (37), (38), (39), (40), (41). Las conclusiones que surgen a partir del análisis de estos gráficos son similares.

Con el objetivo de evaluar si hay ganancias en realizar la técnica de *bagging* sobre los diferentes modelos estimados, decidimos realizar el *test* de Giacomini-Rossi entre cada pronóstico y su versión *bagged*. Para esto, agrupamos los pronósticos en un solo gráfico según el esquema de pronóstico (cantidad de pasos adelante y características de la ventana de estimación). Dada la forma en la que realizamos este *test*, los valores negativos del estadístico reflejan un peor desempeño de la versión

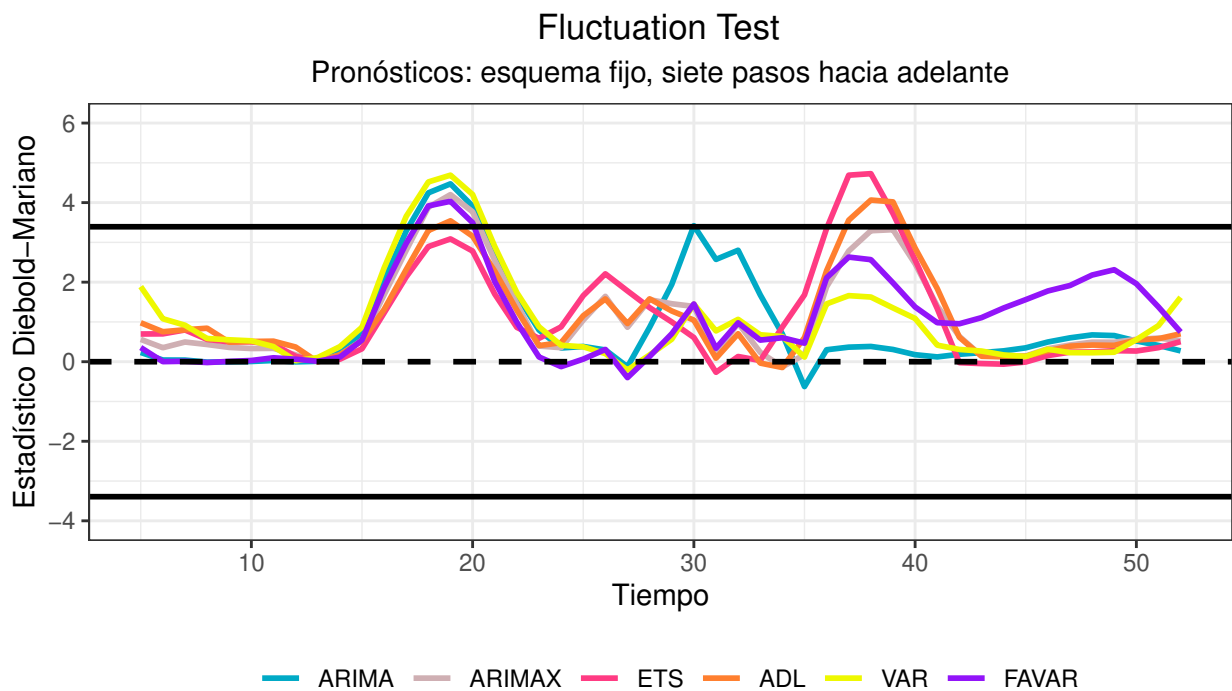


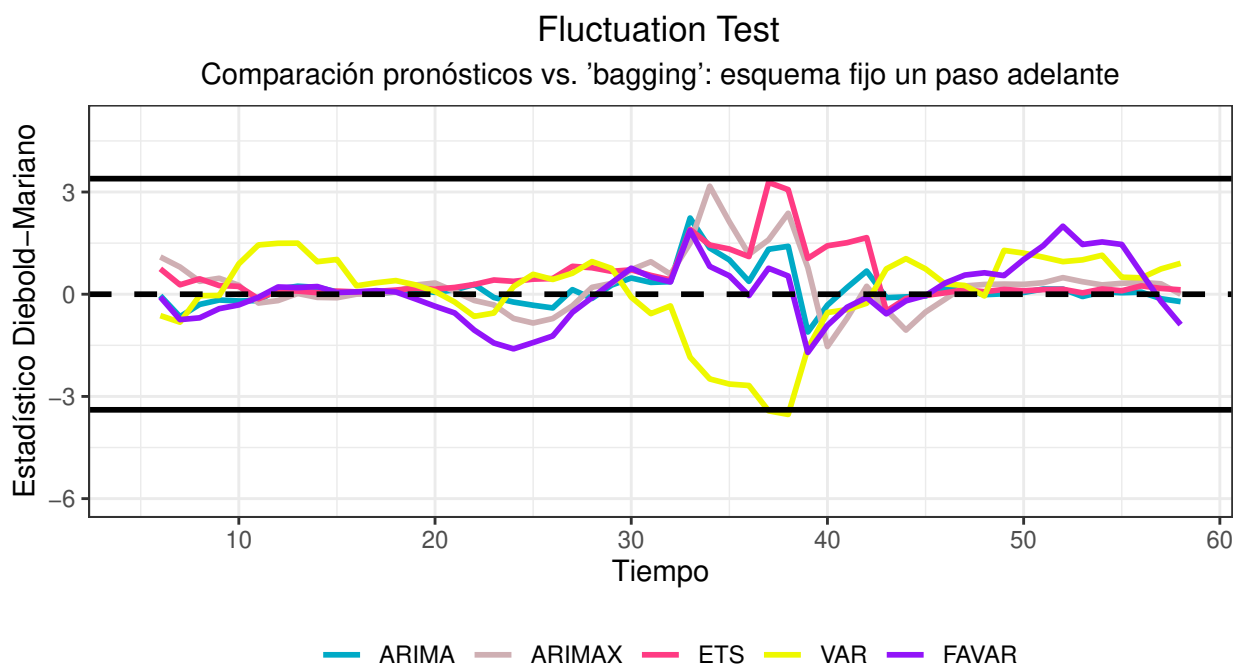
Figura 13: Test de Giacomini Rossi (esquema fijo, siete pasos adelante)

bagged.

La Figura (14) muestra que la versión *bagged* de los pronósticos realizados con el modelo ARIMA, ARIMAX, ETS y FAVAR un paso hacia adelante y con un esquema fijo no resulta superadora a la versión tradicional de cada uno de ellos en ningún momento del período. Para el modelo no mencionado, el VAR, se puede observar que su versión *bagged* tiene aproximadamente dos períodos en los que su rendimiento es peor al del modelo tradicional. En el resto de los períodos, no hay diferencia estadísticamente significativa entre ambas versiones.

Los resultados en los que solo hay un pequeño beneficio de no usar *bagging* en los modelos VAR se replican en los pronósticos con un esquema *rolling* o recursivo (un paso hacia adelante) y su correspondiente versión *bagged*. Los gráficos de estos casos están disponibles en el Apéndice (Figuras 42 y 43).

Ahora bien, cuando consideramos los pronósticos realizados dos pasos adelante, tanto con esquema fijo, *rolling* o recursivo, notamos que, la versión tradicional tiene un desempeño estadísticamente



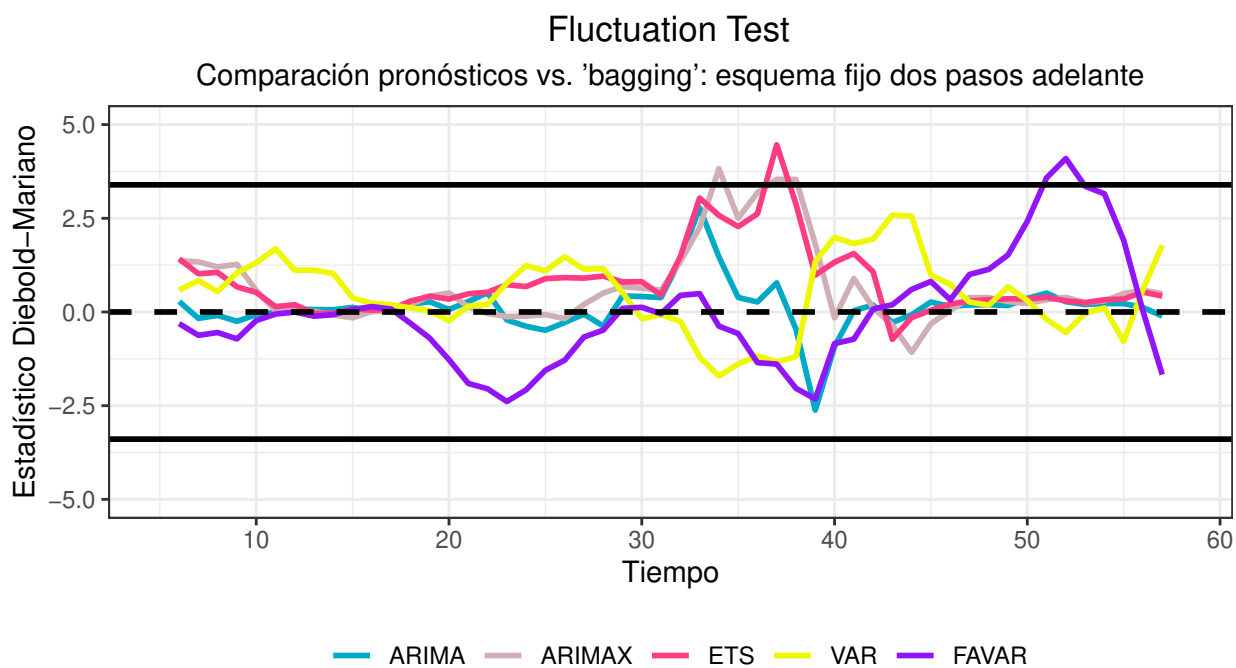
Nota: valores negativos reflejan un peor desempeño de la versión bagging.
Fuente: elaboración propia

Figura 14: Test de Giacomini Rossi (esquema fijo, un paso adelante)

peor a la versión *bagged* para algunos modelos. A continuación, presentamos la Figura (15) que corresponde a la aplicación del *test* de Giacomini-Rossi sobre los pronósticos realizados con un esquema fijo y dos pasos hacia adelante contra su versión *bagged*. Podemos observar que la versión *bagged* de los pronósticos realizados con los modelos ARIMAX, ETS y FAVAR resulta superadora a la versión tradicional.

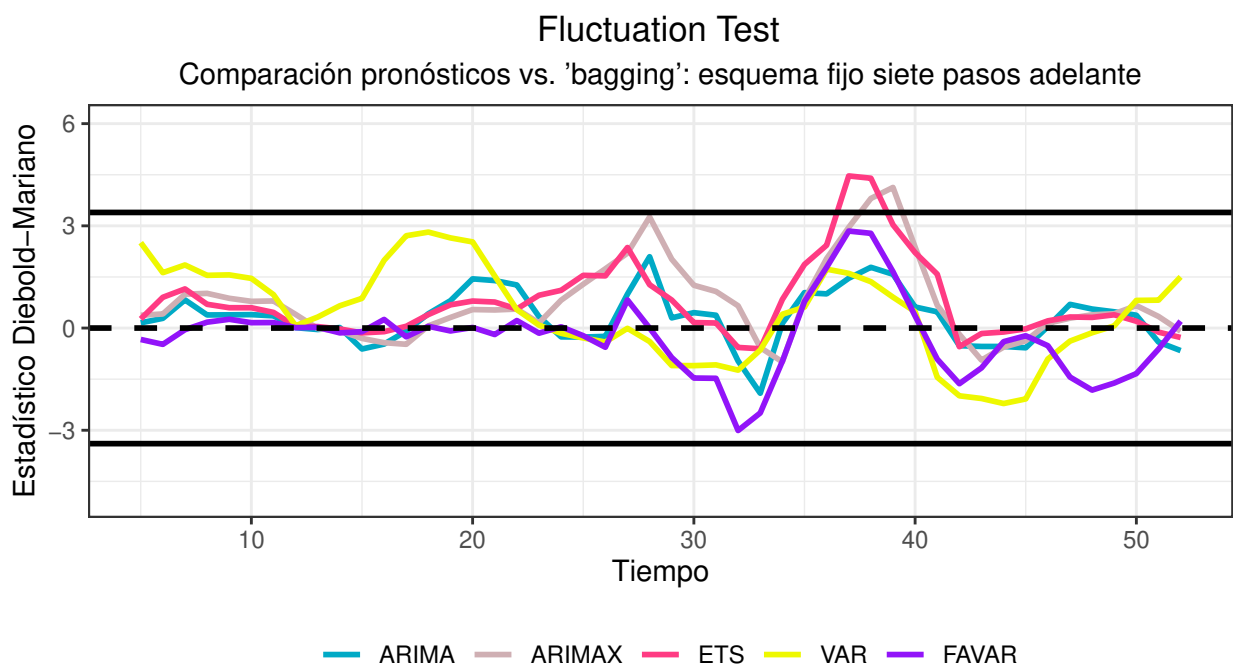
En la Figura (16) aplicamos el *test* de Giacomini-Rossi sobre los pronósticos realizados con esquema fijo siete pasos adelante contra su versión *bagged*. Observamos que en este caso la versión *bagged* de los pronósticos realizados con los modelos ETS y ARIMAX resulta superadora a la versión tradicional. Estos resultados son similares cuando consideramos los pronósticos que resultan de un esquema recursivo o *rolling* (Figuras 44 y 45 en Apéndice).

En resumen, podemos decir que hay ganancia de utilizar la técnica de *bagging* para pronosticar cuando realizamos los pronósticos más de un paso adelante y con determinados modelos. En cambio, cuando realizamos los pronósticos un paso adelante no identificamos una ventaja por utilizar *bagging*



Nota: valores negativos reflejan un peor desempeño de la versión bagging.
Fuente: elaboración propia

Figura 15: Test de Giacomini Rossi (esquema fijo, un paso adelante)



Nota: valores negativos reflejan un peor desempeño de la versión bagging.
Fuente: elaboración propia

Figura 16: Test de Giacomini Rossi (esquema fijo, un paso adelante)

para pronosticar. Estos resultados no varían sustantivamente al utilizar esquemas fijos, recursivos o *rolling*.

6. Conclusiones

A lo largo de este trabajo intentamos pronosticar la evolución del sentimiento en Twitter del presidente Alberto Fernández. Esto lo hicimos planteando distintos modelos incluyendo variables económicas, epidemiológicas y climáticas. Para esto, utilizamos diversos modelos como lo son el ARIMA, ARIMAX, ETS, ADL, VAR y FAVAR. De todos estos modelos hicimos los pronósticos utilizando esquemas fijos, recursivos y *rolling*; uno, dos y siete pasos adelante. A excepción del ADL, repetimos los pronósticos utilizando la técnica *bagging* con mil remuestreos.

Al realizar los pronósticos del período *out-of-sample*, no solo logramos identificar modelos y esquemas superadores sino también claras desventajas en la utilización de otros. En particular, pudimos notar que el aumento de la cantidad de pasos adelante según la cual se realizan los pronósticos no trae consigo ningún beneficio. Además, resulta evidente la imposibilidad de realizar predicciones con un buen ajuste en estos tipos de esquemas. Limitándonos a los pronósticos realizados un paso hacia adelante, podemos decir que no hubo un modelo que se destacó tanto visualmente como al evaluar mediante algunos *tests* su desempeño en comparación al *benchmark* elegido, sino que son el ETS y VAR los que sugieren un desempeño levemente superior.

Tal como se nombró anteriormente, también decidimos incorporar la técnica de *bagging* para realizar los pronósticos. Notamos que en algunos momentos del período *out-of-sample* estos modelos tienen un desempeño estadísticamente peor. Las ventajas de utilización de esta técnica recién aparecen cuando realizamos pronósticos con más pasos adelante. Cabe destacar que, al evaluar algunos modelos en el períodos *in-sample*, llegamos a la conclusión de que había quedado dinámica sin explicar dado que los errores estaban autocorrelacionados. Por lo tanto, remarcamos la importancia de seguir teniendo en cuenta los modelos que tengan estos resultados dada la gran dificultad para incorporar toda la dinámica con datos de esta frecuencia.

Como líneas para futura investigación, creemos que sería interesante entrenar un modelo de predicción del sentimiento y no utilizar uno que ya esté desarrollado. Por ejemplo, se podría entrenar un modelo de predicción utilizando diversos *tweets* que se consideren positivos y otros negativos y, en base a esto, sería relevante analizar si las conclusiones cambian a la hora de pronosticar la serie.

Referencias

- Arias, M., Arratia, A., y Xuriguera, R. (2014). Forecasting with twitter data. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1), 1–24.
- Beigi, G., Hu, X., Maciejewski, R., y Liu, H. (2016). An overview of sentiment analysis in social media and its applications in disaster relief. *Sentiment analysis and ontology engineering*, 313–340.
- Budiharto, W., y Meiliana, M. (2018). Prediction and analysis of indonesia presidential election from twitter using sentiment analysis. *Journal of Big data*, 5(1), 1–10.
- Giacomini, R., y Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, 25(4), 595–620.
- Howarth, E., y Hoffman, M. S. (1984). A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology*, 75(1), 15–23.
- Keller, M. C., Fredrickson, B. L., Ybarra, O., Côté, S., Johnson, K., Mikels, J., ... Wager, T. (2005). A warm heart and a clear head: The contingent effects of weather on mood and cognition. *Psychological science*, 16(9), 724–731.
- Kordonis, J., Symeonidis, S., y Arampatzis, A. (2016). Stock price forecasting via sentiment analysis on Twitter. En *Proceedings of the 20th pan-hellenic conference on informatics* (pp. 1–6).
- Kraaijeveld, O., y De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65, 101188.
- Kranefuss, E., y Johnson, D. K. (2021). Does twitter strengthen volatility forecasts? Evidence from the S&P 500, DJIA and Twitter Sentiment Analysis. *Evidence from the S&P, 500*.

- Liu, R., Yao, X., Guo, C., y Wei, X. (2021). Can we forecast presidential election using Twitter data? an integrative modelling approach. *Annals of GIS*, 27(1), 43–56.
- McCormick, S., y Stuckey, M. (2013). Presidential disfluency: Literacy, legibility, and vocal political aesthetics in the rhetorical presidency. *Review of Communication*, 13(1), 3–22.
- Molteni, L., y De Leon, J. P. (2016). Forecasting with twitter data: an application to usa tv series audience. *International Journal of Design & Nature and Ecodynamics*, 11(3), 220–229.
- Pai, P.-F., y Liu, C.-H. (2018). Predicting vehicle sales by sentiment analysis of Twitter data and stock market values. *IEEE Access*, 6, 57655–57662.
- Rodríguez, S., Allende-Cid, H., Palma, W., Alfaro, R., Gonzalez, C., Elortegui, C., y Santander, P. (2018). Forecasting the chilean electoral year: Using twitter to predict the presidential elections of 2017. En *International conference on social computing and social media* (pp. 298–314).
- Stock, J. H., y Watson, M. W. (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business & Economic Statistics*, 30(4), 481–493.
- Yaqub, U. (2020). Tweeting during the Covid-19 pandemic: Sentiment Analysis of Twitter messages by President Trump. *Digital Government: Research and Practice*, 2(1), 1–7.

7. Apéndice

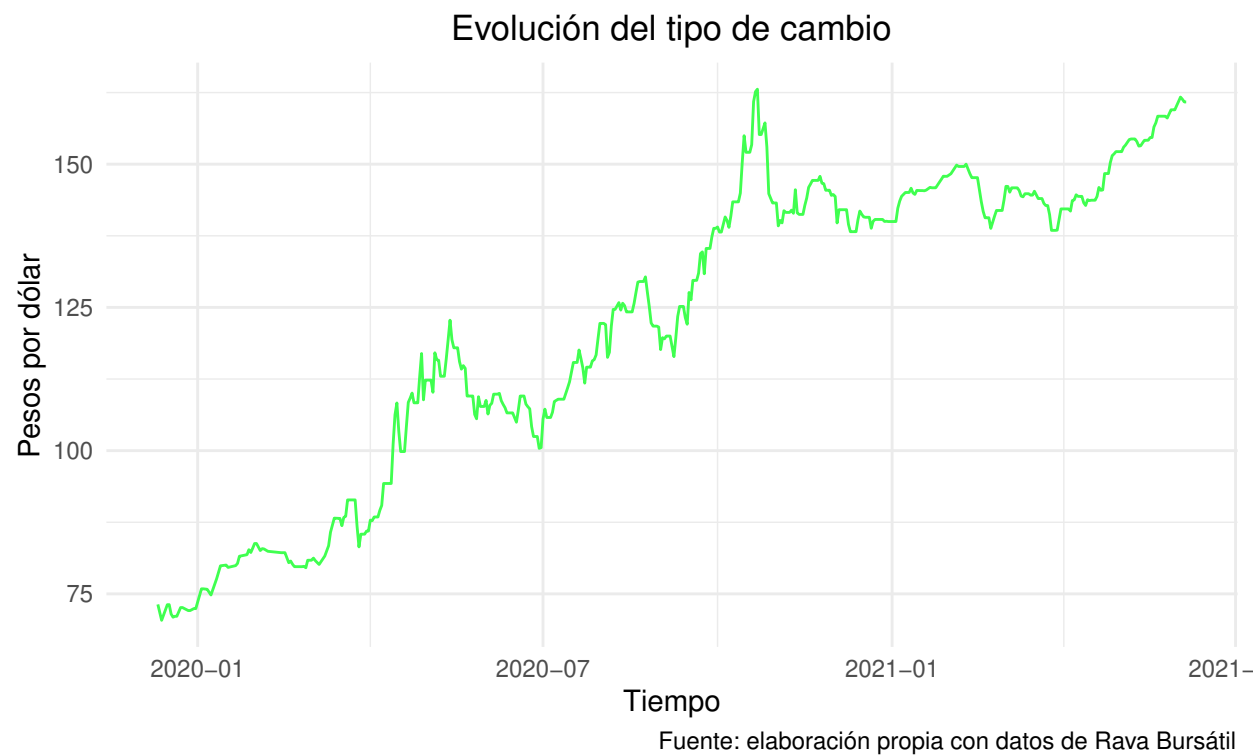


Figura 17: Evolución del tipo de cambio

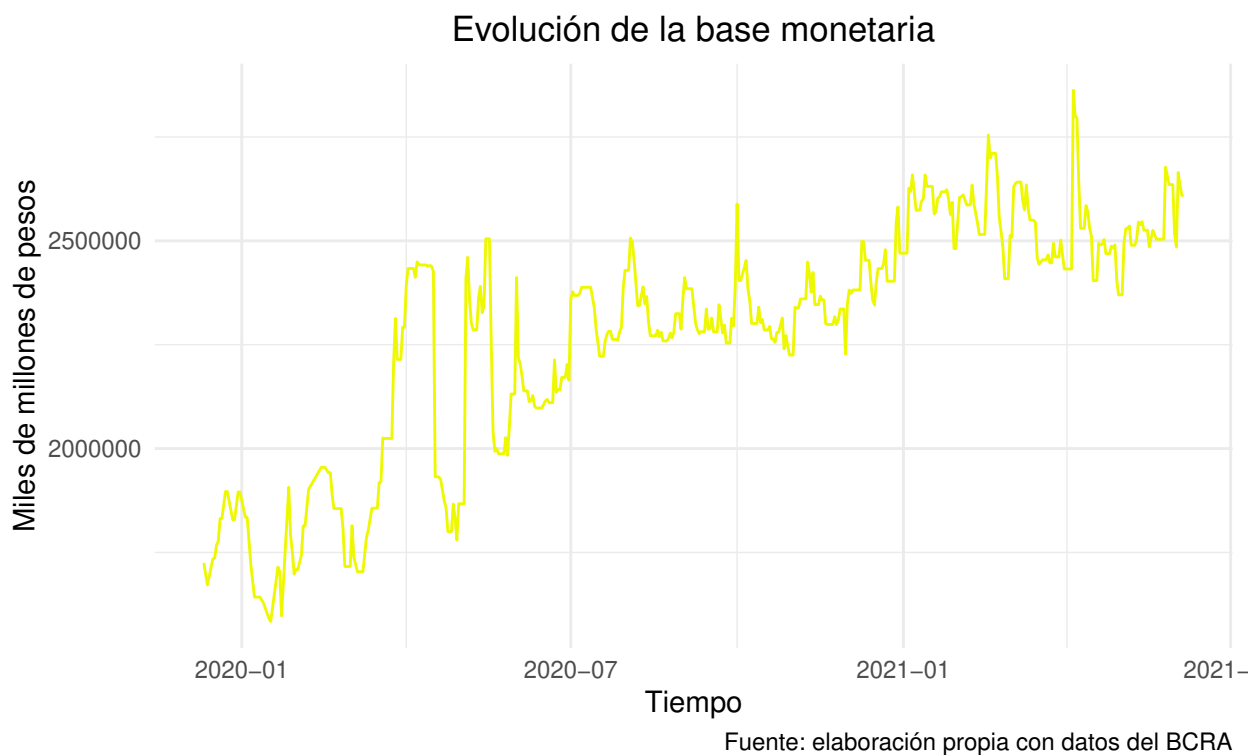


Figura 18: Base monetaria

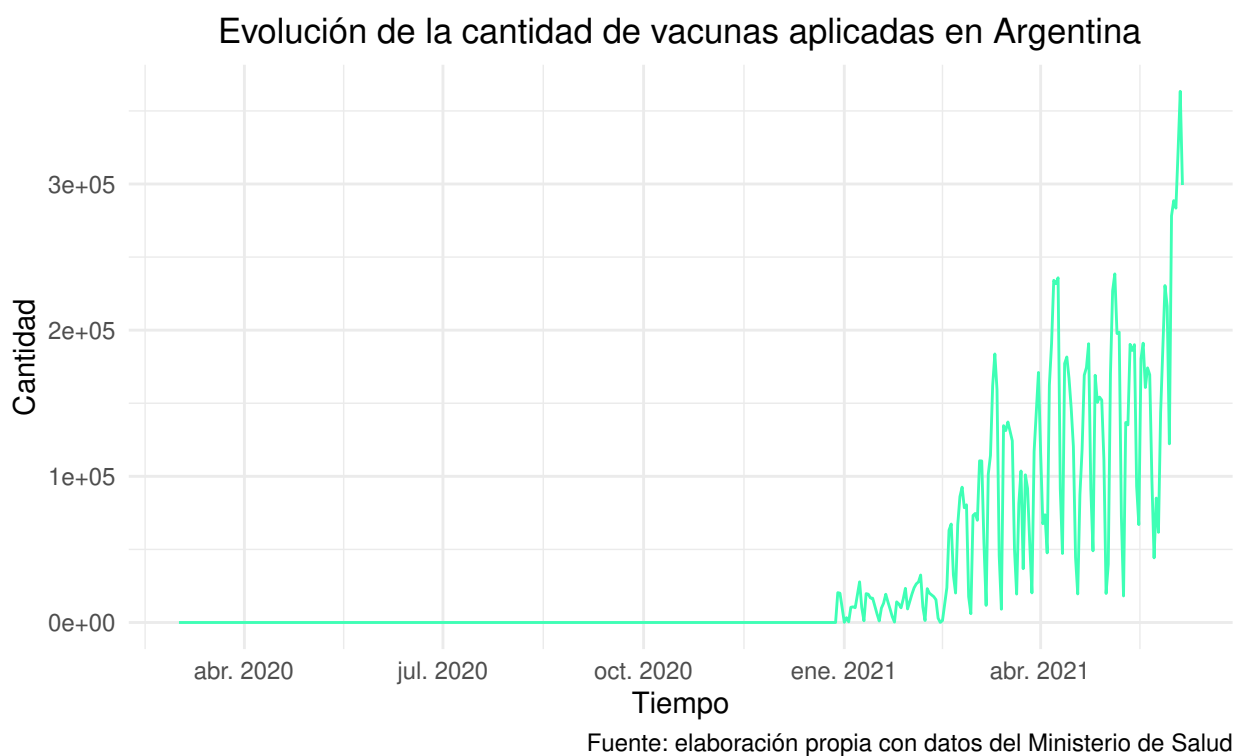


Figura 19: Vacunas aplicadas en Argentina

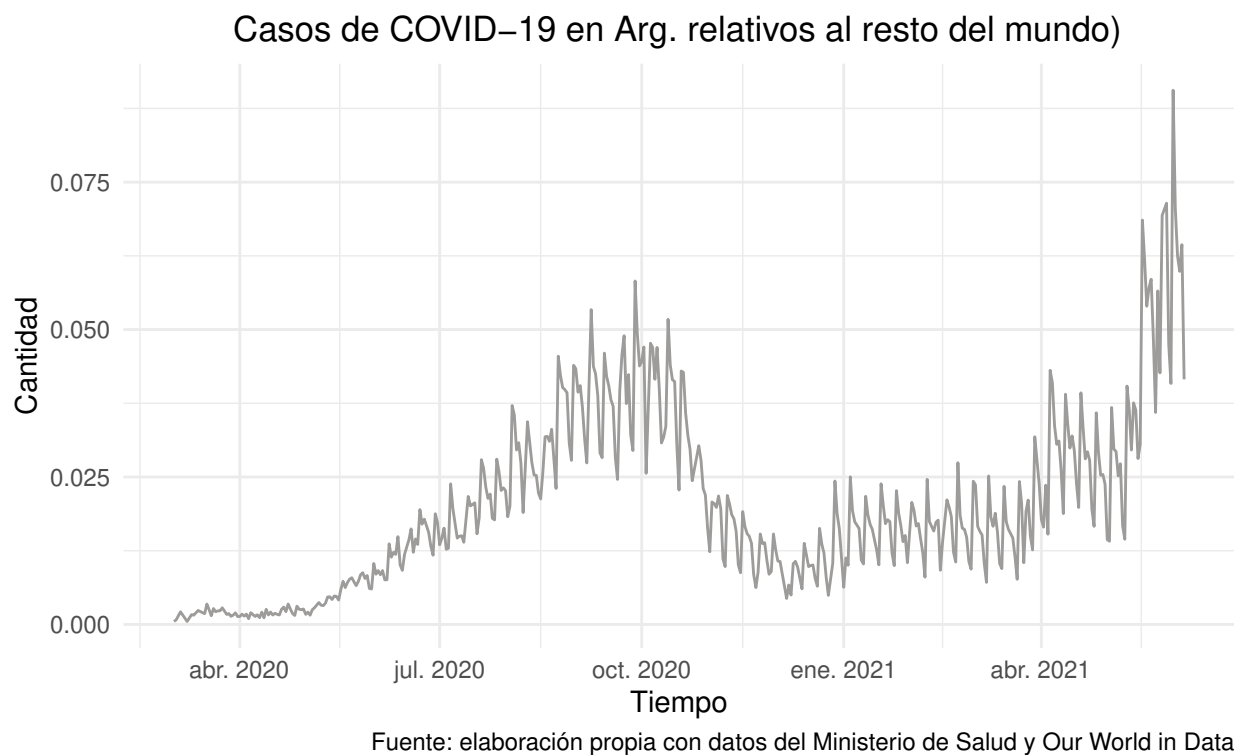


Figura 20: Casos de COVID-19 en Argentina relativos al resto del mundo

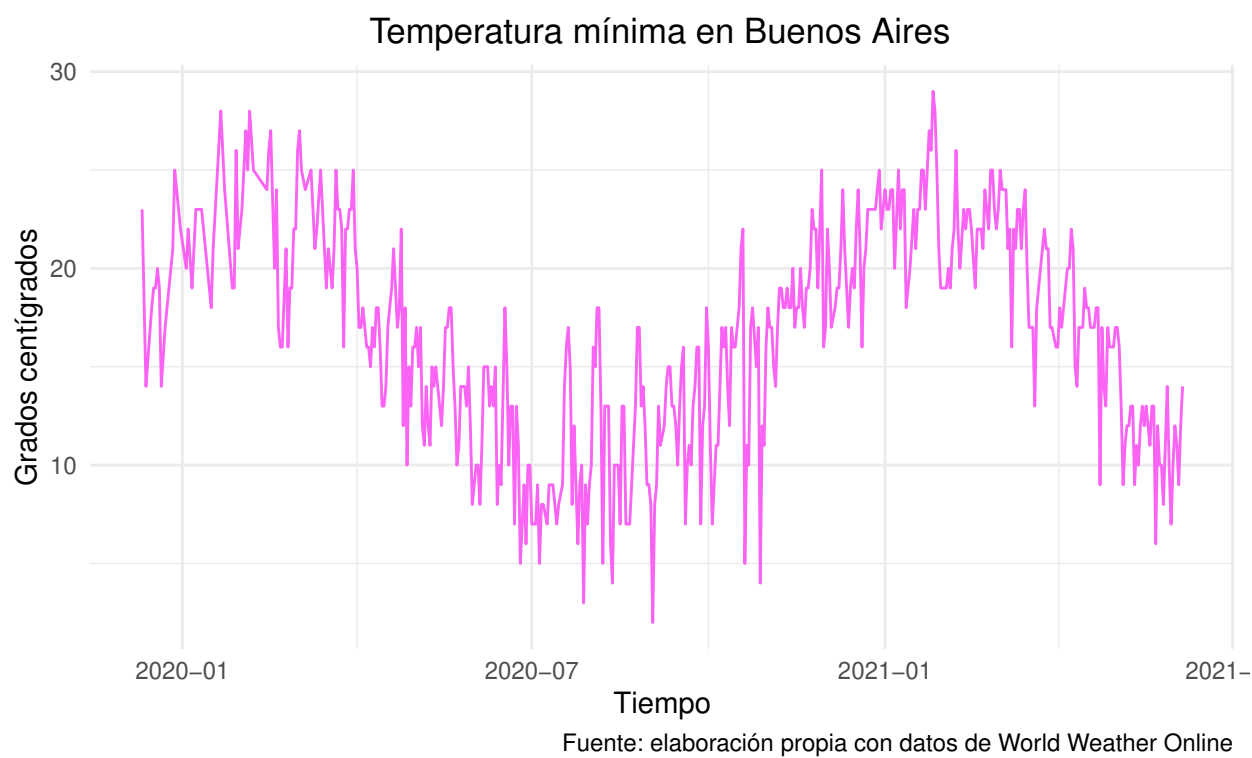


Figura 21: Temperatura mínima en Buenos Aires

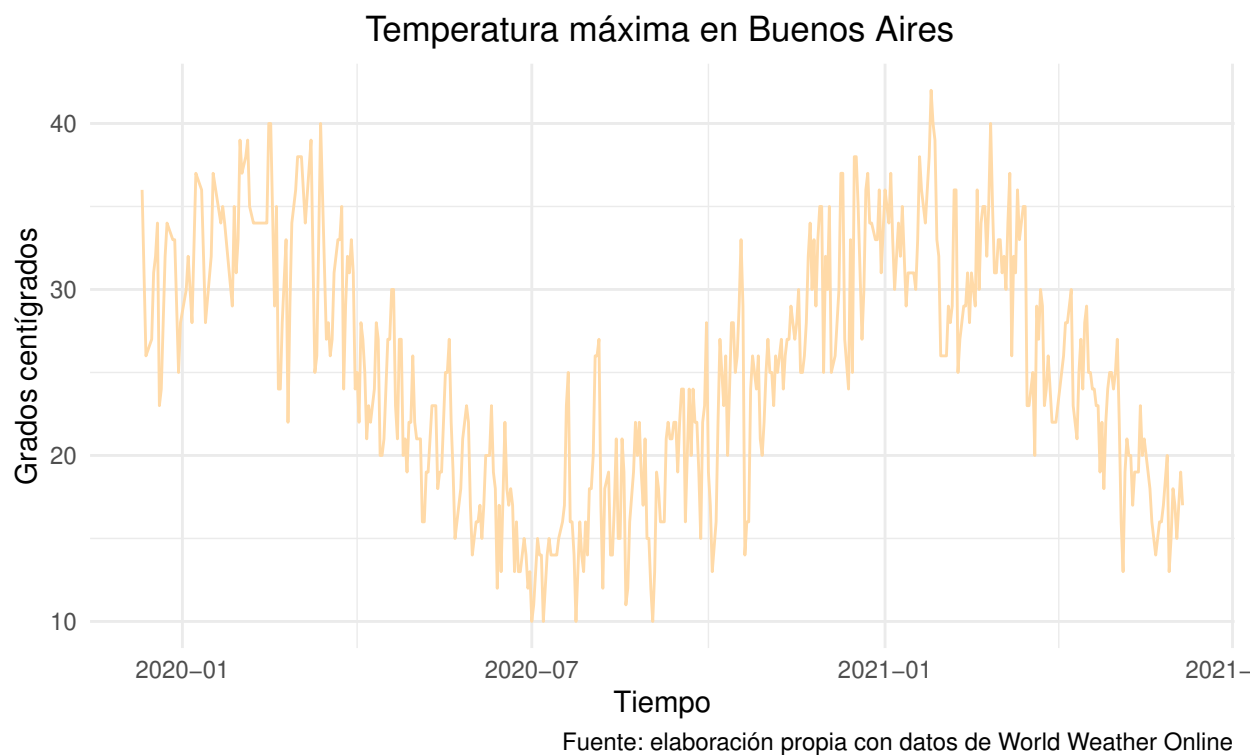


Figura 22: Temperatura máxima en Buenos Aires

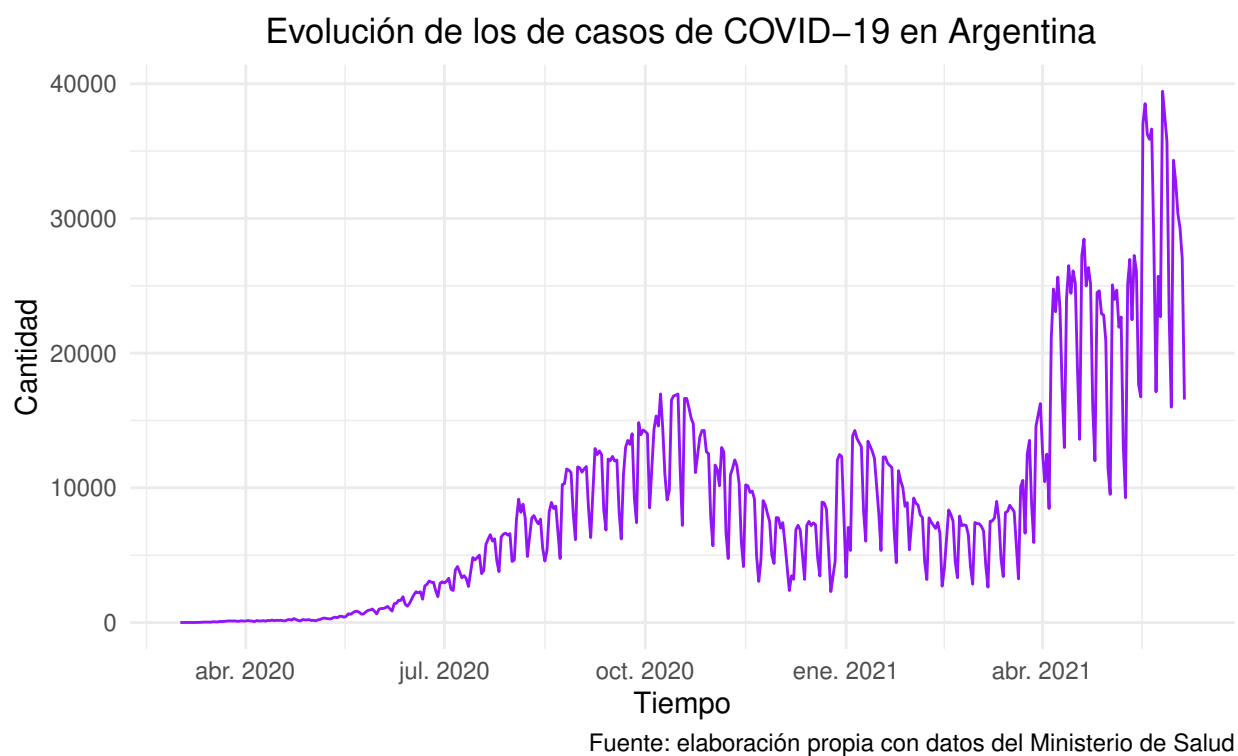


Figura 23: Casos de COVID-19 en Argentina

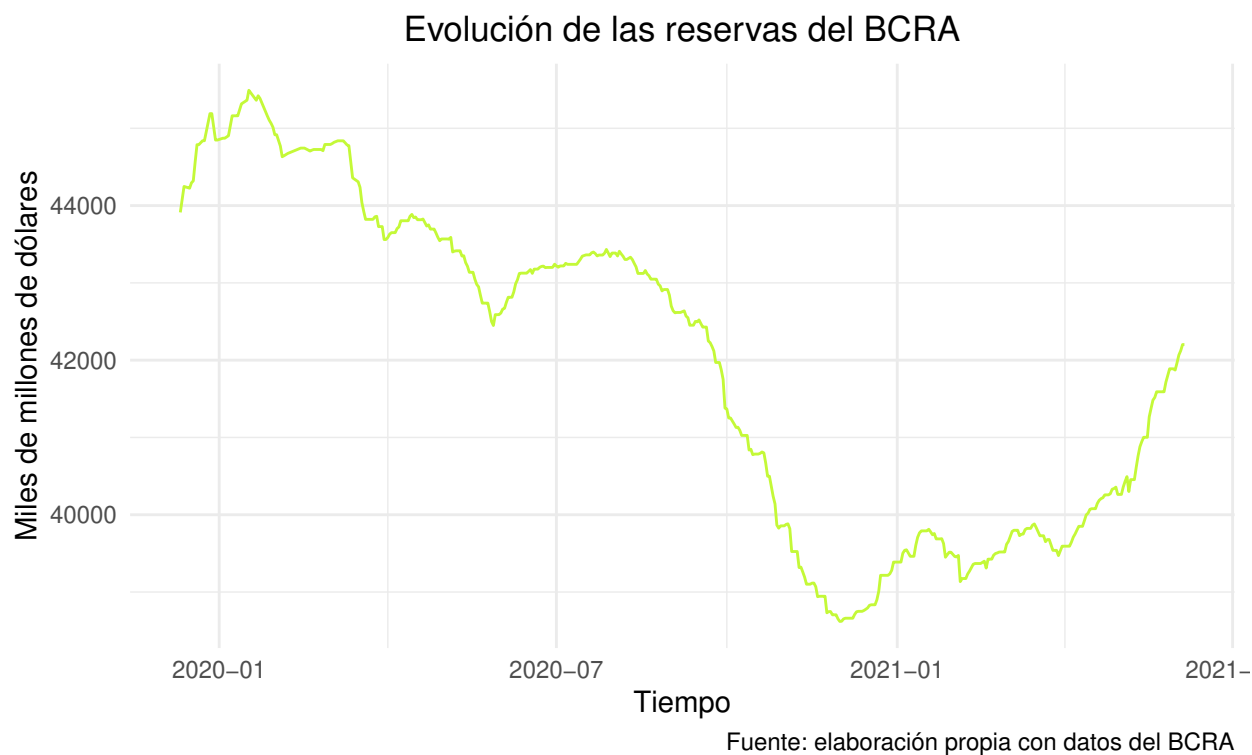


Figura 24: Evolución de reservas BCRA

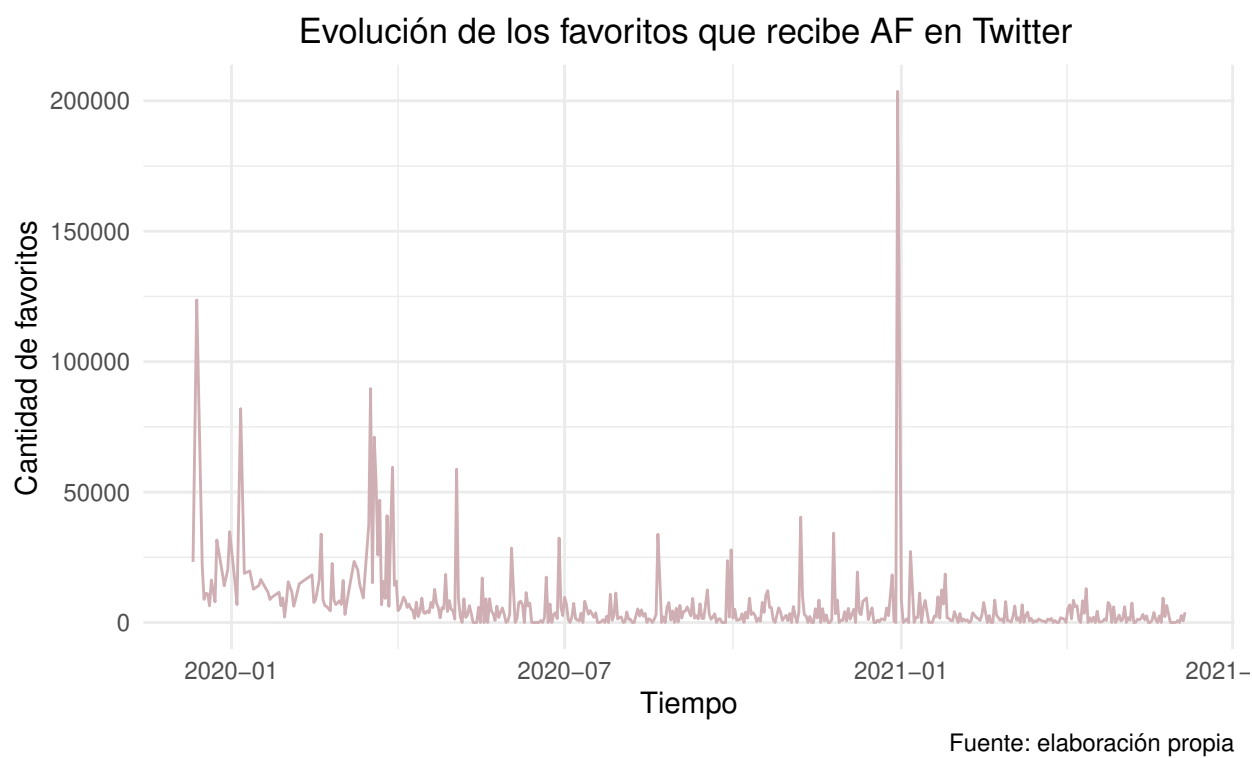


Figura 25: Cantidad de favoritos promedios diarios de AF en Twitter

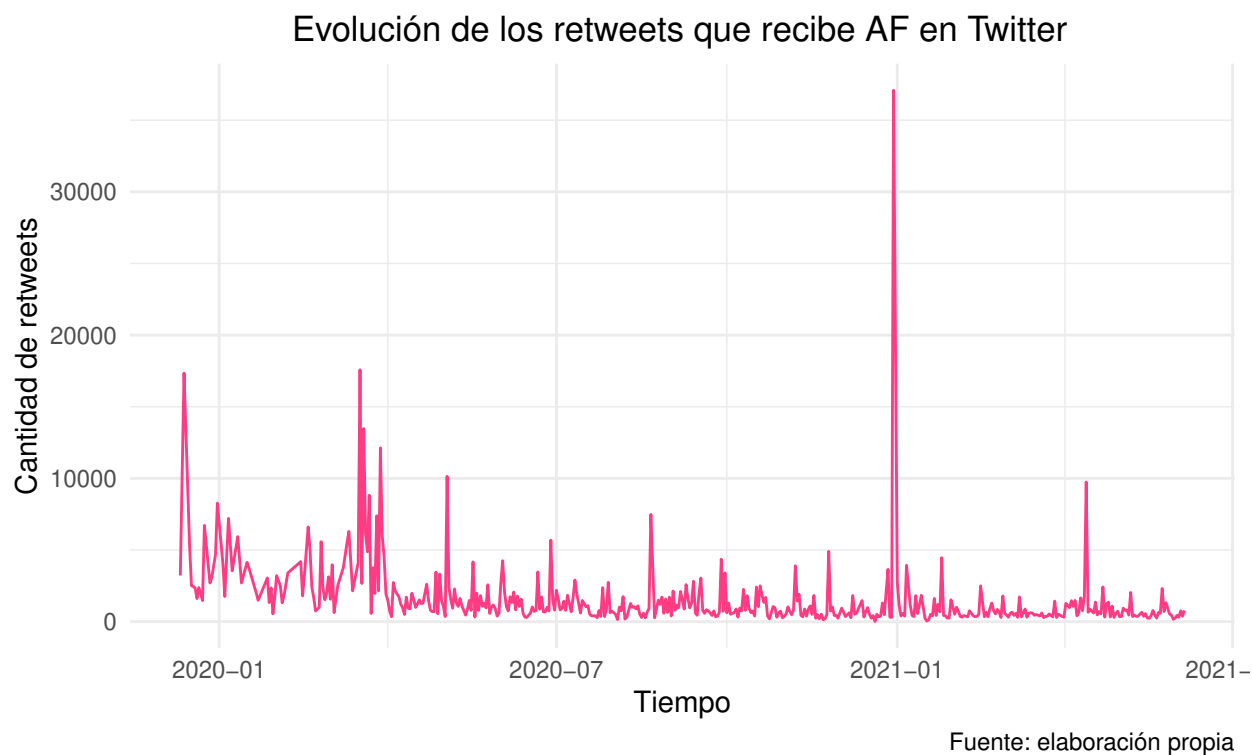


Figura 26: Cantidad de *rtweets* promedios diarios de AF

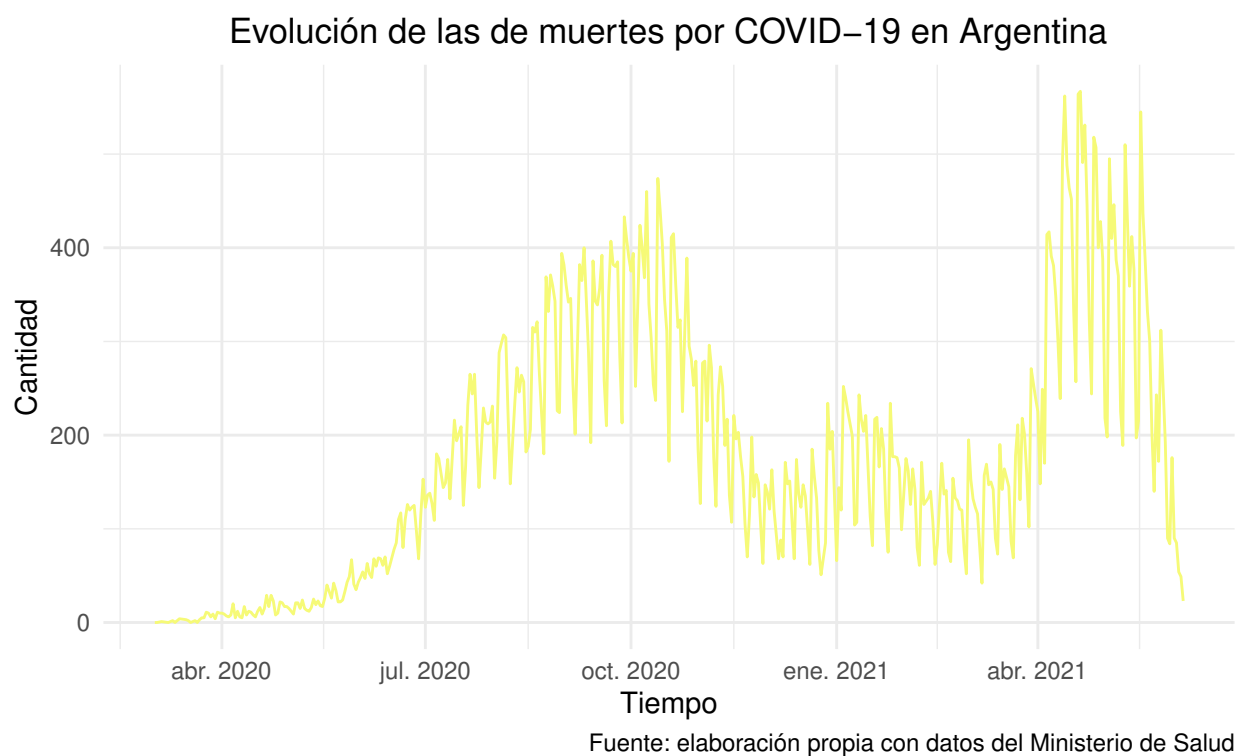
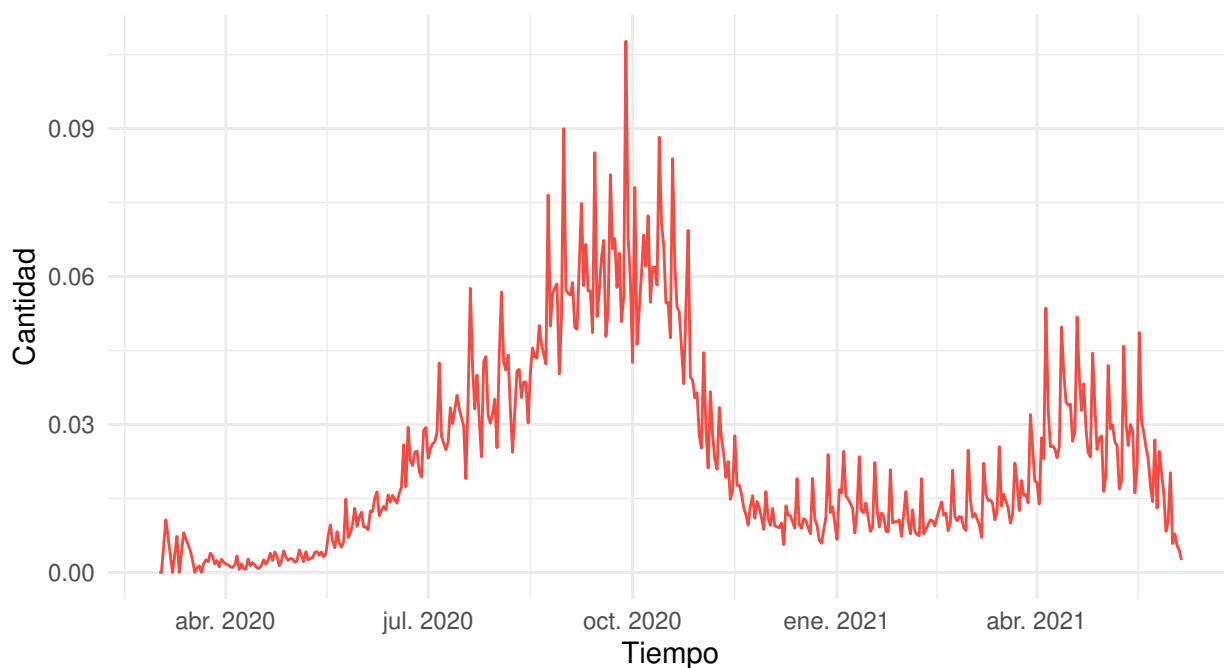


Figura 27: Muertes por COVID-19 en Argentina

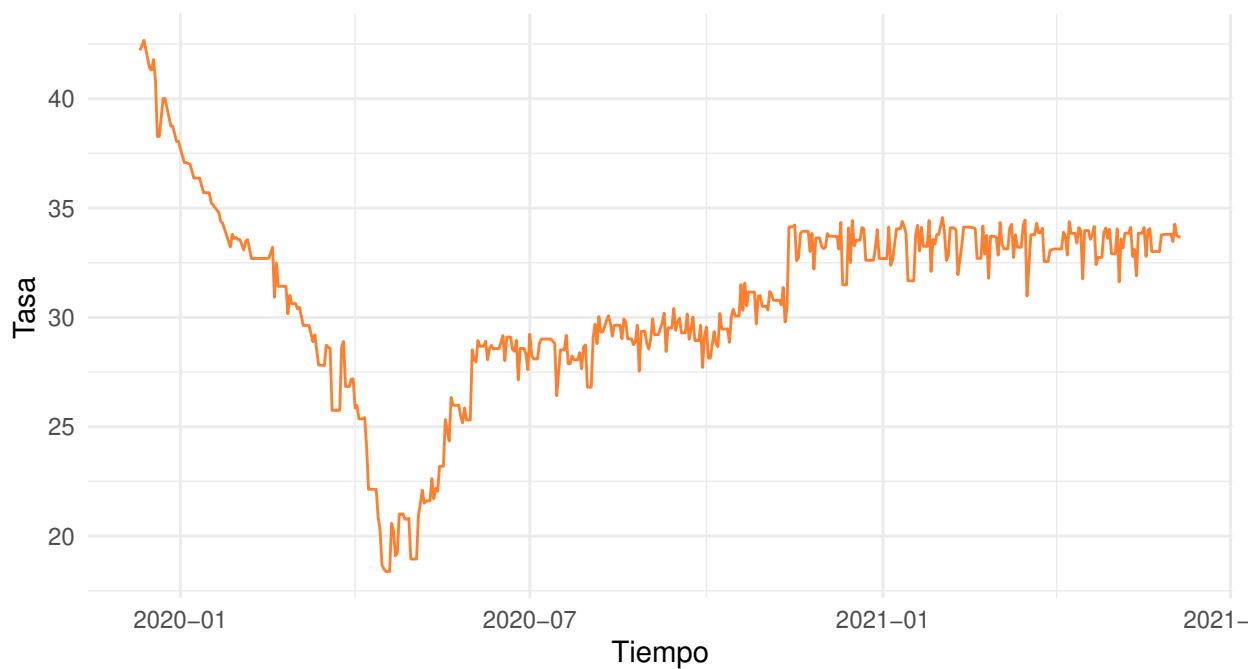
Muertes por COVID-19 en Arg. relativas al resto del mundo



Fuente: elaboración propia con datos del Ministerio de Salud y Our World in Data

Figura 28: Muertes por COVID-19 en Argentina relativas al resto del mundo

Evolución de la tasa de interés



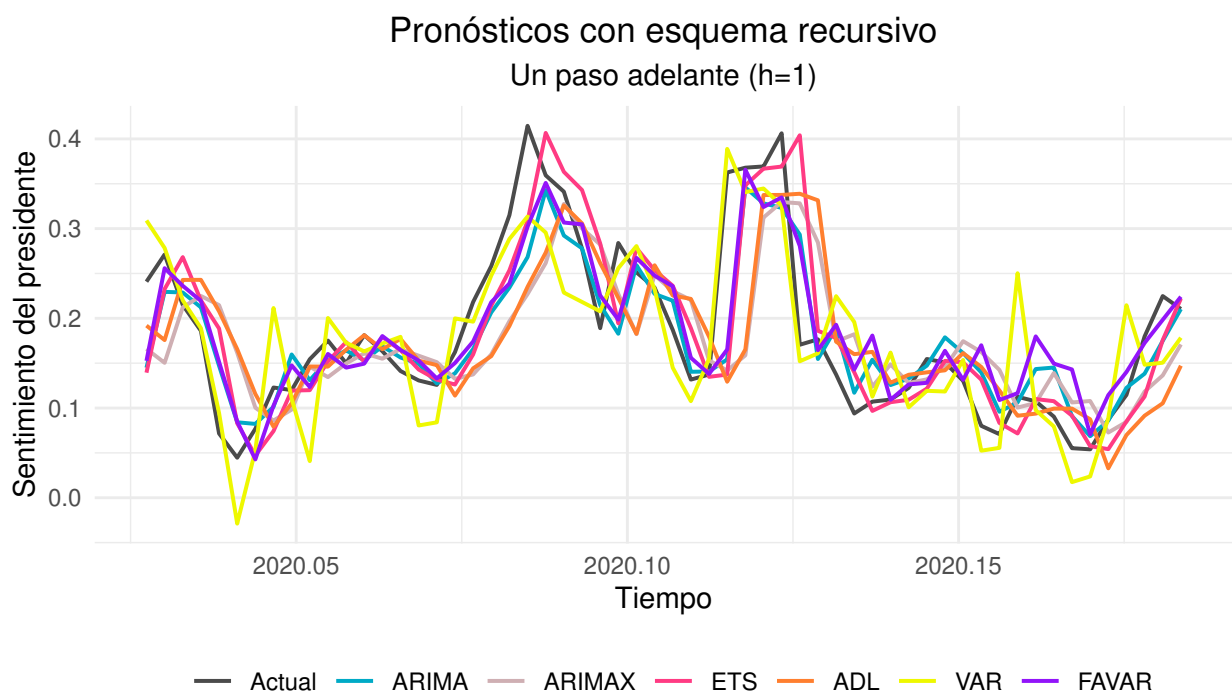
Fuente: elaboración propia con datos del BCRA

Figura 29: Evolución de tasa de interés

Variable	Constant	Trend
SentimientoAF	-8,97***	-8,96***
Favoritos	-21,03***	-20,88***
Retweets	-21,48***	-21,26***
Reservas BCRA	-1.13	0.34
L_1 (Reservas BCRA)	-20,73***	-20,58***
Tasa de interés	-2,86*	-3,95**
Base monetaria	-2,86*	-5,07***
Tipo de cambio	-1.47	-2.39
L_1 (Tipo de cambio)	-19,97***	-19,98***
Casos Arg.	-4,34***	-8,93***
Muertes Arg.	-5,48***	-7,65***
Vacunas	-3,56***	-6,32***
Casos Arg. Rel	-5,1***	-7,43***
Muertes Arg. Rel.	-5,32***	-5,62***
Temp. Min.	-7,42***	-7,42***
Temp. Max.	-6,63***	-6,63***

Nota: *p<0.1; **p<0.05; ***p<0.01

Cuadro 4: Test de raíz unitaria: Phillips-Perrón



Fuente: elaboración propia

Figura 30: Pronósticos

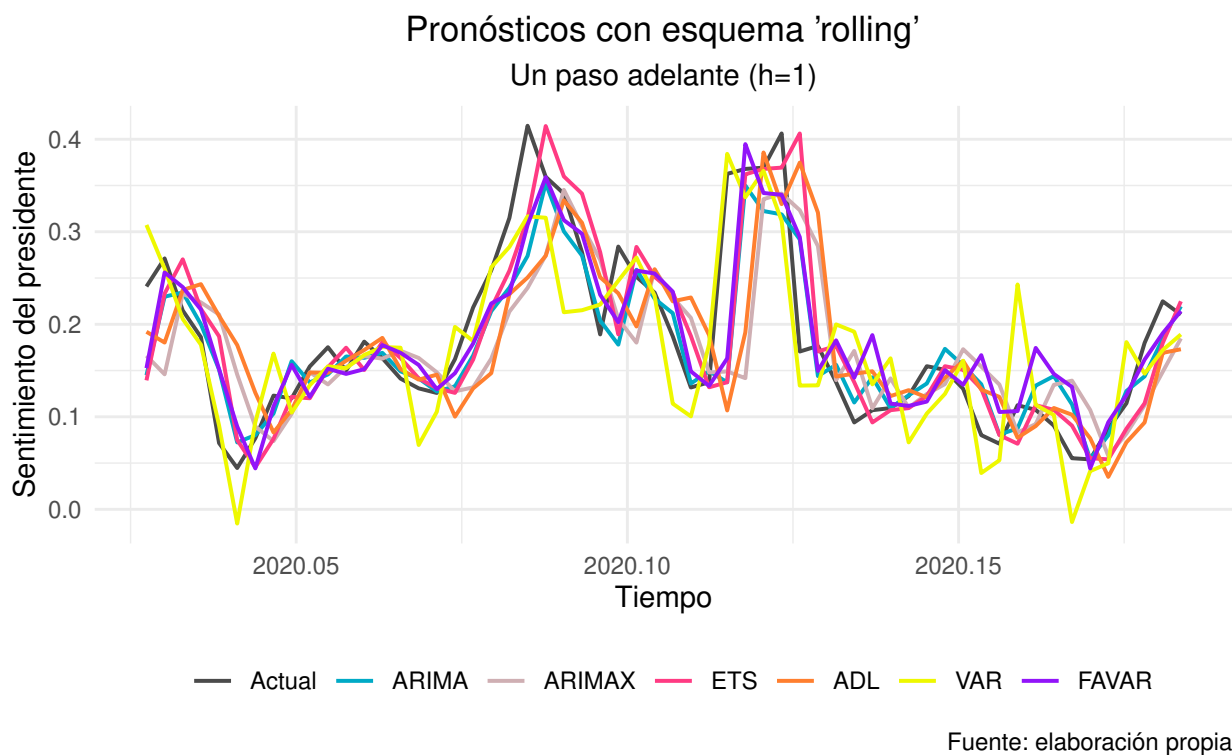


Figura 31: Pronósticos

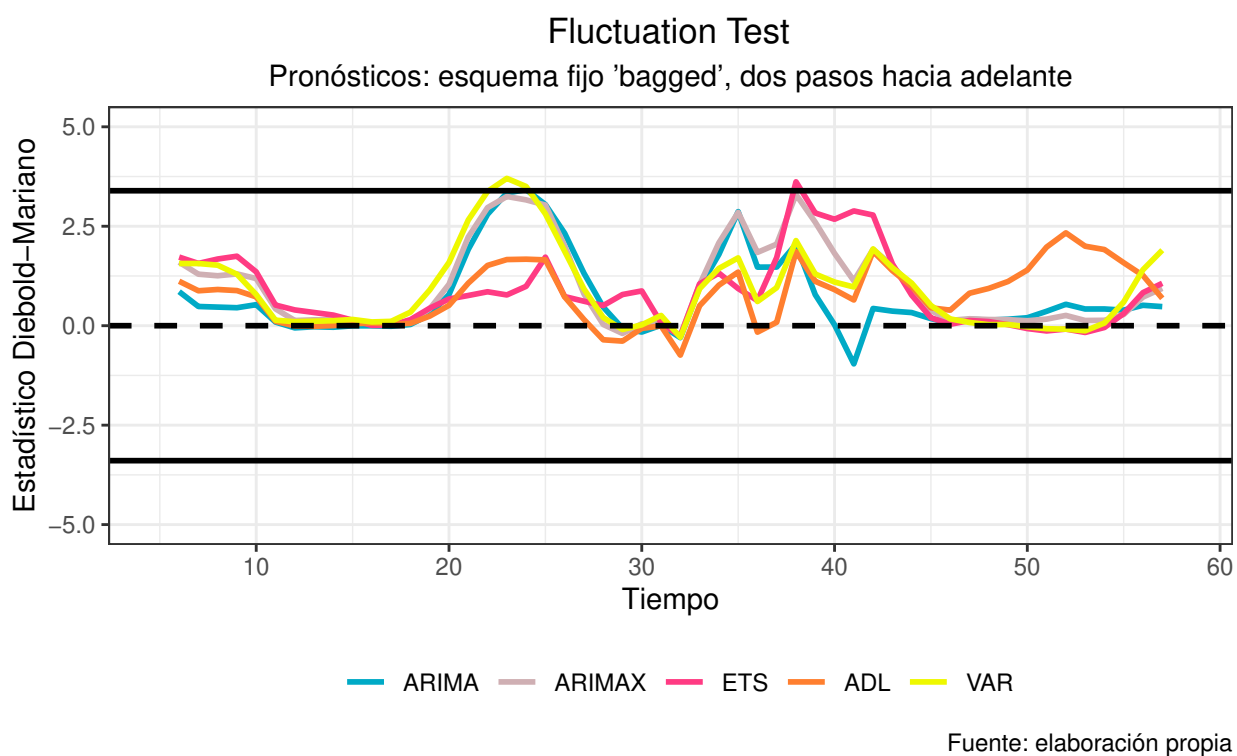


Figura 32: Test de Giacomini Rossi (esquema fijo *bagged*, dos pasos adelante)

Modelo	RMSE	MAE	MAPE	DM	p-val
AR(1)	0.0588	0.0405	25.7162		
ARIMA fijo	0.0762	0.0558	36.0494	2.2663	0.0397
ARIMAX fijo	0.0958	0.0753	51.2643	4.0954	0.0015
ETS fijo	0.0892	0.0683	46.3702	3.4481	0.0005
ADL fijo	0.0976	0.0758	51.3699	4.1011	0.0008
VAR fijo	0.0855	0.0682	46.3931	3.2254	0.0098
FAVAR fijo	0.0755	0.0604	46.1592	2.6934	0.0092
ARIMA recursivo	0.0754	0.0573	38.547	2.3611	0.0224
ARIMAX recursivo	0.0964	0.0759	52.0981	4.3797	0.0016
ETS recursivo	0.0892	0.0682	46.294	3.4358	0.0006
ADL recursivo	0.0981	0.0761	50.9751	4.0649	0.0009
VAR recursivo	0.0843	0.0672	45.4224	3.0595	0.015
FAVAR recursivo	0.0755	0.0604	46.1592	2.6934	0.0092
ARIMA rolling	0.074	0.0526	34.7623	2.0499	0.0354
ARIMAX rolling	0.0961	0.0765	52.2818	4.4071	0.0008
ETS rolling	0.0888	0.0676	45.8008	3.3607	0.0007
ADL rolling	0.0959	0.0729	49.69	3.5447	0.0008
VAR rolling	0.0782	0.0619	42.4499	2.4322	0.0332
FAVAR rolling	0.0739	0.0589	44.6215	2.6359	0.0131
ARIMA fijo bagged	0.0746	0.0557	36.1538	2.3107	0.0547
ARIMAX fijo bagged	0.088	0.0676	45.9714	3.5517	0.0082
ETS fijo bagged	0.0741	0.0568	38.9011	2.9883	0.002
VAR fijo bagged	0.0708	0.0569	42.9266	2.4353	0.0024
FAVAR fijo bagged	0.078	0.0632	41.3869	3.5352	0.0068
ARIMA recursivo bagged	0.0743	0.0576	38.5408	2.4912	0.0313
ARIMAX recursivo bagged	0.0887	0.0689	47.4351	3.7426	0.0053
ETS recursivo bagged	0.0737	0.0565	38.8458	2.9402	0.0021
VAR recursivo bagged	0.0706	0.0567	42.7125	2.3798	0.0031
FAVAR recursivo bagged	0.078	0.0632	41.3869	3.5639	0.0027
ARIMA rolling bagged	0.0721	0.0532	34.9207	2.0672	0.054
ARIMAX rolling bagged	0.0879	0.0679	46.424	3.5582	0.0054
ETS rollingbagged	0.0724	0.0554	38.1238	2.7726	0.0026
VAR rolling bagged	0.0707	0.057	42.2683	2.4873	0.0014
FAVAR rolling bagged	0.0784	0.0631	39.9943	3.4152	0.0089

Cuadro 5: Medidas de desempeño - Dos pasos adelante

Modelo	RMSE	MAE	MAPE	DM	p-val
AR(1)	0.0588	0.0405	25.7162		
ARIMA fijo	0.098	0.0719	42.0135	2.2334	0.2488
ARIMAX fijo	0.1087	0.0875	55.1167	4.0414	0.0158
ETS fijo	0.1266	0.0977	62.5156	4.1225	0.0569
ADL fijo	0.1155	0.0942	60.3013	4.3675	0.0049
VAR fijo	0.1178	0.0945	57.0427	4.5977	0.0129
FAVAR fijo	0.0983	0.0788	56.7441	3.9863	0.0164
ARIMA recursivo	0.1034	0.0805	51.5101	3.8835	0.0238
ARIMAX recursivo	0.1118	0.0902	58.14	4.086	0.0156
ETS recursivo	0.1272	0.0979	62.7083	4.1244	0.0555
ADL recursivo	0.1163	0.0932	60.0919	4.1688	0.0075
VAR recursivo	0.1161	0.0925	55.4165	4.4335	0.0099
FAVAR recursivo	0.0983	0.0788	56.7441	3.9863	0.0164
ARIMA rolling	0.1011	0.0741	44.2576	3.2796	0.0585
ARIMAX rolling	0.1128	0.089	56.8189	4.0274	0.0218
ETS rolling	0.1282	0.0982	62.9109	4.122	0.0508
ADL rolling	0.1196	0.0938	60.413	3.8772	0.008
VAR rolling	0.1117	0.0857	49.0213	3.9422	0.0194
FAVAR rolling	0.1	0.0797	58.3001	4.0155	0.024
ARIMA fijo bagged	0.0978	0.0717	41.8707	3.1176	0.0693
ARIMAX fijo bagged	0.1013	0.0831	52.3967	3.6825	0.0194
ETS fijo bagged	0.1151	0.092	58.4278	4.0115	0.0381
VAR fijo bagged	0.0993	0.081	58.5055	4.5095	0.0103
FAVAR fijo bagged	0.109	0.0865	51.8119	4.0021	0.0087
ARIMA recursivo bagged	0.1032	0.0802	51.1903	3.8369	0.024
ARIMAX recursivo bagged	0.1038	0.086	55.4365	3.8836	0.0141
ETS recursivo bagged	0.1152	0.092	58.4192	4.0104	0.0392
VAR recursivo bagged	0.0995	0.0811	58.6148	4.3147	0.0178
FAVAR recursivo bagged	0.109	0.0865	51.8119	4.0159	0.0128
ARIMA rolling bagged	0.1009	0.0746	44.6077	3.2994	0.0586
ARIMAX rolling bagged	0.1019	0.0836	52.6962	3.7023	0.0181
ETS rolling bagged	0.1155	0.092	58.3634	4.0114	0.0401
VAR rolling bagged	0.0998	0.0819	59.1199	4.245	0.0181
FAVAR rolling bagged	0.1102	0.0864	51.3235	4.0592	0.0093

Cuadro 6: Medidas de desempeño - Siete pasos adelante

Modelo	RMSE	MAE	MAPE
AR(1)	0.0576	0.0385	22.122
ARIMA fijo	0.0538	0.0358	20.4026
ARIMAX fijo	0.0764	0.0568	33.7166
ETS fijo	0.0601	0.0403	23.539
ADL fijo	0.0782	0.0587	34.4651
VAR fijo	0.0519	0.0403	26.9462
FAVAR fijo	0.0538	0.0386	25.3799
ARIMA recursivo	0.0525	0.0355	20.5858
ARIMAX recursivo	0.0754	0.056	33.3447
ETS recursivo	0.0601	0.04	23.3164
ADL recursivo	0.0775	0.0576	33.6215
VAR recursivo	0.0519	0.0403	26.9462
FAVAR recursivo	0.052	0.0374	25.4067
ARIMA rolling	0.0525	0.0336	19.0823
ARIMAX rolling	0.0737	0.0545	33.2633
ETS rolling	0.0596	0.0392	22.8933
ADL rolling	0.0751	0.054	31.6712
VAR rolling	0.0494	0.0398	27.1824
FAVAR rolling	0.0517	0.0366	24.2686
ARIMA fijo bagged	0.051	0.0375	21.5935
ARIMAX fijo bagged	0.0738	0.0559	32.8802
ETS fijo bagged	0.048	0.0354	21.1262
VAR fijo bagged	0.053	0.0406	22.7392
FAVAR fijo bagged	0.0488	0.0369	23.1521
ARIMA recursivo bagged	0.0498	0.0375	22.0026
ARIMAX recursivo bagged	0.074	0.0569	33.9066
ETS recursivo bagged	0.0474	0.0349	20.7527
VAR recursivo bagged	0.0519	0.0402	22.8438
FAVAR recursivo bagged	0.0502	0.0389	25.8429
ARIMA rolling bagged	0.0483	0.0355	20.5609
ARIMAX rolling bagged			
ETS rollingbagged	0.0457	0.0334	19.879
VAR rolling bagged	0.0546	0.0423	23.5172
FAVAR rolling bagged	0.0506	0.0392	25.4762

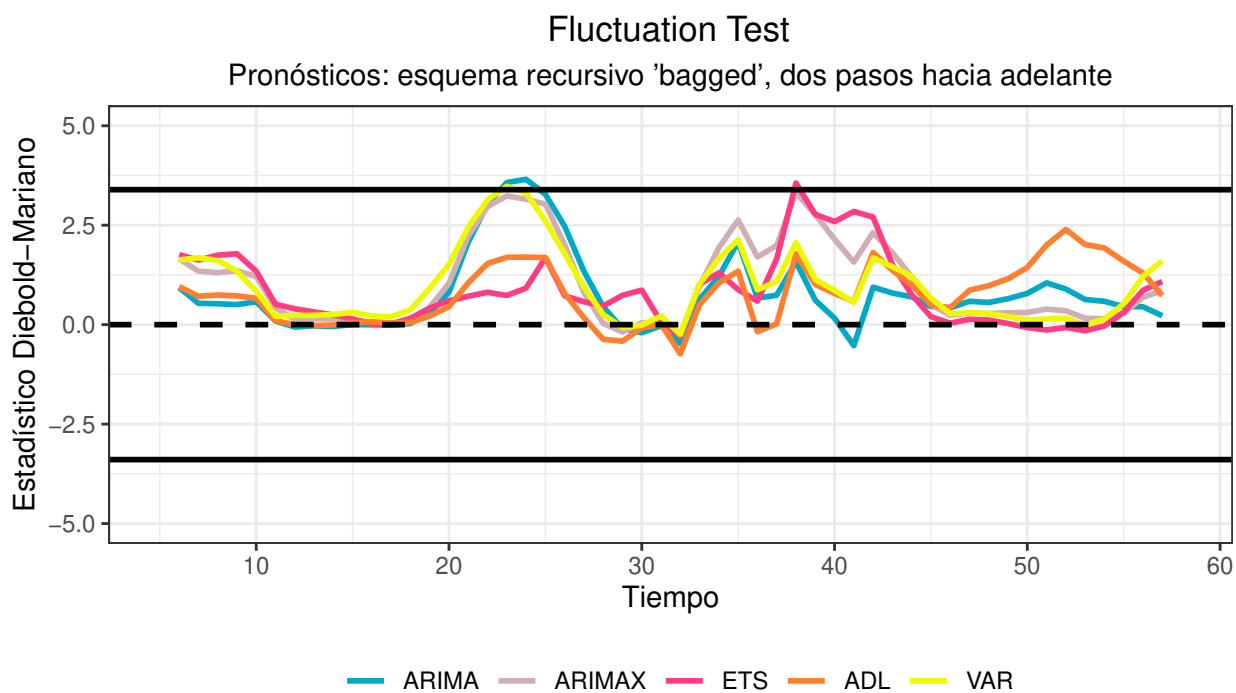
Cuadro 7: Medidas de desempeño comparables - Un paso adelante

Modelo	RMSE	MAE	MAPE
AR(1)	0.0576	0.0385	22.122
ARIMA fijo	0.0754	0.0541	31.2754
ARIMAX fijo	0.0935	0.0725	43.1672
ETS fijo	0.0862	0.0649	38.6871
ADL fijo	0.0958	0.0734	43.2493
VAR fijo	0.0861	0.0679	43.6133
FAVAR fijo	0.0742	0.0585	41.4877
ARIMA recursivo	0.0746	0.0557	33.9699
ARIMAX recursivo	0.0942	0.0732	44.0686
ETS recursivo	0.0862	0.0647	38.6126
ADL recursivo	0.0963	0.0736	42.7459
VAR recursivo	0.0861	0.0679	43.6133
FAVAR recursivo	0.0742	0.0585	41.4877
ARIMA rolling	0.0728	0.0505	29.7048
ARIMAX rolling	0.0938	0.0735	44.2257
ETS rolling	0.0859	0.064	38.2095
ADL rolling	0.0938	0.0703	41.3707
VAR rolling	0.0801	0.0632	41.4544
FAVAR rolling	0.0726	0.0571	39.8927
ARIMA fijo bagged	0.0742	0.0547	31.5781
ARIMAX fijo bagged	0.0867	0.066	39.4641
ETS fijo bagged	0.0724	0.0551	32.8175
VAR fijo bagged	0.07	0.056	38.5859
FAVAR fijo bagged	0.077	0.0605	34.0165
ARIMA recursivo bagged	0.0739	0.0568	34.1535
ARIMAX recursivo bagged	0.0875	0.0674	40.9673
ETS recursivo bagged	0.0719	0.0547	32.7598
VAR recursivo bagged	0.07	0.056	38.6819
FAVAR recursivo bagged	0.0765	0.0613	35.2186
ARIMA rolling bagged	0.0713	0.0518	30.0725
ARIMAX rolling bagged	0.0866	0.0662	39.8672
ETS rollingbagged	0.0705	0.0536	32.0979
VAR rolling bagged	0.07	0.0563	38.1598
FAVAR rolling bagged	0.0779	0.0619	34.6634

Cuadro 8: Medidas de desempeño comparables - Dos pasos adelante

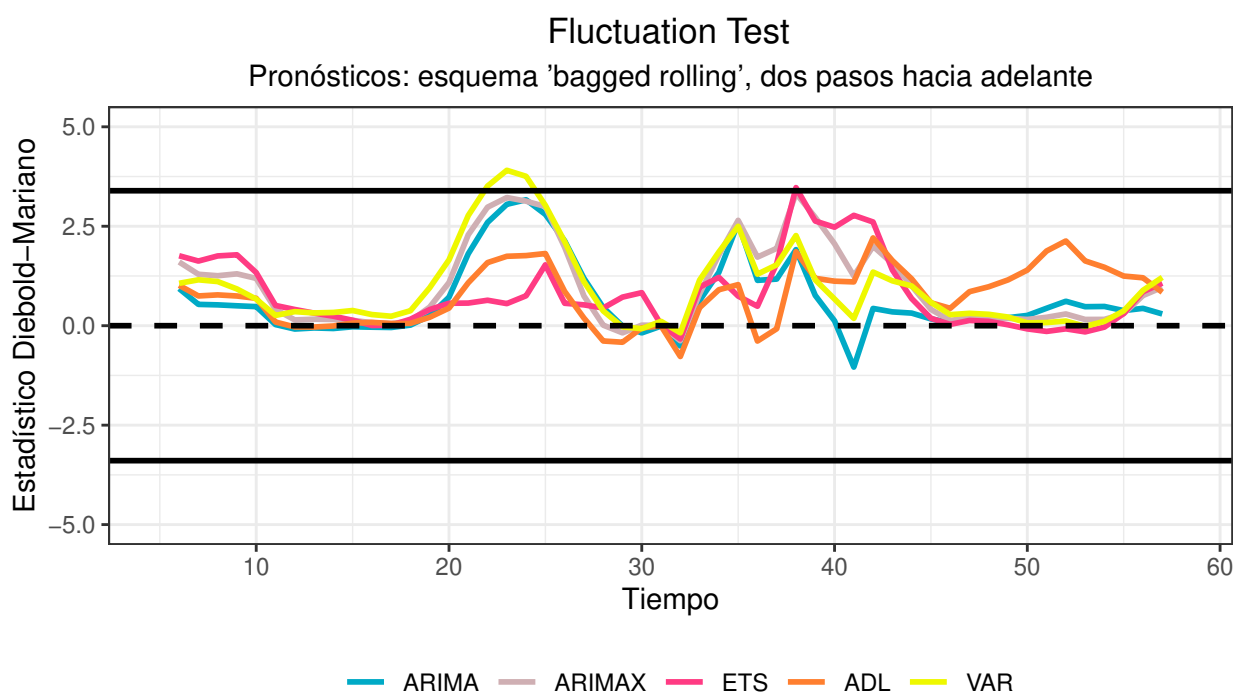
Modelo	RMSE	MAE	MAPE
AR(1)	0.0576	0.0385	22.122
ARIMA fijo	0.098	0.0719	42.0135
ARIMAX fijo	0.1087	0.0875	55.1167
ETS fijo	0.1266	0.0977	62.5156
ADL fijo	0.1155	0.0942	60.3013
VAR fijo	0.1178	0.0945	57.0427
FAVAR fijo	0.0983	0.0788	56.7441
ARIMA recursivo	0.1034	0.0805	51.5101
ARIMAX recursivo	0.1118	0.0902	58.14
ETS recursivo	0.1272	0.0979	62.7083
ADL recursivo	0.1163	0.0932	60.0919
VAR recursivo	0.1161	0.0925	55.4165
FAVAR recursivo	0.0983	0.0788	56.7441
ARIMA rolling	0.1011	0.0741	44.2576
ARIMAX rolling	0.1128	0.089	56.8189
ETS rolling	0.1282	0.0982	62.9109
ADL rolling	0.1196	0.0938	60.413
VAR rolling	0.1117	0.0857	49.0213
FAVAR rolling	0.1	0.0797	58.3001
ARIMA fijo bagged	0.0978	0.0717	41.8707
ARIMAX fijo bagged	0.1013	0.0831	52.3967
ETS fijo bagged	0.1151	0.092	58.4278
VAR fijo bagged	0.0993	0.081	58.5055
FAVAR fijo bagged	0.1091	0.0888	54.8723
ARIMA recursivo bagged	0.1032	0.0802	51.1903
ARIMAX recursivo bagged	0.1038	0.086	55.4365
ETS recursivo bagged	0.1152	0.092	58.4192
VAR recursivo bagged	0.0995	0.0811	58.6148
FAVAR recursivo bagged	0.109	0.0865	51.8119
ARIMA rolling bagged	0.1009	0.0746	44.6077
ARIMAX rolling bagged	0.1019	0.0836	52.6962
ETS rollingbagged	0.1155	0.092	58.3634
VAR rolling bagged	0.0998	0.0819	59.1199
FAVAR rolling bagged	0.1102	0.0864	51.3235

Cuadro 9: Medidas de desempeño comparables - Siete pasos adelante



Fuente: elaboración propia

Figura 33: Test de Giacomini Rossi (esquema recursivo *bagged*, dos pasos adelante)



Fuente: elaboración propia

Figura 34: Test de Giacomini Rossi (esquema *bagged rolling*, dos pasos adelante)

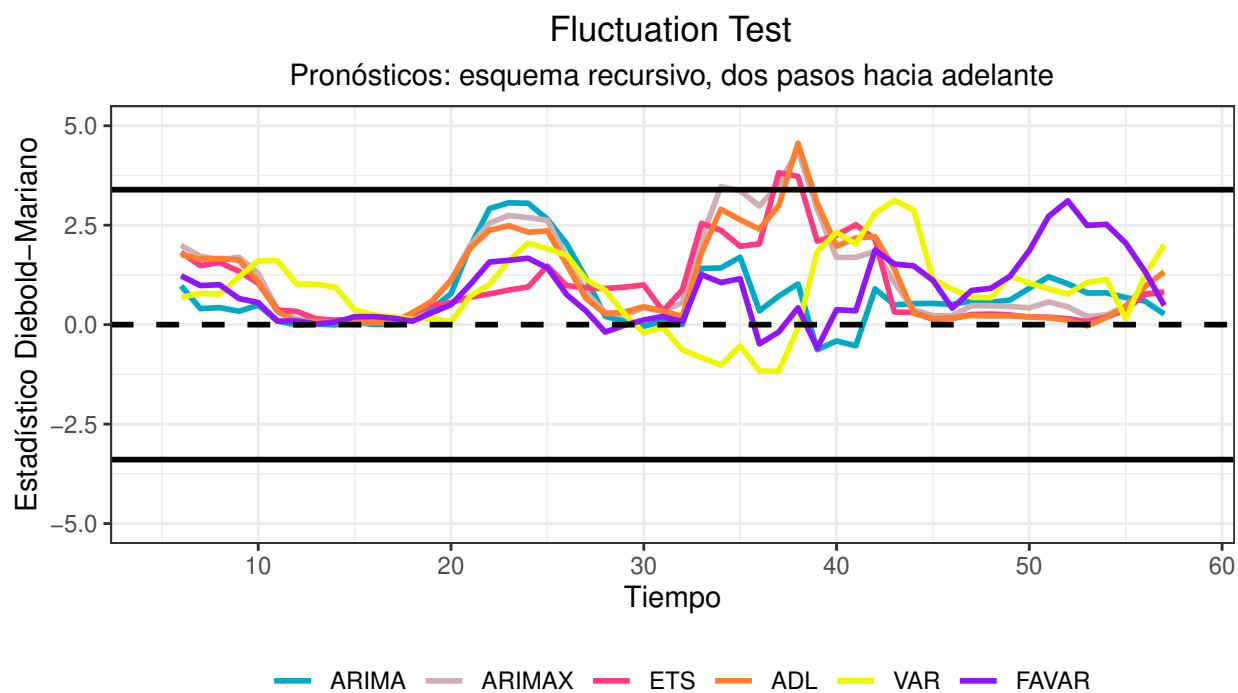


Figura 35: Test de Giacomini Rossi (esquema recursivo, dos pasos adelante)

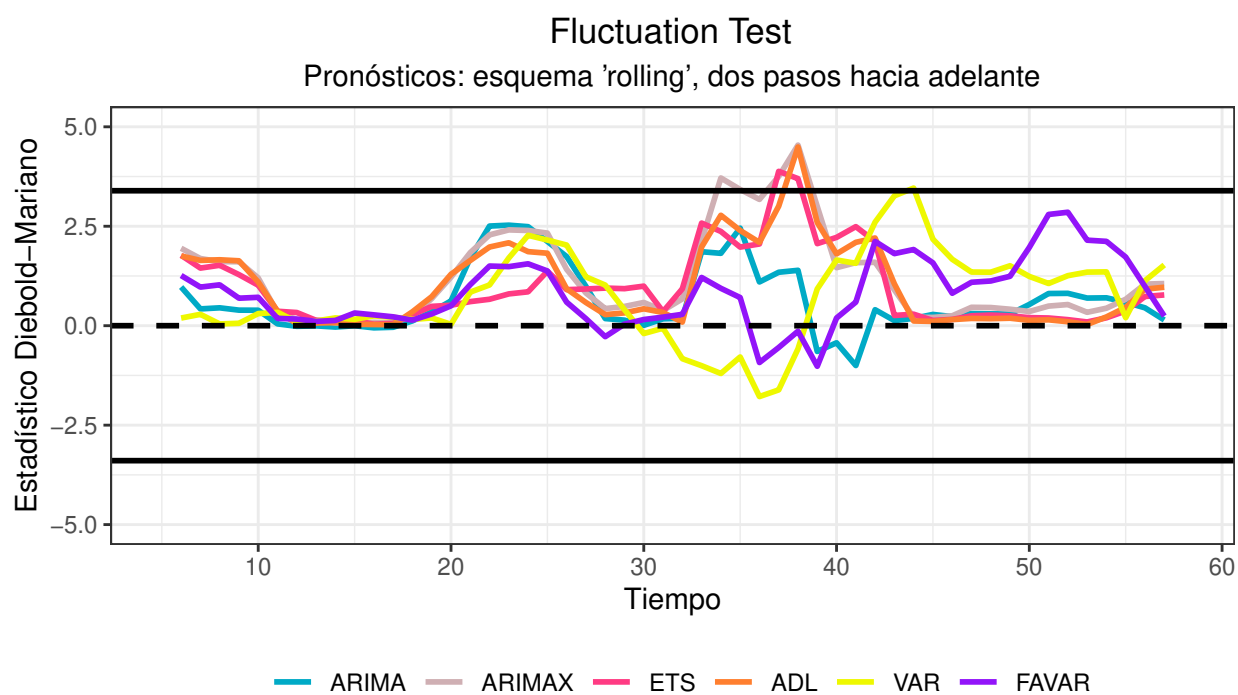
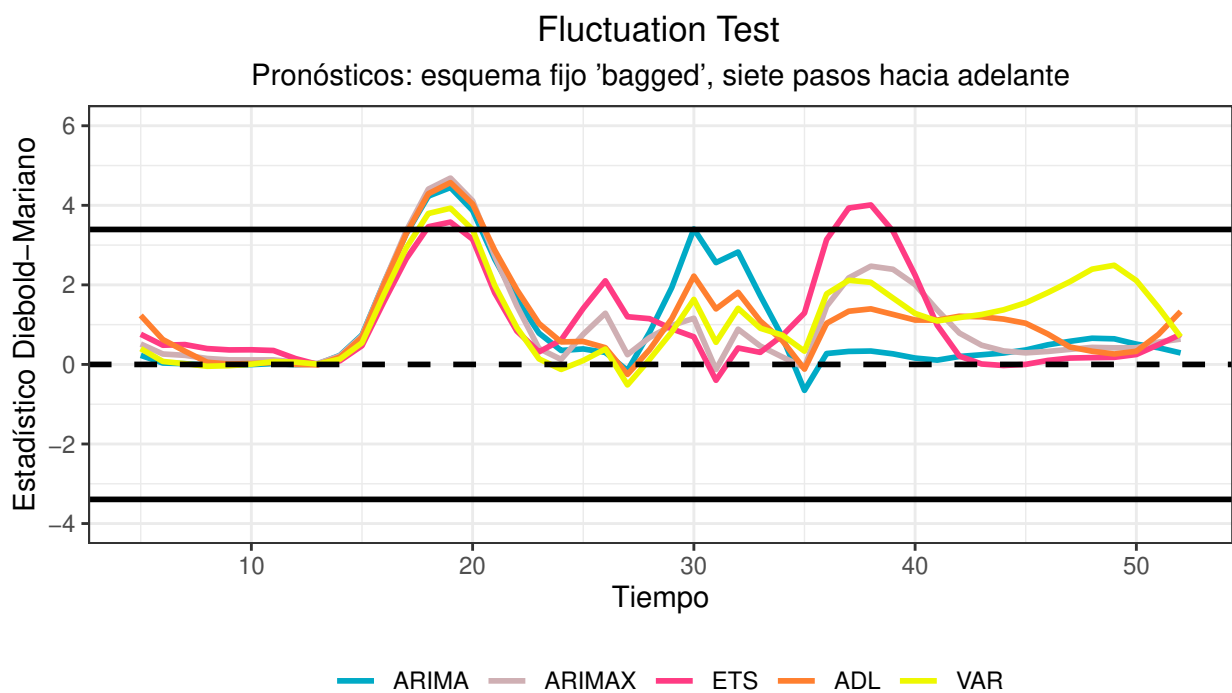
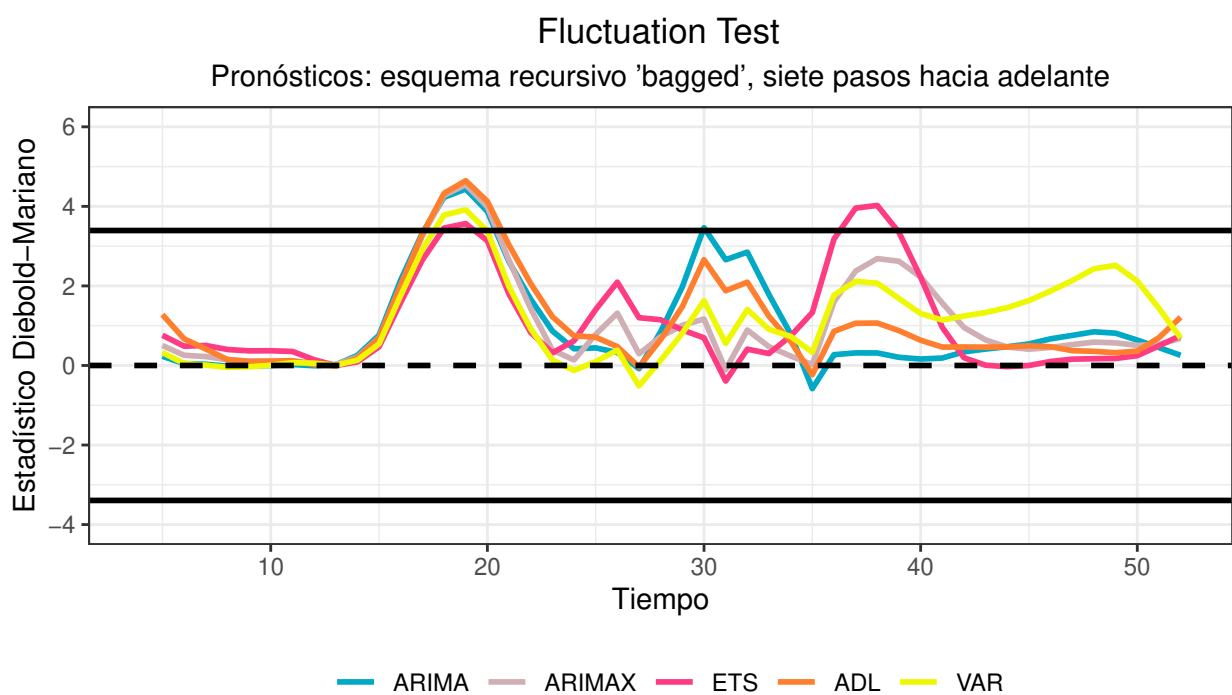


Figura 36: Test de Giacomini Rossi (esquema *rolling*, dos pasos adelante)



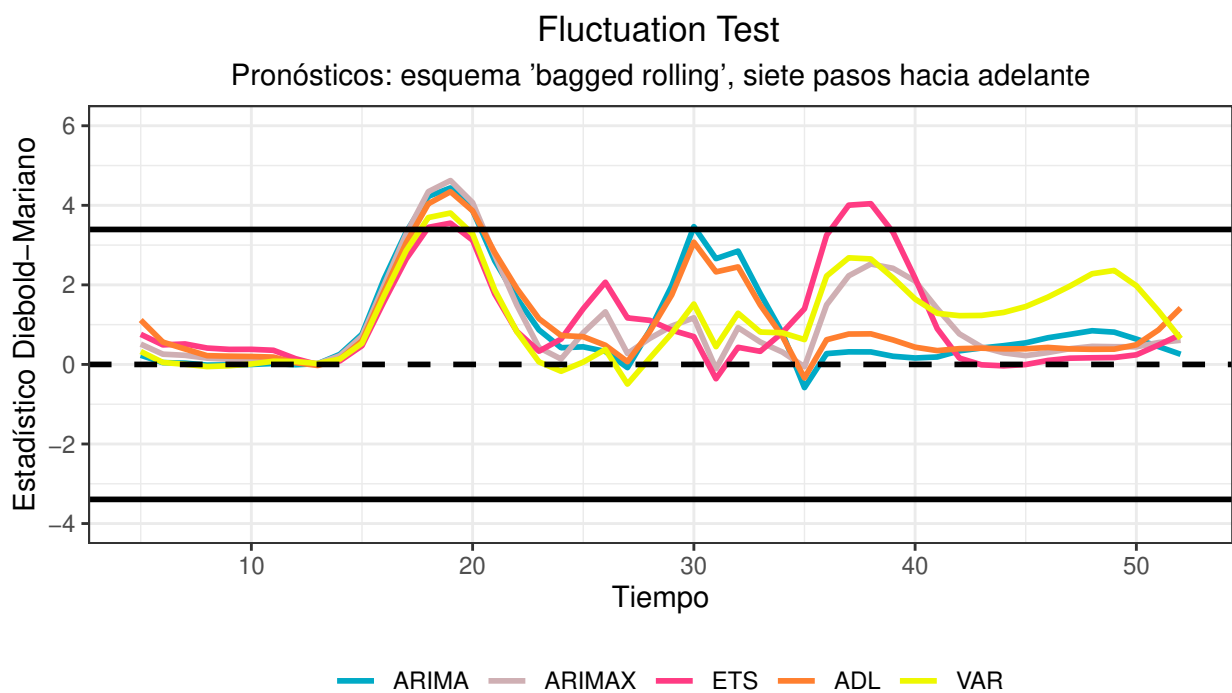
Fuente: elaboración propia

Figura 37: Test de Giacomini Rossi (esquema fijo *bagged*, siete pasos adelante)



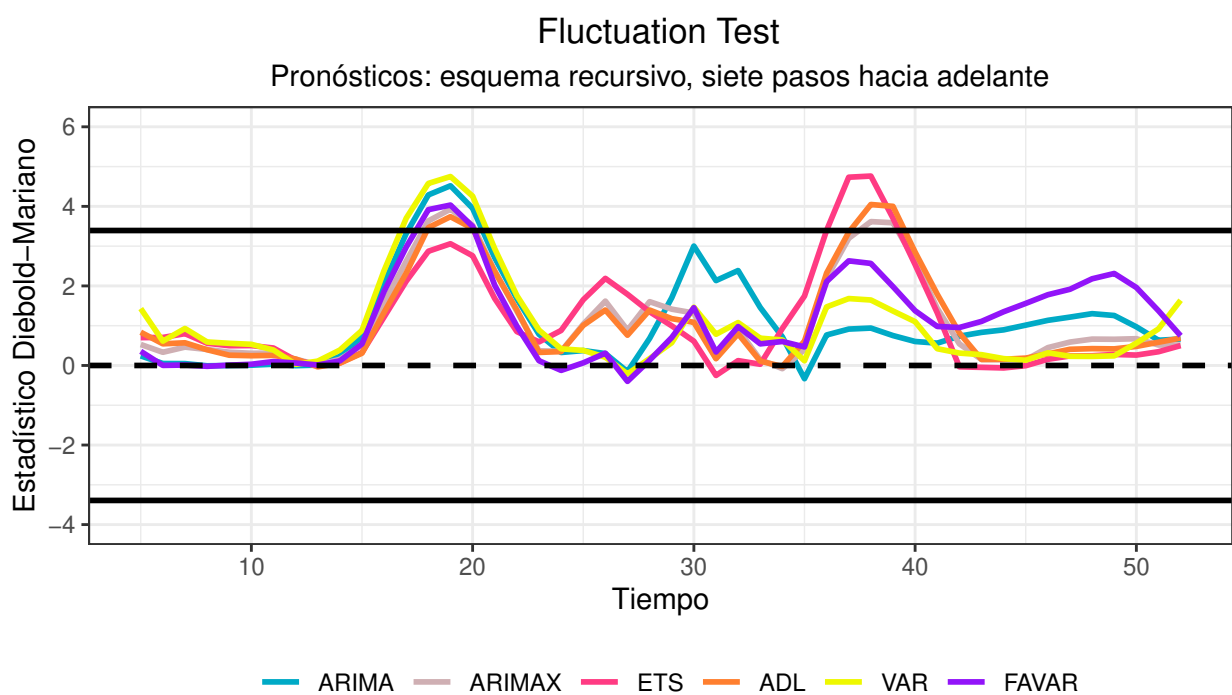
Fuente: elaboración propia

Figura 38: Test de Giacomini Rossi (esquema recursivo *bagged*, siete pasos adelante)



Fuente: elaboración propia

Figura 39: Test de Giacomini Rossi (esquema *bagged rolling*, siete pasos adelante)



Fuente: elaboración propia

Figura 40: Test de Giacomini Rossi (esquema recursivo, siete pasos adelante)

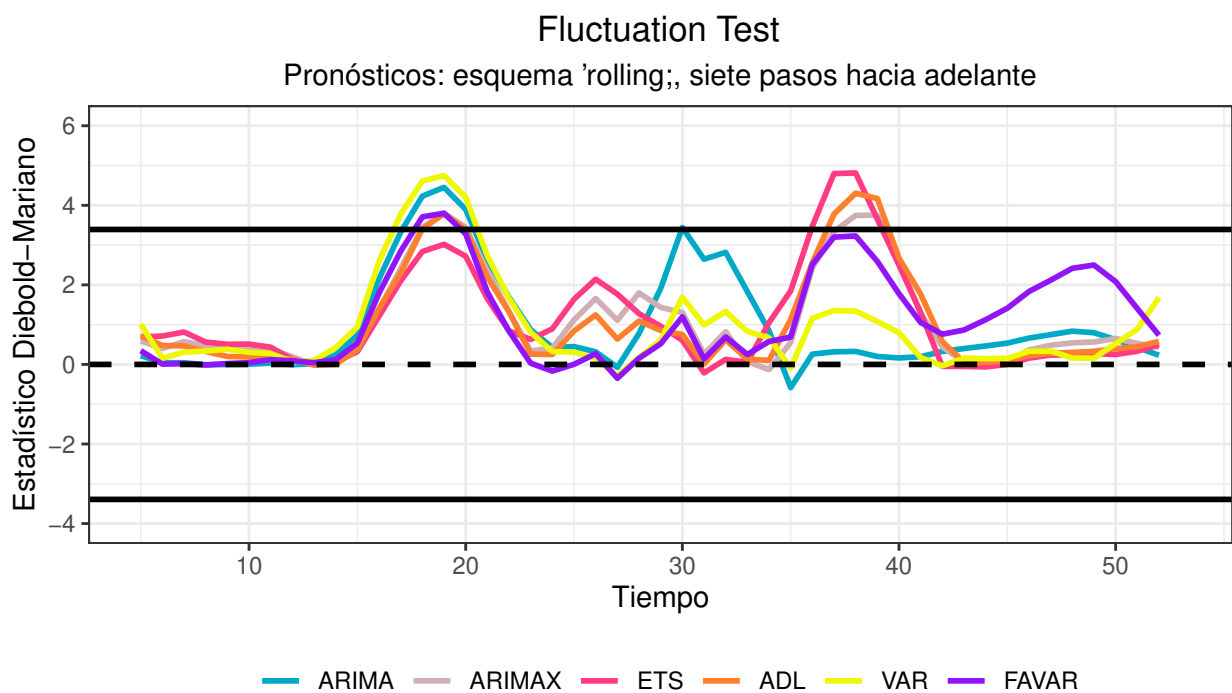


Figura 41: Test de Giacomini Rossi (esquema *rolling*, siete pasos adelante)

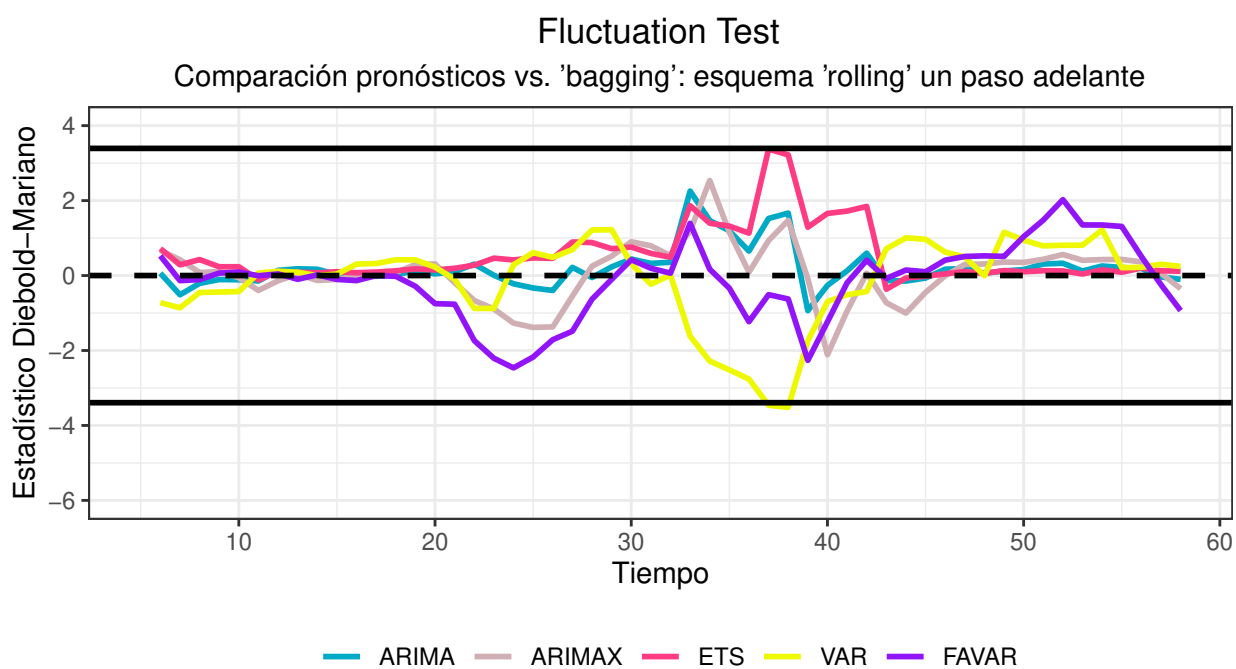
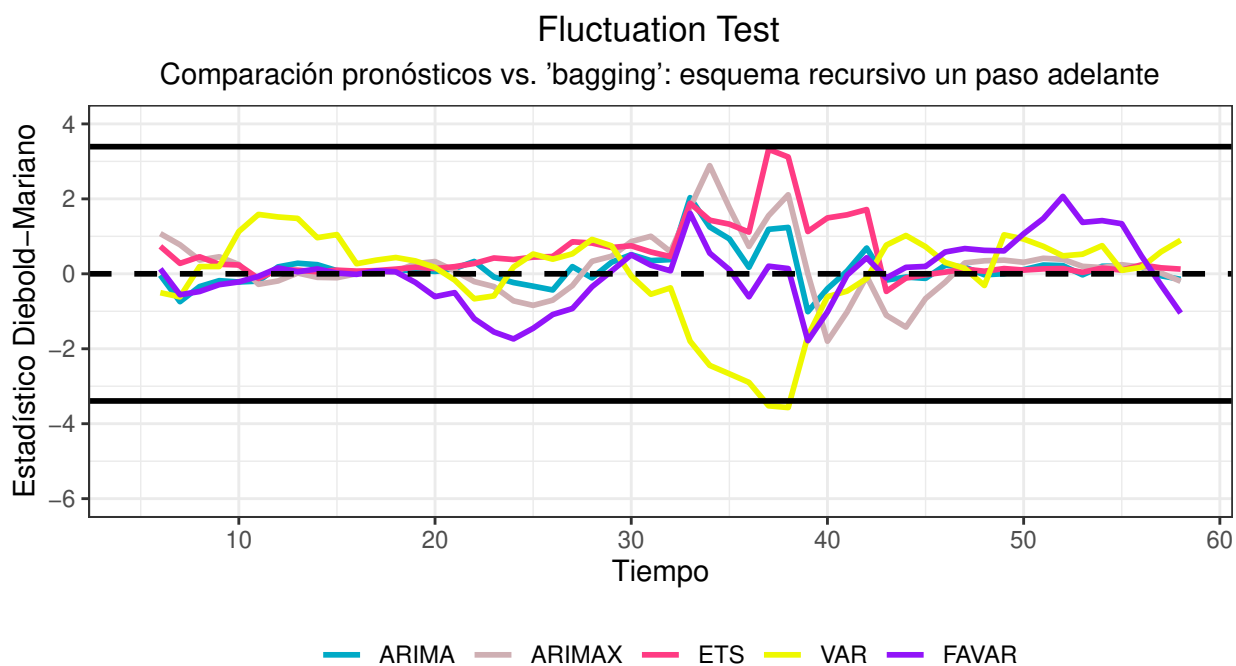
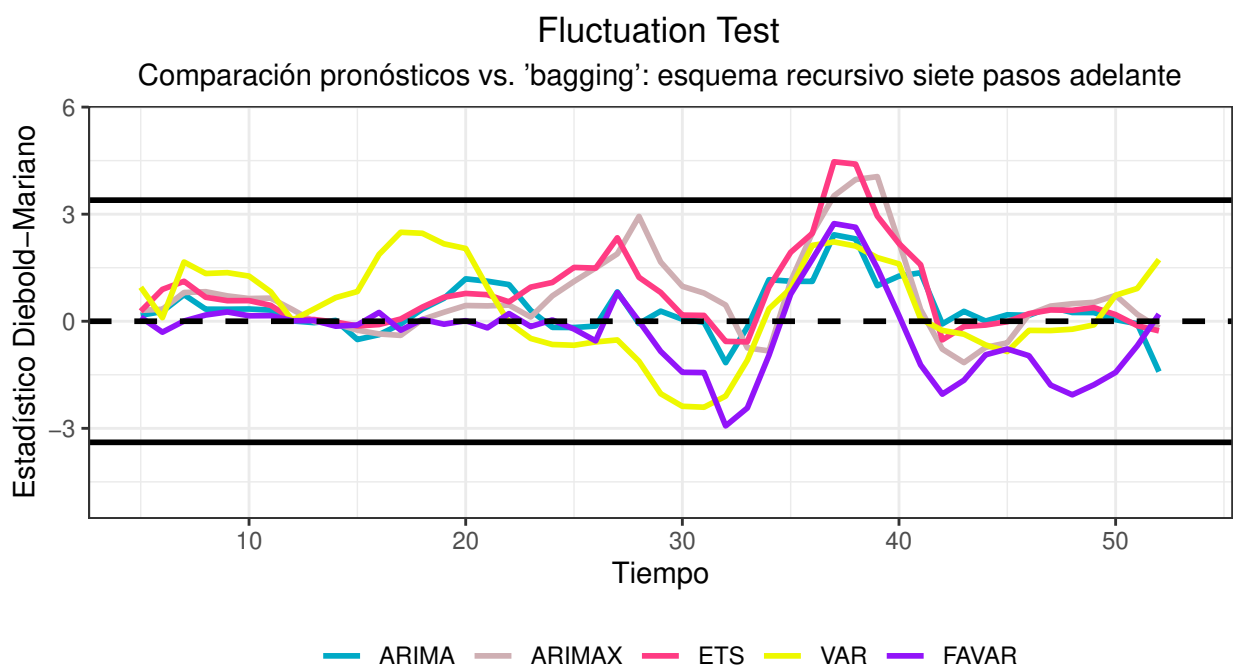


Figura 42: Test de Giacomini Rossi (comparación con modelos *bagged*: *rolling* un paso adelante)



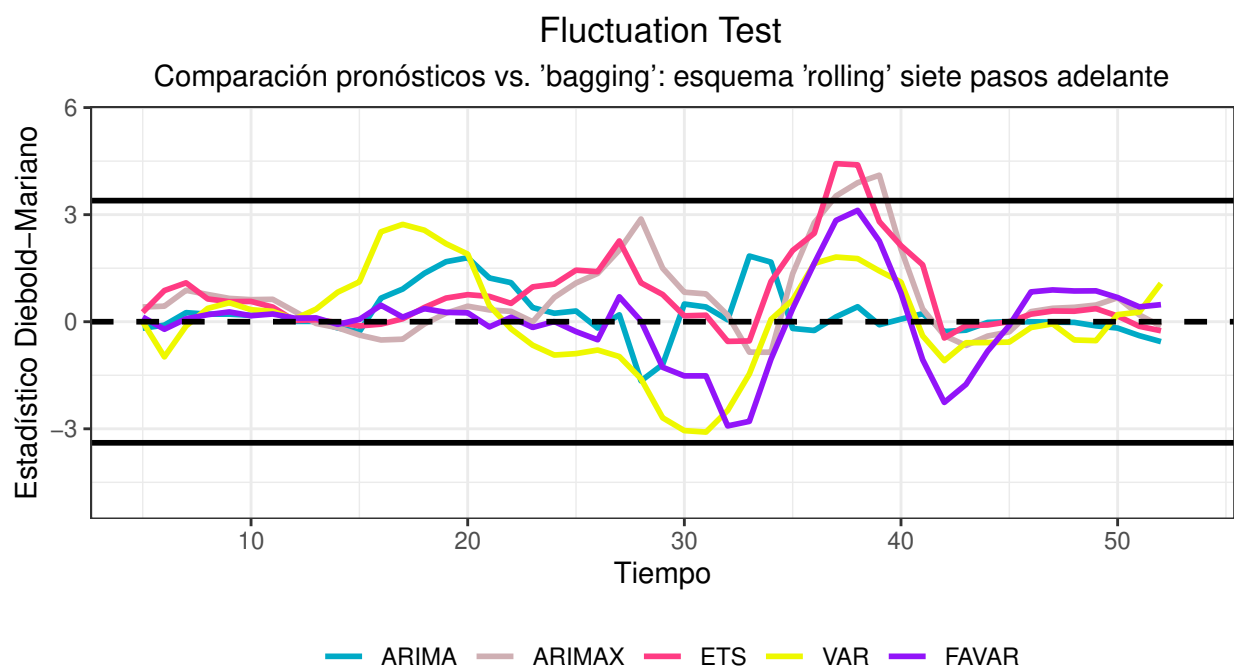
Nota: valores negativos reflejan un peor desempeño de la versión bagging.
Fuente: elaboración propia

Figura 43: Test de Giacomini Rossi (comparación con modelos *bagged*: esquema recursivo un paso adelante)



Nota: valores negativos reflejan un peor desempeño de la versión bagging.
Fuente: elaboración propia

Figura 44: Test de Giacomini Rossi (comparación con modelos *bagged*: esquema recursivo siete pasos adelante)



Nota: valores negativos reflejan un peor desempeño de la versión bagging.
Fuente: elaboración propia

Figura 45: Test de Giacomini Rossi (comparación con modelos *bagged*: esquema *rolling* siete pasos adelante)