

Fundamentos para pronosticar una serie de tiempo estacionaria con información de su propio pasado

WILFREDO BAZÁN RAMÍREZ¹

RECIBIDO: 30/07/2019 ACEPTADO: 23/09/2020 PUBLICADO: 16/10/2020

RESUMEN

Dado que el comportamiento del mercado es volátil, la presente investigación pretende coadyuvar a que inversionistas y organizaciones empresariales puedan realizar pronósticos con certeza y, en consecuencia, con el mínimo error posible, a fin de lograr el éxito en la gestión de sus proyectos y operaciones. Elementos como la tasa de inflación, el tipo de cambio, el precio de las acciones, los resultados económicos financieros, las ventas, entre otras variables, son preocupaciones para los inversionistas. Estos instrumentos financieros, por su estructura de datos, corresponden a las series de tiempo, las cuales toman valores o realizaciones, precisamente, a lo largo del tiempo y, a la vez, están espaciadas cronológicamente. El comportamiento previo es utilizado para pronosticar el valor de la serie, su rendimiento y volatilidad. Y ello debe considerar que pronosticar con las técnicas tradicionales tiene riesgos de imprecisión, por lo que es necesario hacerlo con modelos econométricos por su robustez y precisión, también conocidos como modelos univariados de series de tiempo.

Palabras clave: series de tiempo; estacionariedad; raíz unitaria; ruido blanco; varianza.

INTRODUCCIÓN

Los riesgos negativos que afectan a las organizaciones empresariales requieren ser gestionados, ello con el fin de planificar respuestas que los eviten, transfieran, mitiguen o, de ser necesario, los acepten. Lledó (2017) se interesa en cuantificar la probabilidad de ocurrencia (%) y el impacto (\$) con el objetivo de identificar los riesgos y jerarquizar la atención de los mismos. La preocupación recurrente de las organizaciones, según Berk y DeMarzo (2008), surge cuando invierten en proyectos o instrumentos financieros, como bonos y acciones, o cuando se ven afectadas por la inflación y el tipo de cambio, entre otras variables, por lo cual es importante que los inversionistas y las organizaciones busquen estimar el futuro (Hanke y Wichern, 2010) y, a su vez, sean lo más certeros posible, con miras a planificar sus procedimientos. Esto considerando que un buen pronóstico garantiza la continuidad de sus operaciones en este sistema volátil, competitivo, cambiante y disruptivo.

Por otra parte, Mun (2016) señala lo siguiente:

Pronosticar es el acto de predecir el futuro; ya sea en base a en [sic] datos históricos o en una simple especulación sobre el futuro, en caso de que los datos no existan. Cuando se cuenta con datos históricos, es recomendable hacer una aproximación estadística o cuantitativa; mientras que, si se carece de estos datos, el único recurso es un juicio de valor o un acercamiento cualitativo (p. 429).

Sobre esto, Court y Rengifo (2011) sostienen que, en las finanzas, se trata de pronosticar el rendimiento y la volatilidad esperados de algún instrumento financiero.

¹ Magíster en Finanzas e ingeniero industrial por la Universidad Nacional Federico Villarreal (Lima, Perú). Actualmente, es analista de Telefónica del Perú S. A. A. y docente contratado de la Escuela Profesional de Ingeniería Agroindustrial de la Facultad de Ingeniería Industrial y de Sistemas de la Universidad Nacional Federico Villarreal con certificaciones de PMP y CQRM. (Lima, Perú).
ORCID: <https://orcid.org/0000-0002-2685-8254>
E-mail: wbazan@unfv.edu.pe

En el marco de la era de la digitalización, para este estudio, se buscó pronosticar tanto el rendimiento como la volatilidad de las acciones de Telefónica de España S. A. (TEF), la cual, según Álvarez-Pallete (14 de mayo de 2018), presidente de dicha institución, en una carta dirigida a los accionistas, es una empresa tecnológica de la que analistas financieros reconocen su visión y su gran fundamento. En sintonía con esto, se consideraron los precios de cierre ajustado de las acciones de TEF que se cotizaron en la Bolsa de Nueva York (New York Stock Exchange, NYSE) y se analizó el comportamiento del precio de cierre desde el 2 de enero del año 2000 hasta el 1 de agosto de 2018. El punto aquí es que el precio de cierre de TEF toma un valor o una realización a lo largo de un horizonte de tiempo, por lo que, en las especificaciones financieras, se considera un tipo de muestra de series de tiempo o series cronológicas. Estas, en palabras de Ramón y López (2016), se definen como

[...] una secuencia de datos, observaciones o valores, medidos en determinados momentos del tiempo, ordenados cronológicamente y, normalmente, espaciados entre sí de manera uniforme. El análisis de series temporales comprende métodos que ayudan a interpretar este tipo de

datos, extrayendo información representativa, tanto referente a los orígenes o relaciones subyacentes como a la posibilidad de extrapolar y predecir su comportamiento futuro (p. 12).

En la Figura 1, se observa el comportamiento de los precios de cierre de las acciones de TEF, es decir, se aprecia a niveles, además, aún no está convertido en serie estacionaria. Tomando en cuenta los componentes de las series de tiempo, esta presenta tendencia y tiene un comportamiento volátil o irregular. Vale añadir, siguiendo a Gujarati y Porter (2010), que cuando una serie temporal es no estacionaria, la media, la varianza o ambas son variantes en el tiempo.

Componentes de una serie de tiempo

Newbold, Carlson y Thorne (2008) sostienen que las series temporales tienen cuatro componentes: tendencial, estacional, cíclico e irregular. Cuando se presenta el componente irregular, las series temporales siguen una ruta aleatoria y, como han señalado Gujarati y Porter (2010), del mismo modo se comportan los precios de valores o los instrumentos financieros (por ejemplo, las acciones o las tasas de cambio). Si una serie de tiempo original es

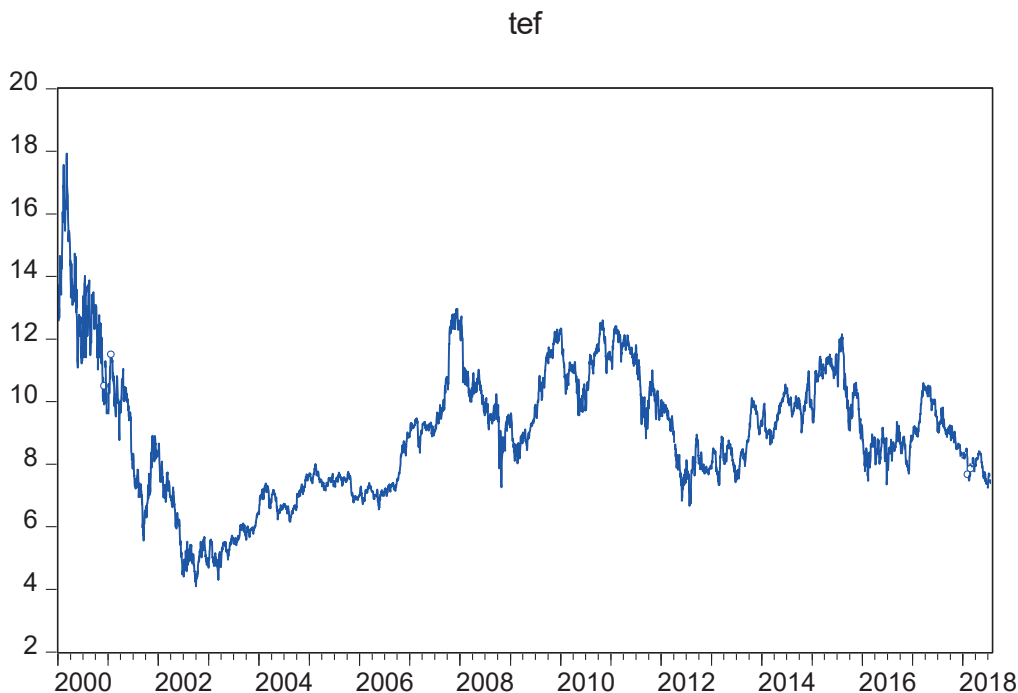


Figura 1. Comportamiento de la serie de precios de cierre de las acciones de TEF.

Fuente: Elaboración propia.

no estacionaria, solo podría estudiarse su comportamiento durante algún periodo en consideración.

Por tanto, cada conjunto de datos perteneciente a la serie de tiempo corresponderá a un episodio particular. En consecuencia, no es posible generalizar para otros periodos. Así, para propósitos de pronóstico, tales series de tiempo (no estacionarias) tienen poco valor práctico (p. 741).

Entonces, es necesario que las series iniciales se conviertan en estacionarias. No obstante, antes de continuar, es importante definir algunos conceptos.

Procesos estocásticos y estacionariedad

Se describe al proceso estocástico como

[...] una secuencia de números aleatorios. El proceso estocástico se escribirá como $\{y_i\}$ para $i = 1, 2, \dots$. Si este índice representa tiempo, el proceso estocástico se llamará serie de tiempo. Si se asigna un posible valor de y por cada i se estará construyendo una posible realización del proceso estocástico (Court y Rengifo, 2011, p. 400).

Herrera (2013) ha cuestionado el uso de los modelos tradicionales con aproximaciones lineales, considerándolos poco eficientes y de aplicabilidad limitada; por ello, resalta el uso del modelo de procesos estocásticos, que permite derivar series temporales con mayor capacidad para identificar los pormenores de datos ocultos. Por otra parte, Ramón y López (2016) señalan que “un proceso estocástico es un concepto matemático que sirve para caracterizar una sucesión de variables aleatorias (Y_t) que evolucionan en función de otra variable, generalmente el tiempo” (p. 63). De acuerdo con estos autores, cada variable del proceso es aleatoria y pueden o no relacionarse entre sí.

Cada una de las variables Y_t que configuran un proceso estocástico tendrá su propia función de distribución con sus correspondientes momentos. Asimismo, cada conjunto de variables tendrá su correspondiente función de distribución conjunta y sus funciones de distribución marginales. Habitualmente, conocer esas funciones de distribución resulta complejo de forma que, para caracterizar un proceso estocástico, basta con especificar la media y la varianza para cada y_t y la covarianza para variables referidas a los

distintos valores de t :

$$\mu_t = E(y_t) \quad (\text{Ec. 1})$$

$$\gamma_0 = \text{Var}(y_t) = E[(y_t - \mu_t)^2] \quad (\text{Ec. 2})$$

$$\gamma_{t,s} = \text{Cov}(y_t, y_s) = E[(y_t - \mu_t)(y_s - \mu_s)] \quad (\text{Ec. 3}) \text{ (p. 63)}$$

Estos autores identifican dos tipos de estacionariedad: proceso estocástico estacionario en sentido fuerte y proceso estocástico estacionario en sentido débil. Para el primer proceso, los cuatro momentos de las distribuciones conjuntas son constantes o invariantes con respecto a un desplazamiento en el tiempo. Para el segundo, solo los dos primeros momentos, vale decir, la esperanza matemática y la varianza de las variables aleatorias, son constantes y no dependen del tiempo, mientras que las covarianzas entre dos variables aleatorias de periodos distintos dependen únicamente del tiempo transcurrido entre ellas mismas.

Villalba y Flores-Ortega (2014), al analizar el comportamiento de la volatilidad del índice de precios y cotizaciones (IPC) del mercado bursátil mexicano, con el fin de estimar la tendencia de los precios de las acciones que lo componen, verificaron la importancia de la estacionariedad en dichas series, las cuales fueron transformadas cuando aplicaron una diferencia logarítmica para convertirlas en rendimientos continuos y estacionarios.

Hossain, Kamruzzaman y Ali (2015) exploraron un modelo estadístico adecuado para resolver la estimación futura del volumen de acciones mediante datos de volúmenes de existencias diarios de la Dhaka Stock Exchange (DSE), utilizando para ello el diagrama de series temporales para explorar los datos. Sin embargo, la trama de la serie de tiempo transgredió la tendencia original y no se pudo eliminar la variación irregular de la serie de datos. Posteriormente, diferenciaron la serie mostrando la media constante, pero no la varianza. Para demostrar el estado estacionario de la serie, utilizaron la prueba de Dickey-Fuller Aumentada (DFA), siendo significativa al 5%. Por tanto, en un primer momento, las series de datos a niveles no son estacionarias y, después de diferenciarlas, se encuentra que las series sí son estacionarias, de acuerdo a las pruebas de DFA para pruebas de raíz unitaria.

Larios-Meño, González-Taranco y Álvarez (2016) se refieren a la estacionariedad de las series de tiempo, y concluyen que, luego de comprobarse su presencia en variables económicas y financieras de uso frecuente en el Perú —como el índice general de

la Bolsa de Valores, la renta de factores, los términos de intercambio y el consumo privado—, las predicciones serán útiles, porque sus resultados tendrán consistencia; en caso contrario, dichos resultados no tendrían credibilidad y, por tanto, serían espurios. Los autores agregan además que

Las series son sometidas a esta verificación mediante correlogramas y el test de raíz unitaria de Dickey-Fuller Aumentado. Para llevar a cabo esta tarea, los datos de las series son ajustados a modelos autorregresivos (AR), de media móvil (MA) y de caminata aleatoria o Random Walk, con el propósito de simular, por ejemplo, condiciones de no estacionariedad que luego son confirmadas por los distintos indicadores obtenidos en esta evaluación. Finalmente, en caso de encontrar series no estacionarias, se propone eliminar esta condición con procesos de diferenciación (s/p).

Este tipo de series, como las producidas por las variables financieras de la TEF, no pueden ser modeladas con las técnicas de pronóstico tradicionales, sino con modelos de series de tiempo univariados o ARIMA (modelos autorregresivos integrados de

medias móviles), conocidos también como metodologías Box-Jenkins; además, Box, Jenkins y Reinsele (2008) sostienen la necesidad de que las series sean estables en el tiempo. La metodología indicada consiste en cuatro pasos: la identificación, la estimación, la validación y el pronóstico.

En la identificación, se verifica si la serie es estacionaria con pruebas de raíz unitaria (RU) y si tiene ruido blanco o memoria para poder pronosticar. Dado que la serie de precios de cierre de las acciones de TEF no es estacionaria, se tiene que diferenciarla para estabilizarla. En la Figura 2, se puede apreciar la serie estabilizada.

Esta investigación tiene por objetivo pronosticar series de tiempo univariadas a partir de su propio comportamiento pasado, para lo cual se necesita estabilizar, vale decir, convertir en estacionarias tanto a la media como a la varianza, a fin de que puedan ser modeladas con las metodologías de Box y Jenkins. Tomando el caso específico de TEF, este estudio busca determinar cómo influye la diferenciación en el rendimiento constante de sus acciones, resolviendo, para este propósito, cómo influye la diferenciación del precio de cierre en el rendimiento constante de sus acciones, en su volatilidad constante y en su autocorrelación.

RTEF01

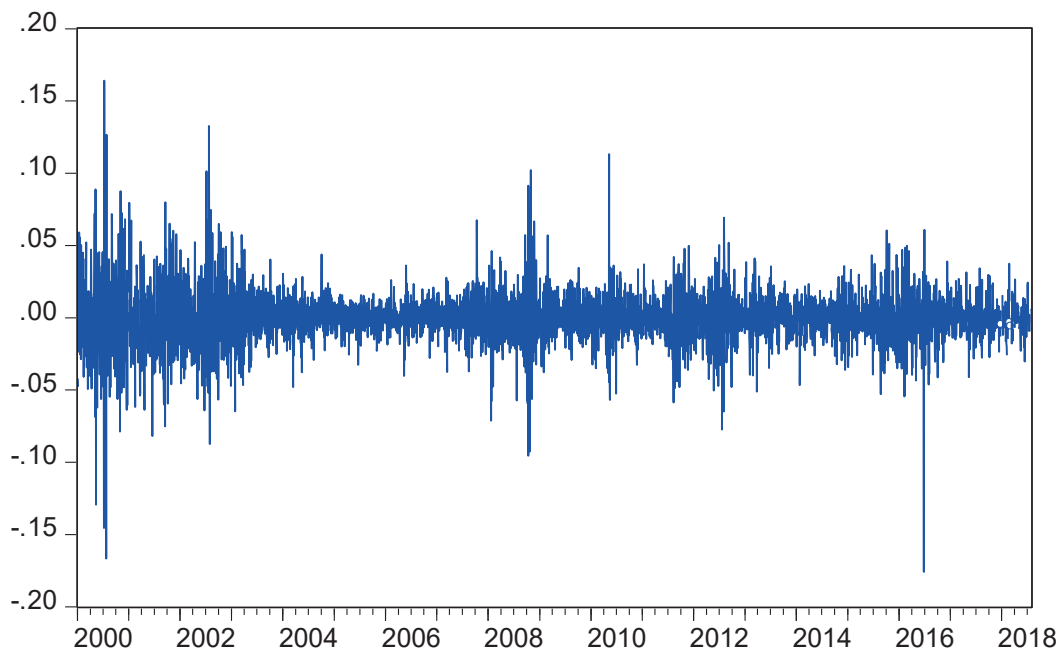


Figura 2. Estabilización de la media y la varianza de las diferencias de los retornos de TEF.

Fuente: Elaboración propia.

El estudio se justifica en su objetivo de evitar o reducir los riesgos negativos que afectan a inversionistas y organizaciones empresariales que invierten en proyectos y/o instrumentos financieros, como, en este caso, las acciones de TEF, una empresa tecnológica que posee gran fortaleza. Por esta razón, es conveniente pronosticar el precio, el rendimiento y la volatilidad de sus acciones que se cotizan en la NYSE, debido a que su movimiento bursátil es mayor comparado al de la Bolsa de Valores de Lima.

METODOLOGÍA

Este análisis pretende explicar el comportamiento de una serie temporal, mediante la estabilización de su media y su varianza, con el fin de pronosticar los retornos y la volatilidad diaria de los precios de cierre ajustado de las acciones de la compañía Telefónica de España S. A. (TEF) que operan en la Bolsa de Nueva York (NYSE).

La unidad de análisis la constituyen las acciones de TEF. La población es representada por las plazas donde la empresa cotiza sus acciones, las cuales pueden ser consultadas en su propio sitio web (https://www.telefonica.com/es/web/shareholders-investors/la_accion/presencia-en-bolsas). De acuerdo con Telefónica (2018), esta tiene presencia en las siguientes plazas: Buenos Aires, Lima, Londres, Madrid y Nueva York. El tamaño de muestra abarca las acciones de TEF entre el 3 de enero del 2000 y el 1 de agosto de 2018, tomadas del sitio web <https://es.finance.yahoo.com/quote/TEF/history?p=TEF>, de la plaza de la NYSE. Para la selección de la muestra, se dificultó el acceso a todas las observaciones, pero investigaciones científicas recomiendan la mayor cantidad de observaciones. En este caso, se investigó a partir de 4848 observaciones del precio de cierre de las cotizaciones diarias que están entre el 3 de enero del 2000 y el 1 de agosto de 2018, obteniendo, por tanto:

- Muestras no aleatorias
- Muestras por conveniencia o intencional

Los criterios para la selección de la muestra fueron considerados por la fácil accesibilidad para la obtención de los datos del sitio web <https://es.finance.yahoo.com/quote/TEF/history?p=TEF>.

Para modelar el comportamiento del activo financiero de TEF, solo bastó realizar una diferencia a la serie original para convertirla en estacionaria. Entonces, la serie estacionaria fue el resultado de la diferencia de los logaritmos de los precios de cierre de un periodo contra el precio de cierre de un periodo anterior, de acuerdo a la siguiente fórmula:

$$r_n = \text{Ln} \left(\frac{P_t}{P_{t-1}} \right) = \text{Ln}(P_t) - \text{Ln}(P_{t-1}). \quad (\text{Ec. 4})$$

Para obtener los rendimientos de este activo financiero, mediante el *software* EViews 10, se generó una variable con el comando `GENR rtef01 = dlog(tef01)`, que es el cálculo de la diferencia logarítmica de la serie original, cuya variable fue denominada *tef01*. En razón de la Figura 2, primero, se realizó un análisis gráfico y se concluyó que la serie *rtef01* es estacionaria, pero fue necesario comprobarlo matemáticamente con las pruebas de raíz unitaria (RU).

Ahora, si $x_t = \alpha x_{t-1} + \varepsilon_t$ es un proceso estacionario,

$$\sigma_{x_t}^2 = \alpha^2 \sigma_{x_{t-1}}^2 + \sigma_{\varepsilon_t}^2 \rightarrow \sigma_{x_t}^2 - \alpha^2 \sigma_{x_t}^2 = \sigma_{\varepsilon_t}^2 \rightarrow \sigma_{x_t}^2 (1 - \alpha^2) = \sigma_{\varepsilon_t}^2 \quad (\text{Ec. 5})$$

$$\sigma_{x_t}^2 = \frac{\sigma_{\varepsilon_t}^2}{(1 - \alpha^2)} \quad (\text{Ec. 6})$$

Si $\alpha = 1$, el proceso no es estacionario (raíz unitaria).

Si $\alpha > 1$, el proceso es explosivo.

Si $\alpha < 1$, el proceso es estacionario.

Entonces, se procedió a realizar las pruebas de raíz unitaria con la prueba de Dickey-Fuller Aumentada (DFA), la que, a decir de Bello (2018), es uno de los exámenes más utilizados. Sobre esta prueba, se agrega, «en un nivel formal, la estacionariedad [la cual] se verifica averiguando si la serie de tiempo contiene una raíz unitaria» (Gujarati y Porter, 2010, p. 768), y, en este caso:

$H_0: \phi = 0$; x_t tiene raíz unitaria (al menos una) $\rightarrow x_t$ es serie no estacionaria.

$H_1: \phi < 0$; x_t no tiene raíz unitaria $\rightarrow x_t$ es serie estacionaria. La serie es $I(0)$.

$I(d)$, expresa la cantidad de veces que una serie temporal deberá diferenciarse para convertirse en estacionaria.

RESULTADOS

Se presentan, a continuación, los resultados, aplicando las pruebas de la metodología Box-Jenkins o ARIMA. Esta metodología, formalizada por George Box y Gwilym Jenkins en 1976, se basa en que las series temporales que intentan pronosticar tienen como punto de apoyo a los procesos estocásticos caracterizados mediante un modelo. Meléndez (2017) sostiene que, al intentar pronosticar con

este modelo, se presentan los siguientes pasos a través de escenarios de ensayo y error:

0. Evaluación de la estacionariedad
1. Identificación
2. Estimación
3. Validación
4. Pronóstico

Interpretación 1

De acuerdo a la Figura 2, se buscó comprobar que la serie diferenciada fuera estacionaria, para esto, se realizó la prueba de RU con la prueba de Dickey-Fuller Aumentada (DFA). En este caso, el objetivo fue rechazar la hipótesis nula, vale decir, que la serie tuviera al menos una raíz unitaria. Las pruebas indicaron la no presencia de raíz unitaria, según la comparación entre el estadístico de la prueba de Dickey-Fuller Aumentada y los distintos

valores críticos de MacKinnon. Se dice que una serie presenta RU si algunos de los valores críticos en valores absolutos de MacKinnon son mayores que el estadístico de la prueba DFA en valores absolutos. En este caso, el valor de la DFA es igual a $|42.39815|$, por lo que es mayor que cualquiera de los valores críticos, al 1% = $|2.565457|$, 5% = $|1.940892|$ y 10% = $|1.616654|$, lo que significa que la serie no presenta raíz unitaria y, por tanto, es una serie estacionaria. Los resultados se muestran en la Figura 3 y se comprueba con la primera diferencia que la serie es estacionaria.

Interpretación 2

Dentro de la fase de identificación, además de la comprobación de que la serie sea estacionaria, se verificó que la serie no tuviera ruido blanco, es decir, que tenga memoria. Este es un proceso aleatorio, ya que su media es igual a cero, su varianza es constante y la autocovarianza es igual a cero; con este proceso que no tiene memoria no podría pronosticarse una serie de tiempo con la

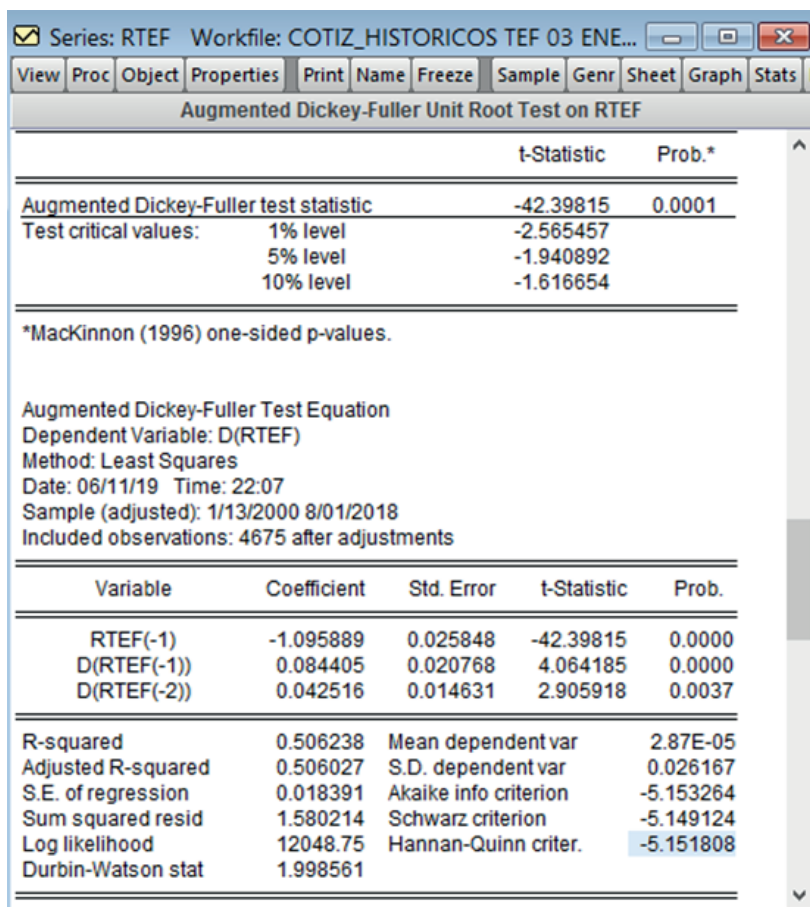


Figura 3. Prueba de raíz unitaria con la prueba de Dickey-Fuller Aumentada (DFA).

Fuente: Elaboración propia.

metodología Box-Jenkins o ARIMA, sino con procesos estocásticos, como el movimiento geométrico browniano (MGB) o con los modelos de la familia ARCH-GARCH. Cuando una serie tiene memoria, se usará su pasado para pronosticar la serie.

$\{\varepsilon_t\}_{t=1}^T \rightarrow$ es un proceso ruido blanco si

$$(\varepsilon_t) = 0, \forall t \quad (\text{Ec. 7})$$

$$E(\varepsilon_t^2) = V(\varepsilon_t) = \sigma_\varepsilon^2, \forall t \quad (\text{Ec. 8})$$

$$E(\varepsilon_t, \varepsilon_{t+\tau}) = \gamma_\tau = 0, \forall t, \tau \neq 0 \quad (\text{Ec. 9})$$

Es este caso, se buscó validar que la serie tuviera memoria con correlogramas, el estadístico Ljung-Box y el p -valor, tal como se aprecia en los resultados contenidos en la Figura 4.

Validación de ruido blanco (PH)

Estadístico Q-prueba conjunta:

$$H_0: \rho_1 = \rho_2 = \rho_3 = \dots = \rho_k = 0; \text{ ruido blanco}$$

$$H_1: \exists \rho_i \neq 0; i = 1, \dots, k; \text{ la serie no es ruido blanco}$$

Estadístico LB, pero para muestras pequeñas Ljung-Box:

$$LB = Q_k = n(n + 2) \sum_{i=1}^k \frac{r_i^2}{n-1} \quad (\text{Ec. 10})$$

El objetivo fue rechazar la hipótesis nula (H_0), que significa ruido blanco, si no se rechaza H_0 , no se podría pronosticar con la metodología Box-Jenkins. Bello (2018) agrega que una serie se considera ruido blanco si los p -valores son mayores que el nivel de significancia de la prueba y que, en la práctica, siempre que no sean los primeros rezagos, se permite que sea

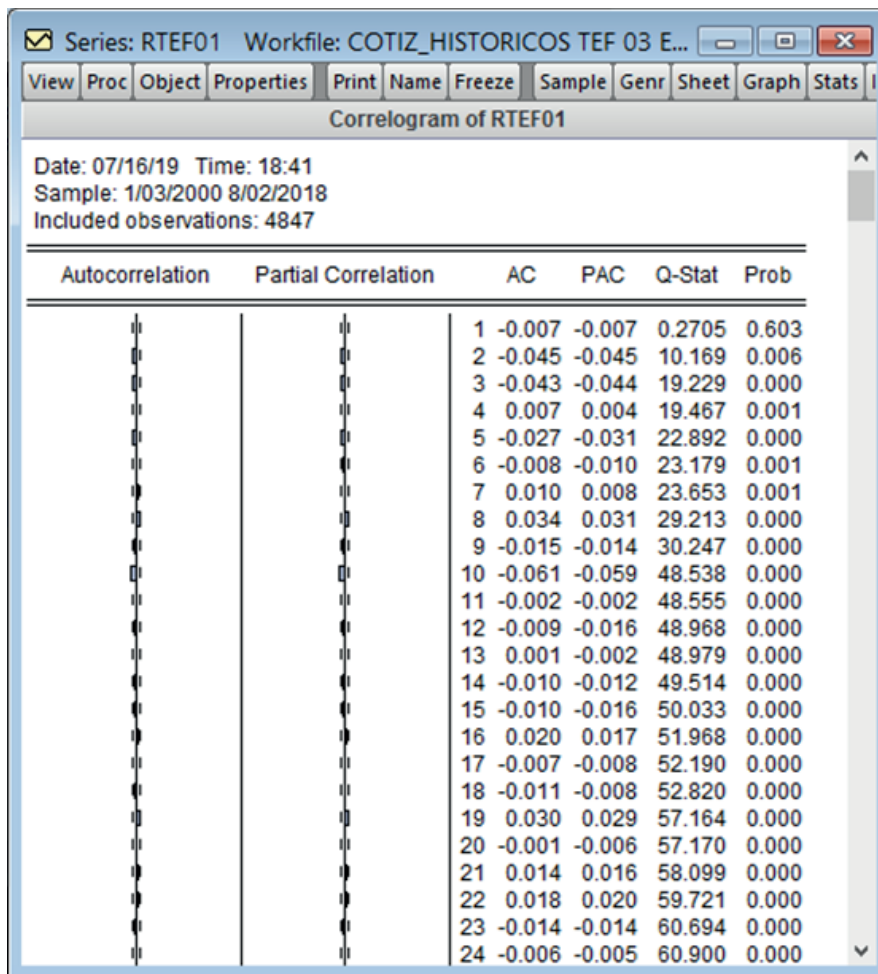


Figura 4. Prueba de correlograma de la serie RTEF01 para la validación de no ruido blanco.

Fuente: Elaboración propia.

mayor el p -valores en uno de cada 20 rezagos para considerar la serie ruido blanco, tal como se dio en el primer rezago. En este caso, no se rechaza la hipótesis nula, pues la varianza no es homocedástica o constante.

Interpretación 3

De los mismos resultados de la Figura 4, se comprobó que la autocovarianza tampoco es constante o invariante en el tiempo, en otras palabras, que la serie no depende de su pasado. Las autocovarianzas entre el valor de una serie y su propio rezago no dependen de la distancia del tiempo que los separa; para corregirla, se procedió a centrar el modelo para determinar si los residuos o errores tienen ruido blanco con la ecuación $dlog(TEF01) C AR(10) MA(2) MA(3)$, cuyos resultados están contenidos en la Figura 5 y muestran que los coeficientes son significativos, pues están por debajo del 5%.

Estos modelos necesitan que los residuos sean ruido blanco, por lo cual se realizó un diagnóstico de los residuos con base en su correlograma y se verificó,

como se ve en la Figura 6, que el comportamiento de los residuos constituye ruido blanco, porque su p -valor es mayor al 5% y, por tanto, pueden pronosticarse por los modelos ARCH-GARCH, que, a decir de Mansilla (2 de abril de 2020), junto a los modelos ARIMA de Box y Jenkins, son de la misma familia.

DISCUSIÓN

Para pronosticar la rentabilidad y volatilidad de la serie financiera de TEF con la metodología Box-Jenkins, primero, es necesario diferenciar su precio de cierre en el tiempo t con su periodo inmediato anterior ($t-1$) para estabilizar o hacer constante tanto el primer como el segundo momento (media y varianza). Se aprecia que la serie es estacionaria a lo largo del espacio comprendido entre el 3 de enero del 2000 y el 1 de agosto de 2018, es decir, su media es constante y se comprueba tomando 522 retornos del periodo 2005-2006, tiempo en el que hubo baja volatilidad en comparación a otras 522 observaciones del periodo 2007-2008, en el que existió gran volatilidad.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000120	0.000227	-0.529692	0.5963
AR(10)	-0.060549	0.008042	-7.528982	0.0000
MA(2)	-0.046635	0.010152	-4.593728	0.0000
MA(3)	-0.046446	0.009414	-4.933813	0.0000
SIGMASQ	0.000333	2.98E-06	111.6110	0.0000

R-squared	0.007820	Mean dependent var	-0.000121
Adjusted R-squared	0.007000	S.D. dependent var	0.018315
S.E. of regression	0.018251	Akaike info criterion	-5.168132
Sum squared resid	1.612899	Schwarz criterion	-5.161441
Log likelihood	12529.97	Hannan-Quinn criter.	-5.165783
F-statistic	9.540690	Durbin-Watson stat	2.020548
Prob(F-statistic)	0.000000		

Figura 5. Resultados de la ecuación $DLOG(TEF01) C AR(10) MA(2) MA(3)$.

Fuente: Elaboración propia.

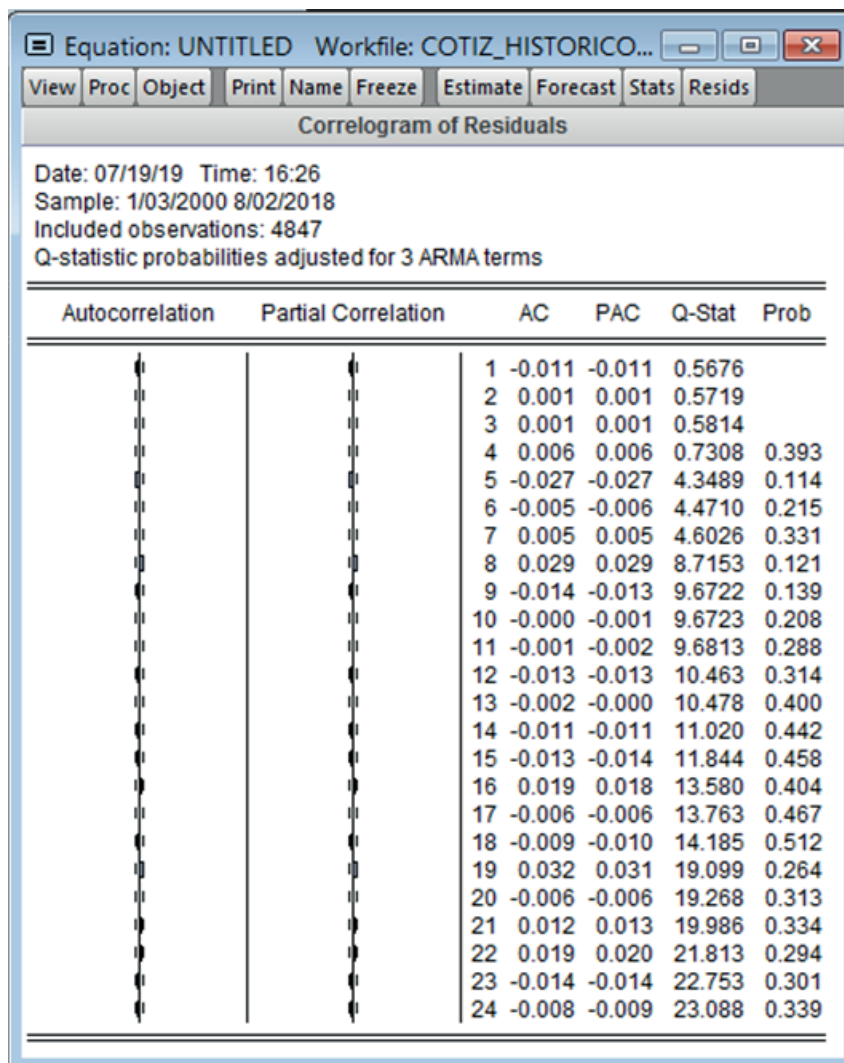


Figura 6. Correlograma de los residuos de la ecuación $DLOG(TEF01) C AR(10) MA(2) MA(3)$.

Fuente: Elaboración propia.

Sin embargo, la varianza no es homocedástica, precisamente porque hubo lapsos de alta volatilidad en los retornos y ya no podrían ser modelados con la metodología Box-Jenkins, sino con los modelos de la familia ARCH-GARCH. En la Tabla 1, se muestran los resultados de la comparación de medias y la prueba de Levene para la igualdad de varianzas.

A continuación, se presentan las siguientes pruebas estadísticas llevadas a cabo.

1. Prueba de hipótesis para la igualdad de medias

H_0 : Las medias de los periodos 2005-2006 y 2007-2008 son iguales.

H_1 : Las medias de los periodos 2005-2006 y 2007-2008 no son iguales.

Nivel de significancia

Nivel de significancia (alfa) $\alpha = 5\%$ o 0.05.

Estadístico de prueba

Prueba t para 2 muestras independientes.

Valor de P

0.0000%.

La prueba demuestra una significancia bilateral de $0.828 > 0.05$, por lo tanto, no se rechaza la hipótesis de igualdad de medias y se acepta la hipótesis nula de que existe igualdad de medias.

Tabla 1. Pruebas de igualdad de medias y de Levene de igualdad de varianzas.

Prueba de muestras independientes										
		Prueba de Levene de igualdad de varianzas		Prueba t para la igualdad de medias						
		F	Sig.	T	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia	
									Inferior	Superior
Retorno	Se asumen varianzas iguales	80.291	0.000	0.217	1042	0.828	0.00019580271	0.00090108264	-0.00157234060	0.00196394602
	No se asumen varianzas iguales			0.217	742.929	0.828	0.00019580271	0.00090108264	-0.00157316870	0.00196477411

Fuente: Elaboración propia.

Toma de decisiones

Las medias de los grupos a comparar no son diferentes.

Interpretación

Las medias de los 2 grupos son constantes.

2. Prueba de hipótesis para la igualdad de varianza

H_0 : Las varianzas de los periodos 2005-2006 y 2007-2008 son homocedásticas o no son diferentes.

H_1 : Las varianzas de los periodos 2005-2006 y 2007-2008 son heterocedásticas o son diferentes.

Nivel de significancia

Nivel de significancia (alfa) $\alpha = 5\%$ o 0.05.

Estadístico de prueba

Prueba de Levene.

Valor de P

0.0000%.

Lectura del valor P: Con una probabilidad de error del 0.000, las varianzas de los grupos a comparar son heterocedásticas o diferentes

Toma de decisiones

Las varianzas de los grupos a comparar no son homocedásticas o no son diferentes.

Interpretación

Las varianzas de los 2 grupos son heterocedásticas o son diferentes.

CONCLUSIONES

1. La serie original del precio de cierre de las acciones de TEF presenta tendencia y, para removerla, fue necesario diferenciarla con el fin de volverla estacionaria, condición necesaria para que, con la metodología Box-Jenkins, el pronóstico sea estable a lo largo del periodo, pues es necesario que tanto el primer como el segundo momento de la estadística sean invariantes, para lo cual se realizó la prueba de raíz unitaria (RU).
2. Con los correlogramas, el estadístico Ljung-Box y el p -valor se intentó validar la variable rtef01, que es el cálculo de la diferencia logarítmica de la serie original, así como que la serie tuviese memoria; pero se concluyó que no es homocedástica o, dicho de otra manera, su varianza no es constante a lo largo del tiempo, por lo que no podrá pronosticarse con la metodología Box-Jenkins.

3. Empero, los residuos sí tienen ruido blanco, es por ello que podrían modelarse según los modelos de la familia ARCH-GARCH, dado que su varianza es heterocedástica.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Álvarez-Pallete, J. (14 de mayo de 2018). [Carta a nuestros accionistas: El ilusionante reto de reinventar Telefónica]. Copia en posesión de Wilfredo Bazán Ramírez.
- [2] Bello, M. (2018). *Modelos econométricos con EViews: Modelos de regresión lineal y series de tiempo*. [Apuntes de clase]. Recuperado de <https://www.software-shop.com/formacion/formacion-info/4404>.
- [3] Berk, J. y DeMarzo, P. (2008). *Finanzas Corporativas*. México D. F., México: Pearson Educación.
- [4] Box, G., Jenkins, G. y Reinsel, G. (2008) *Time Series Analysis: Forecasting and Control*. Nueva York, Estados Unidos: Wiley.
- [5] Court, E. y Rengifo, E. (2011). *Estadísticas y Econometría Financiera*. Buenos Aires, Argentina: Cengage Learning Argentina.
- [6] Gujarati, D. y Porter, D. (2010). *Econometría*. México D. F., México: McGraw-Hill.
- [7] Hanke, J. y Wichern, D. (2010). *Pronósticos en los negocios*. México D. F., México: Pearson Educación.
- [8] Herrera, J. (2013). *Modelo estocástico a partir de razonamiento basado en casos para la generación de series temporales*. (Tesis doctoral). Universidad Nacional de San Agustín de Arequipa, Perú. Recuperado de http://repositorio.concytec.gob.pe/bitstream/20.500.12390/346/6/2013_Herrera_Modelo-estocastico-razonamiento.pdf.
- [9] Hossain, A., Kamruzzaman, M. y Ali, A. (2015). ARIMA with GARCH Family Modeling and Projection on Share Volume of DSE. *Economics World*, 3(7-8), 171-184.
- [10] Larios-Meño, J., González-Taranco, C. y Álvarez, V. (2016). *Investigación en economía y negocios. Metodología con aplicaciones en E-views*. Lima, Perú: Fondo Editorial de la Universidad San Ignacio de Loyola.
- [11] Lledó, P. (2017). *Director de proyectos. Cómo aprobar el examen PMP® sin morir en el intento*. Estados Unidos: Pablo Lledó.
- [12] Mansilla, F. (2 de abril del 2020). *Modelos de volatilidad condicional con EViews*. [Apuntes de clase]. Recuperado de <https://www.software-shop.com/formacion/formacion-info/5391>.
- [13] Meléndez, J. (2017). *Entrenamiento especializado en modelos econométricos de series de tiempo en EViews* [Apuntes de clase]. Recuperado de <https://software-shop.com/formacion/formacion-info/2979>.
- [14] Mun, J. (2016). *Modelación de riesgos. Aplicación de la simulación de Monte Carlo, análisis de opciones reales, pronóstico estocástico, optimización de portafolio, análisis de datos, inteligencia de negocios y modelación de decisiones* (2 volúmenes). California, Estados Unidos: IIPER Press.
- [15] Newbold, P., Carlson, W. y Thorne, B. (2008). *Estadística para Administración y Economía*. Madrid, España: Pearson Educación.
- [16] Ramón, N. y López, J. (2016). *Econometría. Series temporales y modelos de ecuaciones simultáneas*. Elche, España: Universidad Miguel Hernández.
- [17] Telefónica (2018). Presencia en bolsas. *Telefónica*. Recuperado de https://www.telefonica.com/es/web/shareholders-investors/la_accion/presencia-en-bolsas.
- [18] Villalba, F. y Flores-Ortega, M. (2014). Análisis de la volatilidad del índice principal del mercado bursátil mexicano, del índice de riesgo país y de la mezcla mexicana de exportación mediante un modelo GARCH trivariado asimétrico. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 17, 3-22. Recuperado de <https://www.upo.es/revistas/index.php/RevMetCuant/article/view/2191>.

Basics for Forecasting a Stationary Time Series Using Information from Its Past

WILFREDO BAZÁN RAMÍREZ¹

RECIBIDO: 30/07/2019 ACEPTADO: 23/09/2020 PUBLICADO: 16/10/2020

ABSTRACT

Since market behavior is volatile, this research intends to help investors and business organizations make forecasts with certainty and, as a consequence, with the least possible error in order to succeed in the management of their projects and operations. Elements such as inflation rate, exchange rate, stock prices, economic and financial results, sales, among other variables, are causes of concern for investors. Due to their data structure, these financial instruments correspond to time series, which take values or realizations along time and are spaced over time. The previous behavior of the series is used to forecast its value, return and volatility. It must be taken into consideration that forecasting using traditional techniques might result in imprecisions, so it is necessary to forecast using econometric models because of their robustness and precision. These are also known as univariate time series models.

Keywords: time series; stationarity; unit root; white noise; variance.

INTRODUCTION

Negative risks affecting business organizations need to be managed in order to plan responses that avoid, transfer, mitigate or, if necessary, accept them. Lledó (2017) is interested in quantifying the probability of occurrence (%) and impact (\$) in order to identify risks and prioritize them. According to Berk and De-Marzo (2008), the recurrent concern of business organizations arises when they invest in projects or financial instruments, such as bonds and shares, or when they are affected by inflation and exchange rate, among other variables. Thus, it is important that investors and business organizations aim to estimate the future (Hanke & Wichern, 2010) and, in turn, be as accurate as possible, focusing on planning their procedures and considering that a good forecast guarantees the continuity of their operations in this volatile, competitive, changing and disruptive system.

Moreover, Mun (2016) notes the following:

Forecasting is the act of predicting the future, whether it is based on historical data or speculation about the future when no history exists. When historical data exist, a quantitative or statistical approach is best, but if no historical data exist, then a qualitative or judgmental approach is usually the only recourse (p. 429).

About this, Court and Rengifo (2011) maintain that, when it comes to finance, it is a matter of predicting the expected return and volatility of some financial instrument.

In the context of the digital age, this research sought to forecast the return and volatility of Telefónica de España S. A. (TEF) stock shares, which, according to its president Álvarez-Pallete (May

¹ Master in Finance and Industrial Engineering from the Universidad Nacional Federico Villarreal (Lima, Peru). Currently working as analyst for Telefónica del Perú S. A. and professor on a contract basis at the Agro-Industrial Engineering Academic Program of the Industrial and Systems Engineering School of the Universidad Nacional Federico Villarreal. Certified in PMP and CQRM. (Lima, Peru).
ORCID: <https://orcid.org/0000-0002-2685-8254>
E-mail: wbazan@unfv.edu.pe

14, 2018) in a letter to shareholders, is a technology company whose vision and great foundation are recognized by financial analysts. In line with this, the adjusted closing prices of TEF shares listed on the New York Stock Exchange (NYSE) were considered, and the behavior of the closing price from January 2, 2000 to August 1, 2018 was analyzed. The point here is that the closing price of TEF takes a value or realization over a time horizon, so that in financial terms it is considered a sample of time series. These, in the words of Ramón and López (2016), are defined as

"[...] una secuencia de datos, observaciones o valores, medidos en determinados momentos del tiempo, ordenados cronológicamente y, normalmente, espaciados entre sí de manera uniforme. El análisis de series temporales comprende métodos que ayudan a interpretar este tipo de datos, extrayendo información representativa, tanto referente a los orígenes o relaciones subyacentes como a la posibilidad de extrapolar y predecir su comportamiento futuro (p. 12)." [a sequence of data, observations or values, measured at particular moments in time, chronologically arranged and usually evenly spaced.

Time series analysis involves methods that help interpret this type of data, extracting representative information regarding their origins or underlying relationships and the possibility of extrapolating and predicting their future behavior] (p. 12).

In Figure 1, the closing prices behavior of TEF shares is shown at levels and it has not yet been converted into a stationary series. Considering the components of a time series, this series presents a trend and has a volatile or irregular behavior. It is worth adding, following Gujarati and Porter (2010), that when a time series is non-stationary, the mean, the variance, or both, are variable in time.

Components of a time series

Newbold, Carlson and Thorne (2008) state that time series have four components: a trend component, a seasonal component, a cyclical component and an irregular component. When the irregular component is presented, time series follow a random path and, as Gujarati and Porter (2010) have pointed out, so do stock prices or financial instruments (e.g. shares or exchange rates). If an original time series is non-stationary, its behavior could only be studied during a period in consideration.

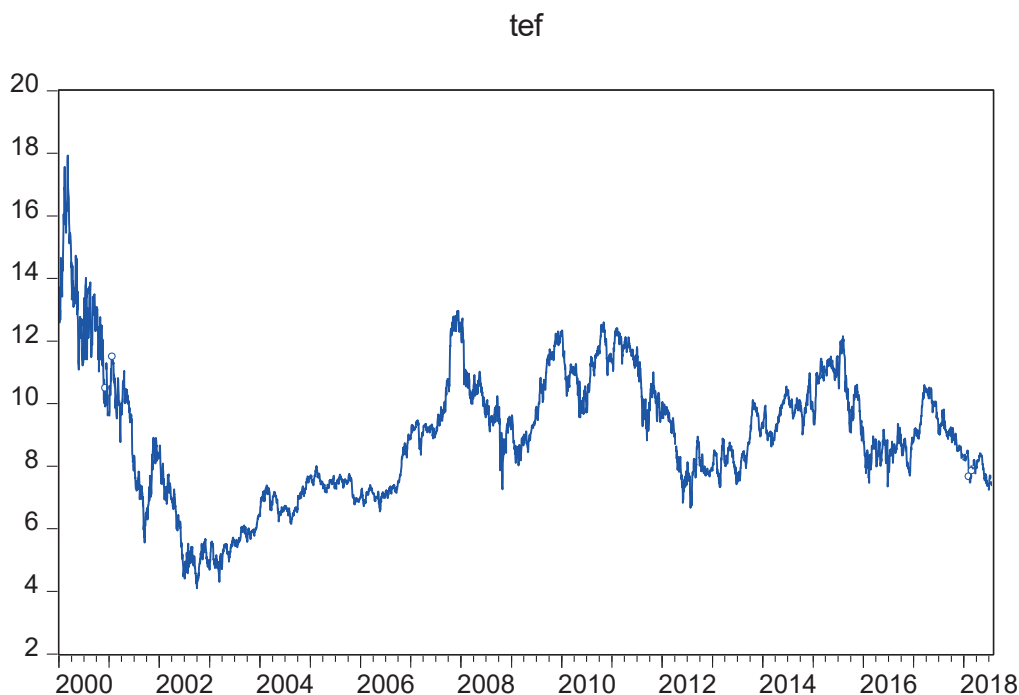


Figure 1. Behavior of the closing prices series of TEF shares.

Source: Prepared by the author.

Each set of time series data will therefore be for a particular episode. As a consequence, it is not possible to generalize it to other time periods. Therefore, for the purpose of forecasting, such (nonstationary) time series may be of little practical value. (p. 741)

Thus, the initial series must become stationary. However, before continuing, it is important to define some concepts.

Stochastic processes and stationarity

The stochastic process is described as

[...] una secuencia de números aleatorios. El proceso estocástico se escribirá como $\{y\}$ para $i = 1, 2, \dots$. Si este índice representa tiempo, el proceso estocástico se llamará serie de tiempo. Si se asigna un posible valor de y por cada i se estará construyendo una posible realización del proceso estocástico” [a sequence of random numbers. The stochastic process will be represented as $y\}$ ($i = 1, 2, \dots$); if this index represents time, the stochastic process will be called a time series. If a possible value is assigned to y for each i , a possible realization of the stochastic process will be constructed (Court & Rengifo, 2011, p. 400)].

Herrera (2013) has questioned the use of traditional models using linear approximations, considering them not very efficient and of limited applicability; therefore, he highlights the use of the stochastic process model, which makes it possible to more ably derive time series to identify hidden data details. On the other hand, Ramón and López (2016) point out that *“un proceso estocástico es un concepto matemático que sirve para caracterizar una sucesión de variables aleatorias (Y_t) que evolucionan en función de otra variable, generalmente el tiempo”* [a stochastic process is a mathematical concept that serves to characterize a succession of random variables (Y_t) that evolve depending on another variable, generally, time] (p. 63). According to these authors, each variable of the process is random and may or may not be related to each other.

“Cada una de las variables Y_t que configuran un proceso estocástico tendrá su propia función de distribución con sus correspondientes momentos. Asimismo, cada conjunto de variables tendrá su correspondiente función de distribución

conjunta y sus funciones de distribución marginales. Habitualmente, conocer esas funciones de distribución resulta complejo de forma que, para caracterizar un proceso estocástico, basta con especificar la media y la varianza para cada y_t y la covarianza para variables referidas a los distintos valores de t ” [Each Y_t variable that makes up a stochastic process will have its own distribution function with its corresponding moments. Moreover, each set of variables will have its corresponding joint distribution function and its marginal distribution functions. Usually, knowing these distribution functions is complex, so in order to characterize a stochastic process, it is enough to specify the mean and variance for each y_t and the covariance for variables referred to the different values of t]:

$$\mu_t = E(y_t) \quad (\text{Eq. 1})$$

$$\gamma_0 = \text{Var}(y_t) = E[(y_t - \mu_t)^2] \quad (\text{Eq. 2})$$

$$\text{Cov}(y_t, y_s) = E[(y_t - \mu_t)(y_s - \mu_s)] \quad (\text{Eq. 3}) \text{ (p. 63)}$$

These authors identify two types of stationarity: strong stationary stochastic process and weak stationary stochastic process. In the first process, the four moments of the joint distributions are constant or invariant to a change in time. In the second, only the first two moments, the expected value and the variance of the random variables, are constant and not time-dependent, while the covariances between two random variables of different periods depend only on the time elapsed between them.

When analyzing the volatility behavior of the Price and Quotation Index (IPC) of the Mexican stock market, with the purpose of estimating the trend of its stock prices, Villalba and Flores-Ortega (2014) verified the importance of the stationarity in said series. These were transformed applying logarithmic differentiation to convert them into continuous and stationary returns.

Hossain, Kamruzzaman and Ali (2015) explored a suitable statistical model to resolve the future estimation of stock volume by using daily stock volume data from the Dhaka Stock Exchange (DSE); for this, the time series chart was used to explore the data. However, the time series plot transgressed the original trend and the irregular variation in the

data series could not be removed. Subsequently, the series were differentiated by showing the mean was constant, but not the variance. To demonstrate the stationary condition of the series, the Augmented Dickey-Fuller (ADF) test was used, which was significant at 5%. Therefore, the data series at levels are non-stationary at first; however, after differentiating them and according to the ADF test for unit root, it is found that the series are stationary.

Larios-Meño, González-Taranco and Álvarez (2016) refer to the stationarity of time series, and after verifying their presence in economic and financial variables frequently used in Peru—such as the General Index of the Lima Stock Exchange, primary income account, terms of trade and private consumption—they conclude that forecasting would be useful, because the results will have consistency; otherwise, such results would not have credibility and, therefore, would be spurious. The authors also add that

“Las series son sometidas a esta verificación mediante correlogramas y el test de raíz unitaria de Dickey-Fuller Aumentado. Para llevar a cabo esta tarea, los datos de las series son ajustados a modelos autorregresivos (AR), de media móvil (MA) y de caminata aleatoria o Random Walk,

con el propósito de simular, por ejemplo, condiciones de no estacionariedad que luego son confirmadas por los distintos indicadores obtenidos en esta evaluación. Finalmente, en caso de encontrar series no estacionarias, se propone eliminar esta condición con procesos de diferenciación.” [The series are verified by using correlograms and the Augmented Dickey-Fuller unit root test. For this, the series data are adjusted to autoregressive (AR), moving average (MA) and random walk models, with the purpose of simulating, for example, non-stationary conditions that are later confirmed by the indicators obtained in this evaluation. Finally, in case of finding non-stationary series, it is proposed to eliminate this condition with differentiation processes.]

This type of series, such as those produced by TEF financial variables, cannot be modelled with traditional forecasting techniques, but with univariate time series models or ARIMA (autoregressive integrated moving average model), also known as Box-Jenkins methods; furthermore, Box, Jenkins and Reinsel (2008) argue that the series need to be stable over time. This method consists of four steps: identification, estimation, validation and prediction.

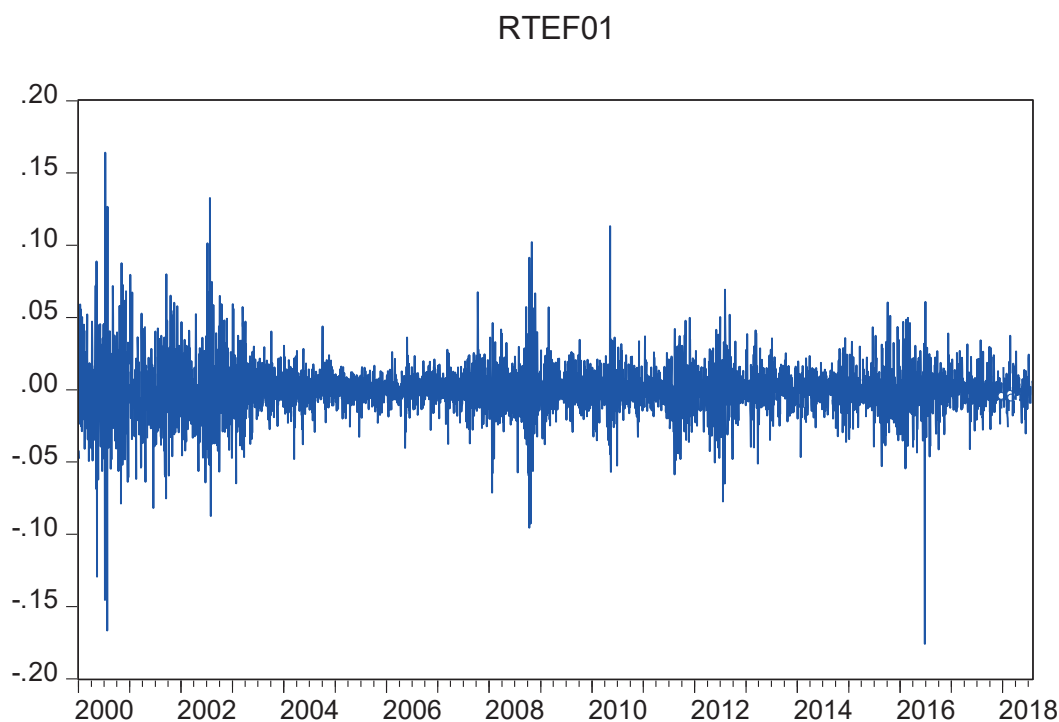


Figure 2. Stabilization of the mean and variance of the differences in TEF returns

Source: Prepared by the author.

Identification consists on verifying if the series is stationary by using unit root (UR) tests and if it is white noise or memory in order to forecast. Since TEF closing prices series is not stationary, differentiation process is necessary to stabilize it. Figure 2 shows the stabilized series.

This research aims to predict univariate time series from their own past behavior. For this, it is necessary to stabilize, that is to say, make stationary both the mean and the variance, so that they can be modeled with Box and Jenkins methodologies. Taking the specific case of TEF, this study seeks to determine how differentiation influences the constant return of its shares, determining, for this purpose, how the closing price differentiation influences its constant performance of its share price, constant volatility and autocorrelation.

The objective that justifies this research is to avoid or reduce the negative risks that affect investors and business organizations that invest in projects and/or financial instruments, such as TEF shares, a well-established technology company. For this reason, it is advisable to forecast the price, return and volatility of its shares listed on the NYSE, given that their trading volume is greater compared to that on the Lima Stock Exchange.

METHODOLOGY

This analysis aims to explain the behavior of a time series by stabilizing its mean and its variance in order to forecast the returns and daily volatility of the adjusted closing prices of Telefónica de España S. A. (TEF) shares on the New York Stock Exchange (NYSE).

The unit of analysis are TEF shares. The population is represented by the stock exchange listings where TEF lists its shares, which can be found on its website (https://www.telefonica.com/es/web/shareholders-investors/la_accion/presencia-en-bolsas). According to Telefónica (2018), the company is present on the following listings: Buenos Aires, Lima, London, Madrid and New York. The sample size covers TEF shares between January 3, 2000 and August 1, 2018, taken from NYSE from the website <https://es.finance.yahoo.com/quote/TEF/history?p=TEF>.

When selecting the sample, access to all the data was difficult, but scientific research recommends the largest number of data. In this case, the research considered 4848 observations of daily closing prices of shares between January 3, 2000 and August 1, 2018, obtaining:

- Non-random samples
- Convenience samples

The criteria for the selection of the sample were considered due to the accessibility to the data from the website <https://es.finance.yahoo.com/quote/TEF/history?p=TEF>.

To model the behavior of TEF financial assets, it was only necessary to differentiate the original series to make it stationary. Then, the stationary series was the result of the logarithm differentiation of the closing price of a period against the closing price of a previous period, according to the following formula:

$$r_n = \text{Ln} \left(\frac{P_t}{P_{t-1}} \right) = \text{Ln}(P_t) - \text{Ln}(P_{t-1}). \quad (\text{Eq. 4})$$

To obtain the returns of this financial asset a variable was generated using EViews 10 software with the command `GENR rtef01 = dlog(tef01)`, which is the calculation of the logarithmic differentiation of the original series, whose variable was called `tef01`. Regarding Figure 2, a graphical analysis was first performed, and it was concluded that the `rtef01` series is stationary, but it was necessary to mathematically prove it with the unit root (UR) tests.

Now, if $x_t = \alpha x_{t-1} + \varepsilon_t$ is a stationary process,

$$\sigma_{x_t}^2 = \alpha^2 \sigma_{x_{t-1}}^2 + \sigma_{\varepsilon_t}^2 \rightarrow \sigma_{x_t}^2 - \alpha^2 \sigma_{x_t}^2 = \sigma_{\varepsilon_t}^2 \rightarrow \sigma_{x_t}^2 (1 - \alpha^2) = \sigma_{\varepsilon_t}^2 \quad (\text{Eq. 5})$$

$$\sigma_{x_t}^2 = \frac{\sigma_{\varepsilon_t}^2}{(1 - \alpha^2)} \quad (\text{Eq. 6})$$

If $\alpha = 1$, the process is not stationary (unit root).

If $\alpha > 1$, the process is explosive.

If $\alpha < 1$, the process is stationary.

Then, the Augmented Dickey-Fuller (ADF) unit root test was performed; according to Bello (2018), it is one of the most used. About this test, it is added that "at the formal level, stationarity can be checked by finding out if the time series contains a unit root." (Gujarati & Porter, 2010, p. 768), and, in this case:

$H_0: \phi = 0$; x_t has unit root (at least one) $\rightarrow x_t$ is a non-stationary series.

$H_1: \phi < 0$; x_t has no unit root $\rightarrow x_t$ is a stationary series. The series is $I(0)$.

$I(d)$, expresses the number of times a time series must be differentiated to become stationary.

RESULTS

The results using ARIMA or Box-Jenkins methodology are shown below. This methodology, formalized by George Box and Gwilym Jenkins in 1976, is based on the fact that the time series they try to forecast are supported by the stochastic processes characterized through a model. Meléndez (2017) argues that, when attempting to predict with this model, the following steps are presented through trial and error scenarios:

0. Evaluation of stationarity
1. Identification
2. Estimation
3. Validation
4. Forecast

Interpretation 1

According to Figure 2, an attempt was made to verify that the differentiated series was stationary; therefore, the UR test was performed with the Augmented Dickey-Fuller (ADF) test. In this case, the objective was to reject the null hypothesis, which is that the series had at least one unit root. The tests indicated the absence of unit roots, according to the comparison between the Augmented Dickey-Fuller test statistic and the different MacKinnon critical values. A series is said to have UR if some of the critical values in absolute MacKinnon values are greater than the ADF test statistic in absolute values. In this case, the value of ADF test is $|42.39815|$, so it is greater than any of the critical values, at 1% = $|2.565457|$, 5% = $|1.940892|$ and 10% = $|1.616654|$, which means that the series does not have unit roots and, therefore, is a stationary series. The results are shown in Figure 3 and the first difference shows that the series is stationary.

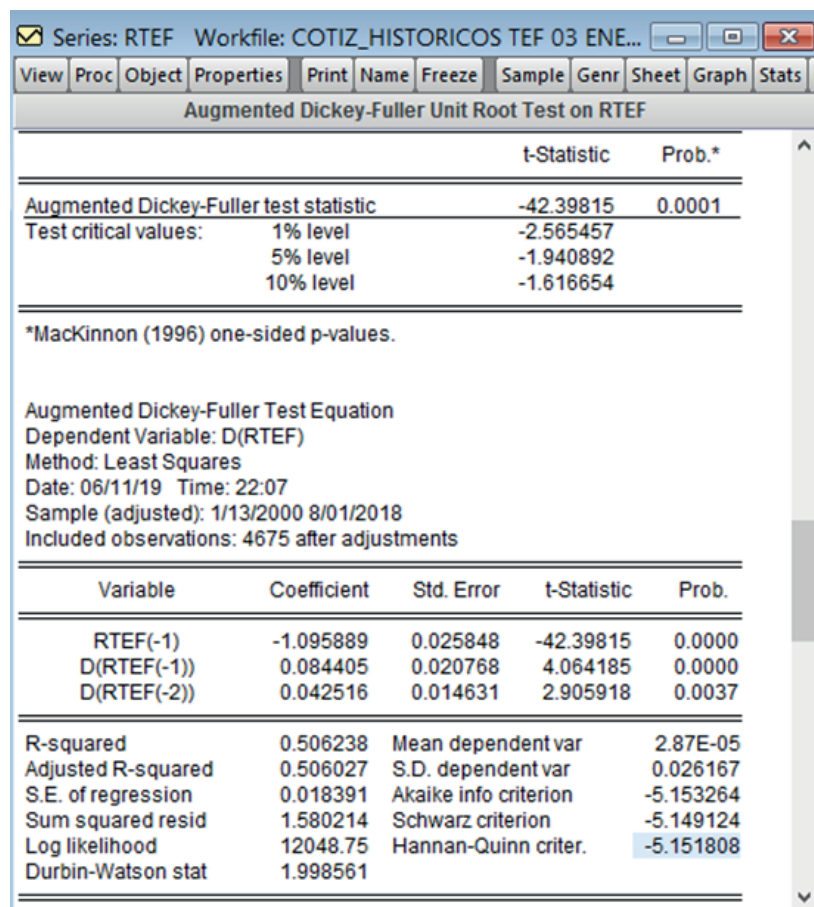


Figure 3. Augmented Dickey-Fuller (ADF) Unit Root Test.

Source: Prepared by the author.

Interpretation 2

Within the identification phase, in addition to checking that the series is stationary, it was verified that the series is not white noise, that is, that it has memory. It was found that this is a random process because its mean is equal to zero, its variance is constant and the autocovariance is equal to zero; with this process, which has no memory, a time series could not be predicted through the Box-Jenkins or ARIMA methodology, but rather with stochastic processes, such as the Geometric Brownian motion (GBM) or with the ARCH and GARCH family models. When a series has memory, its past will be used to forecast the series.

$\{\varepsilon_t\}_{t=1}^T \rightarrow$ is a white noise process if

$$E(\varepsilon_t) = 0, \forall t \quad (\text{Eq. 7})$$

$$E(\varepsilon_t^2) = V(\varepsilon_t) = \sigma_\varepsilon^2, \forall t \quad (\text{Eq. 8})$$

$$E(\varepsilon_t, \varepsilon_{t+\tau}) = \gamma_\tau = 0, \forall t, \tau \neq 0 \quad (\text{Eq. 9})$$

In this case, correlograms, the Ljung-Box statistic and the p -value were used to validate that the series had a memory, as can be seen in the results contained in Figure 4.

White noise validation (hypothesis testing)

Q statistic - joint test:

$$H_0: \rho_1 = \rho_2 = \rho_3 = \dots = \rho_k = 0; \text{ white noise}$$

$$H_1: \exists \rho_i \neq 0; i = 1, \dots, k; \text{ the series is not white noise}$$

LB statistic for Ljung-Box small samples:

$$LB = Q_k = n(n + 2) \sum_{i=1}^k \frac{r_i^2}{n-1} \quad (\text{Eq. 10})$$

The objective was to reject the null hypothesis (H_0), which signifies white noise; if H_0 was not rejected, it could not be forecasted with the Box-Jenkins methodology. Bello (2018) adds that a series is considered white noise if the p -values are greater than the significance level of the test and that, in practice, as

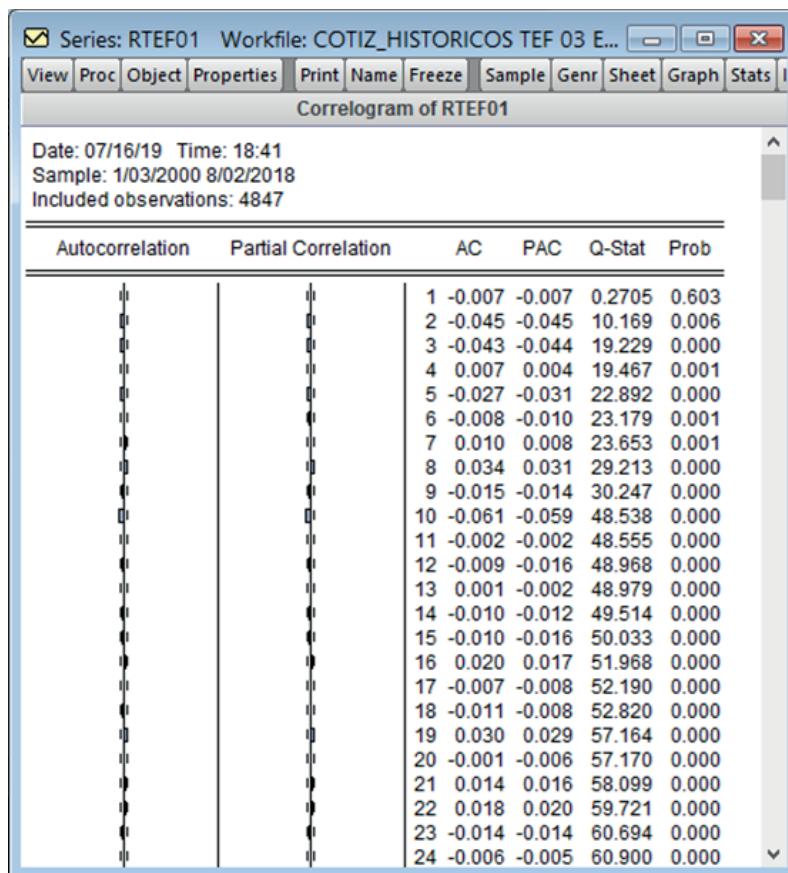


Figure 4. RTEF01 series correlogram test for validation of no white noise.

Source: Prepared by the author.

long as they are not the first lags, the p -values are allowed to be greater in one of every 20 lags to consider the series to be white noise, as was the case in the first lag. In this case, the null hypothesis is not rejected, since the variance is not homoscedastic or constant.

Interpretation 3

From the same results in Figure 4, it was found that autocovariance is neither constant nor invariant in time; in other words, the series does not depend on its past. The autocovariances between the value of a series and its own lag do not depend on the time distance that separates them; to correct it, the model was centered to determine if the residuals or errors have white noise with the equation $dlog(TEF01)$ C AR(10) MA(2) MA(3), whose results can be found in Figure 5 and show that the coefficients are significant as they are below 5%.

These models need the residuals to be white noise, so a diagnosis of the residuals was made based on its correlogram; as shown in Figure 6, it was verified, that the behavior of the residuals constitutes

white noise because its p -value is greater than 5% and, therefore, can be forecasted by the ARCH-GARCH models, which, according to Mansilla (April 2, 2020), together with the ARIMA models of Box and Jenkins, belong to the same family.

DISCUSSION

To forecast the return and volatility of TEF financial series with the Box-Jenkins methodology, it is necessary to differentiate its closing price at time t with its immediate previous period ($t-1$) to stabilize or make constant both the first and second moments (mean and variance). It can be seen that the series is stationary during the time between January 3, 2000 and August 1, 2018, which means that its mean is constant and it is proved by taking 522 returns from the period between 2005-2006, a time in which there was low volatility compared to another 522 observations from the period between 2007-2008, in which there was high volatility. However, the variance is not homoscedastic, because there were periods of high volatility in the returns and they could no longer be modeled with the Box-Jenkins

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000120	0.000227	-0.529692	0.5963
AR(10)	-0.060549	0.008042	-7.528982	0.0000
MA(2)	-0.046635	0.010152	-4.593728	0.0000
MA(3)	-0.046446	0.009414	-4.933813	0.0000
SIGMASQ	0.000333	2.98E-06	111.6110	0.0000

R-squared	0.007820	Mean dependent var	-0.000121
Adjusted R-squared	0.007000	S.D. dependent var	0.018315
S.E. of regression	0.018251	Akaike info criterion	-5.168132
Sum squared resid	1.612899	Schwarz criterion	-5.161441
Log likelihood	12529.97	Hannan-Quinn criter.	-5.165783
F-statistic	9.540690	Durbin-Watson stat	2.020548
Prob(F-statistic)	0.000000		

Figure 5. Results of equation $DLOG(TEF01)$ C AR(10) MA(2) MA(3).

Source: Prepared by the author.

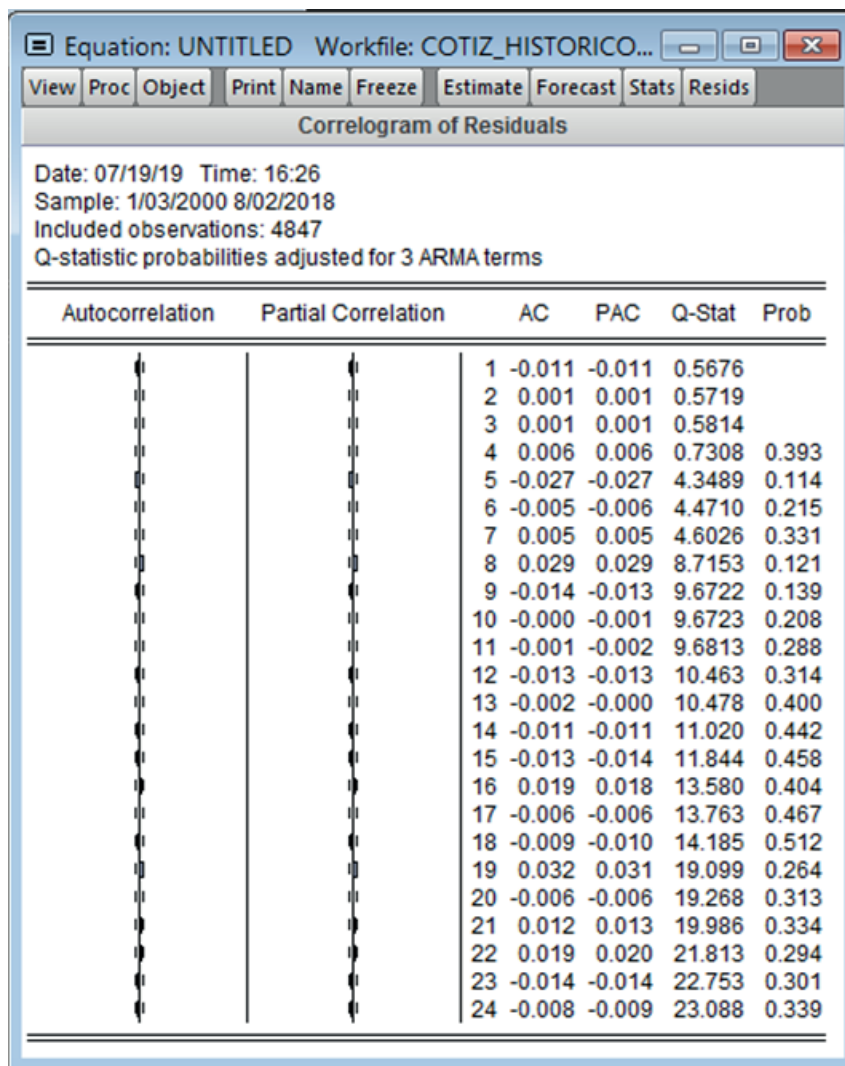


Figure 6. Residuals correlogram of equation DLOG(TEF01) C AR(10) MA(2) MA(3).

Source: Prepared by the author.

methodology, but with the models of the ARCH-GARCH family. The results of the comparison of means and Levene’s test for equality of variances are shown in Table 1.

The statistical tests that have been carried out are shown below.

1. Hypothesis Testing for Equality of Means

H_0 : The means for periods 2005-2006 and 2007-2008 are the same.

H_1 : The means for periods 2005-2006 and 2007-2008 are not the same.

Significance level

Significance level (alpha) $\alpha = 5\%$ or 0.05.

Test statistic

Two independent samples t-test.

P-value

0.0000%.

The test shows a two-tailed significance of $0.828 > 0.05$; therefore, the hypothesis of equality of means is not rejected and the null hypothesis of equality of means is accepted.

Decision making

The means of the groups to be compared are not different.

Interpretation

The means of the two groups are constant.

Table 1. Test for Equality of Means and Levene's Test for Equality of Variances.

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Return	Equal variances assumed	80.291	0.000	0.217	1042	0.828	0.00019580271	0.00090108264	-0.00157234060	0.00196394602
	Equal variances not assumed			0.217	742.929	0.828	0.00019580271	0.00090108264	-0.00157316870	0.00196477411

Source: Prepared by the author.

2. Hypothesis Testing for Equality of Variance

H_0 : The variances of periods 2005-2006 and 2007-2008 are homoscedastic or not different.

H_1 : The variances of periods 2005-2006 and 2007-2008 are heteroscedastic or different.

Significance level

Significance level (alpha) $\alpha = 5\%$ or 0.05.

Test statistic

Levene's Test.

P-value

0.0000%.

P-value reading: With 0.000 probability of error, the variances of the groups to be compared are heteroscedastic or different.

Decision making

The variances of the groups to be compared are not homoscedastic or not different.

Interpretation

The variances of the 2 groups are heteroscedastic or different.

CONCLUSIONS

1. The original closing price series of TEF shares shows a trend, which had to be removed by differentiation to make it stationary. The latter is a necessary condition so that, with the Box-Jenkins methodology, the forecast is stable throughout the period, since it is necessary that both the first and the second moment of the statistic are invariant, for which the unit root (UR) test was performed.
2. Using correlograms, the Ljung-Box statistic and the p -value, an attempt was made to validate: 1) the variable $rtef01$, which is the calculation of the logarithmic difference of the original series and 2) that the series has memory; but it was concluded that it is not homoscedastic or, in other words, its variance is not constant over time, so it cannot be predicted with the Box-Jenkins methodology.
3. However, the residuals are white noise, which is why it could be modelled according to the ARCH-GARCH family models, since its variance is heteroscedastic.

REFERENCES

- [1] Álvarez-Pallete, J. (May 14, 2018). [Carta a nuestros accionistas: El ilusionante reto de

- reinventar Telefónica]. Copy in possession of Wilfredo Bazán Ramírez.
- [2] Bello, M. (2018). *Modelos econométricos con EViews: Modelos de regresión lineal y series de tiempo*. [Lecture notes]. Retrieved from <https://www.software-shop.com/formacion/formacion-info/4404>.
- [3] Berk, J. & DeMarzo, P. (2008). *Finanzas Corporativas*. México D. F., Mexico: Pearson Educación.
- [4] Box, G., Jenkins, G. & Reinsel, G. (2008) *Time Series Analysis: Forecasting and Control*. New York, United States: Wiley.
- [5] Court, E. & Rengifo, E. (2011). *Estadísticas y Econometría Financiera*. Buenos Aires, Argentina: Cengage Learning Argentina.
- [6] Gujarati, D. & Porter, D. (2010). *Econometría*. México D. F., Mexico: McGraw-Hill.
- [7] Hanke, J. & Wichern, D. (2010). *Pronósticos en los negocios*. México D. F., Mexico: Pearson Educación.
- [8] Herrera, J. (2013). *Modelo estocástico a partir de razonamiento basado en casos para la generación de series temporales*. (Doctoral thesis). Universidad Nacional de San Agustín de Arequipa, Peru. Retrieved from http://repositorio.concytec.gob.pe/bitstream/20.500.12390/346/6/2013_Herrera_Modelo-estocastico-razonamiento.pdf.
- [9] Hossain, A., Kamruzzaman, M. & Ali, A. (2015). ARIMA with GARCH Family Modeling and Projection on Share Volume of DSE. *Economics World*, 3(7-8), 171-184.
- [10] Larios-Meño, J., González-Taranco, C. & Álvarez, V. (2016). *Investigación en economía y negocios. Metodología con aplicaciones en E-views*. Lima, Peru: Fondo Editorial de la Universidad San Ignacio de Loyola.
- [11] Lledó, P. (2017). *Director de proyectos. Cómo aprobar el examen PMP® sin morir en el intento*. United States: Pablo Lledó.
- [12] Mansilla, F. (April 2, 2020). *Modelos de volatilidad condicional con EViews*. [Lecture notes]. Retrieved from <https://www.software-shop.com/formacion/formacion-info/5391>.
- [13] Meléndez, J. (2017). *Entrenamiento especializado en modelos econométricos de series de tiempo en EViews* [Lecture notes]. Retrieved from <https://software-shop.com/formacion/formacion-info/2979>.
- [14] Mun, J. (2016). *Modelación de riesgos. Aplicación de la simulación de Monte Carlo, análisis de opciones reales, pronóstico estocástico, optimización de portafolio, análisis de datos, inteligencia de negocios y modelación de decisiones* (2 volúmenes). California, United States: IIPER Press.
- [15] Newbold, P., Carlson, W. & Thorne, B. (2008). *Estadística para Administración y Economía*. Madrid, Spain: Pearson Educación.
- [16] Ramón, N. & López, J. (2016). *Econometría. Series temporales y modelos de ecuaciones simultáneas*. Elche, Spain: Universidad Miguel Hernández.
- [17] Telefónica (2018). *Presencia en bolsas. Telefónica*. Retrieved from https://www.telefonica.com/es/web/shareholders-investors/la_accion/presencia-en-bolsas.
- [18] Villalba, F. & Flores-Ortega, M. (2014). Análisis de la volatilidad del índice principal del mercado bursátil mexicano, del índice de riesgo país y de la mezcla mexicana de exportación mediante un modelo GARCH trivariado asimétrico. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 17, 3-22. Retrieved from <https://www.upo.es/revistas/index.php/RevMetCuant/article/view/2191>.