



Análisis de la Versión Débil de la Hipótesis del Mercado Eficiente en el Perú

FREDDY ESPINO*

El presente trabajo analiza la Versión Débil de la Hipótesis del Mercado Eficiente (HME) para el Perú durante el periodo 2006-2021, contrastando empíricamente si el Índice General de la Bolsa de Valores de Lima (IGBVL) muestra una trayectoria similar a la de un proceso estocástico denominado paseo aleatorio. Las pruebas estadísticas indican que dicha característica no se refleja para datos con frecuencia diaria, semanal y mensual, pero sí para datos trimestrales. De esta manera, se concluye que la Versión Débil de la HME no se cumple en el caso peruano.

Palabras Clave : Bolsa de valores, paseo aleatorio, mercados eficientes.

Clasificación JEL : G14, G32, M21.

1. INTRODUCCIÓN

La HME se basa en el trabajo de Fama (1970), quien indica que un mercado es eficiente si los precios de los activos financieros “reflejan plenamente” toda la información disponible. Malkiel (2003) añade que un mercado es eficiente en tanto no permite a los inversionistas obtener rendimientos superiores a la media sin que asuman riesgos superiores que el promedio de inversionistas. En su versión débil, la HME señala que los precios de los activos financieros reflejan toda la información pasada del precio, así como toda la información relacionada a su transacción a través del tiempo, como el volumen de transacción, tasas de interés, etc.

En términos estadísticos, la versión débil de la HME desemboca en el concepto del proceso estocástico denominado paseo aleatorio, el cual, por definición, no puede ser predecible. La relación entre el concepto de eficiencia de mercado y paseo aleatorio está en que, si los precios de los activos financieros incorporan toda la información disponible, ningún inversionista podrá tomar ventaja de la información que maneja, debido a que los demás ya la tienen incorporada en las transacciones que realizan en el mercado, haciendo imposible predecir un precio de tal manera que tome ventaja sobre el resto de los inversionistas. En todo caso, cualquier ganancia extraordinaria será sólo temporal.

* Espino: Jefe del Departamento de Liquidación y Control de Operaciones Internas del Banco Central de Reserva del Perú (correo electrónico: freddy.espino@bcrp.gob.pe).

El autor agradece los comentarios de Marco Vega. Las opiniones vertidas en este documento son exclusivas del autor y no necesariamente representan a aquellas instituciones a las que pertenece.

Siguiendo las definiciones de paseo aleatorio de [Campbell y otros \(1997\)](#), para el caso modelo de paseo aleatorio tipo 1 (RW1)¹ se realizan las pruebas de (i) secuencias y reversas, (ii) rachas y (iii) de independencia a través de la prueba BDS² mientras que, para el caso del paseo aleatorio tipo 3 (RW3)³ se realizan las pruebas de (i) autocorrelación y (ii) ratios de varianza. Los resultados indican que el mercado financiero en el Perú no es eficiente de acuerdo con la Versión Débil de la HME.

El trabajo está organizado en cinco secciones. En la primera se presenta la introducción, la segunda sección expone la teoría de mercados eficientes y una breve revisión de la literatura empírica, la tercera se ocupa de hacer una revisión de la metodología empírica que se emplea en la presente investigación, la cuarta presenta los resultados y la última, expone las conclusiones.

2. MARCO TEÓRICO

La Hipótesis del Mercado Eficiente (HME) se basa en el trabajo de [Fama \(1970\)](#), quien define un mercado eficiente como aquel en el que los precios proporcionan señales precisas para la asignación de recursos, es decir, uno en el que las empresas puedan tomar decisiones de producción e inversión bajo el supuesto de que los precios en cualquier momento “reflejan plenamente” toda la información disponible.

Por otro lado, [Malkiel \(2003\)](#) añade que un mercado es eficiente en tanto no permite a los inversionistas obtener rendimientos superiores al promedio sin que asuman riesgos superiores a la media, es decir, no hay manera de que un inversionista obtenga mayor rentabilidad que la del mercado asumiendo el mismo riesgo que el resto de los inversionistas.

La HME tiene tres versiones ([Fama, 1970](#)):

- Débil: En esta versión los precios de los activos financieros reflejan toda la información pasada del precio, así como toda la información relacionada a su transacción a través del tiempo, como el volumen de transacción, tasas de interés, etc.
- Semi Fuerte: En esta versión los precios de los activos financieros reflejan, además de lo indicado en la Versión Débil, todos los fundamentos del precio, como son los factores de demanda y de oferta, así como las variables relacionadas a la empresa, como sus balances y estados de ganancias y pérdidas.
- Fuerte: En esta versión, además de incorporar la información indicada en la versión semi fuerte, los precios de los activos incorporan la información confidencial, es decir, aquella que sólo un pequeño grupo de personas relacionadas a la empresa pueden obtener.

La HME tiene los siguientes supuestos: (i) no hay costos de transacción en la negociación de activos financieros, (ii) toda la información está disponible sin costo para todos los participantes del mercado, y (iii) todos están de acuerdo en las implicancias de la información tanto para el precio actual y como para las distribuciones de los precios de cada activo financiero en el futuro. De esta forma, el precio actual de un activo financiero “refleja completamente” toda la información disponible. No obstante, si bien estas

¹ Considera retornos independientes e idénticamente distribuidos. Se desarrolla en la sección 3.

² En referencia a sus autores [Broock y otros \(1996\)](#).

³ Considera retornos no correlacionados. Se desarrolla en la sección 3

condiciones son suficientes para la eficiencia del mercado, no son necesarias⁴.

Esta acepción de eficiencia de mercado toma forma con el concepto de Juego Limpio, el cual es uno en el que ninguno de los participantes tiene ventaja sobre los demás, por lo que cualquier resultado derivado del mismo se considera justo (Campbell y otros, 1997). Aplicado al mercado financiero, significa que ningún agente tiene más o mejor información que otro, por lo que, como resultado de las transacciones que se dan en el mercado, los precios terminan reflejando toda la información disponible. En términos matemáticos, esta idea se traslada a la definición de martingala, que es un proceso estocástico que satisface la siguiente condición:

$$E[P_{t+1} | P_t, P_{t-1}, \dots] = P_t$$

donde P_t representa el precio de un activo en el periodo t . La interpretación es como sigue: el valor esperado condicional del precio para el siguiente periodo P_{t+1} , tomando en cuenta el precio de hoy y toda la información pasada, es igual al precio de hoy P_t . Otra forma de interpretar es que la mejor predicción del precio de mañana, con toda la información histórica, es el precio de hoy. Dicha condición se puede reescribir de la siguiente manera:

$$E[P_{t+1} - P_t | P_t, P_{t-1}, \dots] = 0 \quad (1)$$

Así, el valor esperado hoy del retorno del siguiente periodo, expresado en la diferencia entre el precio esperado para el siguiente periodo y el de hoy, condicionado a toda la información pasada, es cero. Dado que el precio es un proceso estocástico el cual tiene la misma probabilidad de subir que de bajar el siguiente periodo, entonces el retorno esperado condicional es cero.

Sin embargo, tal como lo señalan Campbell y otros (1997), no es una condición necesaria ni suficiente que se cumpla la condición de martingala para la determinación de los precios de los activos, dado que la evidencia muestra que hay una relación directa entre retorno esperado y riesgo, la cual es la base para la teoría de inversión formalizada desde el trabajo de Markowitz (1952). El punto es que la correlación del retorno actual con los pasados es cero o tan bajo que no puede establecerse un patrón identificable y, por lo tanto, predecible.

Tomando en cuenta el siguiente modelo de paseo aleatorio con intercepto:

$$P_t = \mu + P_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.}(0, \sigma^2) \quad (2)$$

donde ϵ_t es un término de error idénticamente e independientemente distribuido (i.i.d.), entonces el valor esperado condicional para $t + 1$ es P_t y el retorno, definido como la diferencia de precios:

$$R_t = \Delta P_t = \mu + \epsilon_t \quad (3)$$

Entonces, los retornos del precio de los activos son independientes entre periodos y además tienen una distribución similar a lo largo del tiempo. Así, una manera de probar la HME de manera empírica es a

⁴ Por ejemplo, tal como señala Fama (1970), incluso grandes costos de transacción que inhiben el flujo de transacciones no implican en sí mismos que los precios no “reflejen completamente” la información disponible. De manera similar, el mercado puede ser eficiente si “un número suficiente” de inversionistas tienen fácil acceso a la información disponible. Y el desacuerdo entre los inversionistas acerca de las implicaciones de la información dada no implica en sí mismo una ineficiencia del mercado a menos que haya inversionistas que puedan hacer consistentemente mejores evaluaciones de la información disponible que las implícitas en los precios del mercado.

través de la medición del grado de autocorrelación del retorno de los activos y definir si es significativo o no, no sólo desde el punto de vista estadístico, si no del valor de los coeficientes estimados de tal manera que se pueda identificar un patrón predecible.

De esta manera, la relación entre el concepto de eficiencia de mercado y paseo aleatorio está en que, si los precios de los activos financieros incorporan toda la información disponible, entonces estos se vuelven impredecibles, debido a que ningún inversionista podrá tomar ventaja de la información que maneja, puesto que los demás ya la habrán incorporado a su conjunto de información, haciendo imposible predecir un precio que le permita obtener ganancias extraordinarias de manera sistemática.

La consecuencia de esto es, tal como lo afirman (Bodie y otros, 2018), es que un mercado sería ineficiente si los precios muestran un patrón predecible, lo cual es una afirmación extrema, puesto que existe de alguna manera una relación entre los precios de hoy y los del pasado, en tanto los inversionistas toman en cuenta la información pasada para sus decisiones de inversión. De hecho, los agentes tienden a comportarse de manera similar en momentos de mucho optimismo o pesimismo, periodos en los que se observa tendencias al alza y a la baja respectivamente.

Por otro lado, contrario a lo que postula la HME, en la práctica se observa que algunos inversionistas acumulan ganancias exorbitantes, lo cual depende del tamaño del fondo que administre y de la tecnología que usen para explotar al máximo cualquier diferencia mínima entre el precio de compra y de venta de activos financieros. Al respecto, Lewis (2014) narra eventos sucedidos en el mercado financiero de los Estados Unidos de América luego de la crisis financiera de 2008, en donde el método de transacciones de alta frecuencia fue usado para anticipar las órdenes colocadas por los inversionistas, lo cual se tradujo en ingente cantidad de ganancias para los bancos de inversión involucrados.

Desde el punto de vista teórico, las principales críticas a la HME provienen de Grossman y Stiglitz (1980), quienes mostraron que es imposible que un mercado sea perfectamente eficiente desde el punto de vista de la información, dado que esta es costosa, por lo que los precios no pueden reflejar perfectamente la información que está disponible, puesto que, si lo hiciera, los inversionistas que gastaron recursos en obtenerla y analizarla no recibirían compensación.

A pesar de que el marco teórico no fue desarrollado hasta mediados de la década de 1960 por Samuelson (1973) y Mandelbrot (1966), los trabajos empíricos que prueban que el precio de los activos tiene una evolución similar a la de un paseo aleatorio empezaron desde 1900 con el trabajo de Bachelier (1900), de acuerdo con Fama (1970). Luego vienen un conjunto de trabajos, siendo los más relevantes el de Kendall (1953), a pesar de que no fue su objetivo, pero si encuentra dicho patrón, y el de Fama (1965). Décadas más tarde, Fama (1991) hace una revisión de la literatura empírica sobre la HME en sus tres versiones, y encuentra tanto evidencia a favor como en contra de la HME en sus tres versiones.

En el caso del Perú son pocos los trabajos realizados en ese tema. Delgado y Humala (1997) prueban la versión débil de la HME en dos partes: primero, a través del análisis de la estacionariedad, autocorrelación, rendimientos por el día de la semana para explorar posibles características estacionales y volatilidad; segundo, mediante el análisis de cointegración del mercado peruano con otros similares. Los autores hallan evidencia estadística en contra de la HME, pero no concluyen que el mercado peruano sea ineficiente.

Duarte y Pérez-Iñigo (2014) prueban la versión débil de la HME en los cinco principales mercados bursátiles de Latinoamérica, entre los cuales está Perú. Concluyen que en general, las cinco principales economías latinoamericanas estudiadas han experimentado una mejora de su eficiencia del mercado en los últimos años, y en el caso de Perú desde 2012. Ruiz y García (2020) también prueban dicha versión para el caso del Mercado Integrado Latinoamericano (MILA), en el cual está el Perú, y encuentran que el mercado peruano no es eficiente, resultado opuesto al de Meneses y Pérez (2020).

3. METODOLOGÍA

En esta sección se discute la metodología para probar la versión débil de la HME. Existen tres versiones sobre el paseo aleatorio, discutidas en [Campbell y otros \(1997\)](#), las cuales pasaremos a revisar, así como las pruebas empíricas que se realizan en cada caso.

3.1. PASEO ALEATORIO 1 (RW1): INCREMENTOS INDEPENDIENTES E IDÉNTICAMENTE DISTRIBUIDOS (I.I.D.)

La versión simple del paseo aleatorio con errores i.i.d. es:

$$P_t = \mu + P_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.}(0, \sigma^2) \quad (4)$$

donde P_t es el precio del activo, μ es el intercepto y ϵ_t representa el término de error i.i.d. con media nula y varianza constante σ^2 . De esta manera, los retornos del precio del activo se expresan de la siguiente manera:

$$R_t = \mu + \epsilon_t \quad (5)$$

La independencia de los retornos implica que el paseo aleatorio es un proceso de juego limpio, pero más estricto que la martingala: la independencia no solo implica que los incrementos no estén autocorrelacionados, sino que también cualquier función no lineal no lo esté también.

La solución por iteración de la versión simple del paseo aleatorio es:

$$P_t = P_0 + \mu t + \sum_{i=1}^t \epsilon_i \quad (6)$$

donde P_0 es el precio inicial. De esta manera, si calculamos la media y varianza incondicional obtenemos:

$$E[P_t] = P_0 + \mu t \quad (7)$$

$$Var[P_t] = \sigma^2 t \quad (8)$$

Por lo que el paseo aleatorio es un proceso no estacionario, dado que la media y la varianza se incrementan a través del tiempo t , condición que se mantiene para las otras versiones de paseo aleatorio que se discuten más adelante. El supuesto más simple sobre la distribución de ϵ_t es que sea normal, por lo que sería un proceso i.i.d. $(0, \sigma^2)$ equivalente a un movimiento aritmético Browniano, distribuido en una muestra separada por unidades. Si bien este supuesto simplifica varios cálculos, padece del mismo problema que afecta a los retornos normalmente distribuidos: si la distribución del proceso P_t es normal, entonces habrá siempre la posibilidad de que $P_t < 0$, lo cual no se ajusta a la realidad. Para evitar dicha posibilidad, se calcula el logaritmo natural de P_t tal que $p_t = \ln(P_t)$, por lo que el precio queda definido como:

$$p_t = \omega + p_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.}\mathcal{N}(0, \sigma^2) \quad (9)$$

De esta manera, el retorno continuo compuesto de los activos se representa como:

$$r_t = \Delta p_t = \omega + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.} \mathcal{N}(0, \sigma^2) \quad (10)$$

La independencia en el incremento considera tanto una relación lineal como no lineal, es decir, que la no autocorrelación se da de manera lineal, $\text{cov}[\epsilon_t, \epsilon_{t-k}] = 0$ para todo $k \neq 0$, como no lineal, $\text{cov}[\epsilon_t^2, \epsilon_{t-k}^2] = 0$ para todo $k \neq 0$. Esta definición implica que la distribución de los datos se mantiene a lo largo de la muestra, no importa qué tan grande sea esta.

Las pruebas empíricas correspondientes a este tipo de paseo aleatorio se discuten a continuación.

Secuencias y reversas

De acuerdo con [Campbell y otros \(1997\)](#), se asume la versión geométrica browniana del RW1 en el cual el logaritmo del precio sigue un proceso aleatorio sin intercepto y con errores i.i.d.:

$$p_t = p_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.} (0, \sigma^2) \quad (11)$$

Se define la función I_t para la siguiente variable aleatoria:

$$I_t = \begin{cases} 1 & \text{si } r_t \equiv p_t - p_{t-1} > 0 \\ 0 & \text{si } r_t \equiv p_t - p_{t-1} \leq 0 \end{cases} \quad (12)$$

La prueba consiste en estimar el número de secuencias (N_s), definido como el número de veces consecutivas de retornos con el mismo signo, y reversas (N_r), definido como el número de veces consecutivas de retornos con signos opuestos:

$$N_s \equiv \sum_{t=1}^n Y_t, \quad Y_t = I_t I_{t+1} + (1 - I_t)(1 - I_{t+1}) \quad (13)$$

$$N_r \equiv n - N_s \quad (14)$$

donde n representa el total de la muestra de retornos.

Si el logaritmo del precio del activo sigue un proceso de paseo aleatorio sin intercepto con errores i.i.d., y si se añade la restricción de que la distribución de los incrementos de ϵ_t son simétricos, entonces si r_t es positivo no negativo, debe tener la misma probabilidad que un experimento de lanzamiento de moneda al aire, es decir, 0,5 en cada caso. Esto implica que, para cada par de retornos consecutivos, una secuencia y una reversa son igualmente probables, lo cual se refleja en el denominado ratio Cowles-Jones:

$$\widehat{\text{CJ}} = \frac{N_s}{N_r} = \frac{\binom{N_s}{n}}{\binom{N_r}{n}} = \frac{\hat{\pi}_s}{1 - \hat{\pi}_s} \quad (15)$$

Donde $\hat{\pi}_s$ es la probabilidad de una secuencia. De esta manera, la ratio debe ser aproximadamente

igual a uno. El hecho de que esta ratio exceda el valor de 1 para varios retornos históricos lleva a pensar que existe un patrón identificable (Campbell y otros, 1997).

No obstante, asumir que el intercepto es cero puede llevar a una conclusión errónea, dado que mientras el intercepto sea distinto de cero, siempre será el caso de que las secuencias serán más probables que las reversas, porque el intercepto induce a una tendencia en el proceso de los datos. Para observar esto, asumamos ahora que:

$$p_t = \omega + p_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2). \quad (16)$$

Entonces, el indicador I_t no es más un experimento de lanzamiento de moneda al aire dado que está sesgado en la dirección del signo del intercepto:

$$I_t = \begin{cases} 1 & \text{con probabilidad } \pi \\ 0 & \text{con probabilidad } 1 - \pi \end{cases} \quad (17)$$

donde $\pi \equiv \Pr(r_t > 0) = \Phi\left(\frac{\omega}{\sigma}\right)$, y $\Phi(\cdot)$ representa la función de distribución acumulada de la distribución normal. Si el intercepto ω es positivo, entonces $\pi > 0,5$ y si es negativo $\pi < 0,5$. Una especificación más general la ratio CJ sería:

$$CJ = \frac{\pi^2 + (1 - \pi)^2}{2\pi(1 - \pi)} \quad (18)$$

El caso del experimento de lanzamiento de moneda al aire sería cuando $\pi = 0,5$ por lo que la ratio CJ sería igual a uno. Para hacer una comparación formal entre el valor estimado bajo la hipótesis de que es un paseo aleatorio, es decir, que $\widehat{CJ} = 1$, usaremos la distribución asintótica de acuerdo con Campbell y otros (1997):

$$\widehat{CJ} \sim \mathcal{N}\left(\frac{\pi_s}{1 - \pi_s}, \frac{\pi_s(1 - \pi_s) + 2(\pi^3 + (1 - \pi)^3 - \pi_s^2)}{n(1 - \pi_s)^4}\right) \quad (19)$$

donde $\pi_s = \pi^2 + (1 - \pi)^2$. Entonces, la prueba estadística consiste en estimar el estadístico \widehat{CJ} bajo la hipótesis nula $H_0 : \widehat{CJ} = 1$.

Rachas

La prueba de rachas es una prueba no paramétrica, es decir, no hace supuestos sobre la distribución de los datos bajo estudio, y consiste en determinar si un proceso estocástico tiene una trayectoria tal que no se pueda identificar un patrón de autocorrelación. Tomando en cuenta el signo del retorno de los activos, bajo la hipótesis de paseo aleatorio, no debería tener un patrón identificable.

De acuerdo con Gujarati y Porter (2009), se define como racha a la secuencia ininterrumpida de valores positivos o negativos en una serie de tiempo⁵. Asimismo, se define como la extensión de la racha al número

⁵ Sea r_t el retorno de un activo, se define el valor 1 si $r_t > 0$ y cero en otro caso. Entonces, en la secuencia de retornos 0001110001 se identificarán dos rachas negativas y dos positivas. En un extremo, sea la secuencia de retornos 01, entonces habrá una racha negativa y otra positiva.

de elementos en ella. De esta manera, examinando cuantas rachas se identifican en una serie de tiempo, se puede derivar una prueba de aleatoriedad de rachas, la cual se elabora de la siguiente manera. Sean:

- N es el número total de observaciones.
- N_1 es el número de observaciones positivas.
- N_2 es el número de observaciones negativas (con lo cual $N = N_1 + N_2$).
- R es el número total de rachas.

Entonces, bajo la hipótesis nula (H_0) de que las observaciones del proceso son i.i.d., y asumiendo que $N_1 > 10$ y $N_2 > 10$, el número de rachas se distribuye asintóticamente normal con media:

$$E(R) = \mu_R = \frac{2N_1N_2}{N} + 1 \quad (20)$$

y varianza:

$$\text{var}(R) = \sigma_R^2 = \frac{2N_1N_2(2N_1N_2 - N)}{N^2(N - 1)}. \quad (21)$$

De esta manera, se construye el intervalo de confianza al α % de significancia:

$$\text{Prob} [\mu_R - N_{(\alpha/2)}\sigma_R \leq R \leq \mu_R + N_{(\alpha/2)}\sigma_R] = 1 - \alpha \% \quad (22)$$

donde $N_{(\alpha/2)}$ representa el valor crítico de la distribución normal estándar a un nivel de $\alpha/2$ % de significancia. De esta manera, si R cae fuera del intervalo, entonces rechazamos la H_0 , mientras que, si cae dentro, fallamos en rechazarla.

Prueba de independencia BDS

Basado en el trabajo de [Broock y otros \(1996\)](#), esta prueba se usa para evaluar si una serie es i.i.d. La independencia puede ser lineal, no lineal o caótica. La idea de la prueba BDS consiste en escoger una distancia entre las observaciones denominada ε . Si las observaciones de la serie son i.i.d., entonces para cualquier par de observaciones la probabilidad de distancia menor o igual a ε debe ser igual desde el punto de vista estadístico.

Sean dos observaciones $\{y_s, y_t\}$, la distancia entre ellas es $|t - s|$. La probabilidad de la distancia entre este par de observaciones se denota como $c_1(\varepsilon)$. A partir de esto, podemos escoger un conjunto de múltiples observaciones que pueden ser consecutivas de acuerdo con la muestra, por ejemplo, dada la observación s y dada la observación t de una serie y_t , se puede construir el siguiente conjunto de observaciones ordenadas de par en par:

$$\{(y_s, y_t), (y_{s+1}, y_{t+1}), \dots, (y_{s+m-1}, y_{t+m-1})\} \quad (23)$$

Donde m es el número consecutivo de observaciones en el conjunto de dimensión incorporada. Se define la probabilidad conjunta de cada dimensión como $c_m(\varepsilon)$. Entonces, bajo el supuesto de independencia, la probabilidad conjunta del conjunto de dimensión incorporada debe ser el producto de la probabilidad de cada par de observaciones individualmente:

$$c_m(\varepsilon) = c_1^m(\varepsilon) \quad (24)$$

Para una serie y_t con n observaciones, la probabilidad de una dimensión particular, referida como la correlación integral, se define como:

$$c_{m,n}(\varepsilon) = \frac{2}{k(k-1)} \sum_{t < s} I_t(y_t^m, y_s^m) \quad (25)$$

donde $y_t^m = (y_t, y_{t+1}, \dots, y_{t+m-1})$, $y_s^m = (y_s, y_{s+1}, \dots, y_{s+m-1})$, $k = n - m + 1$ e $I_t(y_t^m, y_s^m)$ es una función de tal manera que

$$I_t(y_t^m, y_s^m) = \begin{cases} 1 & \text{si } |t - s| < \varepsilon \\ 0 & \text{en otro caso} \end{cases} \quad (26)$$

El estadístico BDS se define como:

$$w_{m,n}(\varepsilon) = \frac{[c_{m,n}(\varepsilon) - c_1^m(\varepsilon)] \sqrt{n}}{\sigma_{m,n}(\varepsilon)} \quad (27)$$

donde $\sigma_{m,n}(\varepsilon)$ es la desviación estándar y puede ser estimado consistentemente. Dicho estadístico se distribuye asintóticamente normal, bajo la hipótesis nula de que el proceso es i.i.d.

Paseo Aleatorio 2 (RW2): Incrementos Independientes

En esta versión se relaja el supuesto de que los incrementos son idénticamente distribuidos a lo largo de la muestra de datos, debido a que es difícil sostener que la distribución de los datos sea la misma a lo largo de una muestra larga como, por ejemplo, 30 o 50 años, sobre todo en economías que sufren constantes cambios en periodos cortos, como las economías en desarrollo.

A pesar de que es menos restrictivo que el RW1, contiene aún la principal característica económica de un paseo aleatorio i.i.d.: cualquier transformación arbitraria del incremento del precio en el futuro es impredecible usando cualquier transformación del incremento en el pasado. Las pruebas estadísticas en este caso se muestran a continuación.

Reglas de Filtros

De acuerdo con [Campbell y otros \(1997\)](#), para probar esta hipótesis se aplica una regla de inversión, que consiste en comprar un activo cuando crece en $x\%$, y vender, incluso en corto, cuando su precio cae $x\%$. La idea es que, si existe tendencias en el mercado de capitales, pero que están ocultos por el rápido movimiento de los precios, entonces, es factible filtrar todos los movimientos pequeños con relación a un nivel y analizar el resto de los movimientos. El retorno total de esta estrategia se considera como prueba de la predictibilidad del retorno de los activos. Se compara el retorno total de este filtro a un portafolio determinado aplicándola al índice compuesto como al IGBVL, y si es mayor, se concluye que existe tendencia en el mercado de capitales.

Sin embargo, los trabajos empíricos que hallan estrategias de inversión que le ganan al mercado, bajo el enfoque de reglas de filtros, lo realizan con datos históricos y asumiendo un costo no coherente con la realidad, por lo que terminan siendo inaplicables en la práctica. De esta manera, consideramos que no es

relevante realizar una prueba en ese sentido, dado que cualquier regla que pueda hallarse quedará siempre en el ámbito teórico.

En general, es difícil encontrar una estrategia de inversión que supere al mercado de manera sistemática y que sea publicada, puesto que, al momento de su publicación, ya todos lo habrán incorporado en su conjunto de estrategias, por lo que las ganancias extraordinarias ser irán desvaneciendo. Como indican [Bodie y otros \(2018\)](#), no es posible encontrar una estrategia como tal publicada dado que el creador estaría perdiendo dinero al hacerlo⁶.

Análisis Técnico

El análisis técnico es esencialmente la búsqueda de patrones recurrentes y predecibles en los precios de los activos ([Bodie y otros, 2018](#)). Se centra en identificar patrones de movimientos de precios, volúmenes de transacción, señales de negociación y varios otros elementos a través del análisis gráfico para evaluar la fortaleza o debilidad de un activo financiero. Los patrones clásicos que tratan de encontrar son doble profundidad, cabeza y hombros, nivel de apoyo y resistencia, etc.⁷ Aunque dicho enfoque no tiene mucho apoyo desde el punto de vista académico, es muy usado por los inversionistas, por lo que viene ganando atención por parte de los investigadores ([Campbell y otros, 1997](#)).

No obstante, al igual que la Regla de Filtros, este esquema se basa en un exhaustivo análisis de datos históricos hasta encontrar una estrategia que sea rentable durante la muestra analizada, pero nada garantiza que eso funcione en el futuro o que sea aplicable. En caso funcione al inicio, ocurrirá lo mismo que la Regla de Filtros: los inversionistas lo tomarán en cuenta en su conjunto de información, haciendo que las ganancias extraordinarias desaparezcan paulatinamente. De esta manera, en este trabajo tampoco se realiza una prueba en ese sentido.

Paseo Aleatorio 3 (RW3): Incrementos No Correlacionados

De acuerdo con [Campbell y otros \(1997\)](#), bajo la versión débil del paseo aleatorio, RW3, la primera diferencia de los niveles del paseo aleatorio no está correlacionada para todos los rezagos o adelantos. Esta es la versión débil del paseo aleatorio que contiene a las versiones RW1 Y RW2 como casos especiales. Entonces, debemos probar la hipótesis nula de que el coeficiente de autocorrelación es cero para varios rezagos.

Un ejemplo particular sobre un proceso que cumple con el supuesto RW3, pero no con RW1 y RW2, es cualquier proceso para el cual $cov[\epsilon_t, \epsilon_{t-k}] = 0$ para todo k , pero $cov[\epsilon_t^2, \epsilon_{t-k}^2] \neq 0$ para algún $k \neq 0$. Este proceso tiene incrementos no correlacionados, pero no son independientes, ya que el cuadrado de los incrementos está correlacionado (efectos ARCH).

Prueba de Autocorrelación

El coeficiente de autocorrelación es una medida de fácil implementación e interpretación. Dada una serie estacionaria, como el retorno de los activos, $\{r_t\}$, la autocovarianza $\gamma(k)$ y la autocorrelación $\rho(k)$ de orden k se definen como:

⁶ Tal como señalan [Bodie y otros \(2018\)](#) “Este es un problema de sesgo de selección; los resultados que podemos observar han sido preseleccionados a favor de intentos fallidos. Por lo tanto, no podemos evaluar de manera justa la verdadera capacidad de los administradores de cartera para generar estrategias ganadoras en el mercado de valores.” En el caso de estudio, no se podría probar dicha hipótesis de manera objetiva.

⁷ Una referencia importante es La Teoría de Olas de Elliot (*Elliott Wave Theory*)

$$\gamma(k) = \text{cov}[r_t, r_{t-k}] \quad (28)$$

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} \quad (29)$$

Nótese que $\gamma(0)$ es la varianza de la serie $\{r_t\}$. Dado que RW1 implica que todas las autocorrelaciones son cero, la prueba estadística más simple, bajo la hipótesis nula de que la serie es RW1, que tiene poder bajo varias hipótesis alternativas, es el estadístico Q de [Box y Pierce \(1970\)](#) que se distribuye con m grados de libertad:

$$Q_m \equiv T \sum_{k=1}^m \rho^2(k) \sim \chi_m^2 \quad (30)$$

[Ljung y Box \(1978\)](#) proveen el siguiente estadístico corregido para muestras finitas:

$$Q'_m \equiv T(T+2) \sum_{k=1}^m \frac{\rho^2(k)}{T-k} \sim \chi_m^2 \quad (31)$$

Sumando todas las autocorrelaciones al cuadrado, la prueba Box-Pierce está diseñada para identificar coeficientes de autocorrelación distintos de cero en cualquier dirección y para cualquier rezago. Por lo tanto, tiene poder contra varias hipótesis alternativas (estadístico de Portmanteau) de ruido blanco. Sin embargo, seleccionar el número de rezagos requiere mucho cuidado: si muy pocos son usados, se puede estar obviando órdenes de autocorrelación mayores, mientras que, si muchos son usados, la prueba pierde poder. De esta manera, si bien esta prueba tiene bastante aceptación en la práctica, debe haber mejores pruebas estadísticas en donde se fije claramente la hipótesis alternativa.

Ratios de Varianza

Una de las propiedades del modelo RW1 es que la varianza se incrementa a través del tiempo. Por ejemplo, bajo la hipótesis nula de RW1, donde el retorno continuo compuesto $r_t \equiv p_t - p_{t-1}$ es i.i.d., la varianza de $r_t + r_{t+1}$ debe ser dos veces la varianza de r_t . Por lo tanto, probar si una secuencia se asemeja a un paseo aleatorio, debe ser el caso de que la varianza de $r_t + r_{t+1}$ debe ser dos veces la varianza de $\{r_t\}$. Bajo la hipótesis nula de RW1, estos deben ser iguales, por lo que se construye la ratio de varianzas.

Sea $r_t(2) = r_t + r_{t+1}$, entonces se construye el estadístico:

$$VR(2) = \frac{\text{Var}[r_t(2)]}{2 \times \text{Var}[r_t]} \quad (32)$$

Dicho ratio se puede descomponer de la siguiente manera:

$$VR(2) = \frac{\text{var}[r_t + r_{t+1}]}{2 \times \text{var}[r_t]} = \frac{2 \times \text{var}[r_t] + 2 \times \text{cov}[r_t, r_{t+1}]}{2 \times \text{var}[r_t]} = 1 + \rho(1). \quad (33)$$

donde $\rho(1)$ es el coeficiente de autocorrelación de primer orden. Para cualquier serie estacionaria, el valor poblacional de la ratio de varianza $VR(2)$ es simplemente uno más el coeficiente de autocorrelación de primer orden. Entonces, bajo la hipótesis nula RW1 el coeficiente de autocorrelación es cero, por lo que $VR(2) = 1$. En la presencia de autocorrelación, $VR(2)$ será mayor a uno para el caso de autocorrelación positiva, por lo que la varianza crecerá rápidamente. De manera alterna, si el coeficiente de autocorrelación es negativo, la varianza va a crecer lentamente.

Para comparaciones mayores a un periodo, la ratio de varianza se representa como:

$$VR(q) = \frac{\text{var}[r_t(q)]}{q \times \text{var}[r_t]} = 1 + 2 \times \sum_{k=1}^{q-1} \left(1 - \frac{k}{q}\right) \times \rho(k). \quad (34)$$

donde $r_t(k) = r_t + r_{t+1} + \dots + r_{t-k+1}$ y $\rho(k)$ es el coeficiente de autocorrelación de orden k . Esto muestra que el estadístico $VR(q)$ es una combinación lineal particular de los primeros $k - 1$ coeficientes de correlación con pesos linealmente decrecientes. Bajo la hipótesis RW1, el ratio $VR(q)$ debe aproximarse a uno, dado que $\rho(k) = 0$. Dicho resultado se mantiene bajo la hipótesis RW2 y RW3 dado que la varianza de r_t es finita y la varianza promedio converge a un número positivo.

El ratio $VR(q)$ se construye bajo la hipótesis nula de que el retorno es RW1:

$$H_0 : r_t = \omega + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.} \mathcal{N}(0, \sigma^2)$$

La hipótesis alternativa es que el retorno no es un proceso RW1. No obstante, es un hecho estilizado el que la volatilidad del retorno de los activos financieros no es constante, tal como lo discuten [Bollerslev y otros \(1986\)](#), por lo que rechazar la hipótesis RW1 bajo la presencia de heteroscedasticidad no es de mucho interés. Entonces, lo que se busca es una prueba estadística bajo la hipótesis nula de que el proceso es un RW3. Mientras los retornos no estén autocorrelacionados, aún en la presencia de heteroscedasticidad, el ratio de varianza se debe aproximar a uno conforme el número de observaciones se incremente ilimitadamente. Esto es, la varianza de la suma de los retornos debe ser igual a la suma de la varianza de los retornos. Sin embargo, la varianza asintótica del ratio de varianza va a depender claramente del tipo de heteroscedasticidad, por lo que el ratio de varianza se construye considerando una forma general de esta.

De esta manera, la prueba desarrollada por [Lo y MacKinlay \(1989\)](#) tiene dos hipótesis: la primera asume que la serie $\{r_t\}$ es un proceso RW1 con errores i.i.d. asumiendo homoscedasticidad; y la segunda asume que es un proceso RW3 asumiendo heteroscedasticidad. Se definen los estimadores de la media de la primera diferencia y la varianza de la diferencia de orden q :

$$\hat{\omega} = \frac{1}{T} \sum_{t=1}^T (p_t - p_{t-1}), \quad (35)$$

$$\hat{\sigma}^2(q) = \frac{1}{Tq} \sum_{t=1}^q (p_t - p_{t-q} - q\hat{\omega})^2, \quad (36)$$

y se construye el correspondiente ratio de varianza:

$$VR(q) = \frac{\hat{\sigma}^2(q)}{\hat{\sigma}^2(1)}. \quad (37)$$

El estimador de la varianza se corrige por sesgo, tal como lo sugiere [Lo y MacKinlay \(1989\)](#)), reemplazando T por $(T - q + 1)$ en el caso que no se considere intercepto, o por $(T - q + 1)(1 - q/T)$ para el caso en el que se le considere. [Lo y MacKinlay \(1989\)](#) muestran que el ratio tiene una distribución asintótica normal para un apropiado estimador de $\hat{\sigma}(q)$:

$$z(q) = \frac{VR(q) - 1}{\hat{\sigma}(q)} \sim \mathcal{N}(0, 1). \quad (38)$$

Bajo la hipótesis de RW1, $\hat{\sigma}(q)$ se define como

$$\hat{\sigma}^2(q) = \frac{2(2q - 1)(q - 1)}{3qT}, \quad (39)$$

y bajo la hipótesis de RW3 o martingala en diferencia, se usa el estimador de Kernel:

$$\hat{\sigma}^2(q) = \sum_{j=1}^{q-1} \left(\frac{2(q-j)}{q} \right)^2 \delta_j \quad (40)$$

donde

$$\hat{\delta}_j = \frac{\sum_{t=j+1}^T (p_{t-j} - \hat{\mu})^2 (p_t - \hat{\mu})^2}{\left(\sum_{t=j+1}^T (p_{t-j} - \hat{\mu})^2 \right)^2}$$

4. RESULTADOS

En esta sección se presentan los resultados del análisis de datos y de las pruebas realizadas bajo la hipótesis nula de que el retorno del IGBVL es un paseo aleatorio. Para ello, se consideran datos diarios desde enero de 2006 hasta setiembre de 2021, extraídos del portal web del Banco Central de Reserva del Perú. El análisis de datos y las pruebas se realizan para datos diarios, semanales, mensuales y trimestrales, con la finalidad de tratar de identificar patrones particulares para cada frecuencia de datos.

4.1. ANÁLISIS DE DATOS

En el Anexo A se muestra el gráfico del IGBVL para las distintas frecuencias. Se parte de datos con frecuencia diaria para luego estimar los datos semanales, mensuales y trimestrales tomando el valor del índice de los viernes, los de fin de mes y de cada trimestre respectivamente.

Como se observa, en logaritmos el índice muestra una evolución similar a un paseo aleatorio. Con respecto al retorno continuo, en todos los casos se muestra una agrupación de volatilidad, no obstante, conforme la frecuencia de datos disminuye, este patrón va diluyéndose, lo cual está influenciado por el número de observaciones. Asimismo, los histogramas muestran una agrupación de los datos alrededor de la media además de presentar observaciones atípicas para el caso de los datos diarios, semanales y mensuales, mientras que en el caso de los trimestrales se observa una menor dispersión. El mismo patrón se observa en los gráficos por cuantiles, en donde se observa que la distribución de los datos no sería normal para el caso de frecuencia diaria, semanal y mensual, mientras que los trimestrales, se aproximan a una distribución normal.

El Cuadro 1 muestra los principales estadísticos del retorno continuo. Como se observa, el valor del coeficiente de asimetría (Skewness) es negativo en todos los casos, lo cual indica que es más probable observar retornos por debajo de la media que por encima. Por otro lado, el valor de la curtosis en los datos diarios, semanales y mensuales, nos indica que tiene una distribución leptocúrtica ($curtosis > 3$). Es decir, que tiene valores extremos que no serían reproducidos por una distribución normal; mientras que, para el caso de datos trimestrales dicho valor es cercano a 3, por lo que estaría más cercano a una distribución normal, lo cual se corrobora en valor p (p -value) de la prueba Jarque-Bera, al calcularse un valor mayor 5% de significancia, por lo que se falla en rechazar la hipótesis nula de que los datos provienen de una distribución normal.

CUADRO 1. Principales Estadísticos del Retorno del IGBVL

	Diaria	Semanal	Mensual	Trimestral
Media	0,03	0,16	0,62	1,81
Mediana	0,02	0,21	0,72	1,48
Máximo	12,82	19,31	32,54	34,62
Mínimo	-13,29	-34,60	-46,65	-46,74
Desviación estándar	1,45	3,62	8,38	16,74
Asimetría	-0,54	-1,10	-0,58	-0,43
Curtosis	14,37	16,45	9,05	3,41
Jarque-Bera	22240,4	6317,9	297,3	2,4
Probabilidad	0,00	0,00	0,00	0,30
Observaciones	4092	816	188	62

4.2. PASEO ALEATORIO 1 (RW1)

En el Cuadro 2 se muestran los resultados de las pruebas de secuencias y reversas. Tal como se observa, en todos los casos el valor del estadístico CJ está por encima de 1, mientras que la probabilidad se encuentra por debajo de 0,05, excepto para los datos trimestrales. Por lo tanto, se rechaza la hipótesis nula de que

los retornos del IGBVL se comportan como un paseo aleatorio, excepto para los datos trimestrales.

CUADRO 2. Prueba de Secuencias y Reversas

	Diaria	Semanal	Mensual	Trimestral
N	4 092	816	188	62
N_s	2 228	442	107	36
N_r	1 864	374	81	26
CJ	1,20	1,18	1,32	1,38
μ	0,03	0,16	0,62	1,81
σ	1,45	3,62	8,38	16,74
π	0,51	0,52	0,53	0,54
π_s	0,50	0,50	0,50	0,50
$CJsd$	0,03	0,07	0,15	0,26
z	6,24	2,59	2,18	1,48
probabilidad	0,00	0,00	0,01	0,07

En el Cuadro 3 podemos ver los resultados de la prueba de rachas. Tal como se observa, se rechaza la hipótesis nula de que los retornos son i.i.d. a un nivel de significancia $\alpha = 5\%$, por lo que no muestra evidencia a favor de la hipótesis de que el IGBVL sea un proceso similar a un paseo aleatorio RW1, excepto para el caso de los datos trimestrales, en el cual fallamos en rechazar la hipótesis nula. Por lo tanto, bajo estos resultados, no hallamos evidencia a favor de la HME, excepto para el caso de los datos trimestrales.

CUADRO 3. Prueba de Rachas

	Diaria	Semanal	Mensual	Trimestral
N	4 092	816	188	62
N1	2 084	435	97	32
N2	2 008	381	91	30
Media	2 046	407	95	32
Desviación estándar	32,0	14,2	6,8	3,9
IC inferior 5 %	1 983,6	379,4	81,5	24,3
IC superior 5 %	2 109,0	435,1	108,3	39,6
Rachas	1 864	374	81	26
	Rechaza	Rechaza	Rechaza	Falla en rechazar

En el Cuadro 4 se muestran los resultados de la prueba BDS. En la prueba se considera la distancia en función de la desviación estándar de un valor de 0,5 y de dimensiones⁸ 2, 3, 4 y 5. Tal como se observa, en todos los casos, rechazamos la hipótesis nula de que el IGBVL se comporta como un paseo aleatorio RW1, incluso para los datos trimestrales. Por lo tanto, bajo estos resultados, no hallamos evidencia a favor de la HME.

⁸ En el Anexo B se muestra más resultados para distancias en función de la desviación estándar de 1,0 y 1,5. Los resultados nos llevan a las mismas conclusiones: en todos los casos se rechaza la hipótesis nula al 10% de significancia.

CUADRO 4. Prueba BDSDistancia: $\sigma = 0,5$

	Dimensión	BDS	Error estándar	Estadístico z	Probabilidad
Diaria	2	0,03	0,00	21,58	0,00
	3	0,06	0,00	26,04	0,00
	4	0,09	0,00	28,35	0,00
	5	0,12	0,00	30,25	0,00
Semanal	2	0,03	0,00	8,56	0,00
	3	0,05	0,01	9,77	0,00
	4	0,07	0,01	10,76	0,00
	5	0,09	0,01	11,59	0,00
Mensual	2	0,01	0,01	1,75	0,08
	3	0,03	0,01	2,92	0,00
	4	0,04	0,01	3,36	0,00
	5	0,05	0,02	3,38	0,00
Trimestral	2	0,04	0,01	4,48	0,00
	3	0,07	0,02	4,44	0,00
	4	0,09	0,02	4,47	0,00
	5	0,09	0,02	4,19	0,00

Los resultados de las pruebas bajo la hipótesis de que el IGBVL es RW1 se rechazan en la mayoría de los casos, siendo la excepción el de los datos trimestrales, en el que la prueba de rachas nos indica que los retornos son i.i.d. Por lo tanto, no encontramos evidencia a favor de la versión débil de la HME bajo estos resultados.

4.3. PASEO ALEATORIO 3 (RW3)

En el Cuadro 5 se muestran los resultados de las pruebas de autocorrelación (AC) de los retornos continuos del IGBVL. Tal como se observa, en la mayoría de los casos, se rechaza la hipótesis nula de que los retornos no están autocorrelacionados, por lo que el IGBVL no se asemejaría a un proceso paseo aleatorio RW3, siendo la excepción los datos trimestrales, aunque en ese caso se rechaza la hipótesis nula en el primer rezago al 5 % de significancia. Por lo tanto, no hallamos evidencia a favor de la versión débil de la HME, excepto para los datos trimestrales.

En el Gráfico 1 se muestra la función de autocorrelación de los retornos hasta el rezago 24 de manera continua y las bandas punteadas representan el nivel de significancia de 5 % de los coeficientes de correlación construido como $\pm 1,96/\sqrt{T}$ donde T es el número de observaciones, por lo que aquellas barras que sobresalgan de esa banda de confianza se consideran como significativos estadísticamente hablando. Como se observa, se identifican coeficientes de correlación significativos, dando indicio de que hay posibles patrones a identificar.

CUADRO 5. Prueba de Autocorrelación

	Rezago	AC	Estadístico Q	Prob
Diaria	1	0,12	60,49	0,00
	3	0,05	71,37	0,00
	6	-0,02	105,91	0,00
	12	0,04	123,10	0,00
	24	0,01	147,99	0,00
Semanal	1	0,13	13,19	0,00
	3	0,06	26,92	0,00
	6	0,03	27,95	0,00
	12	0,01	62,71	0,00
	24	-0,04	75,18	0,00
Mensual	1	0,17	5,35	0,02
	3	0,11	20,88	0,00
	6	0,03	23,65	0,00
	12	-0,04	30,91	0,00
	24	-0,05	39,36	0,03
Trimestral	1	0,27	4,69	0,03
	3	-0,11	5,82	0,12
	6	-0,10	6,99	0,32
	12	-0,02	13,19	0,36
	24	-0,03	17,99	0,80

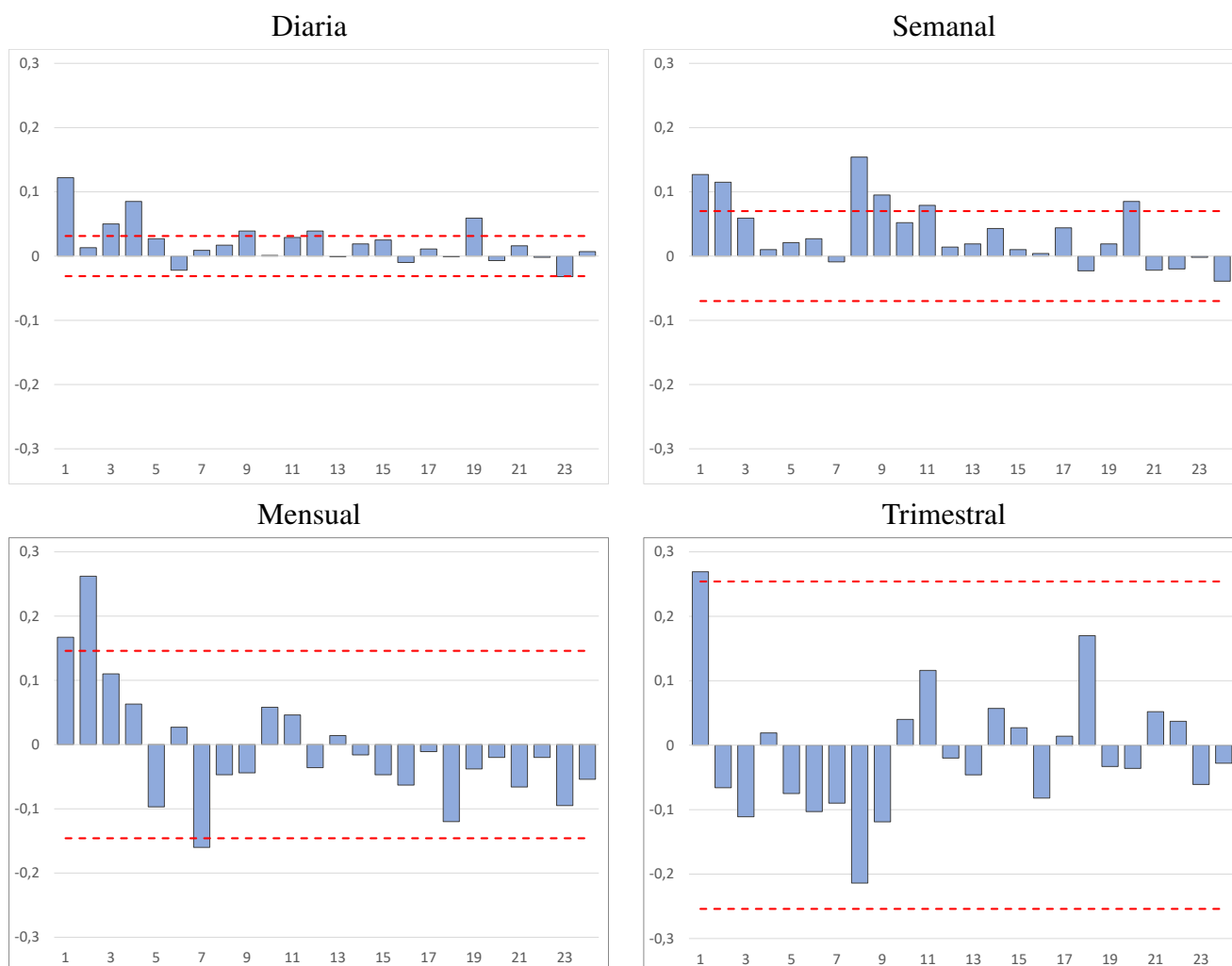
Una manera rápida de identificar posibles patrones es mediante la estimación de modelos ARIMA, los cuales explotan la autocorrelación de las series de tiempo para poder realizar predicciones, bajo el supuesto de que los errores no están autocorrelacionados. Para ello, escogemos el mejor modelo ARIMA en cada caso, a partir de un modelo ARMA(3, 3) usando como criterio de selección de rezagos al de Schwarz:⁹

- En datos diarios, el modelo fue un ARMA(1, 2), pero se rechaza la hipótesis nula de que los errores no estén autocorrelacionados.
- En datos semanales, el modelo fue un ARMA(1, 1) y se encontró que los errores no estarían autocorrelacionados.
- En datos mensuales, el modelo fue un ARMA(2, 0) y se encontró que los errores no estarían autocorrelacionados.
- En datos trimestrales no se probó, debido a que la función de autocorrelación de los retornos no están autocorrelacionados.

Como se observa en dos de los tres casos posibles se halla un modelo ARMA que reúne las condiciones necesarias para realizar predicción, y una búsqueda exhaustiva de un mejor modelo ARIMA debería arrojar mejores resultados en hallar un patrón.

⁹ Los detalles están en el Anexo C.

GRÁFICO 1. Función de Autocorrelación



En el Cuadro 6 se observan los resultados del ratio de varianza de [Lo y MacKinlay \(1989\)](#), asumiendo incrementos heteroscedásticos en el paseo aleatorio, una media distinta de cero y corrigiendo por el sesgo de varianza. Como se observa, en la mayoría de los casos rechazamos la hipótesis nula de que el IGBVL se comporta como un RW3 al 5% de significancia, aunque, como en los casos anteriores, para el caso de datos trimestrales hay evidencia a favor de que el índice sea RW3, y en menor medida en el caso de datos mensuales cuando se compara la varianza de dos periodos consecutivos. Por lo tanto, no hallamos evidencia a favor de la versión débil de la HME, excepto para los datos trimestrales.

Como se observa, en ninguna de las pruebas se encuentra evidencia a favor de la HME, excepto para datos trimestrales. De esta manera, a la luz de los resultados, se puede concluir que el mercado financiero en el Perú no es eficiente de acuerdo con la versión débil de la HME.

CUADRO 6. Prueba de Ratio de Varianza (VR)

	Periodo	VR	Error estándar	Estadístico z	Prob.
Diaria	2	1,12	0,04	3,05	0,00
	3	1,17	0,06	2,92	0,00
	4	1,22	0,07	3,05	0,00
	5	1,29	0,08	3,40	0,00
	6	1,34	0,09	3,61	0,00
Semanal	2	1,14	0,06	2,35	0,02
	3	1,26	0,10	2,72	0,01
	4	1,35	0,12	2,77	0,01
	5	1,40	0,15	2,71	0,01
	6	1,44	0,16	2,70	0,01
Mensual	2	1,17	0,09	1,88	0,06
	3	1,42	0,13	3,11	0,00
	4	1,60	0,17	3,54	0,00
	5	1,74	0,20	3,75	0,00
	6	1,80	0,22	3,62	0,00
Trimestral	2	1,28	0,17	1,64	0,10
	3	1,32	0,25	1,30	0,19
	4	1,27	0,30	0,89	0,37
	5	1,20	0,35	0,59	0,56
	6	1,08	0,38	0,20	0,84

Dados los resultados, se puede inferir que, si los inversionistas tratan de hallar patrones en los datos diarios, semanales y mensuales, deberían encontrarlos.¹⁰ No obstante, esto no implica que al explotarlos vayan a obtener ganancias extraordinarias, puesto que de hecho todos los que actualmente están en el mercado lo vienen haciendo. Por lo tanto, no hay que perder de vista que los resultados hallados sólo se centran en las características de los retornos del mercado bursátil, no en la forma que estos pueden ser explotados, por lo que la prueba empírica de la versión débil de la HME queda de alguna manera incompleta.

El hecho de que en datos trimestrales fallemos en rechazar la HME se puede explicar, en parte, por la forma en que estimamos los datos trimestrales al tomar los datos de fin de cada trimestre, haciendo que los patrones que se puedan hallar en datos de mayor frecuencia se diluyan.

5. CONCLUSIONES

En el presente trabajo se analizó la versión débil de la Hipótesis de Mercado Eficiente (HME) para el Perú durante el periodo 2006 - 2021. La prueba empírica consiste en investigar si el IGBVL tiene una trayectoria similar a la de un proceso de paseo aleatorio. Dicho patrón no se identificó en el caso de datos con frecuencia diaria, semanal y mensual, y sí lo hace para el caso de datos trimestrales. De esta manera, se concluye que la versión débil de la HME no se cumple en el caso peruano.

¹⁰ En realidad, lo más probable es que ya lo vengamos realizando, por lo que la idea va dirigida, principalmente, a los nuevos inversionistas.

Los resultados, interpretados bajo el enfoque teórico señalado, indican que el mercado de capitales peruano no es eficiente en el sentido de que los precios de los activos, a nivel agregado, no estarían reflejando toda la información disponible. En consecuencia, las implicancias de política económica apuntarían a que se deba mejorar el flujo de información en el mercado de capitales peruano, como por ejemplo, un mejor acceso gratuito a los precios históricos de los activos, a los balances de las empresas y a toda la información relevante para la toma de decisiones relacionada a estas. Asimismo, se debe crear los incentivos para que mayores empresas residentes coticen en la bolsa, impulsando así el incremento del número de inversionistas, promoviendo que la información fluya entre más participantes.

Definitivamente las medidas propuestas en este trabajo son insuficientes para la mejora en el flujo de información en el mercado de capitales peruano, teniendo en cuenta que no es el objetivo de este, pero consideramos que es una propuesta básica que se deben tener en cuenta para mejorar el acceso a la información con el objetivo de que los precios de los activos reflejen toda la información disponible de acuerdo con los resultados obtenidos.

Las siguientes investigaciones deberían estar enfocadas en realizar un análisis similar a nivel desagregado. Esto es, por el precio de cada activo financiero que esté registrado en el mercado de capitales peruano. Asimismo, se deberían estudiar otras anomalías en el mercado, como el efecto calendario, la relación entre volatilidad y correlación entre los activos, etc. Asimismo, se debe explorar posibles no linealidades en el patrón de los retornos de los activos en dicho mercado.

REFERENCIAS

- Bachelier, L. (1900). Teoría de la especulación. *Anales científicos de la Ecole Normale Supérieure*, serie 3, 17:21–86.
- Bodie, Z., Kane, A., y Marcus, A. (2018). *Investments*. McGraw Hill Education, 11era edición.
- Bollerslev, T., Engle, R. F., y Nelson, D. B. (1986). ARCH Models. En Engle, R. F. y McFadden, D., editores, *Handbook of Econometrics*, volumen 4 de *Handbook of Econometrics*, capítulo 49, pp. 2959–3038. Elsevier.
- Box, G. E. y Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332):1509–1526.
- Broock, W. A., Scheinkman, J. A., Dechert, W. D., y LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, 15(3):197–235.
- Campbell, J. Y., Lo, A. W., y MacKinlay, A. (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Delgado, L. y Humala, A. (1997). El mercado bursátil peruano y la hipótesis del mercado eficiente. *Revista Estudios Económicos*, 1:73–92.
- Duarte, J. B. D. y Pérez-Iñigo, J. M. M. (2014). Comprobación de la eficiencia débil en los principales mercados financieros latinoamericanos. *Estudios Gerenciales*, 30(133):365–375.
- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.
- Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5):1575–1617.
- Grossman, S. J. y Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3):393–408.
- Gujarati, D. y Porter, D. (2009). *Basic Econometrics 5th edition*. New York: McGraw-Hill/Irwin.
- Kendall, M. (1953). *The Analysis of Economic Time-series: Prices*. Número parte 1 en Reprint series / Research Techniques Unit of the London School of Economics and Political Science). London School of Economics and Political Science.
- Lewis, M. (2014). *Flash Boys*. New York: W.W. Norton & Company.
- Ljung, G. M. y Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.
- Lo, A. W. y MacKinlay, A. C. (1989). The size and power of the variance ratio test in finite samples: A monte carlo investigation. *Journal of econometrics*, 40(2):203–238.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1):59–82.

- Mandelbrot, B. (1966). Forecasts of future prices, unbiased markets, and "martingale" models. *The Journal of Business*, 39(1):242–255.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.
- Meneses, L. y Pérez, C. (2020). Análisis comparativo de eficiencia en mercados emergentes. el caso de Colombia, Chile y Perú. *Apuntes Contables*, pp. 9–24.
- Ruiz, B. y García, G. (2020). Hipótesis de mercados eficientes y estrategias de inversión en el MIB: 2014-2019. *Análisis Económico*, pp. 67–90.
- Samuelson, P. A. (1973). Proof that properly discounted present values of assets vibrate randomly. *The Bell Journal of Economics and Management Science*, 4(2):369–374.

ANEXOS

A. GRÁFICOS

GRÁFICO A-1. Datos a frecuencia diaria

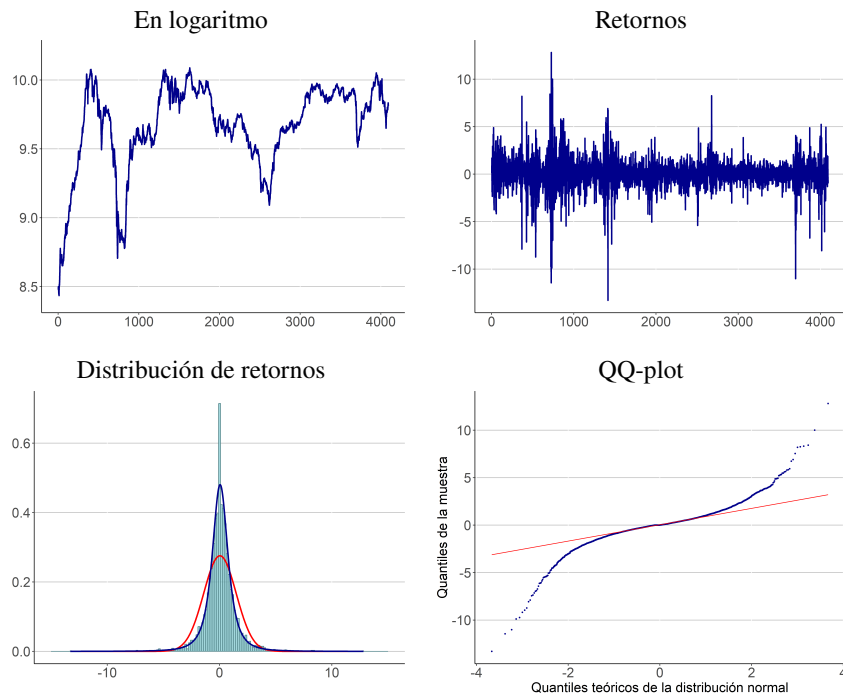


GRÁFICO A-2. Datos a frecuencia semanal

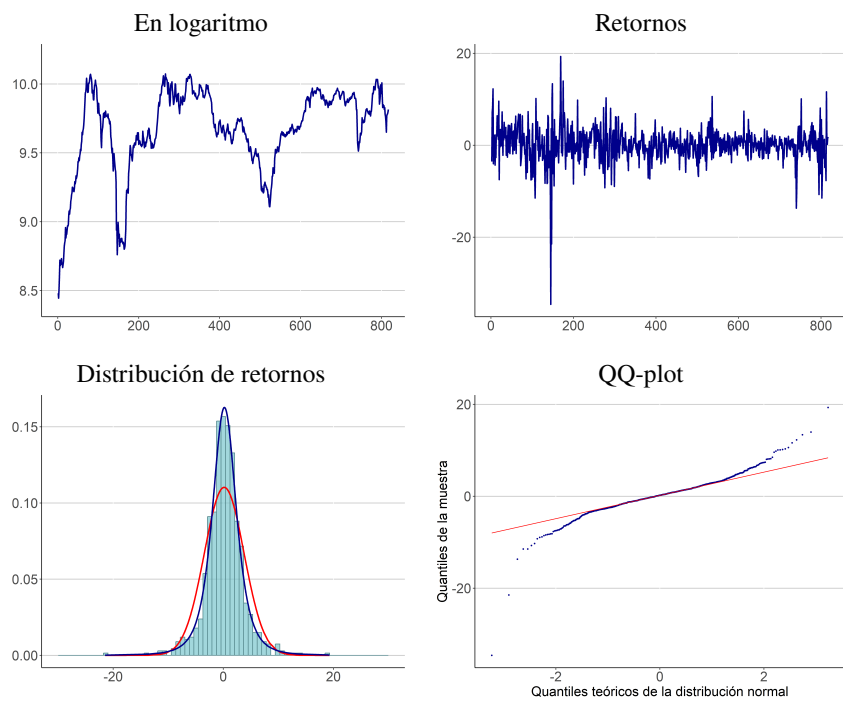


GRÁFICO A-3. Datos a frecuencia mensual

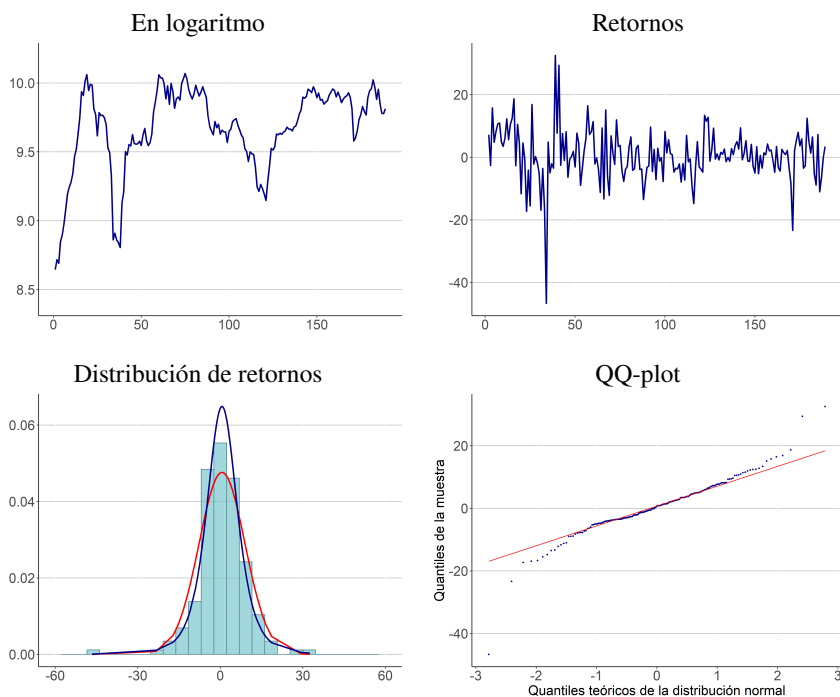
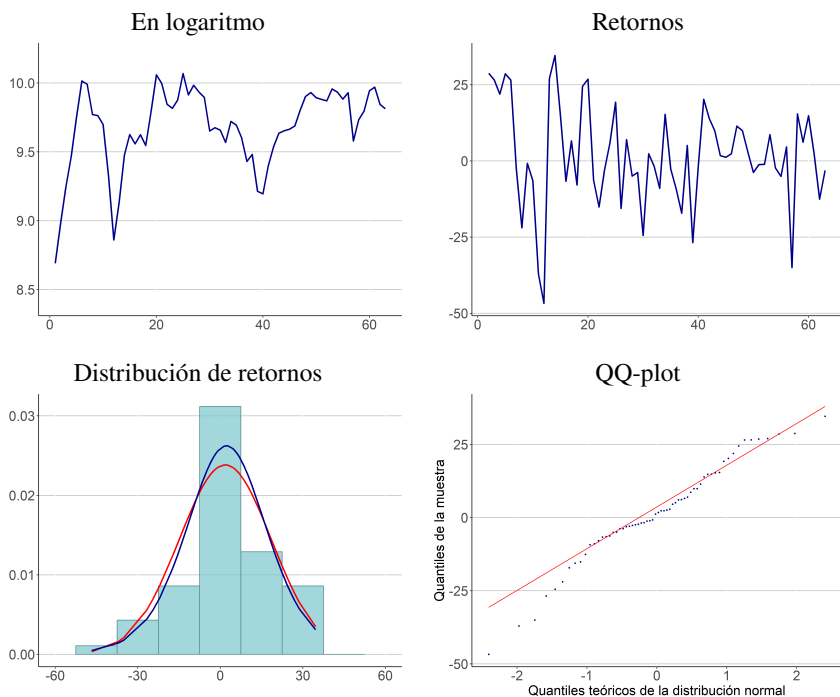


GRÁFICO A-4. Datos a frecuencia trimestral



B. PRUEBAS BDS**CUADRO B-1.** Distancia: $\sigma = 1,0$

	Dimensión	BDS	Error estándar	Estadístico z	Probabilidad
Diaria	2	0,03	0,00	18,62	0,00
	3	0,06	0,00	23,51	0,00
	4	0,07	0,00	27,09	0,00
	5	0,08	0,00	30,32	0,00
Semanal	2	0,03	0,00	8,61	0,00
	3	0,05	0,00	10,44	0,00
	4	0,06	0,00	11,94	0,00
	5	0,06	0,00	13,63	0,00
Mensual	2	0,02	0,01	2,99	0,00
	3	0,03	0,01	3,91	0,00
	4	0,04	0,01	4,28	0,00
	5	0,04	0,01	4,67	0,00
Trimestral	2	0,04	0,01	4,38	0,00
	3	0,05	0,01	4,03	0,00
	4	0,04	0,01	4,11	0,00
	5	0,04	0,01	4,18	0,00

CUADRO B-2. Distancia: $\sigma = 1,5$

	Dimensión	BDS	Error estándar	Estadístico z	Probabilidad
Diaria	2	0,03	0,00	21,58	0,00
	3	0,06	0,00	26,04	0,00
	4	0,09	0,00	28,35	0,00
	5	0,12	0,00	30,25	0,00
Semanal	2	0,03	0,00	8,56	0,00
	3	0,05	0,01	9,77	0,00
	4	0,07	0,01	10,76	0,00
	5	0,09	0,01	11,59	0,00
Mensual	2	0,01	0,01	1,75	0,08
	3	0,03	0,01	2,92	0,00
	4	0,04	0,01	3,36	0,00
	5	0,05	0,02	3,38	0,00
Trimestral	2	0,04	0,01	4,48	0,00
	3	0,07	0,02	4,44	0,00
	4	0,09	0,02	4,47	0,00
	5	0,09	0,02	4,19	0,00

C. MODELOS ARMA

CUADRO C-1. Modelos de series de tiempo

Variable dependiente Variable	Retornos diarios		Retornos semanales		Retornos mensuales	
	Coefficiente	Prob.	Coefficiente	Prob.	Coefficiente	Prob.
C	0,035 (0,05) [0,76]	0,45	0,18 (0,23) [0,78]	0,44	0,64 (1,02) [0,63]	0,53
AR(1)	0,98 (0,01) [116,61]	0,00	0,89 (0,05) [18,79]	0,00	0,13 (0,06) [2,01]	0,05
AR(2)				0,24 (0,07) [3,51]	0,00	
MA(1)	-0,86 (0,01) [-84,83]	0,00	-0,81 (0,06) [-13,59]	0,00		
MA(2)	-0,10 (0,01) [-12,96]	0,00				
SIGMASQ	2,05 (-0,02) [117,78]	0,00	12,71 (0,25) [50,90]	0,00	63,93 (3,73) [17,15]	0,00
R cuadrado	0,02		0,03		0,08	
R cuadrado ajustado	0,02		0,02		0,07	
Error estándar regresión	1,43		3,57		8,08	
Suma de residuos al cuadrado	8 407		10 369		12 019	
Log Verosimilitud	-7 279		-2 195		-658	
Estadístico F	20,75		7,86		5,67	
Prob(Estadístico F)	0		0,00		0,00	
Media de la variable dependiente	0,03		0,16		0,62	
Desv. est. de la variable dependiente	1,45		3,62		8,38	
Criterio Akaike	3,56		5,39		7,04	
Criterio Schwarz	3,57		5,41		7,11	
Criterio Hannan-Quinn	3,56		5,40		7,07	
Estadístico Durbin-Watson	2,00		1,94		2,01	
Número de observaciones	4 092		816		188	

NOTA: Números en paréntesis denotan errores estándar, números en corchetes denotan estadísticos t-student.
Método de estimación = BFGS.

CUADRO C-2. Correlograma del Error Estimado

Rezagos	Datos diarios				Datos semanales				Datos mensuales			
	AC	PAC	Estad. Q	Prob.	AC	PAC	Estad. Q	Prob.	AC	PAC	Estad. Q	Prob.
1	0,00	0,00	0,01	0,92	0,03	0,03	0,67	0,41	-0,01	-0,01	0,01	0,90
2	-0,01	-0,01	0,76	0,68	0,03	0,03	1,39	0,50	0,00	0,00	0,01	0,99
3	0,02	0,02	2,79	0,43	-0,02	-0,02	1,75	0,63	0,08	0,08	1,14	0,77
4	0,06	0,06	18,01	0,00	-0,06	-0,06	5,15	0,27	0,00	0,00	1,15	0,89
5	0,00	0,00	18,05	0,00	-0,04	-0,04	6,72	0,24	-0,11	-0,11	3,56	0,61
6	-0,05	-0,04	26,22	0,00	-0,03	-0,03	7,53	0,28	0,06	0,05	4,16	0,66
7	-0,01	-0,01	26,44	0,00	-0,07	-0,07	11,28	0,13	-0,15	-0,15	8,58	0,28
8	-0,01	-0,01	26,54	0,00	0,12	0,12	22,80	0,00	-0,06	-0,04	9,22	0,32
9	0,02	0,02	28,41	0,00	0,06	0,05	25,33	0,00	-0,02	-0,03	9,32	0,41
10	-0,02	-0,02	30,49	0,00	0,01	-0,01	25,45	0,01	0,09	0,10	10,91	0,36
11	0,01	0,01	30,86	0,00	0,057	0,04	27,27	0,00	0,06	0,08	11,64	0,39
12	0,02	0,02	32,70	0,00	-0,02	-0,01	27,62	0,01	-0,06	-0,09	12,34	0,42
13	-0,02	-0,03	34,82	0,00	-0,01	0,00	27,69	0,01	0,02	0,01	12,44	0,49
14	0,00	0,00	34,84	0,00	0,02	0,03	28,05	0,01	0,01	-0,03	12,46	0,57
15	0,01	0,01	35,27	0,00	-0,01	0,01	28,16	0,02	-0,05	-0,03	12,93	0,61
16	-0,03	-0,03	38,74	0,00	-0,02	-0,02	28,37	0,03	-0,03	-0,04	13,15	0,66
17	0,00	0,00	38,75	0,00	0,03	0,02	29,22	0,03	0,03	0,04	13,33	0,71
18	-0,02	-0,02	40,96	0,00	-0,04	-0,04	30,60	0,03	-0,11	-0,07	16,00	0,59
19	0,05	0,05	50,75	0,00	0,01	0,00	30,67	0,04	-0,01	-0,02	16,02	0,66
20	-0,03	-0,03	54,27	0,00	0,09	0,09	36,92	0,01	0,02	0,00	16,13	0,71
21	0,01	0,01	54,35	0,00	-0,03	-0,03	37,72	0,01	-0,04	-0,05	16,44	0,75
22	-0,01	-0,02	55,14	0,00	-0,02	-0,04	38,21	0,02	0,01	0,03	16,47	0,79
23	-0,05	-0,05	64,00	0,00	-0,00	0,00	38,21	0,02	-0,07	-0,11	17,47	0,79
24	0,00	0,00	64,02	0,00	-0,04	-0,03	39,61	0,02	-0,02	-0,01	17,56	0,82
25	0,00	0,00	64,10	0,00	-0,04	-0,05	40,96	0,02	-0,03	-0,05	17,76	0,85
26	-0,02	-0,02	65,27	0,00	0,03	0,05	41,78	0,03	-0,06	-0,06	18,44	0,86
27	0,01	0,02	65,65	0,00	-0,02	-0,01	42,07	0,03	-0,02	-0,02	18,57	0,89
28	0,04	0,04	72,76	0,00	-0,01	-0,06	42,22	0,04	-0,02	-0,05	18,67	0,91
29	0,02	0,02	74,99	0,00	0,04	0,04	43,30	0,04	-0,04	0,00	18,95	0,92
30	-0,04	-0,04	80,03	0,00	-0,02	-0,01	43,51	0,05	0,00	-0,05	18,95	0,94
31	-0,03	-0,03	83,01	0,00	0,00	-0,02	43,51	0,07	0,01	0,00	18,96	0,96
32	0,01	0,01	83,52	0,00	-0,07	-0,06	47,78	0,04	0,03	0,02	19,18	0,96
33	-0,04	-0,05	90,14	0,00	-0,05	-0,03	49,93	0,03	0,03	0,01	19,39	0,97
34	-0,04	-0,03	95,78	0,00	0,01	-0,01	49,96	0,04	0,01	0,01	19,44	0,98
35	0,01	0,02	96,20	0,00	-0,06	-0,05	52,53	0,03	0,10	0,08	21,73	0,96
36	-0,02	-0,03	98,61	0,00	0,01	0,02	52,55	0,04	-0,08	-0,09	23,27	0,95

NOTA: AC = Autocorrelación, PAC = Autocorrelación parcial.