

ALGORITMOS GENÉTICOS Y MODELOS MULTIVARIADOS RECURSIVOS EN LA PREDICCIÓN DE ÍNDICES BURSÁTILES DE AMÉRICA DEL NORTE: IPC, TSE, NASDAQ Y DJI*

*Antonino Parisi, Franco Parisi
y Edinson Cornejo***

RESUMEN

Con valores de cierre semanales, correspondientes al periodo del 7 de abril de 1998 al 14 de abril de 2003, analizamos la eficiencia de los modelos multivariados dinámicos, elaborados a partir de algoritmos genéticos recursivos, para predecir el signo de las variaciones semanales de los índices bursátiles IPC, TSE, Nasdaq y DJI. Los resultados fueron comparados con los de un modelo AR(1) y de un modelo multivariado elaborado de manera aleatoria. Los mejores modelos producidos por el algoritmo genético obtienen un porcentaje de predicción de signo (PPS) de 59, 60, 59 y 59%, para los índices IPC, Nasdaq, TSE y DJI, respectivamente. La capacidad predictiva resultó significativa en cada uno de los índices, de acuerdo con la prueba de acierto direccional de Pesaran y Timmerman (1992). Al analizar el PPS de los modelos AR(1) se encontró que éstos fueron menores, resultando significativos únicamente en el caso del Nasdaq. Los modelos multivariados dinámicos elaborados de manera aleatoria presentaron el PPS más bajo (excepto en el índice TSE), siendo significativo para el Nasdaq sólo al considerar una significación de 10%. Además, los modelos elaborados por el algoritmo genético generaron el mayor rendimiento acumulado, excepto en el caso del Nasdaq, en el que la rentabilidad más alta fue obtenida por el modelo AR(1). Al efectuar una prueba de solidez por medio del análisis de mil series *bootstrap* se observó que, en promedio, el PPS fue de 51, 53, 49 y 53%, para los índices DJI, IPC, Nasdaq y TSE, respectivamente. Pese a ello los modelos multivariados

* *Palabras clave:* algoritmos genéticos, modelo multivariado dinámico, funcionamiento recursivo, porcentaje de predicción de signo, prueba de acierto direccional. *Clasificación JEL:* G10, G14, G15. Artículo recibido el 14 de julio de 2003 y aceptado el 7 de abril de 2004. Los autores desean agradecer a Washington Macías, por su labor como asistente de investigación, a los comentarios realizados por los académicos José Luis Guerrero y Jorge Gregoire, y a los participantes en el seminario realizado en el Departamento de Administración de la Universidad de Chile.

** A. Parisi es profesor asistente, Departamento de Administración, Universidad de Chile. F. Parisi es profesor asistente, Management Department, Jesse Jones Graduate School of Management, Rice University (correo electrónico: fparisi@rice.edu). E. Cornejo es instructor, Departamento de Administración, Universidad de Chile.

superaron el rendimiento de una estrategia *buy and hold* en 57, 59 y 71% de los casos de los índices DJI, IPC y TSE, respectivamente. En el Nasdaq la frecuencia con que el modelo multivariado superó en rentabilidad a la estrategia pasiva fue de 41 por ciento.

ABSTRACT

Using weekly stock index prices, corresponding to the period between April 07 of 1998 and April 14 of 2003, we analyzed the efficiency of the dynamic multivariate models, from recursive genetic algorithms, to forecast the weekly sign variations of stock-exchange indices IPC, TSE, Nasdaq and DJI. The results were compared with those of a multivariate model AR(1) and a random model. The best models produced by the genetic algorithm threw a sign prediction percentage (SPP) of 59%, 60%, 59% and 59%, for indices IPC, Nasdaq, TSE and DJI, respectively. The forecast capacity was significant in each index, according to Pesaran & Timmerman's directional accuracy test (1992). When analyzing the SPP of the models AR(1), were smaller, being significant only in the case of the Nasdaq. The random dynamic multivariate models presented the lowest SPP (except in index TSE), being significant only in the case of Nasdaq. In addition, the models constructed by the genetic algorithm generated the greater accumulated return, except in the case of the Nasdaq, where the highest yield was registered by the model AR(1). In the test of robustness through the analysis of 1 000 bootstrap series, in average, the SPP was of 50.88%, 52.58%, 49.07%, 52.93%, for indices DJI, IPC, Nasdaq and TSE. The multivariate models surpassed the return of a *buy and hold* strategy in 57%, 59% and 71%, DJI, IPC and TSE, respectively.

INTRODUCCIÓN

Este estudio analiza la eficiencia de los modelos multivariados dinámicos elaborados a partir de algoritmos genéticos para predecir el signo de las variaciones semanales de los índices bursátiles IPC,¹ TSE,² DJI³ y Nasdaq.⁴ El presente artículo busca mostrar la utilidad de los algoritmos genéticos en la determinación de un modelo multivariado dinámico que maximice el porcentaje de predicción de signo, entendiendo que la predicción de la dirección del movimiento del índice accionario es pertinente para elaborar estrategias de transacción efectivas (Leung, Daouk y Chen, 2000).

¹ Índice de precios y cotizaciones (IPC) de la Bolsa Mexicana de Valores (BMV).

² Toronto Stock Exchange, Canadá.

³ Dow Jones Industry, Estados Unidos.

⁴ Nasdaq, índice de acciones tecnológicas, Estados Unidos.

La hipótesis de mercados eficientes (Fama, 1970) plantea que el mercado refleja completa y correctamente toda la información pertinente para la determinación de los precios de los activos. Debido a que el surgimiento de nueva información es de carácter aleatorio, los cambios de los precios accionarios también lo serían. Esto ha llevado a muchos analistas financieros y académicos a señalar que las fluctuaciones de los precios accionarios siguen una caminata aleatoria (*random walk*), en la que el concepto de aleatoriedad se refiere a que las variaciones de precios son generadas a partir de un cierto proceso estocástico. No obstante, varios estudios han concluido que existe evidencia significativa de que los precios accionarios no siguen una caminata aleatoria y muestran que los rendimientos accionarios son predecibles en algún grado (véase, por ejemplo, Lo y MacKinley, 1988; Conrad y Kaul, 1988, 1989; DeBondt y Thaler, 1985; Fama y French, 1988; Poterba y Summers, 1988; Chopra, Lakonishok y Ritter, 1992, y Blume, Easley y O'Hara, 1994). Es decir, se ha mostrado que la idea del carácter azaroso de las fluctuaciones de los precios bursátiles es errónea. En consecuencia, se supone que las regularidades que puedan contener las series históricas de variaciones de precios e índices bursátiles son percibidas por los modelos de series de tiempo, de redes neuronales y de algoritmos genéticos.

No obstante, el grado de predictibilidad de los rendimientos es por lo general considerado económicamente no significativo, por lo que Hodgson y Nicholls (1991) sugieren evaluar la significación económica de predecir la dirección de los cambios en los precios de los activos. En esta materia, Leung, Daouk y Chen (2000) compararon la capacidad predictiva de los modelos de clasificación⁵ con los de estimación de nivel⁶ y concluyeron que los primeros⁷ se desempeñan mejor que los segundos en términos de su tasa de acierto,⁸ y son capaces de generar ganancias más altas. Esto último resulta relevante para los analistas de mercado, agentes de bolsa y para toda la comunidad financiera, ya que indican que deben centrar sus esfuerzos en

⁵ *Linear discriminant analysis, logit model, probit model y probabilistic neural network.*

⁶ *Adaptive exponential smoothing, vector autoregression with Kalman filter, multivariate transfer function y multilayered feedforward neural network.*

⁷ Los cuales son propuestos para predecir el signo o la dirección del rendimiento de los índices bursátiles S&P 500, FTSE 100 y TSE.

⁸ Medida por el número de veces en que la dirección pronosticada es correcta.

predecir con precisión la dirección de los movimientos de los precios e índices bursátiles, en vez de minimizar la desviación de las estimaciones de los valores observados.

Basado en lo anterior este estudio analiza la eficiencia de una técnica novedosa como los modelos multivariados dinámicos elaborados a partir de algoritmos genéticos recursivos para predecir el signo de las variaciones semanales de los índices bursátiles. En este caso los índices seleccionados son el IPC, TSE, DJI y Nasdaq, los cuales fueron escogidos por su relevancia mundial. Los mejores modelos producidos por el algoritmo genético recursivo arrojaron un porcentaje de predicción de signo de 59, 60, 59 y 59%, para los índices IPC, Nasdaq, TSE y DJI, respectivamente. La capacidad predictiva resultó significativa en cada uno de los índices, de acuerdo con la prueba de acierto direccional de Pesaran y Timmermann (1992). Al analizar el porcentaje de predicción de signo de los modelos AR(1), se encontró que éstos fueron menores (50, 59, 45 y 51.5%, respectivamente) y resultaron significativos sólo en el caso del Nasdaq. En cuanto a los modelos multivariados dinámicos elaborados de manera aleatoria, éstos presentaron la capacidad predictiva más baja (excepto en el índice TSE), siendo significativa sólo en los índices Nasdaq y TSE al considerar una significación de 10%. Además, los modelos elaborados por el algoritmo genético recursivo generaron el mayor rendimiento acumulado para los índices IPC, TSE y DJI. Al analizar la rentabilidad acumulada que se habría obtenido siguiendo las recomendaciones de compra-venta de los modelos multivariados en mil series *bootstrap* se encontró que estos superaron el rendimiento de una estrategia *buy and hold* en 57, 59 y 71% de los casos en los índices DJI, IPC y TSE, respectivamente. En el Nasdaq esta frecuencia fue de 41%. Esto sugeriría que sería mejor gestionar la cartera indizada en función de un modelo de proyección que aplicar una estrategia de inversión pasiva, al menos en los casos del DJI, IPC y TSE.

El documento se divide en tres secciones: la primera contiene una breve reseña de los algoritmos genéticos; la sección II describe la metodología utilizada, y la sección III presenta los resultados del modelo de algoritmos genéticos y los compara con los de un modelo Box-Jenkins y un modelo multivariado generado aleatoriamente; al final se presenta las conclusiones del estudio.

I. ALGORITMOS GENÉTICOS

Los algoritmos genéticos, establecidos por Holland (1975), consisten en una función matemática o una rutina que simula el proceso evolutivo de las especies, teniendo como objetivo encontrar soluciones a problemas específicos de maximización o minimización.⁹ Así, el algoritmo genético recibe como entrada una generación de posibles soluciones para el problema de que se trate, y arroja como salida los especímenes más aptos (es decir, las mejores soluciones) para que se apareen y generen descendientes, los que deberían tener mejores características que las generaciones anteriores. El cuadro 1 contiene una explicación de la terminología usada.

CUADRO 1. *Terminología*

<i>Términos</i>	<i>Definiciones</i>
Entorno o medio ambiente	Función objetivo
Individuo o ejemplar	Una de las posibles soluciones
Población	Conjunto de soluciones en un momento determinado
Generación	Nombre que identifica a la población, en un momento determinado
Cromosoma	Código binario
Genes	Los <i>bits</i> que conforman el código binario

Los algoritmos genéticos trabajan con códigos que representan a cada una de las posibles soluciones al problema. Por ello, es necesario establecer una codificación para todo el rango de soluciones, antes de comenzar a utilizar el algoritmo. Al respecto Davis (1994) señala que la codificación más utilizada es la representación de las soluciones por medio de cadenas binarias (conjuntos de ceros y unos).

Según Bauer (1994) este método puede ser utilizado fácilmente en aplicaciones financieras. Davis (1994) muestra una aplicación de algoritmos genéticos en la calificación de créditos bancarios que resultan mejor que otros métodos, como las redes neuronales, debido a la transparencia de los resultados obtenidos. Kingdon y Feldman (1995) usaron algoritmos genéticos para hallar reglas que pronosticaran la bancarrota de las empresas, estableciendo relaciones entre las distintas razones financieras. Bauer (1994) utilizó algoritmos genéticos

⁹ Al momento de aplicar un modelo de algoritmos genéticos se debe identificar correctamente el problema de maximización o minimización por analizar. Si el problema de que se trata no es de ese tipo, se deberá buscar otro método para abordarlo.

para desarrollar técnicas de transacción que indicaran la asignación mensual de montos de inversión en dólares y marcos; Pereira (1996) los utilizó para encontrar los valores óptimos de los parámetros usados por tres reglas de transacción distintas para el tipo de cambio dólar estadounidense/dólar australiano: los parámetros obtenidos mostraron resultados intramuestrales positivos, los cuales disminuyeron al aplicar las reglas fuera de la muestra, aun cuando continuaron siendo rentables. Allen y Karjalainen (1999) usaron algoritmos genéticos para aprender reglas de transacción para el índice S&P 500 y emplearlas como un criterio de análisis técnico y, una vez cubiertos los costos de transacción, encontraron que el exceso de rendimiento calculado sobre una estrategia *buy and hold*, durante el periodo de prueba extramuestral, no era congruente. No obstante, y a diferencia de Allen y Karjalainen, en este artículo se analiza la capacidad de los modelos elaborados por medio de algoritmos genéticos para proyectar el signo de las variaciones semanales de los índices bursátiles ya señalados y, en función de estas proyecciones, desarrollar estrategias de transacción. Kim y Han (2000) mostraron que los algoritmos genéticos pueden ser usados para reducir la complejidad y eliminar factores irrelevantes, lo que resultó mejor que los métodos tradicionales para predecir un índice de precios accionario.

Por otra parte, Feldman y Treleaven (1994) señalaron que la mayor desventaja de los algoritmos genéticos es la dificultad que presentan para escoger una técnica de codificación manejable, y para determinar el tipo de selección y las probabilidades de los operadores genéticos, ya que no hay reglas fijas en esta materia.

II. METODOLOGÍA

El objetivo del algoritmo genético es encontrar el modelo multivariado dinámico que maximice el porcentaje de predicción de signo de las variaciones semanales de los índices bursátiles IPC, TSE, DJI y Nasdaq. Los datos corresponden a valores de cierre semanales¹⁰ del periodo comprendido entre el 7 de abril de 1998 y el 14 de abril de 2003. Los modelos multivariados dinámicos utilizados son modelos

¹⁰ Se usaron valores de cierre semanales debido a que los administradores de fondos de inversión tienden a recomponer sus carteras en función de pronósticos semanales.

Los modelos multivariados dinámicos usados para predecir el signo de las fluctuaciones semanales de los índices bursátiles IPC, TSE, DJI y Nasdaq (IPC_t , TSE_t , DJI_t y NAS_t , respectivamente) se presentan a continuación,

$$\begin{array}{ccccccc}
NAS_t & & NAS_{t-1} & \dots & AR & NAS_t & AR-1 & t-1 & \dots & MA & t & MA \\
1 & & DJI_{t-1} & \dots & X & DJI_t & X & t & & & &
\end{array} \quad (4)$$

¹¹ Se trabajó con un máximo de cuatro rezagos para cada variable incluida en los modelos, debido a las restricciones que se tienen en términos de capacidad de procesamiento de datos y de

los modelos 1, 2 y 4) y 8 (para el modelo 3). Cada variable fue representada con un *bit* y, por tanto, el largo de la cadena binaria resultó igual al máximo número de variables del modelo.¹² De acuerdo con las matemáticas binarias, una cadena de largo L permite representar 2^L soluciones posibles. En consecuencia, contamos con un total de 4 096 modelos para predecir la dirección de las fluctuaciones de los índices IPC, TSE y Nasdaq (derivados de las ecuaciones 1, 2 y 4), y 256 modelos para predecir el comportamiento del DJI (derivados de la ecuación 3), entre los cuales el algoritmo genético realizará la búsqueda de la mejor solución.

2. Evaluación de la predicción

A continuación se evaluó la calidad de cada modelo en función del porcentaje de predicción de signo alcanzado (PPS). La evaluación se realizó sobre la base de un conjunto extramuestral de 200 datos semanales, por medio de un proceso recursivo. La recursividad ha sido empleada para medir el desempeño de modelos de redes neuronales que buscan predecir periodos de recesión en los Estados Unidos (Qi, 2001; Estrella y Mishkin, 1998) y para proyectar el signo de las variaciones de índices bursátiles internacionales (Parisi, Parisi y Guerrero, 2003). Para ello se dividió la muestra total en dos: una de tamaño n (62 datos) para estimar los coeficientes α , β y γ de cada modelo por medio de la minimización de la suma del cuadrado de los residuos del modelo, y otra de tamaño m (conjunto extramuestral de 200 datos) para evaluar la capacidad predictiva de los modelos. Para realizar esto último se comparó el signo de la proyección

tiempo. Incorporar un rezago adicional como variable explicativa significaría que el número de combinaciones o de posibles modelos aumentaría desde 4 096 a 32 768 (2^{15}), mientras que trabajar con un máximo de seis rezagos incrementaría el número de posibles modelos a 262 144 (2^{18}). Esto, dada la actual capacidad de procesamiento de datos, elevaría significativamente la complejidad y el tiempo necesario para buscar una solución al problema que nos interesa resolver, y lo haría inmanejable.

¹² El primer *bit* de la cadena representará al índice, en primera diferencia y rezagado un periodo; el *bit* 2 representará el rezago 2; el *bit* 5 representará el residuo del modelo rezagado un periodo; el *bit* 8 representará el cuarto rezago del residuo; el siguiente *bit* representará el primer rezago del DJI en primera diferencia, y, finalmente, el último *bit* representará el cuarto rezago. Cada *bit* puede tomar el valor de 0 o 1: cuando tome el valor de 0 la variable que representa no será incluida en el modelo, mientras que cuando tome el valor de 1 la variable representada sí formará parte del modelo. Una vez realizada la codificación de los modelos se comenzó a trabajar con el algoritmo genético. La primera generación de modelos (o población inicial), de un tamaño de 80, se obtuvo aleatoriamente desde el rango de interés.

con el signo de la variación observada en cada i -ésimo periodo, en el que $i = 1, 2, \dots, m$. Si los signos entre la proyección y el observado coinciden, entonces aumenta la efectividad del modelo analizado y, en caso contrario, disminuye su capacidad predictiva.

Una vez proyectado el signo de la variación del índice accionario para el periodo $n - 1$, la variación observada correspondiente se incluye en la muestra de tamaño de n con objeto de reestimar los coeficientes del modelo, contando ahora con una observación más.¹³ Así, el mismo modelo pero con sus coeficientes recalculados es utilizado para realizar la proyección correspondiente al periodo $n - 2$. Este procedimiento recursivo se efectuó una y otra vez hasta acabar con las observaciones del conjunto extramuestral. Finalmente, el PPS de cada modelo se calculó de la siguiente forma:

$$\text{Porcentaje} = \frac{\sum_{i=1}^{n-m} p_i}{m} \quad (5)$$

$$p_i = \begin{cases} 1, & \text{si } Index_i * \hat{Index}_i \geq 0 \\ 0, & \text{en otro caso} \end{cases}, i = [n-1, \dots, n-m] \quad (6)$$

en la que Δ representa la variación observada y $\hat{\Delta}$ la variación estimada. De esta manera, los modelos multivariados construidos por medio de algoritmos genéticos se estimaron por medio de mínimos cuadrados ordinarios, pero fueron evaluados en función de su capacidad para predecir el signo de los movimientos del índice bursátil. Además, en esta etapa se aplicó la prueba de acierto direccional de Pesaran y Timmermann (1992), con objeto de medir la significancia estadística de la capacidad predictiva de cada uno de los modelos analizados.

Luego, para analizar si la capacidad predictiva de los modelos se traduce en beneficios económicos, se calculó la rentabilidad acumulada que se habría obtenido si se hubiese comprado o vendido el índice bursátil siguiendo las recomendaciones de compra-venta del modelo de predicción. Para ello, la proyección de una variación positiva del índice (un alza del mercado) fue interpretada como una

¹³ De esta manera, la muestra uno (de tamaño n) se extiende a $n - 1$ observaciones y la muestra dos (de tamaño m) se contrae a $m - 1$ datos.

señal de compra, mientras que el pronóstico de una variación negativa (una caída del mercado) fue interpretado como una señal de venta. Se supuso una inversión inicial de 100 mil dólares, y la rentabilidad acumulada se calculó sobre un conjunto extramuestral de 200 semanas. Al momento de calcular la rentabilidad de la cartera no se realizaron ventas cortas y los costos de transacción no fueron considerados.

Por lo demás, con el objetivo de evitar el problema de *data snooping*,¹⁴ y de despejar las dudas respecto a si la capacidad predictiva se debe a la bondad del modelo, a las características de la muestra de observaciones a la que ha sido aplicado o sencillamente al factor suerte, se tomó el mejor modelo de proyección para cada índice (el de mayor PPS) y se lo evaluó sobre un total de mil conjuntos extramuestrales de 200 datos de cierre semanales cada uno. Estos mil conjuntos extramuestrales fueron generados a partir del conjunto extramuestral original utilizando un proceso de *block bootstrap*.¹⁵

3. Selección y generación de descendencia

Se utilizó el método *window* para asignar las probabilidades de selección a la población, usando como valor de calidad inicial el porcentaje de predicción de signo (PPS) de cada modelo. Este modo de asignación impide que el peor ejemplar de una generación sea elegido para la reproducción, otorgándole una probabilidad de selección nula debido a que el exceso de calidad sobre sí mismo es cero.

La selección de los padres de la nueva generación se realizó con el método de la ruleta truncada, especificando que el 10% más apto de la generación actual se traspasará a la siguiente generación. Se seleccionaron dos cromosomas padres para que generaran dos hijos, uti-

¹⁴ El término *data snooping* (también conocido como *data mining*) “ocurre cuando un determinado conjunto de datos es usado más de una vez para propósitos de inferencia o selección de modelos. Cuando esta reutilización de datos ocurre siempre existe la posibilidad de que cualquier resultado satisfactorio que se haya obtenido pueda deberse sencillamente a la suerte, en vez de a algún mérito inherente al modelo que generó los resultados”. (White, 2000).

¹⁵ Una manera de probar la bondad de los modelos y la validez de sus resultados, independientemente de la muestra de datos a los que han sido aplicados, es utilizar un proceso de *bootstrap*. El *bootstrap* es un proceso de generación de observaciones ficticias a partir de datos históricos, a fin de resolver el problema de escasez de datos y, de este modo, obtener suficiente información para elaborar diferentes conjuntos extramuestrales en los cuales probar la validez de los modelos.

lizando el operador de cruce doble hasta completar el tamaño de población deseado. Se permitió que la mutación ocurriera con una probabilidad de 8.33%.¹⁶ La clonación está implícita en el proceso ya que no se restringió el cruce entre dos modelos iguales. Una vez terminada la nueva generación de modelos se aplicó la función de calidad, continuando con el proceso de selección y generando nuevas descendencias. Este proceso finalizó al lograr 15 generaciones.

Por otra parte, se aplicó la metodología de Box-Jenkins para determinar el modelo ARIMA que describa mejor las fluctuaciones de los índices bursátiles analizados. Observando el correlograma de las variaciones de los índices bursátiles se identificaron tres posibles modelos: AR(1), ARMA(1,1) y MA(1). Para elegir el modelo más indicado se minimizó el criterio de información de Akaike y de Schwarz,

$$\text{Akaike info criterio (AIC)} \quad 2 - \frac{2k}{n} \quad (7)$$

$$\text{Schwarz criterio (SC)} \quad 2 - k \frac{\log n}{n} \quad (8)$$

en los que k es el número de parámetros estimados, n es el número de observaciones y es el valor de la función de verosimilitud utilizando los k parámetros estimados. Ambos criterios coincidieron en identificar al modelo AR(1) como el modelo que describe mejor el comportamiento de las variaciones de los índices estudiados. Además, los resultados fueron comparados con los proporcionados por modelos multivariados dinámicos seleccionados de manera aleatoria.

III. RESULTADOS

Al analizar la serie de valores de cierre, en primera diferencia, de los índices IPC, DJI, Nasdaq y TSE se encontró que, en los tres primeros casos, el coeficiente de asimetría es negativo, por lo que la distribución es desplazada o presenta un sesgo hacia la izquierda, como consecuencia de la existencia de valores extremos y poco usuales. Por su parte, la distribución de las variaciones de los valores de cierre semanales del TSE presenta un sesgo hacia la derecha.

¹⁶ Esto se hace generando un número aleatorio entre 0 y 1, cada vez que se reproduce un hijo. Si el número es menor de 0.01, entonces dicho hijo muta justo antes de formar parte de la población.

Para el IPC, DJI, Nasdaq y TSE el valor de la kurtosis fue superior a 3, por lo que la distribución de las observaciones adoptó una forma leptokurtósica, con las observaciones concentradas en un estrecho rango de valores y, en consecuencia, con una masa relativamente baja en las colas de las funciones de densidad. Al analizar los resultados de la prueba de Jarque-Bera se rechazó la hipótesis de que las variaciones de los valores de cierre semanales de los índices IPC, DJI, Nasdaq y TSE siguen una distribución normal, con una significación de 5%. Además, exceptuando el caso del TSE, los coeficientes de autocorrelación para AR(1) y AR(2) resultaron ser estadísticamente no significativos, al nivel de significación 5 y de 10%, por lo que no se encontró evidencia de autocorrelación entre la variación experimentada por los índices en el periodo t y la ocurrida en los periodos $t-1$ y $t-2$, respectivamente. Por último, las pruebas de Wald-Wolfowitz (de corridas) y Cox-Stuart (de signos no ponderados) no rechazaron la hipótesis nula de aleatoriedad en las variaciones de

CUADRO 2. *Estadísticos descriptivos para las series de valores de cierre semanales, en primera diferencia, de los índices bursátiles de América del Norte^a*

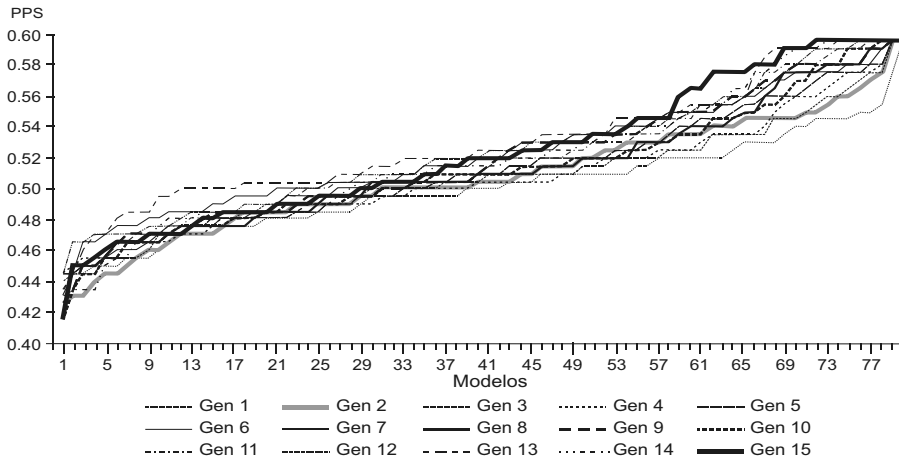
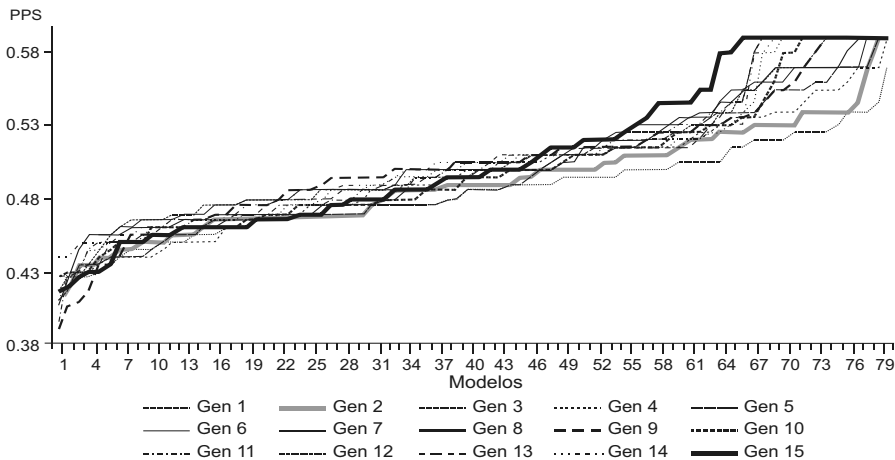
Estadísticos	IPC	TSE	DJI	Nasdaq
Media	5.287786	4.179008	2.814084	1.624122
Mediana	1.570000	6.500000	19.09500	0.315000
Máximo	975.0500	818.8000	666.4000	608.2700
Mínimo	1 224.180	1 163.300	1 370.100	1 125.160
Desviación estándar	250.5790	233.7263	274.9099	142.9195
Skewness	0.229762	0.749851	0.558025	1.792152
Kurtosis	5.761510	7.074174	5.0062633	18.05952
Jarque-Bera ^b	85.55502*	205.7573*	57.53801*	2 616.032*
AR(1) ^b	0.049325 (0.795001)	0.124209 (2.015458)*	0.028941 (0.466280)	0.031840 (0.512714)
AR(2) ^b	0.032715 (0.525897)	0.144456 (2.346084)*	0.028508 (0.458465)	0.044743 (0.719419)
Wald-Wolfowitz ^c	0.12	0.13	0.33	2.35*
Cox-Stuart ^c	0.09	1.49	0.44	0.96

^a Entre paréntesis se presenta la prueba t . Periodo: 7 de abril de 1998 al 14 de abril de 2003. 262 observaciones.

^b Los valores críticos de los estadísticos Jarque-Bera y t de Student son 4.43 y 1.96, respectivamente, para una 5 por ciento.

^c Los valores críticos de los estadísticos Wald-Wolfowitz y Cox-Stuart son de 1.96, para un 5 por ciento.

* Significativo al 5 por ciento.

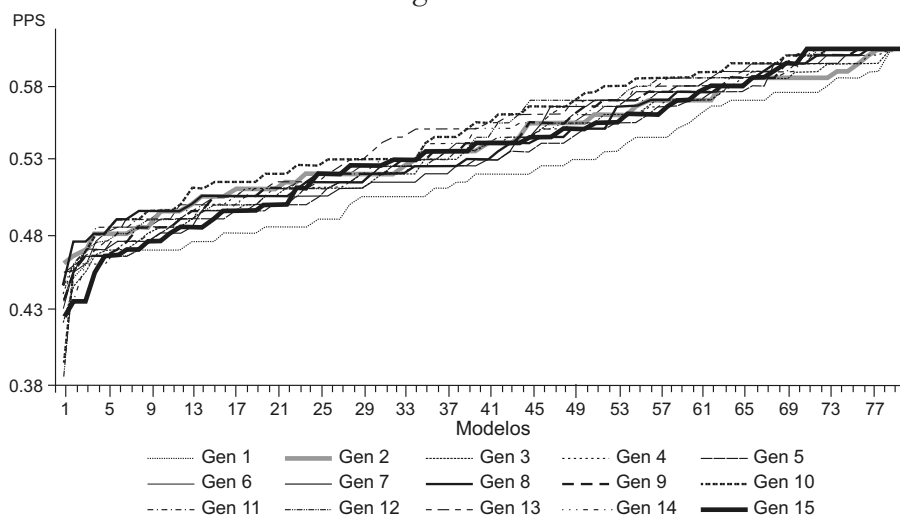
GRÁFICA 1. *IPC: Capacidad predictiva a lo largo de las generaciones***GRÁFICA 2.** *TSE: Capacidad predictiva a lo largo de las generaciones*

los índices accionarios (para el caso del Nasdaq, no existe evidencia concluyente para rechazar dicha hipótesis). Los estadísticos y coeficientes calculados se presentan en el cuadro 2.¹⁷

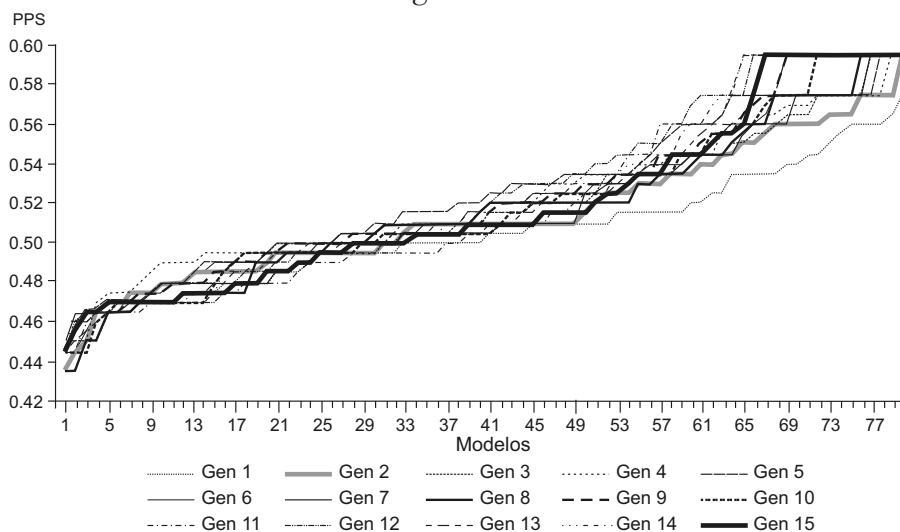
Las gráficas 1-4 muestran la evolución de la capacidad predictiva a lo largo de las generaciones.

¹⁷ En este cuadro se observa que, durante la semana del 10 de abril de 2000, el índice Nasdaq perdió 1 125.16 puntos con respecto al cierre de la semana anterior (3 de abril de 2000), cayendo desde 4 446.45 a 3 321.29 puntos. Este descenso del Nasdaq fue provocado por la decisión judicial en el caso Microsoft. Ante las medidas judiciales estadounidenses, los inversionistas se aterro-

GRÁFICA 3. Nasdaq: Capacidad predictiva a lo largo de las generaciones



GRÁFICA 4. DJI: Capacidad predictiva a lo largo de las generaciones



rizaron por la posibilidad de que dividieran a Microsoft en varias empresas para luchar contra la situación de monopolio en sistemas operativos y navegadores que ostentaba, en la práctica, la compañía de Bill Gates. La situación que afectaba a Microsoft, entre otros factores, provocó que el 14 de abril de 2000 el Nasdaq experimentara un gran retroceso por tercera vez en dos semanas. El 24 de abril de ese año Microsoft cayó 14% en el Nasdaq, consolidando una caída cercana a 40% en tres semanas.

A continuación presentamos los mejores modelos multivariados dinámicos, de acuerdo con el PPS:

$$IPC_t \quad IPC_{t-2} \quad IPC_{t-4} \quad DJI_{t-4} \quad t \quad (9)$$

$$TSE_t \quad TSE_{t-2} \quad TSE_{t-4} \quad t-3 \quad t-2 \quad t-4 \quad DJI_{t-1} \quad DJI_{t-3} \quad DJI_{t-4} \quad t \quad (10)$$

$$DJI_t \quad DJI_{t-2} \quad DJI_{t-3} \quad DJI_{t-4} \quad t-1 \quad t-1 \quad t-4 \quad t \quad (11)$$

$$NAS_t \quad NAS_{t-1} \quad NAS_{t-3} \quad t-1 \quad t-2 \quad t-3 \quad t-4 \quad DJI_{t-2} \quad t \quad (12)$$

Como se observa en el cuadro 3, los mejores modelos producidos por el algoritmo genético (correspondientes a la generación 15) arrojaron un PPS de un 59, 60, 59 y 59%, para los índices IPC, Nasdaq, TSE y DJI, respectivamente. Esta capacidad predictiva, estimada en un conjunto extramuestral de 200 datos semanales, resultó estadísticamente significativa en cada uno de los índices, de acuerdo con la prueba de acierto direccional. Por su parte, al analizar el PPS de los modelos AR(1), se encontró que estos fueron menores (50, 59, 45 y 51.5%, respectivamente) y resultaron significativos sólo en el caso del Nasdaq. En cuanto a los modelos multivariados dinámicos elaborados de manera aleatoria, estos tuvieron la capacidad predictiva más baja (excepto en el índice TSE, en el cual superaron el PPS del modelo AR(1)), siendo significativa sólo en el índice Nasdaq al considerar una significación de 10 por ciento.

Se pudo observar que la capacidad predictiva de los modelos se tradujo en beneficios económicos. Los modelos construidos por el algoritmo genético obtuvieron el mayor PPS y, de manera simultánea, el mayor rendimiento acumulado, excepto en el caso del Nasdaq, en el que la rentabilidad más alta fue obtenida por el modelo AR(1) (19.34%). Además, independientemente de la significación estadística de la capacidad predictiva de los modelos, estos superaron en rentabilidad a la estrategia de inversión pasiva o *buy and hold*, la cual evidenció la rentabilidad más baja en cada uno de los índices

CUADRO 3. Resumen de los resultados
(Porcentaje)

<i>Modelo</i>	<i>PPS</i>	<i>Prueba DA^a</i>	<i>Rentabilidad acumulada^b</i>
IPC			
Multivariado AG	59.0	2.54*	119.86
AR(1)	50.0	0.13	10.65
Multivariado aleatorio	46.5	1.03	22.02
Estrategia pasiva	—	—	10.32
Nasdaq			
Multivariado AG	60.0	2.98*	14.52
AR(1)	59.0	2.70*	19.34
Multivariado aleatorio	56.5	1.74**	11.59
Estrategia pasiva	—	—	45.59
TSE			
Multivariado AG	59.0	2.59*	53.11
AR(1)	45.0	1.49	24.42
Multivariado aleatorio	48.0	0.60	4.38
Estrategia pasiva	—	—	6.84
DJI			
Multivariado AG	59.0	2.55*	24.20
AR(1)	51.5	0.43	0.59
Multivariado aleatorio	51.0	0.28	14.41
Estrategia pasiva	—	—	23.93

^a El valor de *z* crítico es de 1.96 y 1.64, para una significación de 5 y 10%, respectivamente.

^b Rentabilidad acumulada desde el 21 de junio de 1999 hasta el 14 de abril de 2003.

* Significativo al 5 por ciento.

** Significativo al 10 por ciento.

analizados, con la sola excepción del TSE. Por lo anterior, siempre fue mejor gestionar la cartera indizada (representada por el índice bursátil) en función de alguno de los modelos de proyección.

Se probó la solidez de estos resultados a fin de evitar el problema de *data snooping*. Para ello se tomó el mejor modelo de proyección para cada índice y se lo evaluó en un total de mil conjuntos extramuestrales de 200 datos de cierre semanales cada uno. Estos mil conjuntos extramuestrales fueron generados a partir del conjunto extramuestral original utilizando un proceso de *block bootstrap*. Al analizar el PPS generado en los mil conjuntos extramuestrales, no se rechazó la hipótesis de que el PPS de los índices IPC, DJI, Nasdaq y TSE tuvo una distribución normal, con una significación de 5%, de acuerdo con la prueba de Jarque-Bera. Al analizar la rentabilidad acumulada que se habría obtenido siguiendo las recomendaciones de compra-venta

CUADRO 4. Resumen de los parámetros estimados para cada uno de los modelos^a

<i>Variables</i>	<i>Coficiente</i>	<i>D-W</i>	<i>Estadístico F</i>
IPC			
IPC_t 2	0.028292 (0.456566)	1.895213	2.746575
IPC_t 4	0.167957 (2.235067)*	—	—
DJI_t 4	0.119520 (1.741974)**	—	—
Nasdaq			
NAS_t 1	0.127370 (1.402175)	1.990337	1.702300
NAS_t 3	0.892358 (9.688533)*	—	—
DJI_t 2	0.036788 (1.810176)**	—	—
t 1	0.098100 (1.118205)	—	—
t 3	0.884935 (9.578318)*	—	—
t 4	0.010818 (16.88900)*	—	—
TSE			
TSE_t 2	0.069171 (1.507419)	2.186896	2.785186
TSE_t 4	0.740646 (6.510995)*	—	—
t 3	0.073219 (1.646930)**	—	—
t 4	0.771328 (7.117137)*	—	—
DJI_t 1	0.009668 (0.178585)	—	—
DJI_t 3	0.051882 (0.953792)	—	—
DJI_t 4	0.032653 (0.613992)	—	—
DJI			
DJI_t 2	0.034802 (0.813447)	2.020387	2.478574
DJI_t 3	0.022658 (0.532200)	—	—
DJI_t 4	0.736228 (7.095997)*	—	—
t 1	0.022818 (0.699676)	—	—
t 2	0.862587 (11.21233)*	—	—

^a Entre paréntesis se presenta la prueba t . * Significativo al 5%. El valor de z crítico es de 1.96. ** Significativo al 10%. El valor de z crítico es de 1.64.

CUADRO 5. *Resumen estadístico del PPS, considerando los resultados de los mil conjuntos extramuestrales de cada modelo*
(Porcentaje)

<i>Estadísticos</i>	<i>DJI</i>	<i>IPC</i>	<i>Nasdaq</i>	<i>TSE</i>
Media	50.88	52.58	49.07	52.93
Mediana	50.75	52.76	48.74	52.76
Máximo	57.79	60.80	62.31	60.80
Mínimo	42.71	41.71	40.20	43.72
Desviación estándar	2.85	4.31	4.09	4.09
Skewness	0.053684	0.458669	0.245725	0.270067
Kurtosis	2.952336	2.691488	2.971243	2.559184
Jarque-Bera ^a	0.057498	3.902876	1.009792	2.025264

^a El valor crítico del estadístico Jarque-Bera es de 4.43, para un 5 por ciento.

CUADRO 6. *Resumen estadístico de los rendimientos acumulados del modelo multivariado elaborado por medio de algoritmos genéticos y de los rendimientos acumulados de la estrategia buy and hold, considerando los resultados de los mil conjuntos extramuestrales*
(Porcentaje)

	<i>DJI</i>	<i>IPC</i>	<i>Nasdaq</i>	<i>TSE</i>
<i>Modelo multivariado AG</i>				
Media	0.79	190.46	1 236.06	36.27
Mediana	0.93	46.21	2.10	31.85
Máximo	158.05	9 794.09	72 934.55	585.31
Mínimo	278.79	1 353.39	98.97	137.94
Desviación estándar	51.35	1 033.27	8 724.83	83.50
Skewness	1.173044	8.182904	7.942224	3.738914
Kurtosis	11.01892	75.61529	65.16829	23.15278
Jarque-Bera	290.8633	23 086.75	12 180.09	1925.219
<i>Estrategia buy and hold^b</i>				
Media	7.78	29.09	25.46	2.35
Mediana	11.64	11.42	4.47	2.38
Máximo	93.12	336.70	2 777.67	106.46
Mínimo	98.10	97.58	1 396.79	91.32
Desviación estándar	45.28	88.73	425.82	46.40
Skewness	0.282639	1.423305	3.416574	0.158650
Kurtosis	2.483850	5.182781	27.65743	2.348799
Jarque-Bera	2.441457	53.61553	1 936.763	2.186427
Proporción ^a	57	59	41	71

^a Porcentaje de veces en que la rentabilidad del modelo multivariado supera el rendimiento de la estrategia *buy and hold*, considerando las mil series *bootstrap*.

^b El análisis para el índice Nasdaq consideró 710 series *bootstrap*.

de los modelos multivariados se encontró que, de los mil conjuntos extramuestrales, éstos superaron el rendimiento de una estrategia *buy and hold* en 57, 59 y 71% de los casos en los índices DJI, IPC y TSE, respectivamente. En el Nasdaq la frecuencia con que el modelo multivariado superó en rentabilidad a la estrategia pasiva fue de 41%. Los cuadros 5 y 6 muestran los estadísticos para el PPS y la rentabilidad acumulada de la estrategia de inversión.

CONCLUSIONES

Los modelos multivariados dinámicos elaborados a partir de algoritmos genéticos recursivos tuvieron la mayor capacidad para predecir el signo de las variaciones semanales de los índices bursátiles IPC, TSE, Nasdaq y DJI. Los resultados de la prueba de acierto direccional de Pesaran y Timmermann (1992) indicaron que los modelos elaborados a partir de algoritmos genéticos recursivos presentaron una capacidad predictiva estadísticamente significativa. A su vez, estos modelos obtuvieron la mayor rentabilidad acumulada en el periodo extramuestral con los índices IPC, TSE y DJI, superando los resultados de los modelos AR(1) y de los modelos multivariados elaborados de manera aleatoria.

Al efectuar una prueba de solidez mediante el análisis de mil series *bootstrap*, se observó que la rentabilidad acumulada que se habría obtenido siguiendo las recomendaciones de compra-venta de los modelos multivariados de algoritmos genéticos fue superior al rendimiento de una estrategia *buy and hold* en 57, 59 y 71% de los casos de los índices DJI, IPC y TSE, respectivamente. En el Nasdaq la frecuencia con que el modelo multivariado superó en rentabilidad a la estrategia pasiva fue sólo de 41%. Esto sugeriría que sería mejor gestionar la cartera indizada (representada por el índice bursátil) en función de un modelo de proyección que realizar una estrategia de inversión pasiva, al menos en los casos del DJI, IPC y TSE.

De esta manera el artículo presenta evidencia de que los algoritmos genéticos recursivos pueden ser utilizados como otra metodología para mejorar los modelos de proyección de series de tiempo, en función de su capacidad de predicción de signo.

REFERENCIAS BIBLIOGRÁFICAS

- Allen, F., y R. Karjalainen (1999), "Using Genetic Algorithms to Find Technical Trading Rules", *Journal of Financial Economics*, 51, pp. 245-271.
- Bäck, T. (1993), "Optimal Mutation Rates in Genetic Search", S. Forrest (comp.), *Proc. Fifth Int. Conf. on Genetic Algorithms*, San Mateo, CA, Morgan Kaufmann, pp. 2-9.
- Bauer, Richard J., Jr. (1994), *Genetic Algorithms and Investment Strategies*, John Wiley & Sons, Inc.
- Blume, L., D. Easley y M. O'Hara (1994), "Market Statistics and Technical Analysis: The Role of Volume", *Journal of Finance*, 49, pp. 153-182.
- Chopra, N., J. Lakonishok y J. R. Ritter (1992), "Measuring Abnormal Returns: Do Stocks Overreact?", *Journal of Financial Economics* 31, páginas 235-268.
- Conrad, J., y G. Kaul (1988), "Time-Variation in Expected Returns", *Journal of Business* 61, pp. 409-425.
- , y — (1989), "Mean Reversion in Short-horizon Expected Returns", *Review of Financial Studies* 2, pp. 225-240.
- Davis, Lawrence (1991), *Handbook of Genetic Algorithms*, Nueva York, Van Nostrand Reinhold.
- (1994), *Genetic Algorithms and Financial Applications*, Deboeck, páginas 133-147.
- DeBondt, W. F. M., y R. Thaler (1985), "Does the Stock Market Overreact?", *Journal of Finance* 40, pp. 793-805.
- Estrella, A., y F. S. Mishkin (1998), "Predicting US Recessions: Financial Variables as Leading Indicators", *The Review of Economics and Statistics* 80 (1), pp. 45-61.
- Fama, Eugene (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance* 25, pp. 383-417.
- , y K. R. French (1988), "Permanent and Temporary Components of Stock Prices", *Journal of Political Economy* 98, pp. 247-273.
- Feldman, K., y P. Treleaven (1994), "Intelligent Systems in Finance", *Appl. Mathematical Finance* 1, pp. 195-207, Londres, Chapman & Hall.
- Goldberg, D. E., y K. Deb (1991), "A Comparative Analysis of Selection Schemes Used in Genetic Algorithms", G. Rawlins (comp.), *Foundations of Genetic Algorithms*, San Mateo, CA, Morgan Kaufmann, pp. 69-93.
- Hodgson, A., y D. Nicholls (1991), "The Impact of Index Futures Markets on Australian Share Market Volatility", *Journal of Business Finance and Accounting* 18, pp. 267-280.
- Holland, John H. (1975), *Adaptation in Natural and Artificial Systems*, Ann Arbor, The University of Michigan Press.
- (1992a), "Genetic Algorithms", *Scientific American*, julio, pp. 44-50.
- Kim, K., e I. Han (2000), "Genetic Algorithms Approach to Feature Discretiza-

- tion in Artificial Neural Networks for the Prediction of Stock Price Index”, *Expert Systems with Applications*, vol. 19, núm. 2, agosto, pp. 125-132.
- Kingdon, J., y K. Feldman (1995), “Genetic Algorithms and Applications to Finance”, *Applied Mathematical Finance*, vol. 2, núm. 2, junio, pp. 89-116.
- Leung Mark T., Hazem Daouk y An-Sing Chen (2000), “Forecasting Stock Indices: A Comparison of Classification and Level Estimation Models”, *International Journal of Forecasting* (16)2, pp.173-190.
- Lo, A., y A. C. MacKinley (1988), “Stock Market Price Do Not Follow Random Walk: Evidence From a Simple Specification Test”, *Review of Financial Studies* 1, pp. 41-66.
- Michalewicz, Z. (1992), *Genetic Algorithms + Data Structures = Evolutionary Programs*, Berlín, Springer-Verlag.
- Mühlenbein, H. (1992), *How Genetic Algorithms Really Work: Mutation and Hill-Climbing*, R. Männer y B. Manderick (comps.), *Parallel Problem Solving from Nature 2*, Amsterdam, Elsevier Science, pp. 15-26.
- Parisi, A., F. Parisi y J. L. Guerrero (2003), “Modelos predictivos de redes neuronales en índices bursátiles”, *EL TRIMESTRE ECONÓMICO*, vol. LXX (4), núm. 280, octubre-diciembre, pp. 721-744.
- Pereira, Robert (1996), “Selecting Parameters for Technical Trading Rules Using Genetic Algorithms”, *Journal of Applied Finance and Investment*, vol. 1, núm. 3, julio-agosto, pp. 27-34.
- Pesaran, M. H., y A. Timmermann (1992), “A Simple Nonparametric Test of Predictive Performance”, *Journal of Business and Economic Statistics* 10, pp. 461-465.
- Poterba, J. M., y L. H. Summers (1988), “Mean Reversion in Stock Prices: Evidence and Implications”, *Journal of Financial Economics* 22, pp. 27-59.
- Qi, M. (2001), “Predicting US Recessions With Leading Indicators Via Neural Network Models”, *International Journal of Forecasting* 17, pp. 383-401.
- White, H. (2000), “A Reality Check for Data Snooping”, *Econometrica*, vol. 68, núm. 5, pp. 1097-1126.