

USOS Y LIMITACIONES DE LOS PROCESOS ESTOCÁSTICOS EN EL TRATAMIENTO DE DISTRIBUCIONES DE RENDIMIENTOS CON COLAS GORDAS

JOSE CARLOS RAMIREZ SANCHEZ*

ITESM y CIDE, México

Abstract

This paper deals with the main theoretical problems regarding the application of stochastic processes to leptokurtic financial return distributions. A sort of statistical tests based on the stock index Banamex 30 is performed in order to choose the stochastic model that provide the best fit to the fat-tailed empirical distribution, allowing for a better return forecasting or value at risk estimate. In choosing that model the paper points out that any single set of statistical criteria is not appropriate if it is not confronted with the risk manager's experience. Understanding the investor's aversion risk or the transaction costs involved in any trading strategy, among other elements, is very important to justify the use of any stochastic process in risk management techniques.

I. Introducción

El supuesto de normalidad estacionaria es esencial para los modelos de valuación de riesgos o de fijación de precios de activos, debido a que sus principales resultados dependen del análisis de los dos primeros momentos de una distribución normal (Kon, 1984). En los modelos de varianza-covarianza, como el CAPM o el APT, por ejemplo, es crucial que los precios de los activos se comporten como variables aleatorias independientes e idénticamente distribuidas, con una varianza

* Instituto Tecnológico y de Estudios Superiores de Monterrey-Campus Ciudad de México. Director del Programa de Investigación en Regulación Económica, CIGEO, ITESM-CCM y profesor afiliado de la División de Economía del Centro de Investigación y Docencia Económicas, México. E-mail: josecarlos.ramirez@itesm.mx

Se agradecen los comentarios de dos árbitros anónimos.

finita, estable y bien definida para, de esa manera, garantizar que los momentos de la distribución empírica converjan a los de una distribución normal teórica a medida que se incrementa el tamaño de la muestra (conforme al teorema del límite central).¹ Otro tanto puede decirse de las estimaciones de valor en riesgo (VeR), cuyos principales ingredientes de cálculo –factores de riesgo, volatilidades y matrices de correlaciones– requieren que esa convergencia se cumpla para asegurar la existencia de normalidad multivariada y así evitar errores en el cómputo de los requerimientos del capital en riesgo de las instituciones.

La literatura empírica sobre el tema muestra, sin embargo, que la convergencia de los parámetros ha probado ser alta sólo para intervalos de tiempos relativamente largos, en particular para series de rendimientos de un mes, por lo que la hipótesis de normalidad resulta realmente heroica de sostener en análisis de series diarias o de alta frecuencia (véase Andersen *et al.*, 2001). Si a esto agregamos que, debido a las variaciones experimentadas por las medias y las varianzas de las series de rendimientos a través del tiempo, es poco lo que se puede argumentar en favor del supuesto de estacionariedad, entonces queda claro que hay algo erróneo en los modelos de VeR o de pronóstico de rendimientos que asumen normalidad estacionaria (Campbell y Siddique, 2000).

Las dificultades asociadas con la infracción del supuesto de normalidad estacionaria para series de rendimientos de tiempo corto han llevado a algunos autores a explorar diversos instrumentos analíticos que enfrenten el problema de las llamadas *colas gordas o pesadas* que es, con mucho, el problema principal asociado con la violación de ese supuesto (Kupiec, 1995 y Hull y White, 1998).² Los esfuerzos más notables en esta dirección han estado dirigidos a adoptar los procesos de Itô, cuya peculiar estructura formal ha permitido, por un lado, modelar series con parámetros cambiantes en el tiempo y, por otro, calcular por repetición los precios de subyacentes y derivados al asumir mercados dinámicamente completos (Nielsen, 1999).³

La primera de esas características, esto es su cualidad de modelar series con parámetros cambiantes, es inherente a su definición. Como se sabe, un proceso de Itô es la suma de un valor inicial, una integral de tiempo y una integral estocástica. Las integrales tienen incrementos instantáneos cuyas medias (en el caso de la integral de tiempo), varianzas y covarianzas (en el caso de la integral estocástica) pueden fluctuar estocásticamente en el tiempo. De aquí que si se desea atacar el problema de *colas gordas* ocasionado, digamos, por una mayor volatilidad de los rendimientos, una solución posible es adaptar un proceso de Itô a las series históricas mediante la manipulación de la integral del tiempo, llamada *drift*, o de la integral estocástica, llamada dispersión o coeficiente del proceso de Wiener.⁴ La manipulación puede abarcar desde un cambio del *drift* hasta una mezcla de procesos (Baxter y Rennie, 1996).

La segunda característica que ha hecho atractivo el uso de los procesos de Itô es igualmente importante, pues permite simular los precios de una obligación contingente a partir de una estrategia de negociación cuyos valores iniciales y de dispersión sean iguales a los de un proceso de Itô. En presencia de mercados dinámicamente completos, el precio de cualquier obligación contingente puede ser replicado con la información generada por el proceso de Wiener.⁵ Este es un resultado

que se desprende del hecho de que el supuesto de mercados completos es válido si y sólo si los precios de las obligaciones son procesos de Itô (Nielsen, 1999).⁶ Las fluctuaciones extremas de los precios pueden ser, entonces, modeladas a discreción buscando que el proceso de Wiener simule, por ejemplo, la trayectoria leptocúrtica de los datos originales (Baxter y Rennie, 1996).

De la variedad de procesos de Itô aplicados en finanzas, los movimientos brownianos son los más utilizados en los modelos de administración de riesgos y de ingeniería financiera, debido a sus características inigualables para medir los cambios en los valores de los portafolios cuando los precios de los activos subyacentes se mueven en forma extrema (Neftci, 1996).⁷ Su gran aceptación ha llevado incluso a diversos autores a expandir los movimientos brownianos mediante la inclusión de ecuaciones importadas de la Física, como la de Langevin, o la utilización de brownianos fraccionarios y no ideales a fin de modelar mejor las imperfecciones del mercado (véase Rogers, 1996 y Takahashi, 1996). Estas extensiones, junto con las mezclas gaussianas, las caminatas aleatorias, las cadenas de Markov y las martingalas por parte de la familia de procesos de Itô, más los procesos de Poisson, las distribuciones de Student y los modelos de heterocedasticidad condicional, forman el acervo de técnicas más utilizado por los analistas financieros y que nosotros identificaremos, de ahora en adelante, como procesos estocásticos alternativos a la distribución normal o, simplemente, como procesos estocásticos.⁸

Ahora bien, aunque la incorporación de los procesos estocásticos en finanzas obedece, para nuestros fines, a su probada utilidad en el tratamiento de series afectadas por sesgos, excesos de curtosis o variaciones en los dos primeros momentos de una distribución normal, su uso indiscriminado ha llevado a pasar por alto algunos inconvenientes. Entre ellos, se puede mencionar la gran dificultad que existe para identificar procesos que enfrenten *simultáneamente* algunos problemas relacionados con las series no normales estacionarias o la falta de criterios estadísticos para asegurar que las series no normales sigan conservando en el futuro los mismos patrones de comportamiento que los pronosticados con los datos pasados.

Los inconvenientes surgen porque el uso de los procesos estocásticos no garantiza en todos los casos la solución completa al problema de falta de normalidad estacionaria, lo que provoca que la elección de cualquier proceso tenga asociado siempre su grado de penalización. Hay procesos que son más eficientes que otros en capturar algunas distorsiones evidentes de las distribuciones empíricas, como el sesgo, pero que no explican, por ejemplo, la volatilidad grupal (*clustering volatility*). Las virtudes de unos pueden aparecer como defectos en otros y viceversa, por lo que encontrar procesos estocásticos adecuados que describan el comportamiento de distribuciones empíricas cuyos parámetros no son normales y estacionarios representa, ante todo, un problema conceptual.

El objetivo de este documento consiste en dar cuenta de algunos de esos problemas y, en particular, de aquellos que tienen que ver con la selección adecuada de los procesos estocásticos para el estudio de las distribuciones de rendimientos de acciones que exhiben *colas gordas o pesadas*. La importancia de estos problemas es a todas luces evidente ya que, como lo sugiere la conclusión principal del documento, sin una reflexión adecuada de los problemas conceptuales que están

asociados con la aplicación de un proceso estocástico no es posible entender sus límites en el análisis financiero.

Con el fin de darle una explicación intuitiva a estos problemas, la exposición incluye un análisis estadístico del paquete accionario *Banamex 30*, en el que se discute la secuencia de pruebas conducente a la aplicación de un proceso estocástico, así como las dificultades que surgen en los pronósticos de rendimientos y en el cálculo del VeR cuando se viola el supuesto de normalidad estacionaria. El análisis es un recurso heurístico para enfatizar algunos puntos de las series de rendimientos que son comunes a la mayoría de los mercados financieros y no, como es característico de algunos estudios (véase, por ejemplo, De la Luz, 2002), para explicar las particularidades del caso mexicano. Las acciones mexicanas aquí consideradas cotizan simultáneamente en bolsas de valores de mercados desarrollados y emergentes y eso permite hacer puntualizaciones sobre diferentes aspectos de las distribuciones con *colas gordas* en ambos tipos de mercados, sin ninguna pérdida de generalidad. Los resultados de las series de tiempo muestran que las características del paquete *Banamex 30* son, de hecho, similares a las documentadas para otros índices accionarios de los principales mercados de Norteamérica, Europa, Latinoamérica y Asia.

El documento se divide en dos partes. En la primera se exponen los principales usos de los procesos estocásticos en el tratamiento de distribuciones de rendimientos de acciones que acusan *colas gordas* (Sección II). En la segunda se hace mención de los problemas conceptuales más importantes que se relacionan con la utilización de los procesos estocásticos en la elaboración de un pronóstico o en el cálculo del VeR (Sección III). Las conclusiones incorporan una pequeña discusión de los resultados alcanzados (Sección IV).

II. Los Usos de los Procesos Estocásticos en el Tratamiento de Distribuciones de Rendimientos con *Colas Gordas*

La aplicación de los procesos estocásticos al análisis de los rendimientos de las acciones se justifica, en principio, cuando no es posible asociar una o a varias distribuciones normales a las series de rendimientos o cuando se observan cambios en los valores estacionarios de la media y la varianza. En ambos casos, la decisión de optar por un proceso implica, primero, llevar a cabo diversas pruebas de normalidad y luego, en caso de violación del supuesto, determinar si los datos pasados son útiles para extrapolar su comportamiento futuro.

En lo que corresponde a las pruebas de normalidad, éstas deben incluir tratamientos estadísticos univariados y multivariados de las series para garantizar el análisis correcto de los portafolios de acciones, pues no hay que olvidar que series que son univariadamente normales pueden no resultar multivariadas, por efecto de la correlación cruzada de los precios de las acciones. De la misma manera, series que no son univariadamente normales nunca podrán ser multivariadamente normales (Richardson y Smith, 1993). En este último caso es importante comprobar si por medio de una transformación, por ejemplo, del tipo Box Cox o

de Fisher, las series univariadas pueden recuperar su comportamiento normal (Johnson y Wichern, 1992). Cuando esto no es posible quedan entonces dos alternativas antes de elegir un proceso estocástico particular. La primera consiste en ensayar pruebas basadas en mezclas de distribuciones para descartar si dos o más normales o una *t* de Student son capaces de estimar el sesgo y el exceso de curtosis que no son captados por una sola distribución normal. De no ser así, se vuelve necesario, como segunda alternativa, comprobar si las series históricas son útiles para predecir el comportamiento futuro de los rendimientos de las acciones. En caso afirmativo, la tarea se reduce a adaptar un proceso estocástico particular a las series originales o corregidas. En caso contrario, las series se comportarían como una martingala y, por lo tanto, ninguna serie histórica serviría para elaborar un pronóstico.

Para ilustrar estas posibilidades, el Cuadro 1 muestra los resultados de las pruebas de normalidad para las series de rendimientos diarios del paquete *Banamex 30* corregidas con potencias del tipo Box-Cox.⁹ De acuerdo con los estadísticos de prueba, en especial los que se refieren a las pruebas conjuntas y de Shapiro y Wilk, ninguna serie, real o simulada por el método de Monte Carlo, resultó normal para cualquier nivel de significancia.¹⁰

CUADRO 1
PRUEBAS DE NORMALIDAD UNIVARIADA PARA SERIES
DE RENDIMIENTOS DIARIOS

Activo	Sesgo	Valor p	Monte Carlo Valor p	Curtosis	Valor p	Monte Carlo Valor p	Prueba conjunta Valor p	Shapiro-Wilk Valor p
ALFA.A	1.170	0.000	0.000	13.346	0.000	0.000	0.000	0.00
APASCO	0.777	0.000	0.000	6.196	0.000	0.000	0.000	0.000
BANACCIO	0.295	0.000	0.000	2.979	0.000	0.000	0.000	0.000
BIMBO.A	0.267	0.000	0.000	3.304	0.000	0.000	0.000	0.000
CEMEX.CPO	0.247	0.000	0.000	6.159	0.000	0.000	0.000	0.000
CIFRA.V	0.548	0.000	0.000	5.893	0.000	0.000	0.000	0.000
COMERCI.UBC	0.636	0.000	0.000	11.479	0.000	0.000	0.000	0.000
DESC.B	0.624	0.000	0.000	4.537	0.000	0.000	0.000	0.000
ELEKTRA.CPO	0.667	0.000	0.000	8.225	0.000	0.000	0.000	0.000
FEMSA.UBD	0.501	0.000	0.000	12.466	0.000	0.000	0.000	0.000
GCARSO.A1	-0.016	0.000	0.000	4.851	0.799	0.578	0.000	0.000
GFNORTE.O	0.151	0.000	0.000	4.232	0.015	0.015	0.000	0.000
GMEXICO.B	0.765	0.000	0.000	3.643	0.000	0.000	0.000	0.000
ICA	-0.144	0.000	0.000	9.436	0.000	0.000	0.000	0.000
SAVIA.A	0.533	0.000	0.000	8.404	0.000	0.000	0.000	0.000
SORIANA.B	0.342	0.000	0.000	2.579	0.000	0.000	0.000	0.000
TAMSA	0.518	0.000	0.000	6.220	0.000	0.000	0.000	0.000
TELECOM.A1	0.389	0.000	0.000	4.292	0.000	0.000	0.000	0.000
TELMEX.L	0.626	0.000	0.000	3.221	0.000	0.000	0.000	0.000
TELEVISA.CPO	0.620	0.000	0.000	5.487	0.000	0.000	0.000	0.000
VITRO.A	0.905	0.000	0.000	9.101	0.000	0.000	0.000	0.000
BANAMEX-30	0.222	0.000	0.000	6.456	0.000	0.000	0.000	0.000

El rechazo a las hipótesis nulas de normalidad con valores p iguales a cero confirma la tesis, ampliamente respaldada por la experiencia de los mercados desarrollados, de que también en los mercados emergentes, como el mexicano, ningún análisis de valor en riesgo (VeR) o pronóstico de los rendimientos diarios es, en principio, correcto si se hace sobre la base de los métodos tradicionales de media-varianza (o covarianza). La razón estriba en que, de obviarse la falta de normalidad en las series individuales, el riesgo calculado se subestimaría con respecto a los verdaderos valores de una normal teórica (Zangari, 1996). Esta conclusión ha sido profusamente documentada desde hace largo tiempo por Mandelbrot (1963) y Fama (1965), quienes reportaron por primera vez la presencia de *colas gordas* en las series del índice *Dow Jones* o, más recientemente, por Affleck-Graves y McDonald (1989) y Ojah y Karemera (1999), quienes presentaron resultados similares para el caso de Europa, Estados Unidos y Latinoamérica.

La ausencia de normalidad ha dado pie a que se utilicen otras herramientas, como las mezclas de distribuciones, que se caracterizan por capturar eficientemente la falta de simetría y el exceso de curtosis de las distribuciones empíricas y, sobre todo, por mejorar los cálculos, por ejemplo, del VeR en intervalos de confianza del 95% y 98% (Venkataraman, 1997). La eficiencia de las mezclas reside en el hecho de que ataca el problema de la falta de normalidad estacionaria considerando, a diferencia de los métodos de estimación del riesgo basados en una distribución normal, que los sesgos y excesos de curtosis obedecen a variaciones en la media y la varianza de las distribuciones empíricas. En general, se supone que los sesgos y los excesos de curtosis que presentan esas distribuciones son debidos a desplazamientos de sus parámetros entre un número finito de valores, ocasionados por cambios en la estructura de capital, eventos exógenos al mercado (*desplazamientos de tiempo ordenado*) o por la mayor varianza de los rendimientos en algunos días de la semana (*desplazamientos cíclicos*). La persistencia con que los primeros desplazamientos explican el sesgo observado y los segundos el exceso de curtosis, demanda, pues, la utilización de modelos como el de mezclas de distribuciones que incorporen las variaciones en la media y la varianza en el tiempo.¹¹

Para medir el poder descriptivo de la mezcla de distribuciones, las pruebas estadísticas incluyen usualmente una comparación con la distribución t de Student. La justificación radica en el alegato de que la distribución de Student captura más valores extremos que una normal estacionaria pero no la combinación de los desplazamientos cíclicos y estructurales, o de tiempo ordenado, de los parámetros de la distribución (Kim y Kon, 1994). Las pruebas contenidas en el Cuadro 2 no parecen, sin embargo, confirmar los resultados encontrados para otros países pues, mientras que la distribución de Student describe mejor las series diarias de 15 activos, la mezcla de dos y tres normales sólo lo hace para 6 activos y el índice Banamex-30.

La conclusión se desprende al observar el signo del estadístico del logaritmo de la tasa de verosimilitud $\Lambda_{NS} = \frac{\ell(\theta_N / r)}{\ell(\underline{\delta} / r)}$, entre una y más distribuciones

CUADRO 2

COMPARACION ENTRE UNA DISTRIBUCION STUDENT Y LA
MEZCLA DISCRETA DE NORMALES GENERALIZADA

Acción	·1S	·2S	·3S
ALFA.A	- 0.659	295.965	ND
APASCO	- 0.906	187.764	ND
BANACCL.O	- 0.841	ND	ND
BIMBO.A	- 1.124	ND	ND
CEMEX.CPO	- 0.770	ND	ND
CIFRA.V	- 0.858	147.306	184.021
COMERCI.UBC	- 0.646	- 23.913	ND
DESC.B	- 0.848	ND	ND
ELEKTRA.CPO	- 0.245	ND	ND
FEMSA.UBD	- 0.580	ND	ND
GCARSO.A1	- 0.872	123.809	ND
GFNORTE.O	- 0.817	ND	ND
GMEXICO.B	- 1.006	ND	ND
ICA	- 0.648	- 15.246	ND
SAVIA.A	- 1.100	ND	ND
SORIANA.B	- 1.100	ND	ND
TAMSA	- 0.767	ND	ND
TELECOM.A1	- 0.922	149.136	ND
TELMEX.L	- 1.290	ND	ND
TELEvisa.CPO	- 0.795	ND	ND
VITRO.A	- 0.678	ND	ND
BANAMEX- 30	- 2.242	2.159	ND

ND: No Disponible.

normales (N) y una distribución Student (S). Si el logaritmo de Λ_{NS} es positivo, entonces es más probable que la mezcla de normales genere los datos observados de una manera más confiable que la distribución de Student, contrario a lo que sucedería si el signo de Λ_{NS} fuera negativo (para mayor información consúltese Kon, 1984). Los valores negativos del estadístico en los renglones del Cuadro 2 indican que la distribución de Student es superior a la descripción hecha por una normal, Λ_{1S} , en todas las series de activos, pero no así al incluir la mezcla de dos normales: en este caso los signos positivos de Λ_{2S} revelan una mayor capacidad descriptiva de la mezcla sobre la distribución Student en las series diarias de alfa, apasco, cifra, comerci, gcarso, ica, telecom y el índice Banamex-30. La inclusión de tres normales no mejora la potencia descriptiva del modelo de mezclas sobre la distribución de Student.¹²

Esta ambigua situación, en la que no queda clara la superioridad de la mezcla de dos o más normales sobre la distribución de Student, induce a pensar que la inclusión de otros procesos estocásticos puede arrojar una mejor especificación en la modelación de las series. Para tal efecto es necesario determinar, antes, en qué medida los datos pasados son útiles para predecir el comportamiento futuro de los

rendimientos. Y para ello no hay nada mejor que utilizar pruebas basadas en los mismos procesos de Itô para decidir si los rendimientos de los activos son pronosticables o no. En concreto, se sostiene que los rendimientos son pronosticables si se rechaza la hipótesis nula de que las series de tiempo se comportan como una caminata aleatoria o una martingala. En caso contrario, las series de rendimientos no servirían de referencia para llevar a cabo pronósticos fiables.

Las pruebas sobre las diferentes versiones de las caminatas aleatorias (RW) se reducen a rechazar o aceptar las hipótesis de que los incrementos experimentados por las variables aleatorias son independientes e idénticamente distribuidos (IID) con media cero y varianza σ^2 en una ecuación con la siguiente forma estructural:

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

donde P_t es el cambio de precio del activo en el tiempo t y μ es el cambio esperado del precio.

La ecuación (1) es una caminata aleatoria del tipo 1 (RW1) si cumple con los dos supuestos sobre los incrementos, es decir, si $\epsilon_t \sim IID(0, \sigma^2)$; una caminata aleatoria del tipo 2 (RW2) si los incrementos son independientes pero no idénticamente distribuidos y, finalmente, una caminata del tipo 3 (RW3) si los incrementos son dependientes pero no correlacionados. Debido a las dificultades inherentes a las estimaciones no paramétricas de RW2 y al poco realismo de los supuestos de RW1 en mercados que se caracterizan por tener rendimientos altamente correlacionados, las pruebas de caminatas aleatorias se concentran básicamente en evaluar los supuestos de RW3 (Campbell *et al.*, 1997). El énfasis puesto en verificar si los rendimientos están correlacionados o si conservan un patrón de reciprocidad entre ellos ha llevado a poner en práctica pruebas basadas en índices de autocorrelación “reescalados” (como la Q de Pierce o de Ljung-Box), de razones de varianzas o de *lead-lag*, que son igualmente aplicadas para determinar la existencia de martingalas.¹³ La comprobación de que los rendimientos conservan un patrón de correlación específico entre ellos no sólo equivale a rechazar la hipótesis nula de RW3 sino también la de martingala, pues de esa manera se supone que el mejor pronóstico del rendimiento futuro no es, como lo supone un juego justo, el basado en el rendimiento actual de los activos sino también en el de períodos pasados.

Los resultados expuestos en los Cuadros 3 y 4 rechazan las hipótesis nulas de RW3 y de martingalas para el caso de México, con lo que se confirma que la ausencia de normalidad en los rendimientos diarios es en parte explicada por su alta correlación.¹⁴ De acuerdo con el Cuadro 3, la volatilidad de los rendimientos no crece linealmente con el tiempo sino, como lo comprueban las razones de varianza (RV), con el horizonte de inversión inmediato. La explicación obedece al hecho de que los rendimientos están positivamente correlacionados, así como al rezago de tiempo que toma al mercado homogeneizar los rendimientos.¹⁵ Este patrón ha sido encontrado por Ojah y Karemera (1999) para series mensuales de rendimientos en distintos mercados latinoamericanos, pero no para series semanales. En este

CUADRO 3
COEFICIENTES DE AUTOCORRELACION Y RAZONES DE VARIANZAS
PARA VARIOS REZAGOS

Activo	Rezagos				No. k de observaciones restadas para formar RV		
	1	2	4	8	2	4	8
ALFA.A	0.125 (0.000)	-0.074 (0.003)	-0.023 (0.386)	-0.014 (0.990)	1.124 (105.856)	1.072 (34.797)	1.007 (4.739)
APASCO	0.011 (0.686)	-0.024 (0.456)	-0.025 (0.240)	0.018 (0.455)	1.011 (8.320)	0.982 (-5.222)	0.952 (-13.512)
BANACCIO	0.199 (0.000)	-0.029 (0.010)	-0.002 (0.602)	0.002 (0.513)	1.200 (98.908)	1.268 (148.673)	1.317 (118.335)
BIMBO.A	0.049 (0.081)	0.011 (0.828)	-0.043 (0.094)	0.011 (0.711)	1.050 (54.433)	1.067 (42.410)	0.979 (-4.195)
CEMEX.CPO	0.101 (0.000)	-0.003 (0.799)	0.016 (0.378)	0.034 (0.172)	1.100 (84.975)	1.132 (65.244)	1.165 (58.204)
CIFRA.V	-0.008 (0.789)	-0.001 (0.958)	0.040 (0.095)	-0.027 (0.324)	0.992 (-5.320)	0.976 (-13.262)	0.992 (-2.728)
COMERCI.UBC	0.013 (0.640)	0.000 (0.770)	0.045 (0.121)	-0.008 (0.643)	1.013 (10.216)	1.028 (13.045)	1.125 (42.126)
DESC.B	0.137 (0.000)	0.006 (0.970)	0.012 (0.428)	-0.051 (0.036)	1.138 (140.035)	1.188 (110.666)	1.183 (73.893)
ELEKTRA.CPO	0.154 (0.000)	0.021 (0.715)	0.030 (0.547)	-0.059 (0.005)	1.154 (137.757)	1.269 (134.553)	1.397 (130.580)
FEMSA.UBD	0.044 (0.105)	-0.038 (0.138)	-0.015 (0.521)	-0.014 (0.575)	1.045 (30.322)	1.011 (3.694)	0.912 (-25.685)
GCARSO.A1	0.098 (0.000)	0.009 (0.691)	0.026 (0.236)	-0.025 (0.256)	1.097 (90.603)	1.150 (76.906)	1.169 (60.807)
GFNORTE.O	0.224 (0.000)	0.050 (0.866)	-0.003 (0.986)	-0.036 (0.144)	1.223 (216.291)	1.382 (216.054)	1.439 (175.853)
GMEXICO.B	0.054 (0.020)	0.003 (0.869)	-0.027 (0.334)	0.037 (0.107)	NA NA	NA NA	NA NA
ICA	0.043 (0.130)	0.009 (0.705)	0.051 (0.069)	0.005 (0.935)	1.044 (52.800)	1.083 (54.207)	1.198 (87.774)
SAVIA.A	0.074 (0.004)	-0.049 (0.074)	0.001 (0.941)	0.055 (0.042)	1.075 (78.409)	1.049 (30.415)	1.043 (21.946)
SORIANA.B	0.051 (0.083)	-0.017 (0.474)	-0.016 (0.700)	-0.056 (0.089)	1.051 (61.341)	1.045 (29.583)	0.991 (0.207)
TAMSA	0.150 (0.000)	0.005 (0.496)	0.007 (0.834)	-0.024 (0.447)	1.151 (156.459)	1.244 (137.589)	1.274 (103.586)
TELECOM.A1	0.073 (0.003)	0.042 (0.078)	0.043 (0.082)	-0.010 (0.701)	1.073 (78.717)	1.139 (80.842)	1.220 (91.043)
TELMEX.L	0.011 (0.571)	0.006 (0.641)	0.015 (0.484)	-0.048 (0.065)	1.011 (12.201)	0.994 (-1.689)	0.944 (-18.537)
TELEVISA.CPO	0.075 (0.002)	-0.005 (0.741)	0.006 (0.624)	-0.014 (0.592)	1.074 (84.218)	1.087 (55.834)	1.067 (27.208)
VITRO.A	0.026 (0.224)	-0.070 (0.011)	-0.018 (0.263)	-0.007 (0.632)	1.026 (198.908)	0.974 (148.673)	0.918 (118.335)
BANAMEX 30	0.157 (0.000)	0.002 (0.702)	0.031 (0.403)	-0.015 (0.407)	1.157 (131.434)	1.230 (109.714)	1.329 (110.062)

Nota: De las columnas dos a la cuatro aparecen los coeficientes de autocorrelación con sus respectivos valores p. Las tasas de varianza aparecen en las columnas 5-7. Los estadísticos robusto-heterocedásticos (Vz) aparecen debajo de su respectiva tasa de varianza. Vz se distribuye asintóticamente como una normal estándar.

CUADRO 4

ASIMETRIA DE LAS MATRICES DE AUTOCOVARIANZAS
(ANALISIS DE EXISTENCIA DE LEAD-LAG)

$$\hat{Y}(1) - \hat{Y}'(1) = \begin{bmatrix} 0 & -0.047 & -0.150 \\ 0.047 & 0 & -0.152 \\ 0.150 & 0.152 & 0 \end{bmatrix}$$

$$\hat{Y}(2) - \hat{Y}'(2) = \begin{bmatrix} 0 & -0.000 & -0.028 \\ 0.000 & 0 & -0.021 \\ 0.028 & 0.021 & 0 \end{bmatrix}$$

$$\hat{Y}(3) - \hat{Y}'(3) = \begin{bmatrix} 0 & 0.000 & -0.064 \\ 0.000 & 0 & -0.048 \\ 0.064 & 0.048 & 0 \end{bmatrix}$$

Nota: El cuadro muestra las diferencias entre las matrices de autocorrelación y sus transpuestas para el vector de rendimiento de portafolio $Y_{(i)} \cdot [R_{1t}, R_{2t}, R_{3t}]'$ donde R_{it} es el rendimiento en el día t del i -ésimo portafolio; $i=1,2,3$ (el portafolio 1 contiene las acciones de menor bursatilidad y así sucesivamente).

último caso, Ojah y Karemera (1999), Grieb y Reyes (1999) y Long, Payne y Jang (1999) aceptan la hipótesis de la caminata aleatoria, o de que la varianza de los rendimientos crece linealmente con el tiempo en casi todos los mercados emergentes de Asia y Latinoamérica.

Los coeficientes de autocorrelación son más altos en los primeros períodos que en los posteriores y, por tanto, los índices “reescalados” de dependencia de largo plazo son menos significativos a partir de cierto período. La evidencia es, sin embargo, irregular, pues mientras que esos índices resultaron significativos para algunos activos durante todo el período considerado, para otros apenas mostraron correlación serial durante el primer período. En cualquier caso, la dependencia entre los rendimientos de los activos no es homogénea ya que, como lo muestra el Cuadro 4, hay un esquema de *lead-lag* en el que la correlación entre los rendimientos actuales de los activos de menor capitalización y los rendimientos pasados de los activos de mayor capitalización es siempre mayor que la correlación entre los rendimientos actuales de los activos de mayor capitalización y los rendimientos pasados de los activos de menor capitalización.

El resultado se observa al comparar las matrices de autocovarianzas que fueron construidas para tres tipos de portafolios de diferente bursatilización. Con excepción de un periodo, todas las entradas debajo de la diagonal de las matrices resultaron mayores que los valores arriba de las diagonales, lo que indica la existencia del citado patrón (véase Campbell *et al.*, 1997).

Las pruebas tomadas en conjunto revelan, pues, que con excepción de contados casos los rendimientos son *pronosticables* y que un modelo estadístico adecuado para describir la dependencia entre los activos es, en principio, uno que capture la mayor varianza actual de los rendimientos. Para probar este último resultado estimamos tres tipos de modelos comúnmente citados en la literatura: el modelo de difusión con saltos y dos modelos de heterocedasticidad condicional temporendependientes que privilegian la varianza actualizada del último rezago, GJR(1,1) y EGARCH(1,1). El primero de ellos incluye una ecuación en la que se supone que cada vez que hay un arribo de información anormal existe un salto instantáneo en el precio de las acciones de tamaño J que es independiente de Z :

$$\log(r_t) = \left(\alpha - \frac{1}{2} \sigma^2 \right) t + \sigma Z(t) + \sum_{n=1}^{N(t)} J_n \quad (2)$$

donde: $Z(t)$ = proceso estandarizado de Wiener; $N(t)$ = proceso de Poisson no homogéneo; J_n = una variable aleatoria normal que representa el logaritmo de uno más el porcentaje de cambio en el precio del activo causado por el n -ésimo salto; α = la tasa instantánea condicional esperada de rendimiento por unidad de tiempo del proceso de Wiener; y σ^2 = varianza condicional instantánea de los rendimientos por unidad de tiempo del proceso de Wiener.

De acuerdo con los datos del Cuadro 5, el modelo resultó adecuado para explicar el comportamiento de 20 de 21 de los activos y del índice *Banamex 30*, ya que en esos casos los valores p de los estimadores del proceso de Wiener, α y σ^2 , y del proceso de Poisson, δ^2 , fueron iguales a cero.¹⁶ El rechazo a la hipótesis nula de igualdad de medias y varianzas para cualquier nivel de significancia confirma que la presencia de información anormal sí afecta el precio de esos 20 activos, alterando con saltos la volatilidad de sus rendimientos.

Con relación a los modelos EGARCH (1,1) y GJR (1,1), el procedimiento consistió en suponer una misma ecuación de la media condicional, m_t , del tipo ARMA (1,1) para los dos modelos (ecuación 3) y dos especificaciones diferentes para la varianza condicional, h_t^2 , en las que se incluyen los parámetros γ y k , usados para estimar la volatilidad grupal y apalancamiento, respectivamente (ecuaciones 4 y 5). Con el propósito de incorporar mayor variedad en la especificación de las varianzas se incluyeron los dos parámetros por separado en la ecuación (4) del modelo EGARCH (1,1) y en forma conjunta a través del parámetro η_1 , en la ecuación (5) del modelo GJR (1,1).¹⁷

El orden de los modelos fue seleccionado para hacer hincapié en que el rendimiento y el riesgo de las series dependen crucialmente del valor esperado y de la varianza acumuladas en el último rezago (véase Sánchez Cerón, 2001):

$$m_t = c_0 + b r_{t-1} + a \varepsilon_{t-1} \quad (3)$$

$$\log(h_t^2) = \varpi + \beta \log(h_{t-1}^2) + \gamma \left| \frac{\varepsilon_{t-1}}{h_{t-1}} \right| + k \frac{\varepsilon_{t-1}}{h_{t-1}} \quad (4)$$

CUADRO 5

VALORES DE LOS PARAMETROS DEL PROCESO DE DIFUSION CON SALTO

No.	.	α	σ^2	.2
ALFA.A	0.6443 0.0000	- 0.1720 0.0000	0.3751 0.0000	0.1879 0.0000
APASCO	1.4688 0.0000	- 0.2806 0.0000	0.2322 0.0000	0.4920 0.0000
BANACCI.O	1.2568 0.0000	- 0.1995 0.0000	0.4754 0.0000	0.3061 0.0000
BIMBO.A	1.4953 0.0000	- 0.3141 0.0000	0.4055 0.0000	0.5243 0.0000
CEMEX.CPO	0.6994 0.0000	- 0.1018 0.0000	0.4171 0.0000	0.2575 0.0000
CIFRA- V	1.0512 0.0000	- 0.2389 0.0000	0.3187 0.0000	0.2057 0.0000
COMERCI.UBC	0.9904 0.0000	0.0454 0.0000	0.3533 0.0000	0.1882 0.0000
des.B	0.4661 0.0000	0.0319 0.2293	0.3034 0.0000	0.3397 0.0000
ELEKTRA.CPO	3.4168 0.0000	- 3.1226 0.0000	- 0.4299 0.0000	0.5208 0.0000
FEMSA.UBD	0.1393 0.0000	0.2676 0.0000	0.2875 0.0000	0.4303 0.0000
GCARSO.A1	- 1.3239 0.0000	1.5506 0.0000	0.2608 0.0000	0.3012 0.0000
GFNORTE.0	0.0157 0.0000	0.8644 0.0000	0.8033 0.0000	0.4983 0.0000
GMEXICO.B	0.6700 0.0000	- 0.1833 0.0000	0.3188 0.0000	0.4117 0.0000
ICA	0.6606 0.0000	- 0.0590 0.0000	0.3764 0.0000	0.2344 0.0000
SAVIA.A	- 1.1981 0.0000	0.9658 0.0000	0.2522 0.0000	0.3070 0.0000
SORIANA.B	NC	NC	NC	NC
TAMSA	0.5592 0.0000	0.1527 0.0905	0.3653 0.0000	0.2499 0.0000
TELECOM.A1	0.2281 0.0000	0.6146 0.0000	0.3416 0.0000	0.2437 0.0000
TELMEX.L	0.5972 0.0000	0.0680 0.0000	0.3602 0.0000	0.2330 0.0000
TELEVISA.CPO	0.6874 0.0000	0.0844 0.0000	0.4407 0.0000	0.3410 0.0000
VITRO.A	0.0982 0.0000	0.6886 0.0000	0.3180 0.0000	0.2546 0.0000
BANAMEX- 30	3.0245 0.0000	- 3.2211 0.0000	- 0.6177 0.0000	0.3535 0.0000

NC: No Converge. Los valores debajo de los estimadores de cada acción son los valores p.

$$h_t^2 = \varpi + \zeta \varepsilon_{t-1}^2 + \eta_1 d_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1}^2 \quad (5)$$

El Cuadro 6 revela una situación mixta para ambos modelos, ya que si bien capturan la dependencia lineal en la media y la varianza a la vez que abaten la correlación serial en residuales y residuales al cuadrado (Q(12) y Q(12) al cuadrado), sus especificaciones no eliminan por completo el sesgo y exceso de curtosis de las series de rendimientos. En efecto, al comparar los Cuadros 1 y 6 se puede observar que mientras los dos nuevos modelos disminuyen la curtosis (b_2) en 12 casos, las estimaciones sobre el sesgo ($\sqrt{b_1}$) resultaron menores en sólo 9 series que las registradas por las distribuciones de rendimientos originales. No obstante esto, los valores de los coeficientes de sus varianzas condicionales rezagadas, β , son significativamente positivos y menores que uno, lo que muestra que las series de rendimientos tienen un alto factor de persistencia (o larga memoria) sobre la varianza condicional del período anterior. En el caso concreto del EGARCH (1,1), el modelo comprobó la existencia de volatilidad grupal en todas las series con valores p iguales a cero, pero no dio cuenta de la presencia del efecto de apalancamiento en 21 activos. Los valores positivos de k en la mayoría de las series, así como su pobre significancia estadística en 4 activos que presentaron signo negativo, confirman el rechazo a la hipótesis de que las malas noticias del mercado accionario disminuyen el precio de las acciones. El mismo resultado puede observarse en el modelo GJR, donde los valores positivos de η_1 para 11 series de activos no resultaron significativos para niveles menores al 5%.¹⁸

La inclusión de estas tres últimas especificaciones no permite llegar a una conclusión definitiva sobre la mayor capacidad predictiva de un determinado modelo pues, al utilizar el criterio de Schwarz (CS) de probabilidad posterior, se encontró, en primera instancia, que la *t* de Student era la mejor hipótesis alternativa contra la normal estacionaria.¹⁹ Pero sólo en primera instancia, ya que, al comparar las celdas del Cuadro 7, se observa que hay una escasa diferencia entre los valores de CS de esta especificación y los de EGARCH y GJR, lo que impide concluir que la distribución de Student sea la única alternativa posible para llevar a cabo el pronóstico.²⁰ Cambios en la especificación de los dos últimos modelos ya sea en su media o en su varianza, pueden mejorar su potencia descriptiva con relación a la *t* de Student y, por ende, ofrecer mejor base para el pronóstico del rendimiento-riesgo de los activos.

III. Los Problemas Conceptuales en la Selección Adecuada de los Procesos Estocásticos

La exposición anterior dejó en claro que los criterios de selección de un proceso estocástico son, en buena medida, de orden empírico (Kon, 1984). Dos especificaciones distintas, como la *t* de Student y la mezcla de difusión con saltos, pueden arrojar resultados estadísticos satisfactorios aun cuando ambas privilegien parámetros diferentes. Decidir cuál de ellos es mejor parece, en este contexto, un asunto que tiene que ver más con la parsimonia y la verosimilitud del modelo (y

CUADRO 6

MODELOS DE HETEROCEDASTICIDAD CONDICIONAL TEMPO-DEPENDIENTES

°	GJR								EGARCH							
	ω	ζ	η	β	$\sqrt{b_1}$	b_2	Q	Q2	ω	γ	k	β	$\sqrt{b_1}$	b_2	Q	Q2
1	0.208 0.000	0.146 0.000	0.031 0.171	0.822 0.000	0.679	5.770	8.999 0.532	8.291 0.406	-0.126 0.000	0.244 0.000	0.013 0.284	0.969 0.000	0.712	5.868 0.375	10.78 0.364	8.7441 0.364
2	0.489 0.000	0.119 0.000	0.089 0.000	0.796 0.000	0.658	5.596	2.917 0.983	18.971 0.041	-0.089 0.000	0.324 0.000	-0.046 0.001	0.928 0.000	0.682	5.833 0.969	3.4543 0.041	18.924 0.041
3	0.046 0.000	0.150 0.000	-0.143 0.000	0.907 0.000	0.513	5.091	12.465 0.255	6.908 0.734	-0.135 0.000	0.193 0.000	0.091 0.000	0.987 0.000	0.47	4.969 0.646	7.8206 0.622	8.0672 0.622
4	0.172 0.000	0.133 0.000	-0.074 0.000	0.891 0.000	0.493	5.122	20.242 0.027	11.717 0.304	-0.100 0.000	0.190 0.000	0.049 0.000	0.981 0.000	0.455	4.64 0.290	20.018 0.157	14.354 0.157
5	0.133 0.000	0.050 0.000	0.020 0.140	0.916 0.000	0.389	5.009	12.120 0.277	12.946 0.227	-0.039 0.000	0.102 0.000	0.013 0.195	0.980 0.000	0.431	5.319 0.242	12.689 0.032	19.743 0.032
6	0.145 0.000	0.108 0.000	-0.100 0.000	0.919 0.000	0.670	5.986	6.521 0.770	13.082 0.219	-0.093 0.000	0.166 0.000	0.057 0.000	0.983 0.000	0.588	5.552 0.795	6.237 0.011	22.826 0.011
7	0.242 0.000	0.095 0.000	0.026 0.102	0.863 0.000	0.412	6.229	7.053 0.720	8.312 0.598	-0.095 0.000	0.193 0.000	-0.013 0.208	0.975 0.000	0.433	6.295 0.655	7.629 0.136	14.897 0.136
8	0.482 0.000	0.129 0.000	0.003 0.858	0.814 0.000	0.151	8.026	16.230 0.093	28.583 0.001	-0.070 0.000	0.304 0.000	-0.015 0.225	0.925 0.000	0.111	8.733 0.057	17.872 0.014	22.19 0.014
9	0.393 0.000	0.132 0.000	-0.013 0.533	0.836 0.000	0.684	5.943	15.368 0.119	9.186 0.515	-0.065 0.001	0.263 0.000	0.020 0.158	0.939 0.000	0.669	5.924 0.150	14.535 0.722	7.272 0.699
10	0.129 0.000	0.104 0.000	0.014 0.464	0.889 0.000	0.846	8.523	16.078 0.097	5.451 0.859	-0.108 0.000	0.245 0.000	0.007 0.637	0.970 0.000	0.939	9.871 0.072	17.086 0.877	5.204 0.877
11	0.240 0.000	0.117 0.000	-0.072 0.000	0.893 0.000	0.798	7.297	8.575 0.573	13.691 0.188	-0.101 0.000	0.180 0.000	0.062 0.000	0.984 0.000	0.763	7.076 0.550	8.814 0.056	17.926 0.056
12	0.164 0.000	0.163 0.000	-0.121 0.000	0.875 0.000	0.343	4.316	6.627 0.760	14.290 0.160	-0.109 0.000	0.217 0.000	0.085 0.000	0.968 0.000	0.359	4.391 0.678	7.496 0.048	18.459 0.048
13	0.067 0.000	0.059 0.000	-0.056 0.000	0.961 0.000	-0.066	6.777	5.673 0.842	41.498 0.000	0.101 0.000	0.162 0.000	0.095 0.000	0.902 0.000	-0.324	9.28 0.916	4.598 0.083	16.611 0.083
14	0.986 0.000	0.139 0.000	0.109 0.000	0.679 0.000	0.635	5.173	11.515 0.319	16.214 0.094	-0.003 0.924	0.321 0.000	-0.022 0.153	0.876 0.000	0.645	5.201 0.319	10.585 0.048	18.459 0.048
15	0.147 0.000	0.090 0.000	0.000 0.996	0.905 0.000	-0.228	13.587	8.874 0.544	3.704 0.960	-0.087 0.000	0.168 0.000	0.015 0.024	0.985 0.000	-0.087	11.711 0.485	9.505 0.918	4.573 0.918
16	0.032 0.000	0.096 0.000	0.010 0.382	0.901 0.000	0.587	8.406	10.617 0.388	10.193 0.424	-0.120 0.000	0.202 0.000	0.003 0.715	0.982 0.000	0.563	7.871 0.417	10.27 0.327	11.409 0.327
17	0.587 0.000	0.130 0.000	-0.064 0.006	0.815 0.000	0.415	4.703	15.804 0.105	5.052 0.888	-0.019 0.335	0.190 0.000	0.044 0.009	0.936 0.000	0.416	4.596 0.096	16.136 0.684	7.43 0.684
18	0.164 0.000	0.073 0.000	0.007 0.561	0.905 0.000	0.572	6.626	8.141 0.615	7.702 0.658	-0.015 0.013	0.145 0.000	0.021 0.028	0.957 0.000	0.607	7.66 0.447	9.926 0.057	17.882 0.057
19	0.172 0.000	0.111 0.000	-0.080 0.000	0.902 0.000	0.676	6.289	9.255 0.508	9.173 0.516	-0.075 0.000	0.167 0.000	0.062 0.000	0.972 0.000	0.685	6.506 0.584	8.462 0.242	12.689 0.242
20	0.095 0.000	0.092 0.000	-0.065 0.000	0.915 0.000	0.522	4.568	13.779 0.183	12.693 0.241	-0.083 0.000	0.154 0.000	0.046 0.001	0.975 0.000	0.532	4.575 0.289	11.944 0.236	12.792 0.236
21	0.148 0.000	0.075 0.000	-0.103 0.000	0.951 0.000	0.515	6.268	18.071 0.054	14.607 0.147	-0.074 0.000	0.138 0.000	0.033 0.001	0.986 0.000	0.594	5.445 0.127	15.135 0.105	15.804 0.105
22	0.550 0.000	0.110 0.000	0.010 0.624	0.834 0.000	0.870	7.897	6.965 0.729	12.821 0.234	-0.025 0.208	0.244 0.000	-0.006 0.660	0.931 0.000	0.858	7.939 0.705	7.211 0.071	17.137 0.071

CUADRO 7
COMPARACION DE LOS MODELOS CON BASE EN EL
CRITERIO DE SCHWARZ

Modelo	EGARCH (1,1)	GJR (1,1)	Student	Difusión con salto	Mezcla generalizada
EGARCH (1,1)	...	8	16	6	0
GJR (1,1)	15	...	17	6	0
Student	6	5	...	6	0
Difusión con salto	16	16	16	...	0
Mezcla	22	22	22	22	...

que es capturada por criterios como el de Schwarz) que con la naturaleza del fenómeno financiero bajo estudio. Y esto no deja de preocupar, porque la información que captura el modelo estocástico eventualmente rechazado, y que puede ser relevante para el análisis, no siempre es capturada por el modelo que se adopta.

La aceptación de un proceso estocástico no necesariamente significa, pues, que se adopte la mejor especificación *per se* o que la información que provea sea superior a la del proceso rechazado. Es muy probable que la explicación se halle, también, en el hecho de que el proceso seleccionado haya resultado menos castigado por el bajo volumen de transacciones realizadas o por la falta de información requerida para la estimación eficiente de sus parámetros.²¹ Las consecuencias de estas irregularidades en el registro de las series tienden a afectar más a algunos activos que a otros, sobre todo en periodos de crisis, por lo que la deficiente elección de un portafolio o de cierto periodo de análisis puede disminuir la capacidad descriptiva de los modelos orientados a capturar los saltos en los precios de los instrumentos financieros, como seguramente sucedió aquí con el modelo de difusión con saltos.

Muy cercano a este problema se encuentra otro de igual envergadura que tiene que ver con la pretendida superioridad de estos modelos estocásticos lineales sobre otras mezclas de procesos no lineales que se han ensayado con bastante éxito en mercados desarrollados. Entre esas destacan las mezclas de cadenas de Markov de alto orden con caminatas aleatorias, que se caracterizan por no requerir del supuesto de normalidad en las distribuciones de las series ni por asumir estructuras que privilegian los cambios en la varianza del último rezago. Las pruebas de estas mezclas tratan, en general, de determinar si la probabilidad de ocurrencia de un rendimiento bajo es mayor después de observar una secuencia de períodos de ingresos altos o una secuencia de ingresos bajos. El resultado más recurrente es que los rendimientos no siguen la tendencia de una caminata aleatoria debido a la presencia de un marcado patrón de rachas de rendimientos que se suceden con diferente regularidad (véase McQueen y Thorley, 1991).

El análisis basado en el paquete *Banamex 30* reveló consistentemente un patrón no lineal en las series de rendimientos. De acuerdo con el Cuadro 8, las ocurrencias

y probabilidades de transición de las series de rendimientos exhiben una fuerte correlación positiva en los rendimientos cada dos períodos, lo cual significa que la probabilidad de obtener un rendimiento alto luego de haber tenido una secuencia de dos rendimientos altos es mayor a la probabilidad de obtener un rendimiento bajo luego de dos altos o de cualquier otro ordenamiento, esto es $\lambda_{11} > \lambda_{00} > \lambda_{10} > \lambda_{01}$, para cualquier valor de λ_{ij} ($i, j = 0, 1$). La significancia de esta correlación es confirmada en el Cuadro 9, en donde queda claro que la propuesta de la caminata aleatoria es enfáticamente rechazada en favor de la existencia de una cadena de Markov de segundo orden, para todas las hipótesis de igualdad de λ_{ij} con valores p iguales a cero (para una versión sintética del modelo véase el Apéndice).

El resultado plantea una encrucijada, pues tanto este modelo no lineal como los modelos lineales ya vistos ofrecen ajustes estadísticamente significativos para valores p muy bajos y no hay, en principio, manera de optar por alguno de ellos, dada su distinta naturaleza. La decisión de elegir un tipo de modelo no es anodina si se toma en cuenta que ambos pueden arrojar, por ejemplo, diferentes estimaciones del VeR o del pronóstico de los rendimientos. En particular, la elección requiere tener en mente dos tipos de problemas.

CUADRO 8
OCURRENCIAS Y PROBABILIDADES DE TRANSICION

Estados anteriores		Ocurrencias Estado actual		Probabilidades Estado Actual	
		0	1	0	1
0	0	160	107	0.5993	0.4007
0	1	67	105	0.3895	0.6105
1	0	106	104	0.6199	0.3801
1	1	104	576	0.1529	0.8471
$\sigma(\lambda_{00}) = 0.02999$		$\sigma(\lambda_{01}) = 0.03718$		$\sigma(\lambda_{10}) = 0.03712$	
				$\sigma(\lambda_{11}) = 0.01380$	

CUADRO 9
PRUEBAS DE HIPOTESIS SOBRE LAS PROBABILIDADES DE TRANSICION DEL PROCESO DE MARKOV DE SEGUNDO ORDEN

H ₀	RV	p-value	Wald	p-value	$\chi_1^2(5\%)$	$\chi_1^2(1\%)$
$\lambda_{00} = \lambda_{11}$	179.53	0.00	182.75	0.00	3.84	6.63
H ₀	RV	p-value	Wald	p-value	$\chi_3^2(5\%)$	$\chi_3^2(1\%)$
$\lambda_{00} = \lambda_{01} = \lambda_{10} = \lambda_{11}$	253.31	0.00	7.82	0.00	7.82	11.34

Nota: RV y Wald son las pruebas de razón de verosimilitud y de Wald, respectivamente.

El primero tiene que ver con el uso indiscriminado de algunos procesos estocásticos que se ponen de moda en el análisis financiero, como los brownianos o los modelos tipo GARCH, sin un conocimiento pormenorizado del mercado. Y es que los movimientos brownianos *puros* pueden, por ejemplo, resultar muy valiosos para explicar series afectadas por cambios en sus parámetros estacionarios, pero son de poca utilidad teórica para explicar la existencia de arbitraje o, en determinadas ocasiones, para capturar la leptocurtosis de las distribuciones empíricas (Schoutens, 2003). Para eso, tal vez sea necesario combinarlos con algunas ecuaciones *ad hoc*, como las de fricción o de Langevin (Takahashi, 1996), incorporarlos en un proceso de Levy (Eberlein *et al.*, 1998) o, directamente, sustituirlos por algunas distribuciones multivariadas, como la *t* de Student, que posean colas más pesadas que la normal (Glasserman *et al.*, 2002).

Sobre este último punto Eberlein *et al.* (1998) sostienen que una variante de los procesos de Levy,²² el modelo hiperbólico, ajusta mejor la distribución leptocúrtica de los rendimientos que los modelos basados en las distribuciones log-normales, como los brownianos *puros*.²³ La razón reside no sólo en que la log-densidad del modelo de Levy es una hipérbola, y no una parábola como la de la distribución normal, lo cual permite una mejor perspectiva para el análisis de distribuciones de rendimientos con formas picudas, sino también porque las trayectorias muestrales del modelo hiperbólico son discontinuas, tales como lo son los movimientos de los precios que uno observa en las transacciones diarias de cualquier mercado accionario. Glasserman *et al.* (2002) muestran, a su vez, que algunas generalizaciones de la distribución *t* multivariada (las elípticamente configuradas) arrojan mejores resultados que los procesos de Levy, en el análisis de series leptocúrticas, debido a su mayor capacidad para modelar cambios en los factores de riesgo cuando las distribuciones presentan colas excesivamente pesadas. En concreto, los autores afirman que algunas propiedades específicas de estas generalizaciones permiten considerar a las distribuciones *t* multivariadas como un instrumento más flexible que los procesos de Levy en el análisis de series financieras leptocúrticas. Entre estas propiedades destacan: la posibilidad de regular la pesadez de las colas mediante la manipulación de un parámetro a una tasa polinomial (y no exponencial como en los procesos de Levy); la mayor facilidad en el cómputo de las probabilidades de pérdidas sin asumir linealidad entre el valor del portafolio y los cambios en los factores de riesgo y, sobre todo, la conservación de ciertas propiedades de las distribuciones normales multivariadas en presencia de colas pesadas.

Cualquiera sea el caso, lo importante es tener claro que no hay ningún modelo estándar y universal para resolver, a manera de prescripción, los problemas que supone la violación del supuesto de normalidad estacionaria, pues es claro que la bondad de ajuste de un determinado proceso estocástico puede ser muy distinta cuando se aplica a acciones que cuando se aplica a activos afectados por un número mayor de factores de riesgo.

El segundo problema está relacionado con el carácter pronosticable de las series de rendimientos o, específicamente, con la existencia de patrones de dependencia entre ellos. De acuerdo con Malkiel (1999), el rechazo a la hipótesis de la caminata aleatoria no significa la violación automática del supuesto de

eficiencia débil de los mercados, debido a que la existencia de correlaciones entre los rendimientos de los activos puede obedecer también a ajustes extemporáneos del mercado o a ejercicios de *data mining* por parte del analista. Por estas razones hay que tener mucho cuidado en no sobreestimar el significado estadístico de una serie pronosticable. En el ejercicio hecho para el paquete *Banamex 30* concluimos, por ejemplo, que las series siguen un patrón de correlación positiva simplemente porque es el que sugieren los modelos que mejor ajustan los datos. Pero no hay nada que indique que ese patrón sea el resultado de movimientos temporales o permanentes descritos por uno o varios factores explicativos, o que ese sea el derrotero forzoso de los rendimientos en el futuro. El pronóstico tan sólo dice que, de mantenerse la tendencia pasada, los rendimientos futuros se escalonarían positivamente cada dos períodos, pero sin que nada lo garantice ni nada, más allá de los parámetros del modelo, lo explique.

Entre los argumentos económicos o financieros que explican la correlación positiva de las series de acciones de *Banamex 30* destacan dos hipótesis ampliamente constatadas en los mercados desarrollados. La primera es que esa correlación no es más que el fruto de una reacción tardía por parte de las acciones de menor capitalización a las buenas noticias registradas por las acciones de mayor capitalización. Como se sabe, las acciones pequeñas tienden a internalizar las buenas noticias un tiempo después de que las grandes acciones hayan registrado cambios favorables en sus rendimientos, por lo que es posible observar portafolios ganadores compuestos por esas pequeñas acciones posterior al movimiento favorable producido por las grandes (tal como lo sugiere el análisis de *lead-lag* aquí desarrollado).

La segunda hipótesis propone, por su parte, que los movimientos de los precios de las acciones, y en consecuencia de los rendimientos, son ajustes momentáneos a las variaciones experimentadas por las tasas de interés. Así tenemos que cuando estas últimas suben, los precios de las acciones tienden comúnmente a bajar, *ceteris paribus*, para proporcionar mayores rendimientos esperados en el futuro ya que, de otra manera, las acciones no serían competitivas frente a los bonos de alto rendimiento. Del mismo modo, cuando las tasas de interés bajan, las acciones tienden a subir su precio para competir en mejores términos con los bonos de bajo rendimiento. En ambas situaciones, la existencia de estos patrones de dependencia son simples ajustes temporales y no deben considerarse como una fuente permanente de arbitraje.

El tema del pronóstico de los rendimientos es, pues, un asunto sin conclusión, ya que aun cuando hay un número creciente de estudios que evidencian el rechazo a la hipótesis nula de la caminata aleatoria, también hay una importante resistencia a aceptar que la presencia de otras hipótesis alternativas, como la del componente estacionario de la reversión a la media, justifique la naturaleza pronosticable de las series (Fama, 1991). Lo uno no implica lo otro. Y es que a las investigaciones que aseguran que el pronóstico es una función de algunas variables claves (en particular de los rendimientos pasados, los dividendos, la razón entre ganancias de mercado y precio y la estructura intertemporal de tasas) se les atribuye una gran variedad de deficiencias, tales como: su marcada manipulación de los datos, o *data mining*, la pobre robustez de sus pruebas estadísticas ante la ausencia de

normalidad estacionaria y la incorrecta inferencia de sus resultados cuando incluyen muestras pequeñas (Xia, 2001).

Una situación parecida parece acontecer con los métodos alternativos de cálculo del VeR que buscan atacar el problema de las colas gordas con técnicas estadísticas diversas y no necesariamente compatibles. Los métodos paramétricos, no paramétricos y semiparamétricos orientados a rehabilitar el método original del VeR, han tenido que enfrentar los mismos problemas que los modelos de pronósticos al momento de elegir un proceso estocástico alternativo. En concreto, los procedimientos sugeridos por Venkataraman (1997), Hull y White (1998), Eberlein *et al.* (1998) y Li (1999) han tenido que recurrir a una variedad de procesos (como las mezcla de normales, los modelos GARCH y los hiperbólicos de Levy, entre otros) para comparar el número de observaciones de una distribución empírica que está fuera de ciertos intervalos de confianza con relación al que es esperado bajo una normal teórica. No obstante la diferente naturaleza de los procesos, los resultados confirman que todos los procedimientos permiten capturar mejor esas observaciones y, por ende, disminuir el porcentaje de violaciones del VeR en relación al estimado por una distribución normal.²⁴ El hecho no deja de llamar la atención, sobre todo si consideramos que las metodologías empleadas son muy diversas e incluyen desde mapeos entre una distribución de probabilidad alternativa y una normal (Hull y White, 1998) hasta la construcción de intervalos de confianza basados en el sesgo y la curtosis de las distribuciones empíricas (Li, 1999).

El mejoramiento introducido en la estimación del VeR por estos procedimientos no es, sin embargo, automático: depende crucialmente de la adecuada selección que el administrador de riesgos haga del método y del proceso estocástico en cuestión. Sin un juicio pertinente sobre la naturaleza de las series, cualquier aplicación estadística puede incluso subestimar aún más el riesgo que con los métodos tradicionales. Por eso es importante introducir criterios de evaluación sobre la adaptación de un determinado proceso. Y en este punto el tipo de trabajo desarrollado por Kupiec (1995), en el que se fundamentan las propiedades estadísticas de un análisis de cociente de fallas, puede resultar muy importante. De acuerdo con dicho análisis, el método más eficiente de estimación del VeR es aquel que incorpore el proceso con mejor ajuste del exceso de curtosis, ya que tanto el sesgo como la desviación estándar de los estimadores de los valores críticos 0,01 y 0,05 se incrementan a medida que la distribución subyacente sea más leptocúrtica.

Pero aquí hay que aclarar que ningún pronóstico o análisis de VeR puede hacerse sobre la base de un mero tratamiento estadístico. Nada puede sustituir el juicio del administrador de riesgos en la decisión de elegir un determinado proceso estocástico, pues ningún criterio de evaluación es completo sin el entendimiento del grado de aversión al riesgo del inversionista, de los costos de transacción o del comportamiento de los factores de riesgo de los instrumentos que componen la cartera. El conocimiento de los elementos idiosincrásicos y ambientales (como son las políticas internas de las empresas o su entorno legal y macroeconómico) es esencial en cualquier estrategia de balanceo de portafolio y, de ninguna manera, puede estar al margen de un análisis estadístico, por más sofisticado y completo que éste parezca.

IV. Conclusiones

El documento hace énfasis en el hecho, muchas veces soslayado por la práctica financiera, de que la evaluación del supuesto de normalidad estacionaria es esencial para formular un análisis adecuado del riesgo-rendimiento de un activo o de un portafolio. El reconocimiento a la recurrente violación de ese supuesto en las distribuciones empíricas de los instrumentos financieros ha dado lugar a un buen número de técnicas orientadas a la detección y corrección de las distribuciones de rendimientos con colas gordas o pesadas.

Los resultados basados en el paquete *Banamex 30* confirman que los rendimientos diarios no son normales estacionarios y que, por tanto, los análisis de VeR o de pronóstico de rendimientos tanto para mercados emergentes como desarrollados no pueden formularse correctamente con el uso de métodos tradicionales. El rechazo a la hipótesis nula de la caminata aleatoria, RW3, para esas series sugiere que dichos análisis pueden ser mejorados sustancialmente con el apoyo de algunos procesos estocásticos. Los modelos más descriptivos de las series anormales de rendimientos fueron, de acuerdo con el criterio de Schwartz, los modelos de Student, EGARCH (1,1) y GJR (1,1). La conclusión sugerida en el texto es que la adopción de cualquiera de las tres especificaciones es hasta cierto punto arbitraria, ya que aun cuando la distribución de Student registró la más alta probabilidad de explicar las series no normales, los otros dos modelos podrían mejorar su capacidad descriptiva al combinarlos con otros modelos de media y varianza. La selección de determinado modelo debe hacerse, pues, sobre bases conceptuales.

La contribución del trabajo no se reduce, sin embargo, a mostrar la combinación de técnicas que mejor ajuste las series de rendimientos no normales sino, fundamentalmente, a llamar la atención sobre la infinidad de problemas conceptuales que están relacionados con la selección adecuada de un proceso estocástico. Los argumentos discutidos en la segunda parte del documento ofrecen elementos para una discusión más avanzada sobre estos delicados aspectos, ya que insisten en que no hay reglas fijas para la selección de determinados modelos estocásticos ni algoritmos únicos para corregir la no normalidad de las distribuciones empíricas.

La selección de cualquier proceso estocástico debe ir más allá de los simples criterios estadísticos, pues no está claro qué modelo rechazar cuando las especificaciones son distintas y la información contenida en los parámetros del modelo rechazado no se puede incluir en el modelo escogido. El juicio del administrador de riesgos es esencial para decidir el mejor modelo y, sobre todo, para entender el significado estadístico de una serie pronosticable o la corrección de una distribución leptocúrtica mediante un método alternativo de cálculo del VeR. Su conocimiento sobre el grado de aversión al riesgo del inversionista o sobre los costos de transacción incluidos en la estrategia de una inversión *estadísticamente exitosa*, son imprescindibles para justificar el empleo adecuado de los procesos estocásticos en las políticas de administración de riesgos de una institución.

Notas

- 1 La ecuación básica de estos modelos de fijación de precios de activos es $P_{it} = E_t[M_{t+1}X_{i,t+1}]$; donde P_{it} es el precio del activo i en el tiempo t , E_t es el operador de la esperanza condicional a la información en el tiempo t , $X_{i,t+1}$ es el rendimiento aleatorio del activo i en el tiempo $t+1$ y M_{t+1} es el factor de descuento estocástico (FDE). Las aplicaciones de esta ecuación parten del análisis de los dos primeros momentos de una distribución normal e incluyen desde la elaboración de modelos del FDE, que ajustan tanto los datos de $X_{i,t+1}$ como los de P_{it} , hasta el uso de modelos microeconómicos en los que se hacen explícitos las preferencias del inversionista y los patrones que rigen a $X_{i,t+1}$ y al FDE. La violación del supuesto de normalidad estacionaria conduce en cada caso a reformulaciones muy profundas de la ecuación básica (véase Campbell, 2000).
- 2 La presencia de colas gordas o pesadas es debido al exceso de curtosis de una distribución empírica con relación a la supuesta por una normal teórica. Este exceso incrementa la probabilidad de presenciar valores grandes y pequeños a costa de disminuir la probabilidad de ocurrencia de los valores moderados. Como sabemos, la curtosis de una distribución normal es 3 (mesocúrtica). Si la distribución empírica tiene curtosis menor que 3 se dice que ésta es platocúrtica y si tiene valores mayores que tres, entonces, se dice que es leptocúrtica o de colas gordas.
- 3 Dado que existe una cuenta de mercado de dinero, se dice que un mercado es dinámicamente completo si hay al menos tantos activos riesgosos como fuentes de riesgo y, además, esos activos no están perfectamente correlacionados en cada instante del tiempo. En caso de que haya menos activos riesgosos no habría flexibilidad para negociar y replicar su precio, como, por ejemplo, sucedería en caso de que hubiera correlación instantánea (ya que en esa situación uno de los activos sería combinación de los otros y, por lo tanto, habría menos activos que factores de riesgo).
- 4 La naturaleza adaptada a la estructura de información de los parámetros de la integral estocástica o integral de Itô, garantiza que los valores de éstos sean dependientes únicamente de la trayectoria de Wiener (Kunita, 1990).
- 5 En caso de que la obligación dependa de más información que la ofrecida por la trayectoria de Wiener, entonces ya no es posible asegurar que el precio pueda ser replicado por la estrategia de negociación.
- 6 Este tipo de replicación requiere a veces de varios pasos sucesivos. Por ejemplo, para encontrar la estrategia autofinanciable en el modelo de Black y Scholes es necesario, primero, convertir la ecuación diferencial estocástica del subyacente en una martingala y, segundo, usar el teorema de representación de martingalas para crear una estrategia de replicación para cada obligación.
- 7 Nielsen (1999) asegura que los movimientos brownianos son la piedra fundamental del 98% de toda la teoría financiera en tiempo continuo. La mezcla de los brownianos con los procesos de Poisson, conocida como mezcla de difusión con saltos, es, de hecho, el instrumento analítico más socorrido para analizar los mercados financieros que se caracterizan por tener alta inestabilidad, incertidumbre e información asimétrica.
- 8 Otros trabajos elaborados por físicos han utilizado distribuciones de Levy y de leyes de potencias, análisis de atractores caóticos de baja dimensión, modelos de multifractales y de turbulencias, entre otros, para modelar el comportamiento de precios cuyas distribuciones son afectadas por sesgo y exceso de curtosis (para una revisión completa del tema, ver Farmer, 1999). Una corriente más reciente de autores ha mostrado las bondades de algunas variantes de las distribuciones t multivariadas para analizar las series de rendimientos que exhiben leptocurtosis (véase, por ejemplo, Glasserman *et al.*, 2002)
- 9 El paquete *Banamex-30* es un portafolio que consta originalmente, como su nombre lo indica, de igual número de acciones. Desafortunadamente, la falta de información nos obligó a reducir el tamaño de la muestra a 21 series accionarias. Aun así ésta es muy grande, pues abarca a series de 1,541 observaciones diarias para cada una de las acciones entre el 3 de enero de 1994 y el 24 de febrero de 2000. Los rendimientos fueron calculados de acuerdo con la práctica habitual:

$$R_t = \ln\left(\frac{X_t}{X_{t-1}}\right)$$
, donde R_t es el rendimiento en el tiempo t de un activo diario y X_t es el precio del activo al cierre de cada día. Adicionalmente, y para fines de comparación, se incluye la serie del índice *Banamex 30*, tal como es computado por dicho banco mexicano. En el cálculo de

cada serie de rendimientos diarios se hicieron los ajustes correspondientes para distintos años calendarios, así como para los lunes de cada semana.

- ¹⁰ Los valores de sesgo y curtosis de la tabla indican que la distribución de rendimientos tiene una cola más larga a la derecha que a la izquierda (sesgo positivo) así como un pico más alto que la normal (leptocurtosis). El exceso de curtosis indica que las colas de la distribución normal se acercan a cero mucho más rápido que las de la distribución empírica.
- ¹¹ En estos modelos el sesgo observado es explicado por desplazamientos en la media y las colas gordas por los desplazamientos en la varianza.
- ¹² Las pruebas se hicieron estimando las series de cada activo bajo el supuesto de que las distribuciones se comportaban como una distribución de Student y una mezcla de normales, de acuerdo con las funciones de verosimilitud: $L_T(\delta) = \sum_{t=1}^T \log g(r_t / \delta)$ (para la distribución t de Student); y $\ell\left(\frac{\theta}{r}\right) = \prod_{t=1}^T \left[\sum_{i=1}^N \lambda_i p(r / \gamma_i) \right]$ (para la mezcla de normales) donde $\gamma_i = (\alpha_i, \sigma_i^2)$; y $p(r/\gamma_i)$ es la función de densidad normal con media α_i y varianza σ_i^2 .
- ¹³ Muy brevemente decimos que una martingala es un proceso estocástico que satisface la condición de juego justo o equivalentemente que: $E[P_{t+1} - P_t / P_t, P_{t+1}, \dots] = 0$. Para mayor información consúltese Ross (1996).
- ¹⁴ Los resultados de otros trabajos hechos para México revelan que la evidencia es mixta, incluso para datos relativamente largos (un mes). En particular el trabajo de De la Luz (2002), que utiliza datos semanales para un período mas o menos cercano al que aquí usamos, sostiene que la mayoría de las acciones de alta y media bursatilidad (con un porcentaje de no operación del 30%) que seleccionó en su muestra se comportó como una martingala
- ¹⁵ Las razones de varianza son significativas para 8 rezagos. Cuando los rendimientos están correlacionados positivamente, la varianza de la suma de los rendimientos de dos períodos individuales será mayor que la suma de la varianza de los rendimientos de un período, con lo que las varianzas crecerán más que linealmente.
- ¹⁶ Para computar la densidad de este modelo se requirió, primero, estimar la función de densidad de difusión $f(r) = (1 - \lambda)\phi(r; \alpha, \sigma^2) + \lambda\phi(r; \alpha, \sigma^2, \delta^2)$ y, luego, incorporar un proceso de Poisson truncado con densidad $\sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} \phi(r; \alpha, \sigma^2 + n\delta^2)$. Los parámetros del modelo mixto fueron estimados por máxima verosimilitud de acuerdo con el método BFGS. Para una justificación teórica de la forma de estimación de estos modelos consúltese Ball y Torous (1985).
- ¹⁷ Los parámetros de la ecuación (4) indican que si $\gamma > 0$ entonces la volatilidad condicional tenderá a aumentar o decrecer en razón proporcional al valor absoluto del residual estandarizado, con lo que se puede afirmar que cambios significativos en los precios de las acciones son precedidos por otros cambios significativos en los precios (volatilidad grupal). Asimismo, un valor negativo de k indica la existencia del efecto de apalancamiento. En el modelo GJR ambos parámetros son captados mediante la agregación del término de asimetría n_1 a un modelo GARCH en el que $d_{t-1} = 1$ si $\varepsilon_{t-1} < 0$ y $d_{t-1} = 0$ de otra manera. El impacto del residual cuadrado sobre la volatilidad condicional variará dependiendo si el residual de rezago 1 es positivo o negativo. Por esa razón el efecto de apalancamiento se detecta en este modelo si $\eta_1 > 0$.
- ¹⁸ La explicación a esta aparentemente extraña conclusión puede hallarse en el irregular comportamiento de las acciones durante el período bajo estudio. En concreto, después de los movimientos irregulares de los primeros meses de 1995 el mercado accionario entró en una fase relativamente estable que, con sus excepciones, se mantuvo hasta el final del periodo (de hecho, las acciones de la muestra siguen cotizando hoy en día), lo cual probablemente inhibió el efecto apalancamiento de las empresas durante 1995. La división del periodo puede, seguramente, ayudar a hacer más notorio ese efecto en más número de acciones.
- ¹⁹ En conformidad con el criterio de Schwarz, la especificación de un modelo es más probable, o tiene la probabilidad posterior más alta, de predecir el comportamiento de una serie, si el valor de CS es más pequeño, o equivalentemente, si el logaritmo de las probabilidades posteriores del modelo competidor con relación a una normal estacionaria es negativo.
- ²⁰ El Cuadro 7 muestra, por celda, los casos en que cada modelo predice mejor que su modelo competidor el comportamiento de las series accionarias, según los valores del CS.

- 21 Kim y Kon (1994) sostienen que la asincronía en la negociación de títulos es una explicación de la correlación serial y de la menor capacidad descriptiva de los modelos.
- 22 Un proceso de Lévy es un proceso estocástico que se inicia en cero y que tiene incrementos independientes y estacionarios regulados por una distribución infinitamente divisible, esto es: que la distribución de un incremento sobre $(s, s+t)$, $s, t \geq 0$ tiene $(\phi(u))^t$ como su función característica. Una distribución es infinitamente divisible si para cada entero n , $\phi(u)$ es también la n -ésima potencia de la función característica. Un proceso de Lévy consta de tres partes independientes: una parte determinística lineal, una parte Browniana y una de salto puro. Ejemplos de estos procesos son los procesos homogéneos y compuestos de Poisson, los gamma, los inversos gaussianos y los inversos gaussianos generalizados (Schoutens, 2003).
- 23 El ajuste que hacen estos autores de las series de rendimientos de algunas acciones alemanas con base en el modelo hiperbólico es casi perfecto.
- 24 En un ejercicio hecho para el paquete *Banamex 30*, y que no reproducimos aquí por falta de espacio, comprobamos que la t de Student, el modelo GJR, el EGARCH y el modelo mixto de difusión con saltos capturaron eficientemente las observaciones que estaban fuera de los intervalos de confianza de acuerdo con los modelos sugeridos por Li (1999) y Hull y White (1998). No encontramos criterio alguno para discriminar el mejor modelo.
- 25 En Ramírez y Sandoval (2003) se justifica la elección del orden del proceso de Markov.

Referencias

- AFFLECK-GRAVES, J. y B. MCDONALD (1989). "Non-normalities and Test of Asset Pricing Theories". *The Journal of Finance*, 44 (4), pp. 889-908.
- ANDERSEN, T.; T. BOLLERSLEV y A. DAS (2001). "Variance-ratio and High-frequency Data: Testing for Changes in Intraday Volatility Patterns". *The Journal of Finance*, 56 (1), pp. 305-327.
- BALL, C. y W. TOROUS (1985). "On Jumps in Common Stock Prices and their Impact on Call Option Pricing". *The Journal of Finance*, 40 (1), pp. 73-155.
- BAXTER, M. y A. RENNIE (1996). *Financial Calculus. An Introduction to Derivative Pricing*. Cambridge University Press, Cambridge.
- CAMPBELL, J.; A.W. LO y A.C. MACKINLAY (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- CAMPBELL, J. (2000). "Asset Pricing at the Millennium". *The Journal of Finance*, 55 (4), pp. 1515-1567.
- CAMPBELL, J. y A. SIDDIQUE (2000). "Conditional Skewness in Asset Pricing Tests". *The Journal of Finance*, 55 (3), pp. 1263-1295.
- DE LA LUZ, N. (2002). "La Hipótesis de Martingala en el Mercado Bursátil Mexicano". *Estudios Económicos*, 17 (33), pp. 91-119.
- HULL, J. y A. WHITE (1998). "Value at Risk When Daily Changes in Market Variables are not Normally Distributed". *The Journal of Derivatives*, 5 (3), pp. 9-19.
- EBERLEIN, E.; U. KELLER y K. PRAUSE (1998). "New Insights into Smile, Mispricing, and Value at Risk: The Hyperbolic Model". *The Journal of Business*, 71 (3), pp. 371-405.
- FAMA, E. (1965). "The Behavior of Stock Market Prices". *Journal of Business*, 38 (1), pp. 34-105.
- FAMA, E. (1991). "Efficient Capital Markets: II". *The Journal of Finance*, 45 (5), pp. 1575-1615.
- FARMER, J. (1999). "Physicists Attempt to Scale the Ivory Towers of Finance". *Computing in Science & Engineering*, Noviembre-Diciembre, pp. 26-39.
- GLASSERMAN P.; P. HEIDELBERGER y P. SHAHABUDDIN (2002). "Portfolio Value-At-Risk with Heavy-Tailed Risk Factors". *Mathematical Finance*, 12 (3), pp. 239-269.
- GRIEB, T. y M. REYES (1999). "Random Walks Tests for Latin American Equity Indexes and Individual Firms". *Journal of Financial Research*, 22 (4), pp. 371-383.
- JOHNSON, R. y D. WICHERN (1992). *Applied Multivariate Statistical Analysis*. Englewood Cliffs, Prentice Hall, New Jersey.

- KIM, D. y S. KON (1994). "Alternative Models for the Conditional Heteroscedasticity of Stock Returns". *The Journal of Bussines*, 67 (1), pp. 63-98.
- KON, S.J. (1984). "Models of Stock Returns. A Comparison". *The Journal of Finance*, 39 (1), pp. 147-165.
- KUNITA, H. (1990). *Stochastic Flows and Stochastic Differential Equations*. Cambridge University Press, Cambridge.
- KUPIEC, P.H. (1995). "Techniques for Verifying the Accuracy of Risk Measurement Models". *The Journal of Derivatives*, 3, pp. 73-84.
- LI, D.X. (1999). *Value at Risk Based on the Volatility, Skewness and Kurtosis*, Riskmetrics Group, New York.
- LONG, M.; J. PAYNE y F. JANG (1999). "Information Transmission in the Shangai Equity Market". *The Journal of Financial Research*, 22 (1), pp. 29-45.
- MALKIEL, B. (1999). *A Random Walk Down Wall Street*. WW Norton & Company Inc., New York.
- MANDELBROT, B. (1963). "The Variation of Certain Speculative Prices". *Journal of Business*, 36 (4), pp. 394-419
- MCQUEEN, G. y S. THORLEY (1991). "Are Stock Returns Predictable? A Test Using Markov Chains". *The Journal of Finance*, 46 (1), pp. 239-263.
- NEFTCI, S.N. (1996). *An Introduction to the Mathematics of Financial Derivatives*. Academic Press, London.
- NIELSEN, L. (1999). *Pricing and Hedging of Derivative Securities*. Oxford University Press, New York.
- OJAH, K. y D. KAREMERA (1999). "Random Walks and Market Efficiency Tests of Latin America Emerging Equity Markets: A Revisit". *The Financial Review*, 34 (2), pp. 57-72.
- RAMÍREZ, J.C. y R. SANDOVAL (2003). "Una Propuesta para Evaluar Pronósticos de Rendimientos de Acciones cuando las Distribuciones no son Normales Estacionarias". *Estudios Económicos*, 18 (36), pp. 237-277.
- RICHARDSON, M. y T. SMITH (1993). "A Test for Multivariate Normality in Stock Returns". *The Journal of Bussines*, 66 (2), pp. 295-321.
- ROGERS, L.C. (1996). "Arbitrage with Fractional Brownian Motion". *Mathematical Finance*. 7 (1), pp. 95-105.
- ROSS, S. (1996). *Stochastic Processes*. John Wiley and Sons, Inc., New York.
- SANCHEZ-CERON, C. (2001). *Valor en Riesgo y otras Aproximaciones*. Valuación, Análisis y Riesgo, A.C., México.
- SCHOUTENS, W. (2003). *Lévy Processes in Finance. Pricing Financial Derivatives*. John Wiley and Sons, New York.
- TAKAHASHI, M. (1996). "Non-ideal Brownian Motion. Generalizad Langevin Equation and its Application to the Security Market". *Financial Engineering and the Japanese Markets*. 3 (2), pp. 87-119.
- VENKATARAMAN, S. (1997). "Value at Risk for a Mixture of Normal Distributions: The Use of Quasi-Bayesian Estimation Techniques". *Economic Perspectives*, 21 (2), Federal Reserve Bank of Chicago, March/April, pp. 2-13.
- XIA, Y. (2001). "Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation". *The Journal of Finance*, 56 (1), pp. 205-245.
- ZANGARI, P. (1996). "Test of Conditional Normality". *RiskMetricsTM -Technical Document*, J.P. Morgan/Reuters.

APENDICE

El modelo de cadenas de Markov y caminatas aleatorias expuesto a continuación está basado en McQueen y Thorley (1991) y en Ramírez y Sandoval (2003). Su estructura requiere, primero, de ordenar la serie de rendimientos de acuerdo con una secuencia I_t definida de la siguiente forma:

$$I_t = \begin{cases} 1 & \text{si } r_t > R \\ 0 & \text{si } r_t < R \end{cases}$$

donde $r_t = \ln(P_t/P_{t-m})$ es el rendimiento del activo ($m=1$ para rendimientos diarios o $m=5$ para semanales), P_t es el precio del activo en el tiempo t , y R una medida de la esperanza matemática que divide a los rendimientos en altos y bajos. Una vez establecida la secuencia hay que considerar a $S_T = \{i_1, i_2, \dots, i_T\}$ como una realización de $\{I_t\}$ para, de esa manera, definir a la función de verosimilitud de la cadena de Markov de segundo orden²⁵ como:

$$L(S_T) = P(I_2 = i_2, I_1 = i_1) \prod_{s=0}^{T-3} P(I_{T-s} = i_{T-s} | I_{T-s-2} = i_{T-s-2}, I_{T-s-1} = i_{T-s-1})$$

la cual, una vez incorporadas las probabilidades de transición y las del estado inicial,

$$P(I_k = 0 | I_{k-2} = i, I_{k-1} = j) = \lambda_{ij}, \quad i, j = 0, 1$$

$$P(I_2 = i_2, I_1 = i_1) = \pi_0$$

la podemos expresar de la siguiente manera:

$$L(S_T, \Lambda, \pi_0) = \pi_0 \prod_{i,j=0}^1 \lambda_{ij}^{N_{ij}} (1 - \lambda_{ij})^{M_{ij}}, \quad i, j = 0, 1$$

$$\Lambda = [\lambda_{00} \quad \lambda_{01} \quad \lambda_{10} \quad \lambda_{11}]'$$

donde N_{ij} representa el número de secuencias de *ceros* y *unos* tales que $(I_t=0/I_{t-2}=i, I_{t-1}=j)$, $i, j = 0, 1$ y $t = 3, 4 \dots T$, y M_{ij} el número de secuencias con el patrón $(I_t=1/I_{t-2}=i, I_{t-1}=j)$, $i, j = 0, 1$ y $t = 3, 4 \dots T$. La forma logarítmica (o función de log-verosimilitud) de la ecuación es:

$$\ln\{L(S_T, \Lambda, \pi_0)\} = \ln(\pi_0) + \sum_{i,j=0}^1 \{N_{ij} \ln(\lambda_{ij}) + M_{ij} \ln(1 - \lambda_{ij})\}$$

El problema de incluir las probabilidades del estado inicial en la función de log-verosimilitud es que se necesita utilizar métodos numéricos engorrosos para obtener los estimadores máximo verosímiles. Por esa razón conviene suponer que el número de observaciones es grande, como en efecto es el caso de este trabajo, para obviar el cálculo de esas probabilidades en la estimación. De esta manera, al resolver las condiciones de primer orden del problema de maximización de $\ln(L)$, ignorando p_0 , se obtienen los siguientes estimadores más simplificados:

$$\lambda_{ij} = \frac{N_{ij}}{N_{ij} + M_{ij}}, \quad i, j = 0, 1$$

Las matrices de ocurrencias y de probabilidades de transición del proceso de Markov de segundo orden resultantes de estos estimadores se resumen en el Cuadro A1.

CUADRO A1
OCURRENCIAS Y PROBABILIDADES DE TRANSICION

Estados Anteriores		Ocurrencias Estado Actual		Probabilidades Estado Actual	
		0	1	0	1
0	0	N_{00}	M_{00}	λ_{00}	$1 - \lambda_{00}$
0	1	N_{01}	M_{01}	λ_{01}	$1 - \lambda_{01}$
1	0	N_{10}	M_{10}	λ_{10}	$1 - \lambda_{10}$
1	1	N_{11}	M_{11}	λ_{11}	$1 - \lambda_{11}$

La pruebas orientadas a determinar la existencia de un patrón en las series de rendimientos, y que tienen como hipótesis nula el comportamiento de una caminata aleatoria, se reducen en el modelo a hacer un contraste de igualdad entre las distintas combinaciones de las probabilidades de transición.