

ESTUDIO DE LA VOLATILIDAD REALIZADA APLICADO AL ÍNDICE DE PRECIOS Y COTIZACIONES DE MÉXICO

Arturo Lorenzo Valdés.*

Departamento de Contabilidad y Finanzas

Tecnológico de Monterrey, Campus Ciudad de México

(Recibido 2 de septiembre 2004, aceptado 3 de diciembre 2004)

Resumen

El objetivo de este trabajo es motivar el uso de datos de alta frecuencia en los mercados financieros mediante el estudio de la volatilidad realizada. En los análisis financieros y la administración de riesgos, un elemento importante es la estimación de la volatilidad. Para lo anterior se sugiere el uso de la volatilidad realizada diaria calculada con datos intradía y aplicada al IPC. Se presentan sus propiedades y hechos estilizados así como un estudio empírico para contrastarlos, en el que se encuentra evidencia de no normalidad pero sí de memoria larga. Posteriormente se propone un modelo ARFIMAX para describir y predecir la volatilidad comparándolo con los modelos GARCH tradicionales.

Abstract

The objective of this work is to motivate the use of high frequency data on financial markets by means of the study of realized volatility. In financial analysis and risks management, an important element is the estimation of the volatility. To achieve this it is suggested the use of the realized daily volatility calculated with intraday information and applied to the IPC. We present realized volatility properties and stylized facts as well as an empirical study to contrast them, finding evidence of not normality but of long memory. Later we propose an ARFIMAX model to describe and to predict the volatility comparing it with the traditional models GARCH.

Clasificación JEL: C22, C53, G12

Palabras clave: Alta frecuencia, Volatilidad realizada, ARFIMAX, GARCH

* Departamento de Contabilidad y Finanzas. Tecnológico de Monterrey, Campus Ciudad de México. Calle del Puente no. 222, oficinas 3, segundo piso. Col. Ejidos de Huipulco. Del. Tlalpan, c.p. 14380, México, D. F. Teléfono:(52) 5483-2259. Correo electrónico: arvaldes@itesm.mx

1. Introducción

Un elemento importante en el análisis de derivados y la administración financiera de riesgos es la determinación del precio y para ello, la estimación de la volatilidad de rendimientos de activos es fundamental. Sin embargo, la volatilidad es una variable no observable de manera directa, lo cual dificulta su estimación.

Tradicionalmente se han empleado los modelos de volatilidad estocástica y los modelos GARCH con relativo éxito para modelar el proceso de volatilidad, la desventaja que estos procedimientos conllevan es que se vuelven muy restrictivos en tanto sólo son válidos bajo supuestos específicos.

La estimación de la volatilidad podría facilitarse empleando datos disponibles de manera intradía de activos financieros. De esta manera, al emplear enfoques económicos o estadísticos, se eliminan los supuestos necesarios y como se cuenta con un registro de rendimientos para cada día, se pueden calcular las varianzas diarias sin incurrir en errores.

Merton (1980) nota que la varianza (sobre un intervalo fijo) de una variable aleatoria independiente e idénticamente distribuida (iid) puede estimarse como la suma de realizaciones al cuadrado. Empleando un ajuste en el tiempo continuo, Anderson, Bollerslev, Diebold y Labys (2001) han mostrado que, cuando el proceso sigue una semimartingala, la suma de los rendimientos al cuadrado lleva a estimaciones consistentes de la varianza del proceso del rendimiento.

Por otro lado, existen estudios empíricos que utilizan datos intradía para estimar la volatilidad diaria simplemente sumando los rendimientos intradía al cuadrado. A esta volatilidad se le conoce en la literatura como volatilidad realizada. Pero en este procedimiento se corre el riesgo de sobre o subestimar la volatilidad diaria promedio si solamente se suman los rendimientos al cuadrado y no se pone cuidado en la elección de la frecuencia de la serie.

El objetivo de este trabajo es entonces describir la volatilidad realizada como una medida de riesgo en finanzas y motivar su uso mediante su aplicación en el Índice de Precios y Cotizaciones de México (IPC). Para lo anterior, en las secciones dos y tres se define la volatilidad realizada así como su forma de cálculo y su aplicación en los datos y se presentan algunos hechos estilizados de la misma. En la sección cuatro se contrastan dos hechos estilizados, a saber la normalidad del logaritmo de la varianza empleada en los análisis de derivados y el supuesto de memoria larga. En la última parte, se propone un modelo para describir y predecir la volatilidad, así como las estimaciones y resultados de evaluación de desempeño.

2. Volatilidad Realizada

Una medida común para la volatilidad diaria es calcularla mediante los rendimientos diarios al cuadrado. Anderson *et. al.* (1998) afirman que es más exacto estimar la volatilidad realizada con los rendimientos intradía al cuadrado. Sea $P_{t,j}$ el precio de un activo en el periodo j del día t . La tasa de rendimiento, en porcentaje, se define como:

$$r_{t,j} = 100(\ln P_{t,j} - \ln P_{t,j-1}).$$

Se supondrá que:

$$\begin{aligned} E[r_{t,j}] &= 0 \\ E[r_{t,j}r_{s,i}] &= 0 \quad \forall t, s, j, i, \quad t \neq s, j \neq i, \\ E[r_{t,j}^2 r_{s,i}^2] &< \infty \quad \forall t, s, j, i. \end{aligned}$$

Es decir, los rendimientos tienen media cero y no están correlacionados; además, las varianzas y covarianzas de los rendimientos al cuadrado existen y son finitas. Con lo anterior, se define la volatilidad realizada al tiempo t a partir de datos históricos como

$$\hat{\sigma}_t^2 = \sum_{j=1}^N r_{t,j}^2,$$

donde N es el número de rendimientos intradía. El cálculo anterior mide sólo la volatilidad durante las horas en las que se pueden hacer transacciones. Para evitar el sesgo anterior Martens (2002) sugiere utilizar

$$\hat{\sigma}_t^2 = (1 + c) \sum_{j=1}^N r_{t,j}^2,$$

con c una constante positiva. Hol y Koopman (2002) sugieren que c sea la razón entre la varianza del rendimiento en las horas de transacción y la varianza del rendimiento en el periodo de cierre del día anterior y la apertura (*overnight*).

3. Datos

Los datos de nuestro estudio empírico consisten en precios de transacción del Índice de Precios y Cotizaciones del periodo del 4 de enero de 1999 al 31 de diciembre del 2002. De la base de datos original que incluye precios obtenidos en cada transacción, se construyó una serie de precios con intervalos de cinco minutos. Se tomó el precio de la última transacción antes del tiempo de cinco minutos. Posteriormente se calculó el rendimiento por cinco minutos de la siguiente manera

$$r_{t,j} = 100(\ln P_{t,j} - \ln P_{t,j-1}),$$

donde $r_{t,j}$ denota el rendimiento del periodo intradía j en un día t . La Bolsa Mexicana de Valores (BMV) abre a las 8:30 a.m. y cierra a las 3:00 p.m. por lo que un día consta de 78 rendimientos cada cinco minutos.

De acuerdo a la literatura, existen algunos hechos estilizados con respecto a la serie de la volatilidad realizada

- la distribución incondicional de rendimientos diarios es no sesgada pero exhibe exceso de curtosis.
- la distribución incondicional de la volatilidad realizada (calculada como la suma de rendimientos intradía al cuadrado a frecuencias entre 5 y 20 minutos dependiendo de los datos usados) y la varianza de la desviación estándar se acerca a la normal.
- el logaritmo de la volatilidad realizada muestra un grado de autocorrelación (positiva) alto que decrece en forma muy lenta.

- la volatilidad realizada no parece tener raíces unitarias pero existe evidencia de integración fraccional de orden 0.4.
- los rendimientos diarios estandarizados por la medida de volatilidad realizada son cercanos a una normal.

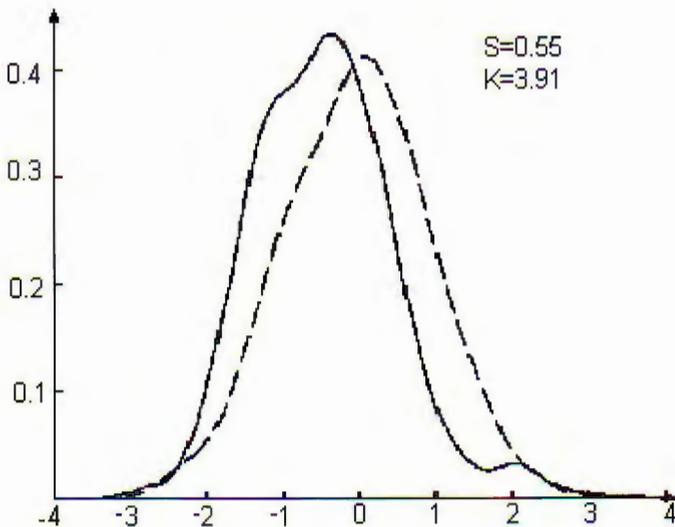
4. Propiedades de la Volatilidad Realizada

Se estudiarán algunas propiedades de la volatilidad realizada de los rendimientos del instrumento en estudio. Nos enfocamos principalmente en el logaritmo de la varianza y estudiamos su distribución así como su persistencia.

4.1 Distribución

En la Figura 1 se grafica la función de densidad del logaritmo de la varianza en línea continua, así como la función de densidad de una normal estándar que se presenta en líneas cortadas. El sesgo (S) y curtosis (K) se presentan en el lado superior derecho.

Figura 1. Logaritmo de la Varianza



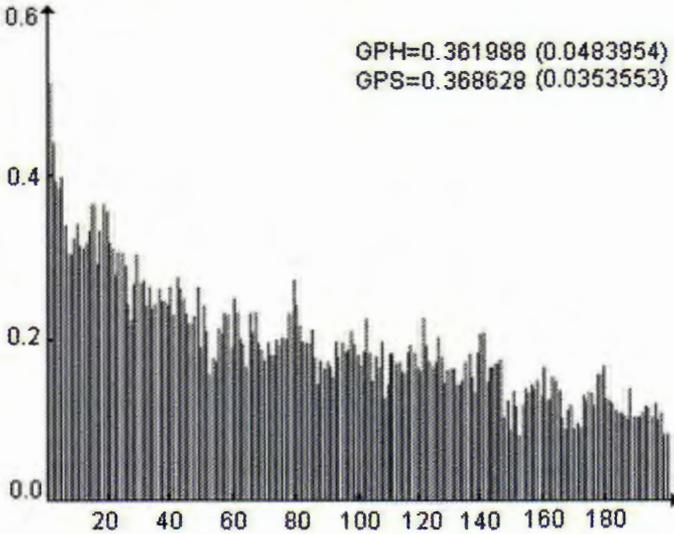
Esta distribución parece aproximarse a una normal, sin embargo las pruebas estándar la rechazan.

4.2 Memoria Larga

Un proceso estacionario en covarianza y fraccionalmente integrado (de orden d) tiene la propiedad de memoria ya que la autocorrelación disminuye a una tasa lenta cuando $0 < d < 0.5$. En la Figura 2 se presenta la función de autocorrelación muestral así como la estimación de d (orden de integración) mediante el estadístico de Geweke y Porter-Hudak (GPH, 1983), así como el estadístico de Robinson y Henry (GPS, 1998) en la parte superior derecha. Entre paréntesis se presenta el error estándar. La función de autocorrelación, así como la

estimación de d más dos veces su error estándar menor a 0.5, sugieren que la volatilidad realizada es estacionaria en covarianza y fraccionalmente integrada.

Figura 2. Función de Autocorrelación Muestral (FAC)



Lo anterior sugiere que la volatilidad se puede modelar con procesos fraccionalmente integrados como los modelos ARFIMA.

5. Modelos

En esta sección se emplearán los modelos tratados en este trabajo. En primer lugar se discutirán los modelos que estudian directamente la volatilidad realizada contrastándolos con los modelos que consideran a la volatilidad como segundo momento de los rendimientos. En segundo lugar se presentan los resultados empíricos.

5.1 Modelos

Consideramos un modelo de memoria larga ARFIMA para la volatilidad, tomando en cuenta los rendimientos positivos y negativos del día anterior, por lo que el modelo en este caso sería ARFIMAX(p, d, q) y se representa de la siguiente manera:

$$(1 - L)^d(1 - \beta(L_p))\ln(\sigma_t^2) = \omega_0 + \omega_1 r_{t-1}^- + \omega_2 r_{t-1}^+ + (1 + \alpha(L_q))\varepsilon_t,$$

donde $\varepsilon_t \sim i.i.dN(0, \sigma^2)$, $\alpha(L_q) = \sum_{i=1}^q \alpha_i L^i$ y $\beta(L_p) = \sum_{i=1}^p \beta_i L^i$. Los rendimientos con un signo (-) consideran el rendimiento cuando es negativo y cero en otro caso, de la misma manera con el positivo. Es similar a un modelo ARMA(p, q) con la diferencia que se le agrega los rendimientos y un parámetro de integración d que representa memoria larga.

Además de considerar el modelo anterior se estima el modelo tradicional GARCH(1,1) representado como

$$r_t = \varepsilon_t + \theta\varepsilon_{t-1},$$

$$\sigma_t^2 = \omega_0 + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2.$$

A este último modelo se le incluye la información de la volatilidad realizada, obteniendo un modelo GARCHX(1,1) expresado como

$$r_t = \varepsilon_t + \theta\varepsilon_{t-1},$$

$$\sigma_t^2 = \omega_0 + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\hat{\sigma}_{t-1}^2.$$

5.2 Resultados

Se estimaron los tres modelos con 800 de las observaciones que comprende del 4 de enero de 1999 al 12 de marzo de 2002, dejando las 202 observaciones restantes para evaluar los resultados.

En la Tabla 1 se reportan las estimaciones de los parámetros junto con los intervalos de confianza al 95% para los tres modelos. El Criterio de Información de Akaike (AIC) se calculó como $-2\ln L + 2p$, con p el número de parámetros estimado. Asimismo se presenta el estadístico de Ljung-Box para los errores observados y entre paréntesis el p -value del estadístico indicando que los modelos son adecuados.

Tabla 1

	ARFIMAX		GARCH(1,1)		GARCHX(1,1)
ω_0	-0.535 (-1.045, -0.024)	θ	0.201 (0.128, 0.275)	θ	0.194 (0.122, 0.267)
ω_1	-0.124 (-0.178, -0.070)	ω_0	0.048 (0.000, 0.097)	ω_0	0.065 (0.002, 0.128)
ω_2	0.097 (0.049, 0.145)	α	0.036 (0.016, 0.056)	α	0.019 (-0.004, 0.042)
β	-0.137 (-0.235, -0.039)	β	0.948 (0.919, 0.978)	β	0.935 (0.894, 0.976)
σ_ε^2	0.554			γ	0.069 (0.001, 0.137)
d	0.339 (0.273, 0.406)				
L	-898.59	L	-1591.585	ln L	-1586.84
AIC	1809.181	AIC	3191.17	AIC	3183.675
Q(12)	12.402(0.259)	Q(12)	11.403(0.410)	Q(12)	10.698(0.470)

Como criterios de evaluación de predicción, se considera a la volatilidad realizada (conocida) como criterio de evaluación y se realizan predicciones sobre los siguientes 202 días, con horizontes de predicción de uno y hasta cinco días.

Se reportan en la Tabla 2, los resultados de estimar por mínimos cuadrados ordinarios la ecuación

$$\text{Realizado}_t = a + b \text{ Pronosticado}_t + \varepsilon_t$$

y realizar la prueba de hipótesis conjunta de que la constante sea cero y la pendiente sea uno. Entre paréntesis se presenta el *p*-value. Como era de esperar el modelo ARFIMAX sale bien librado, ya que el modelo considera directamente el valor realizado.

Tabla 2

Modelo		Horizonte de predicción				
		N=1	N=2	N=3	N=4	N=5
ARFIMAX	a	0.05612	0.0620	0.0425	0.0405	0.0218
	b	1.0214	1.0004	1.0563	1.0416	1.1025
	H_0	1.5356	1.2533	1.5268	1.1672	1.3288
	$a=0, b=1$	(0.2179)	(0.2878)	(0.2197)	(0.3133)	(0.2671)
GARCH(1,1)	a	0.0403	1.1035	0.7192	0.3833	0.0774
	b	0.1860	-0.2916	-0.1168	0.0614	0.2149
	H_0	879.3866	1070.0105	379.6307	403.1286	429.288
	$a=0, b=1$	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
GARCHX	a	-0.1102	1.1814	1.2825	0.3479	0.1258
	b	0.2797	-0.3501	-0.4450	0.0862	0.2074
	H_0	694.4465	884.1489	377.0744	309.5583	347.536
	$a=0, b=1$	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

También se reportan (Tabla 3) dos estadísticos de error usados por Martens (2002) como son el error cuadrático medio ajustado por heterocedasticidad (HRMSE) y el error absoluto medio (HMAE) calculados como

$$\text{HRMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(1 - \frac{\text{Realizado}_t}{\text{Pronosticado}_t} \right)^2}$$

$$\text{HMAE} = \frac{1}{T} \sum_{t=1}^T \left| 1 - \frac{\text{Realizado}_t}{\text{Pronosticado}_t} \right|$$

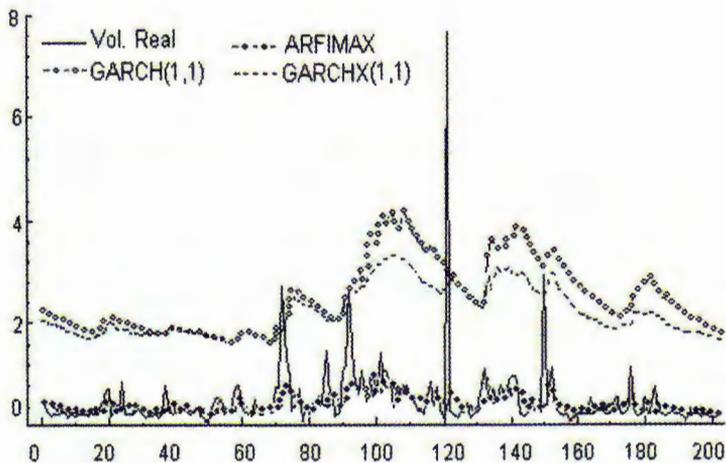
Tabla 3

Modelo		Horizonte de predicción				
		N=1	N=2	N=3	N=4	N=5
ARFIMAX	HRMSE	0.8195	0.8836	0.8646	0.8814	0.9027
	HMAE	0.5609	0.5633	0.5737	0.5635	0.5928
GARCH(1,1)	HRMSE	0.8300	0.8191	0.8077	0.8149	0.8118
	HMAE	0.8172	0.7805	0.7744	0.7834	0.7855
GARCHX	HRMSE	0.8187	0.8119	0.8037	0.8057	0.8026
	HMAE	0.8024	0.7716	0.7637	0.7682	0.7701

De acuerdo al estadístico HMAE el modelo ARFIMAX es el mejor. Es necesario señalar que de los dos modelos que toman la volatilidad como segundo momento de los rendimientos, GARCH y GARCHX, el último tiene mejores resultados.

Cabe hacer notar que los modelos GARCH sobreestiman la varianza de los rendimientos por lo que al compararlos con la volatilidad realizada son demasiado grandes como se muestra en la Figura 3, donde se grafican las predicciones para un horizonte de predicción de un día.

Figura 3



6. Conclusiones

En este trabajo se utilizaron los rendimientos al cuadrado con frecuencia de cinco minutos para estimar la volatilidad realizada diaria. Se estudiaron las

propiedades de esta última y se encontró que su distribución en logaritmos es sesgada y leptocúrtica contradiciendo este hecho estilizado. Se afirma el hecho de que es estacionaria en covarianza así como que es fraccionalmente integrada por lo que existe memoria larga.

Esto último nos lleva a la posibilidad de analizar la volatilidad realizada por medio de modelos ARFIMA. Se analizó este modelo y se encuentra buen ajuste. Utilizando horizontes de predicción de uno a cinco días se encontró buena precisión, que al compararla con los modelos GARCH es mejor. Estos últimos son modelos de naturaleza distinta que incluyen más supuestos, pero al incluir la información de la volatilidad realizada, a este tipo de modelos para construir los modelos GARCHX, se llega a un mejor ajuste.

Por último, hay que considerar que en la actualidad se cuenta con la tecnología y la información financiera necesarias para poder calcular la volatilidad realizada y desarrollar investigación al respecto ya sea para describir las propiedades teóricas y empíricas, así como diseñar distintos tipos de modelos o incluirla en metodologías conocidas.

Bibliografía

- Anderson, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens (2001). The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics*, 61, pp. 43-76.
- Anderson, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association*, 96, pp. 42-55.
- Anderson, T.G. and T. Bollerslev (1998). Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review*, 39, pp. 885-905.
- Areal, N.M. and S.J. Taylor (2002). The Realized Volatility of FTSE-100 Futures Prices. *Journal of Futures Markets*, 22, pp. 627-648.
- Blair, B.J., S. Poon, and S.J. Taylor (2001). Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High Frequency Returns. *Journal of Econometrics*, 105, pp. 5-26.
- Dacorogna, M., R. Gençay, U. Müller, R. Olsen, and O. Pictet (2001). An Introduction to High-Frequency Finance. Academic Press.
- Ebens, H. (1999). Realized Stock Volatility. Working Paper. John Hopkins University.
- Geweke, J. and S. Porter-Hudak (1983). The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, 4, pp. 221-238.
- Hol, E. and S. Koopman (2002). Stock Index Volatility Forecasting with High Frequency Data. Discussion Paper. Tinbergen Institute. Amsterdam-Rotterdam, The Netherlands.
- Martens, M. (2002). Measuring and Forecasting S&P 500 Index-Futures Volatility Using High-Frequency Data. *Journal of Futures Markets*, 22, pp. 497-518.
- Merton, R. C. (1980). On Estimating the Expected Return on the Market; an Explanatory Investigation. *Journal of Financial Economics*, 8, pp. 323-361.
- Robinson, P. M. and M. Henry (1998). Long and Short Memory Conditional Heteroscedasticity in Estimating the Memory Parameter of Levels. Discussion Paper. London School of Economics and Political Science.
- Taylor, S.J. and X. Xu (1997). The Incremental Volatility Information in One Million Foreign Exchange Quotations. *Journal of Empirical Finance*, 4, pp. 317-340.