LONG MEMORY IN VOLATILITY OF KOREAN STOCK MARKET RETURNS

Jihyun Lee, Tong Suk Kim, and Hoe-Kyung Lee

Graduate School of Management
Korea Advanced Institute of Science and Technology
207-43 Cheongryangri-dong, Dongdaemun-gu, Seoul 130-012, Korea

ABSTRACT

We examine the volatility process of the KOSPI200 returns using the FIGARCH (Fractionally Integrated GARCH) model which includes GARCH and IGARCH process as special cases. The robustness of this approach is investigated by considering aggregated data. To check possible spuriousness we selected weekly returns in the KOSPI and 10 individual stock returns in the KOSPI200 to see the effect of time and cross-sectional aggregation respectively. The results from this aggregation were compared with the original returns. We found long-term dependencies in the volatility process, which does not seem to be caused by spuriousness due to aggregation.

KEYWORDS

Long Memory; Fractionally Integrated GARCH; Volatility; Stock; KOSPI200

1. Introduction

The presence of long memory, or long-term dependencies can be defined in terms of the persistence of autocorrelations. The presence of long memory in a certain time series implies that there exist dependencies between distant observations. In contrast, if correlations among observations become negligible at long lags, the series is said to exhibit short memory. The autocorrelation function of a long memory process decays hyperbolically and eventually dies out. The autocorrelation function for an I(0) process, however, shows an exponential decay and for an I(1) process it shows an infinite persistence.

Research on the long memory processes arose from the examination of data in the physical sciences. The study conducted by Hurst[12] is a well-known example in which the persistence in streamflow data has been found using the R/S analysis. The R/S method was then applied to the analysis of returns on financial assets by Mandelbrot[15] and modified later by Lo[13] on controlling for the effect of short-term correlations.

Granger and Joyeux[10] proposed the fractionally integrated autoregressive moving average (ARFIMA) model to explain the long memory property that exists in the conditional mean. The ARFIMA model is obtained by extending the integration order of the conventional ARMA model to a non-integer value between 0 and 1. By applying this method, considerable analyses have been conducted on the data of stock prices, exchange rates, and futures prices.

Besides looking at the conditional mean of a time series, the long memory in the volatility process has also been investigated. Ding *et al.*[5]and Dacorogna *et al.*[4] showed that autocorrelation coefficients of the squared daily returns for stock prices and exchange rates decay very slowly. Motivated by the presence of apparent long-memory in the autocorrelations of squared or absolute returns of various financial asset prices, Baillie *et al.*[2] proposed the fractionally integrated generalized autoregressive conditional heteroskedasticity (FIGARCH) model by combining the fractionally integrated process for the mean with the regular GARCH process for the conditional variance. They found that the long memory property exists in the volatility of the Deutschmark-U.S.dollar exchange rate.

The purpose of this study is to examine the long memory property of the volatility of returns for the KOSPI200 which is one of the Korean Stock Market indices. The robustness of the results is also investigated by considering the aggregation of data. We examine the long memory property of weekly data extracted from original daily data and the property of individual stocks which are included in the KOSPI200.

The remainder of this paper is organized as follows. We briefly review the volatility process models in Section 2. The next section describes the estimation and testing procedures for long memory, and empirical results are discussed in Section 4. The final section provides a brief conclusion.

2. The FIGARCH model

Following Engle[7], the process $\{\varepsilon_t\}$ is said to follow an ARCH model if

$$\varepsilon_t = \sigma_t \, z_t \tag{1}$$

where $E_{t-1}(z_t)=0$ and $Var_{t-1}(z_t)=1$, and σ_t is measurable with respect to the time t-1 information set. Conditional variance of the ARCH(q) model can be written as a linear function of past squared values of the process,

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^{\ 2} + \Lambda + \alpha_q \varepsilon_{t-q}^{\ 2} \tag{2}$$

where $\alpha_0 > 0$ and $\alpha_1, \Lambda, \alpha_q \geq 0$. This model captures the tendency for volatility clustering, that is, for large (small) price changes to be followed by other large (small) price changes.

Bollerslev[3] extended the ARCH class of models to the generalized ARCH, or GARCH which has an autoregressive moving average form for the conditional variance σ_t^2. Conditional variance of the GARCH(p,q) model is expressed as

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^{\ 2} + \Lambda + \alpha_q \varepsilon_{t-q}^{\ 2} + \beta_1 \sigma_{t-1}^{\ 2} + \Lambda + \beta_p \sigma_{t-p}^{\ 2} \\ &= \omega + \alpha(L)\, \varepsilon_t^{\ 2} + \beta(L)\, \sigma_t^2 \end{aligned} \tag{3}$$

where $\omega > 0$ and $\alpha_1, \Lambda, \alpha_q, \beta_1, \Lambda, \beta_p \geq 0$. $\alpha(L)$ and $\beta(L)$ are the lag polynomials with orders of q and p respectively. That is, by adding a parameter p to the ARCH model, longer lags can be considered in the GARCH model with low orders. The GARCH(p,q) process in Eq.(3) can be rewritten as an ARMA process in $\varepsilon_t^{\ 2}$,

$$[1-\alpha(L)-\beta(L)]\,\varepsilon_t^{\ 2} = \omega + [1-\beta(L)]\, v_t \tag{4}$$

where $v_t \equiv \varepsilon_t^{\ 2} - \sigma_t^2$. The $\{v_t\}$ process is interpreted as the 'innovations' for the conditional variance. To ensure conditional variance to be nonnegative, it is assumed that all the roots of the polynomial $[1-\beta(L)]$ lie outside the unit circle.

To take account of a unit root in the autoregressive polynomial $[1-\alpha(L)-\beta(L)]$, Engle and Bollerslev[8] introduced the IGARCH model. The IGARCH(p,q) process is given by

$$\phi(L)\,(1-L)\,\varepsilon_t^{\ 2} = \omega + [1-\beta(L)]\, v_t \tag{5}$$

where $\phi(L) = [1-\alpha(L)-\beta(L)]\,(1-L)^{-1}$. In the IGARCH process, current information remains important for the forecast of the conditional variance for all horizons.

Similar to the ARFIMA process for the mean by Granger and Joyeux[10], the FIGARCH process can be obtained by extending Eq.(5). A FIGARCH process of order (p,d,q) is defined by

$$\phi(L)\,(1-L)^d\,\varepsilon_t^{\ 2} = \omega + [1-\beta(L)]\, v_t \tag{6}$$

where the parameter d is allowed to be any real number between 0 and 1 and $\phi(L) = 1 - \phi_1 L - \Lambda - \phi_q L^q$, and $\beta(L) = \beta_1 L + \Lambda + \beta_p L^p$. All the roots of $\phi(L)$ and $[1-\beta(L)]$ lie outside the unit circle. In Eq.(6), persistence of shocks to the conditional variance, or the degree of long-term dependencies is measured by the fractional differencing parameter d. From the fact that Eq.(6) is identical to the GARCH(p,q) model for d=0 and to the IGARCH(p,q) model for d=1, we can see that the FIGARCH process includes the GARCH and IGARCH processes as special cases.

3. Model Specification and Estimation

As the simplest case, we choose the FIGARCH(1,d,0) model to investigate the long memory property of volatility of stock index returns. Previous studies using the GARCH(p,q) model show that most financial time series are well modeled by the low order of (1,1). The model specification is the following:

$$y_t = \mu + \varepsilon_t \quad (8)$$

$$\varepsilon_t = \sigma_t z_t, \quad z_t \sim N(0,1), \quad \varepsilon_t / \psi_{t-1} \sim N(0, \sigma_t^2) \quad (9)$$

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1 - \beta(L)] v_t \quad (10)$$

In the previous three equations: y_t is the stock return at time t, defined by $ln(S_t/S_{t-1})$ and S_t is the stock price at t; ψ_{t-1} is the information set at time t-1; and, $\phi(L)$ and $\beta(L)$ are lag polynomials of order 0 and 1, respectively. Eq.(10) can then be rewritten as:

$$\sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + [1 - \beta_1 - (1-L)^d] \varepsilon_t^2 \quad (11)$$

The fractional differencing operator, $(1-L)^d$, is given by

$$(1-L)^d = \sum_{k=0}^{\infty} \Gamma(k-d) L^k / \Gamma(k+1)\Gamma(-d) \quad (12)$$

with $\Gamma(\bullet)$ denoting the gamma function. We truncate the lags of the fractional differencing operator at 1000, which we think is large enough to examine the long memory process.

Parameters of the above model are estimated by the maximum likelihood estimation. The corresponding log likelihood function to be maximized is,

$$\log L(\theta; y_1, \Lambda, y_T) = -\frac{1}{2} T \log(2\pi) - \frac{1}{2} \sum_{t=1}^{T} [\log(\sigma_t^2) - \frac{\varepsilon_t^2}{\sigma_t^2}] \quad (13)$$

where $\theta' = (\mu, \omega, \beta_1, \phi_1, d)$ and ϕ_1 is fixed to zero. And the standard error of $\hat{\theta}_T$ is from the following asymptotic distribution:

$$\hat{\theta}_T \sim N(\theta_0, T^{-1} A(\hat{\theta}_T)^{-1} B(\hat{\theta}_T) A(\hat{\theta}_T)^{-1}) \quad (14)$$

In the above equation, $A(\cdot)$ is the Hessian matrix and $B(\cdot)$ is the outer product of the gradient.

4. Empirical Results

In Table 1 we report results for the estimation of parameters of the FIGARCH(1,d,0) model as well as the GARCH(1,1) and IGARCH(1,1) models. Daily log returns for the KOSPI200 between January 1990 and September 1999 were used. The sample size is 2824.

The FIGARCH(1,d,0) model is tested against the short-term models; the GARCH(1,1) and IGARCH(1,1) models. These short-term models are nested in the FIGARCH model. It is assumed in the FIGARCH(1,d,1) model that d=0 for the GARCH(1,1) and that ϕ_1= 0 and d =1 for the IGARCH(1,1). In the GARCH(1,1) model ϕ_1 is estimated as 0.986, which is very close to unity. This means that the level of persistence of volatility shocks is very high with the value near 1 and coincides with previous works. Since we cannot reject the null hypothesis that ϕ_1 equals 1, the IGARCH(1,1) model is investigated. All the estimated parameters are

Table 1

FIGARCH models for daily returns for the KOSPI200.

(p,d,q)	GARCH(1,1) (1,0,1)	IGARCH(1,1) (1,1,0)	FIGARCH(1,*d*,0) (1,*d*,0)
μ	0.007	0.010	-0.004
	(0.024)	(0.023)	(0.023)
ω	0.048*	0.037*	0.106*
	(0.011)	(0.009)	(0.025)
β_1	0.860*	0.862*	0.356*
	(0.016)	(0.016)	(0.050)
ϕ_1	0.986	1.000	0.000
	(0.016)	---	---
d	0.000	0.000	0.447*
	---	---	(0.048)
Q(20)	80.995	63.98	66.8525
$Q^2(20)$	30.51	36.22	24.437

Note: Standard errors are reported in parentheses.
* indicates significant at 5% level.
Q(20) and $Q^2(20)$are Ljung-Box Q statistics of order 20 for the standardized residual, or $z_t = \varepsilon_t \sigma_t^{-1}$ and squared standardized residual respectively. Critical values at significant level of 5% are 30.14 for the FIGARCH(1,*d*,0) and 28.87 for others. The Q statistic is used to determine whether the estimated model is appropriate.

significant at the 5 percent level but the value of $Q^2(20)$ is 36.22. This is greater than the critical value of 30.14. Therefore, the IGARCH(1,1) model is not appropriate. For the FIGARCH(1,*d*,0) model, the estimated value of *d*, which measures the degree of persistence, is 0.477 and the null hypotheses of both d=0 and d=1cannot be accepted. This means that the true volatility model is neither the GARCH(1,1) nor the IGARCH(1,1) and that the FIGARCH(1,*d*,0) model is different from the conventional GARCH or IGARCH class of models. Hence we can see that the volatility of daily returns of the KOSPI200 has a long memory property.

Drost and Nijman[6] demonstrated that the data generated from an IGARCH process at high frequencies shows weak IGARCH behaviors at lower frequencies. Since our long memory model assumes conditional heteroskedasticity as in the IGARCH, we are concerned with how the time-aggregation of data during the same period affects the long memory properties. We created weekly data from the original daily data and then performed the same analysis conducted with the daily data. If the results had been significantly different from each other, we could have said that the revealed long memory property is affected by time-aggregation and, therefore, the result indicating a long memory property for the daily data might be a spurious one. Unfortunately, the number of weekly data observations from the KOSPI200 is 509, which is too small to analyze long-term behaviors. Hence we used the data for the KOSPI as a proxy which is expected to show similar behaviors with the KOSPI200.

Table 2 shows the estimated parameters of three volatility models for each data set. Daily and weekly returns for the KOSPI between January 1980 and September 1999 are used. The sample size is 5751 for daily data and 1021 for weekly data. The results obtained here show similarities with those for the KOSPI200. It is especially worthwhile to note that all the daily and weekly data appear to have long memory properties. The values of *d* for the daily and weekly data are 0.378 and 0.379 respectively, which are very close with each other. From the above, it can be inferred that the long memory property is not a result of the time-aggregation of data. And estimated parameters of the FIGARCH model for daily data for the KOSPI200 can be trusted.

Our next concern is another type of spuriousness resulting from cross-sectional aggregation. Robinson[17] and Granger[11] showed that aggregation of weakly dependent series can produce a strong dependent series. Analyzing the S&P500 index and 30 individual stocks between 1962 and 1994, Lobato and Savin[14] found that squared returns on the index follow a long memory process whereas 6 individual stocks among those examined do not. Based on this result, they pointed out the possibility of spurious long memory appearing in the index. Since the KOSPI200 is also a composite index calculated from individual stock prices, the long memory property found in the index may be spurious.

Table 2
FIGARCH models for daily and weekly returns for the KOSPI

(p,d,q)	GARCH(1,1) (1,0,1)		IGARCH(1,1) (1,1,0)		FIGARCH(1,d,0) (1,d,0)	
	Daily	Weekly	Daily	Weekly	Daily	Weekly
μ	0.025	0.185*	0.023	0.192*	0.023	0.172
	(0.019)	(0.090)	(0.021)	(0.089)	(0.017)	(0.090)
ω	0.059*	0.297*	0.051*	0.169*	0.059	0.904*
	(0.032)	(0.142)	(0.020)	(0.087)	(0.033)	(0.295)
β_1	0.797*	0.864*	0.790*	0.882*	0.230*	0.259*
	(0.032)	(0.035)	(0.036)	(0.035)	(0.063)	(0.095)
ϕ_1	0.976	0.982*	1.000	1.000	0.000	0.000
	(0.013)	(0.029)	---	---	----	----
d	0.000	0.000	0.000	0.000	0.378*	0.397*
	---	---	---	---	(0.053)	(0.082)
Q(20)	133.4	37.218	131.579	28.070	138.65	29.599
$Q^2(20)$	8.452	18.250	10.181	20.062	7.179	22.966

Note; Standard errors are reported in parentheses.
* indicates significant at 5% level.
Q(20) and $Q^2(20)$are Ljung-Box Q statistics of order 20 for the standardized residual, or $z_t = \varepsilon_t \sigma_t^{-1}$ and squared standardized residual respectively. Critical values at significant level of 5% are 30.14 for the FIGARCH(1,d,0) and 28.87 for others.

We examine individual stocks that are included in the KOSPI200. Ten stocks that have enough data for a long memory analysis and are random in the firm-size measured by the market value of equities at the end of 1998 are studied. Daily returns for the individual stocks between January 1980 and December 1998 were used. The sample size is 5564. The results of analysis are reported in Table 3.

For two stocks, Hyundai Motors and Kumho Tires, the FIGARCH(1,*d*,0) model appears inappropriate since the Q-statistics are greater than the critical value of 30.14. For other stocks the value of d ranges between 0.256 and 0.371. These figures are not far from the one obtained from the KOSPI200. So the long memory that appeared in the volatility of the index seems hardly a spurious one.

One thing we have to mention is that both the IGARCH and the FIGARCH models might appear to be appropriate. If we consider only the integer order of integration, the IGARCH model would be chosen. Explanations from various perspectives are possible for it. Though a considerable amount of studies of high frequency financial data with ARCH class of models have found considerable support for the integrated GARCH model, there is less theoretical motivation for truly integrated behavior in the conditional variance. In the IGARCH process, the occurrence of a shock to volatility persists for an infinite time horizon so that the pricing functions for long-term contracts are sensitive to the initial conditions. Engle and Mustafa[9] argued that the IGARCH model is not compatible with the persistence observed after large shocks such as on Black Monday in 1987. Also, there are issues of temporal aggregation. As pointed out in the work of Baillie and Bollerslev[1], while studies of daily asset returns data have found IGARCH behavior, studies with higher-frequency data over shorter time spans have often uncovered less persistence. Baille *et al.*[3] found out that the FIGARCH volatility process may be easily mistaken for IGARCH behavior. By simulations they estimated parameters of the IGARCH(1,1) model with data generated from the FIGARCH(1,*d*,0) volatility process. Among those, about 80 percent of the cases were judged to follow IGARCH volatility processes. Therefore, it seems that FIGARCH rather than IGARCH is closer to a true model for the volatility process of the KOSPI200 returns.

5. Conclusions

In this paper, we employed a FIGARCH model, which has a fractional integration order between 0 and 1

Table 3
FIGARCH(1,*d*,0) model for 10 individual stocks.

	Estimated parameters

Stock	μ	ω	β_1	d	Q(20)	$Q^2(20)$
Hyundai Motors	0.002 (0.025)	0.568* (0.067)	0.107* (0.028)	0.285* (0.022)	80.88	35.23
Samsung Corporation	-0.015 (0.025)	0.529* (0.067)	0.126* (0.030)	0.332* (0.026)	58.28	17.31
Kumho Tires	-0.049 (0.041)	1.842* (0.196)	0.090* (0.024)	0.272* (0.022)	77.20	81.91
Daesung Industry	0.012 (0.027)	0.757* (0.078)	0.180* (0.039)	0.350* (0.034)	48.32	1.32
Donga Construction	-0.078 (0.027)	0.814* (0.088)	0.078* (0.029)	0.305* (0.025)	49.85	3.25
Seotong	0.036 (0.028)	0.652* (0.084)	0.129* (0.027)	0.297* (0.021)	37.29	19.38
Se-A wires	-0.023 (0.025)	0.737* (0.081)	0.082* (0.028)	0.275 * (0.022)	42.01	12.43
Sungchang Enterprise	0.004 (0.035)	1.167* (0.140)	0.070* (0.024)	0.256* (0.017)	54.94	15.73
Sepoong	-0.021 (0.030)	0.648* (0.075)	0.218* (0.030)	0.371* (0.024)	73.50	16.00
Kohap	-0.035 (0.028)	0.672* (0.086)	0.149* (0.030)	0.312* (0.024)	33.89	14.03

Note; Standard errors are reported in parentheses.
* indicates significant at 5% level.
Q(20) and $Q^2(20)$are Ljung-Box Q statistics of order 20 for the standardized residual, or $z_t = \varepsilon_t \sigma_t^{-1}$ and squared standardized residual respectively. Critical values at significant level of 5% are 30.14 for the FIGARCH(1,d,0) and 28.87 for others.

as an alternative to the knife-edge classification between I(0) and I(1), in order to examine long memory properties of the volatility process for the KOSPI200 returns. We found out that there exists a long memory property in the index. The results obtained here are robust in that these are not affected by time and cross-sectional aggregation of data.

We only estimated parameters of the FIGARCH model to see the existence of long-term dependencies. However, to be meaningful, performance evaluation is needed to see how much it improves the performance of forecasting compared with conventional models by introducing the idea of long memory into the model. Also, it will be helpful to extend the idea of fractional integration order exploited in the FIGARCH model to the exponential GARCH(EGARCH), which is known to incorporate characteristics of stock prices very well, in order to understand more about long memory properties.

REFERENCES

1. Baillie, R. T. and Bollerslev, T. (1991). Intra day and inter market volatility in foreign exchange rates. Review of Economic Studies, 58, 565-585
2. Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 74, 3-30
3. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31, 307-327
4. Dacorogna, M. M., Muller, U. A., Nagler, R. J., Olsen, R. B., and Pictet, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. Journal of International Money and Finance, 12, 413-438
5. Ding, Z., Granger, C. W. J., and Engle, R. F. (1993). A long memory property of stock returns and a new model. Journal of Empirical Finance, 1, 83-106
6. Drost, F. C. and Nijman, T. (1993). Temporal aggregation of GARCH processes. Econometrica, 61, 909-927
7. Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K.

inflation. Journal of Econometrics, 50, 987-1008

8. Engle, R. F. and Bollerslev, T. (1986). Modeling the persistence of conditional variance. Journal of Econometrics Reviews, 5, 1-50
9. Engle, R. F. and Mustafa, C. (1992), Implied ARCH models from options prices. Journal of Econometrics, 52, 289-311
10. Granger, C. W. J. and Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. Journal of Time Series Analysis, 1, 15-39
11. Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. Journal of Econometrics, 14, 227-238
12. Hurst, H. E. (1951). Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers, 116, 770-799
13. Lo, A. W. (1991). Long-term memory in stock market prices. Journal of Econometrics, 59, 1279-1313
14. Lobato, I. and Savin, N. E. (1997). Real and spurious long memory properties of stock market data. Journal of Business and Economic Statistics, 16, 261-279
15. Mandelbrot, B. B. (1971). When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models. Review of Economics and Statistics, 53, 225-236
16. Mandelbrot, B. B. (1972). Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. Annals of Economic and Social Measurement, 1, 259-290
17. Robinson, P. M. (1978). Statistical inference for a random coefficient autoregressive model. Scandinavian Journal of Statistics, 5, 163-168