Table of Contents

I. Introduction

Following the seminal work of Mandelbrot (1963) and Fama (1965), many researchers have found that the empirical distribution of stock returns is significantly non-normal, such as Hsu *et al.* (1974), Hagerman (1978), Lau et al. (1990), Kim and Kon (1994). They found that (1) the kurtosis of the stock returns time series is obviously larger than the kurtosis of the normal distribution, in other words, the time series of stock returns are leptokurtic; (2) the distribution of stock returns is skewed, either to the right (positive skewness) or to the left (negative skewness); (3) the variance of the stock returns is not constant over time or the volatility is clustering. Some researchers regarded this as the persistency of the stock market volatility and the financial analyst called this uncertainty or risk. This uncertainty is crucially important in modern financial theory. Before the seminal paper by Engle (1982), the uncertainty of speculative prices, changing over time (Mandelbrot, 1963; Fama, 1965) measured by the variances and covariance has been accepted for decades.

Many conventional time series and econometric models work only if the variance is constant. Until lately, the financial and economic researchers have started modeling time variation in second- or higher-order moments. Engle (1982) has characterized the changing variances using the Autoregressive Conditional Heteroskedasticity (ARCH) model and its extensions as well as its modifications. Since then, hundreds of researchers have applied these models to financial time series data. In many applications, the linear ARCH (p) model requires a long length of p. The alternative and more flexible lag structure is the generalized ARCH (GARCH) introduced by Bollerslev (1986). It is proven that a small lag as GARCH (1, 1) is sufficient to model the variance changing over long sample periods (French *et al.*, 1987; Franses and Van Dijk, 1996). According to the paper of Choo Wei Chong (1999), who studies the performance of the GARCH model by using the rate of returns from daily stock market indices of the Kuala Lumpur Stock Exchange (KLSE) including Composite Index, Tins Index, Plantation Index, Properties Index and Finance Index, a very high order ARCH model is needed to model the heteroskedasticity. The basic ARCH (p) model is short-memory process in that only the most recent p-squared residuals are used to estimate the changing variance. Contrary to the short-memory ARCH (p), the GARCH model allows long-memory processes, which use all the past squared residuals to estimate the current variance. The statistical criteria that he utilizes, Q-statistic and LM test, suggest the use of the GARCH model instead of the ARCH model. The efficiency of a small lag in describing daily data of stock return time series is the basic reason for the selection of the GARCH (1, 1) in our study.

Even though the vast majority of earlier studies relied on the GARCH framework, there is recently a large and diverse time series literature on volatility modeling. Almost universally, reported results towards a very high degree of intemporal volatility persistence. In spite of highly significant in-sample parameter estimates, numerous studies find that standard volatility models explain little of the variability in ex-post squared returns. This has led to the suggestion that these models may be of limited practical value.

West, Edison and Cho (1993) evaluate the out-of sample forecasting performance of some univariate models for exchange rate volatility, using bilateral weekly data for the dollar versus the currencies of other five countries, from 1973-1989. They suggest that in the first three weeks horizon of this period, GARCH (1, 1) model seems to do a poorer job, judging by a mean squared prediction error criterion. During the next four weeks, it is hard to say which of the models perform best. Philip Hans Franses and Dick van Dijk report the skeptics that follow the forecasting performance of the GARCH models. According to the findings from forecast competitions, GARCH models seem to provide seemingly poor volatility forecasts and explain only little of the variability of asset returns in the sense that the MSPE (or any other measure of forecast accuracy) is very large. Kenneth D. West (1994) compares the out-of sample forecasting performance of univariate homoskedastic, GARCH, autoregressive and nonparametric models for condition variances, using five bilateral weekly exchange rates for the dollar versus currencies of other countries, for the period 1973-1989. He concludes to the fact that the GARCH models tend to provide slightly more accurate forecasts, but with disappointment he reports the results for longer time periods, where it is not so obvious which model performs best.

The purpose of this study is to examine the forecasting performance of the GARCH (1, 1) model in an attempt to answer to the findings of several forecasts competitions that present the GARCH models as poor forecast predictors. We compare the forecast accuracy of the GARCH (1, 1) model with that of a homoskedastic one, by using as statistical criterion the mean squared prediction error. We utilize bilateral daily data for the dollar versus the currencies of other ten countries and bilateral daily data of stock market returns for ten financial markets. We compare the out of sample performance realization of the squared of the daily change in an exchange of stock return rate with the value predicted by a model of the conditional variance

The results reported by our empirical application are slightly confusing. At a one-day till one-week horizon, it seems that for some time series the GARCH model has a slight edge over the homoskedastic one. For longer periods, it is not apparent which model performs best. On the other hand, there are series where the forecasting performance of the homoskedastic model appears to be superior for the entire time period.

Motivated by these findings and due to the existing skeptics, we decide to investigate the forecasts of the GARCH (1, 1) model in juxtaposition with those of a homoskedastic one, through a Monte Carlo experiment. The results are of great interest; the basic advantage of the Monte Carlo experiments is the availability of the true conditional volatility, h_{t+j}. Its utilization in the computation of the squared prediction error, instead of an estimator such as the squared shock, provides us with a direct measure for judging which of the two models performs better. We conclude to the fact that the seemingly poor forecasting accuracy of the GARCH (1, 1) model is not attributed to its inequality to perform good forecasts for the true volatility, a necessary and significant measure of financial risk. The Monte Carlo results lead us to the conclusion that the problem is pinpointed in the statistical criterion that we utilize.

Section II describes the autoregressive conditional heteroskedasticity model (ARCH) and its generalization (GARCH). Section III gives the necessary conditions for stationarity in the GARCH (1, 1) model and Section IV presents a survey of application and extensions. In Section IV we describe data and methodology (a) and the results of the empirical application (b). Section VI includes the Monte Carlo experiment and section VII its results. Finally section VIII concludes.

II. The ARCH, GARCH And IGARCH Models

Given the importance of predicting volatility in many asset pricing and portfolio management problems, many approaches of forecasting volatility have been proposed in the literature. The most popular one is the class of autoregressive conditional heteroskedasticity (ARCH) models originally introduced by Engle (1982). In a survey by Bollerslev *et al.* (1992) more than 200 papers are cited applying ARCH and related models to financial time series.

Let R_t be the rate of return of a particular stock from time t-1 to time t. Also, let F_{t-1} be the past information set containing the realized values of all relevant variables up to time t-1. Since investors know the information in F_{t-1} when they take their investment decision at time t-1, the relevant expected return and volatility to the investors are the conditional expected value of R_t, given F_{t-1}, and the conditional variance of R_t, given F_{t-1}. We denote these by m_t, and h_t respectively. That is, $m_t \equiv E(R_t \backslash F_{t-1})$ and $h_t \equiv Var(R_t \backslash F_{t-1})$. Given these definitions, the unexpected return at time t is $e_t \equiv y_t - m_t$. We treat e_t as a collective measure of news at time t. A positive e_t (an unexpected increase in price) suggests the arrival of good news, while a negative e_t (an unexpected decrease in price) suggests the arrival of bad news. Further, a large value of $|e_{t-1}|$ implies that the news is "significant" or "big" in the sense that it produces a large unexpected change in price.

We assume that the conditional variance of R_t and e_t varies over time, so the e_t is conditionally heteroskedastic. A convenient way to express this in general is

$$e_t = z_t \sqrt{h_t} \qquad (1)$$

where z_t is the dependent and the identically distribution with zero mean and unit variance. For convenience, we assume that z_t has a standard normal distribution.

From (1) and the properties of z_t it follows immediately that the distribution of e_t conditional upon the history F_{t-1} is normal with mean zero and variance h_t. also note that the unconditional variance of e_t is still assumed to be constant. By using the law of iterated expectations,

$$\sigma^2 = E[e_t^2] = E[E[e_t^2 / F_{t-1}]] = E[h_t] \qquad (2)$$

Hence, we assume that the unconditional expectation of h_t is constant. To complete the model, we need to specify how the conditional variance of e_t evolves over time.

Engle (1982) introduced the class of the Autoregressive Conditionally heteroskedastic (ARCH) models to capture the volatility clustering of financial time series (even though the first empirical applications did not deal with high frequency financial data). In the basic ARCH model, the conditional variance of the stock that occurs at time t is a linear function of the squares of past shocks. For example, in the ARCH model of order 1, h_t is specified as

$$h_t = \alpha_o + \alpha_1 e_{t-1}^2 \qquad (3)$$

Obviously, the conditional variance needs to be nonnegative. In order to guarantee that this is the case for the ARCH (1) model, the parameters in (3) have to satisfy the conditions $\alpha_o > 0$ and $\alpha_1 \geq 0$. Where $\alpha_1 = 0$, the conditional variance is constant, hence the series e_t is conditionally homoskedastic.

To understand why the ARCH model can describe the volatility clustering, observe that model (1) with (3) basically states that the conditional variance of e_t is an increasing function of the square of the shock that occurred in the previous time period. Therefore, if e_{t-1} is large (in absolute value), e_t is expected to be large (in absolute value) as well. In other words, large (small) shocks tend to be followed by large (small) shocks, of either sign.

An alternative way to see the same thing is to note that the ARCH (1) model can be written as an AR (1) model for e_t^2. Adding e_t^2 to (3) and subtracting h_t from both sides gives

$$e_t^2 = \alpha_o + \alpha_1 e_{t-1}^2 + v_t \tag{4}$$

where $v_t = e_t^2 - h_t = h_t(z_t^2 - 1)$. Notice that $E[v_t/F_{t-1}] = 0$. Using the theory for AR models, it follows that (4) is covariance-stationary if $\alpha_1 < 1$. In that case the conditional mean of e_t^2, or the conditional variance of e_t, can be obtained as

$$\sigma^2 = E[e_t^2] = \alpha_o/1-\alpha_1 \tag{5}$$

furthermore, (5) can be rewritten as

$$\begin{aligned} e_t^2 &= (1-a_1)\frac{a_o}{1-a_1} + a_1 e_{t-1}^2 + v_t \\ &= (1-a_1)\sigma^2 + a_1 e_{t-1}^2 + v_t \\ &= \sigma^2 + a_1(e_{t-1}^2 - \sigma^2) + v_t \end{aligned} \tag{6}$$

Assuming that $0 \leq \alpha_1 < 1$, (6) shows that if e_{t-1}^2 is larger (smaller) than its unconditional expected value σ^2, e_t^2 is expected to be larger (smaller) than σ^2 as well.

The ARCH model cannot only capture the volatility clustering of financial data, but their excess kurtosis. From (1) it can be seen that the kurtosis of e_t always exceeds the kurtosis of z_t,

$$E[e_t^4] = E[z_t^4]E[h_t] \geq E[z_t^4]E[h_t]^2 = E[z_t^4]E[e_t^2]^2 \tag{7}$$

Which follows from Jensen's inequality. As shown by Engle (1982), for the ARCH (1) model with normally distributed z_t the kurtosis of e_t is equal to

$$K_e = \frac{E[e_t^4]}{E[e_t^2]^2} = \frac{3(1-a_1^2)}{1-3a_1^2} > 3 \tag{8}$$

which is finite if $3a_1^2 < 1$. Clearly, K_e is always larger than the normal value of 3.

Another characteristic of the ARCH (1) model, which is worthwhile noting, is the implied autocorrelation function for the squared shocks e_t^2. From the AR (1) representation in (4), it follows that the *k*th order autocorrelation of e_t^2 is equal to a_1^k. The small first order autocorrelation would imply a small value of α_1 in the ARCH (1) model, but this in turn would imply that the autocorrelations would become close to zero quite quickly. Thus it appears that the ARCH (1) model cannot describe the two characteristic features of the empirical autocorrelations of the return series simultaneously.

To cope with the extended persistence of the empirical autocorrelation function, one may consider generalizations of the ARCH (1) model. One possibility to allow for more persistent autocorrelations is to include additional lagged squared shocks in the conditional variance function. The general ARCH (p) model is given by

$$h_t = a_o + a_1 e_{t-1}^2 + a_2 e_{t-2}^2 + ... + a_p e_{t-p}^2 \qquad (9)$$

to guarantee nonnegativeness of the conditional variance, it is required that $\alpha_o > 0$ and $\alpha_i \geq 0$ for all i = 1,…, p. The ARCH (p) model can be written as an AR (p) model for e_t^2 in exactly the same fashion as writing (3) as (4), that is,

$$e_t^2 = a_o + a_1 e_{t-1}^2 + ... + a_p e_{t-p}^2 \qquad (10)$$

It follows that the unconditional variance of e_t is equal to

$$\sigma^2 = \frac{a_o}{1 - a_1 - a_2 - ... - a_p} \qquad (11)$$

while the ARCH (p) model is covariance stationary if all roots of the lag polynomial $1-\alpha_1 L-…-\alpha_p L^p$ are outside the unit circle.

To capture the dynamic patterns in conditional volatility adequately by means of an ARCH (p) model, p is often needs to be taken quite large. It turns out that it can be quite cumbersome to estimate the parameters in such a model, because of the nononegative and stationarity conditions that need to be imposed. To reduce the computational problems, it is common to impose some structure on the parameters in the ARCH (p) model, such as $\alpha_i = \alpha(p+1-i)/(p(p+1)/2)$, i = 1,…, p, which implies that the parameters of the lagged squared shocks decline linearly and sum to α (see Engle, 1982, 1983). As an alternative solution, Bollerslev (1986) suggested adding lagged conditional variances to the ARCH model instead. For example, adding h_{t-1} to the ARCH (1) model results in the Generalized ARCH (GARCH) model of order (1, 1)

$$h_t = a_o + a_1 e_{t-1}^2 + \beta_1 h_{t-1} \qquad (12)$$

The parameters in this model should satisfy $\alpha_o > 0$, $\alpha_1 \geq 0$ and $\beta_1 \geq 0$ to guarantee that $h_t \geq 0$, while α_1 must be strictly positive for β_1 to be identified. To see why the lagged conditional variance avoids the necessity of adding many lagged squared residual terms to the model, notice that (12) can be rewritten as

$$h_t = a_o + a_1 e_{t-1}^2 + \beta_1 (a_o + a_1 e_{t-2}^2 + \beta_1 h_{t-2}) \quad (13)$$

or, by continuing the recursive substitution, as

$$h_t = \sum_{i=1}^{\infty} \beta_1^i a_o + a_1 \sum_{i=1}^{\infty} \beta_1^{i-1} e_{t-i}^2 \quad (14)$$

this show that the GARCH (1, 1) model corresponds to an ARCH (∞) model with a particular structure for the parameters of the lagged e_t^2 terms.

Alternatively, by adding e_t^2 to both sides of (12) and moving h_t to the right hand side, the GARCH (1, 1) model can be written as an ARMA (1, 1) model for e_t^2 as

$$e_t^2 = a_o + (a_1 + \beta_1) e_{t-1}^2 + v_t - \beta_1 v_{t-1} \quad (15)$$

where again v_t = e_t^2 - h_t. Using the theory for ARMA models, it follows that the GARCH (1, 1) model is covariance stationary if and only if $a_1 + \beta_1 < 1$. In that case the unconditional mean of e_t^2 -or equivalent the unconditional variance of e_t- is equal to

$$\sigma^2 = \frac{a_o}{1 - a_1 - \beta_1} \quad (16)$$

The ARMA representation in (15) also makes clear why α_1 needs to be strictly positive for identification of β_1. If $\alpha_1 = 0$, the AR and MA polynomials both are equal to 1-β_1L. Rewriting the ARMA (1, 1) model for e_t^2 as an MA (∞), these polynomials cancel out,

$$e_t^2 = \frac{1 - \beta_1 L}{1 - \beta_1 L} v_t = v_t \quad (17)$$

which shows that β_1 then is not identified.

As shown by Bollerslev (1986), the unconditional fourth moment of e_t is finite if $(a_1 + \beta_1)^2 + 2a_1^2 < 1$. If in addition the z_t are assumed to be normally distributed, the kurtosis of e_t is given by

$$K_e = \frac{3[1 - (a_1 + \beta_1)^2]}{1 - (a_1 + \beta_1)^2 - 2a_1^2} \quad (18)$$

which again is always larger than the normal value of 3. Notice that if β_1 = 0, (18) reduces to (8).

The autocorrelations are derived in Bollerslev (1988) and are found to be

$$\rho_1 = a_1 + \frac{a_1^2 \beta_1}{1 - 2a_1\beta_1 - \beta_1^2} \tag{19}$$

$$\rho_k = (a_1 + \beta_1)^{k-1} \rho_1 \text{ for k = 2,3,…} \tag{20}$$

even though the autocorrelations still decline exponentially, the decay factor in this case is $a_1 + \beta_1$. If the sum is close to one the autocorrelations will decrease only very gradually. When the fourth moment of e_t is not defined, the autocorrelations of e_t^2 are time-varying. Of course, one can still compute the sample autocorrelations in this case. As shown by Ding and Granger (1996), if $a_1 + \beta_1 < 1$ and $(a_1 + \beta_1)^2 + 2a_1^2 \geq 1$, such that the GARCH (1, 1) model is covariance stationary but with infinite fourth moment, the autocorrelations of e_t^2 behave approximately as

$$\rho_1 \approx a_1 + \beta_1/3 \tag{21}$$

$$\rho_k \approx (a_1 + \beta_1)^{k-1} \rho_1 \text{, for k=2,3,…} \tag{22}$$

The parameter restriction $(a_1 + \beta_1)^2 + 2a_1^2 = 1$ is equivalent to $1 - 2a_1\beta_1 - \beta_1^2 = 3a_1^2$, from which it follows that (21) is identical to (19) where this restriction is satisfied. Therefore, the autocorrelation of e_t^2 can be considered as continuous functions of a_1 and β_1, in the sense that their behavior does not suddenly change when these parameters take values for which the condition for existence of the fourth moment is no longer satisfied.

The general GARCH (p,q) model is given by

$$\begin{aligned} h_t &= \alpha_o + \sum_{i=1}^{p} a_i e_{t-1}^2 + \sum_{i=1}^{q} \beta_i h_{t-i} \\ &= \alpha_o + \alpha(L)\, e_t^2 + \beta(L) h_t \end{aligned} \tag{23}$$

where $\alpha(L) = \alpha_1 L + \ldots + \alpha_p L^p$ and $\beta(L) = \beta_1 L + \ldots + \beta_q L^q$. Assuming that all the roots of $1-\beta(L)$ are outside the unit circle, the model can be rewritten as an infinite order ARCH model

$$\begin{aligned} h_t &= \frac{a_o}{1 - \beta(L)} + \frac{a(L)}{1 - \beta(L)} e_t^2 \\ &= \frac{a_o}{1 - \beta_1 - \ldots - \beta_q} + \sum_{i=1}^{\infty} \delta_i e_{t-i}^2 \end{aligned} \tag{24}$$

For nonnegativeness of the conditional variance it is required that all δ_i in (24) are nonnegative. Nelson and Cao (1992) discuss the conditions this implies for the parameters α_i, i = 1, …, p and β_i, i = 1, …, q, in the original model (23).

Alternatively, the GARCH (p, q) can be interpreted as an ARMA (m, q) model for e_t^2 given by

$$e_t^2 = a_o + \sum_{i=1}^{m}(a_i + \beta_i)e_{t-i}^2 - \sum_{i=1}^{q}\beta_i v_{t-i} + v_t \tag{25}$$

where m= max (p, q), $\alpha_i = 0$, for i > p and $\beta_i = 0$ for i > q. It follows that the GARCH (p, q) model is covariance-stationary if all the roots of 1-α (L) –β (L) are outside the unit circle.

To determine the appropriate orders p and q in the GARCH (p, q) model, one can use a general- to-specific procedure by starting with a model with p and q set equal to large values, and testing down using likelihood-ratio-type restrictions (see Akgiray, 1989; Cao and Tsay, 1992). Alternatively, one can use modified information criteria, as suggested by Brooks and Burke (1997, 1998). Even though the general GARCH (p, q) model might be of theoretical interest, the GARCH (1, 1) often appears adequate in practice (see also Bollerslev, Chou and Kroner, 1992).

In applications of the GARCH (1, 1) model (12) to high-frequency financial time series, it is often found that the estimates of α_1 and β_1 are such that their sum is close or equal to one. Following Engle and Bolerslev (1986), the model that results when $\alpha_1 + \beta_1 = 1$ is commonly referred to as Integrated GARCH (IGARCH). The reason for this is that the restriction $\alpha_1 + \beta_1 = 1$ implies a unit root in the ARMA (1, 1) model for e_t^2 given in (15), which then can be written as

$$(1-L)e_t^2 = a_o + v_t - \beta_1 v_{t-1} \tag{26}$$

The analogy with a unit root in an ARMA model for the conditional mean of a time series is however rather subtle. For example, from (16) it is seen that the unconditional variance of e_t is not finite in this case. Therefore, the IGARCH model is not covariance-stationary. However, the IGARCH (1, 1) model may still be strictly stationary, as shown by Nelson (1990). This can be illustrated by rewriting (12) as

$$\begin{aligned} h_t &= a_o + (a_1 z_{t-1}^2 + \beta_1)h_{t-1} \\ &= a_o + (a_1 z_{t-1}^2 + \beta_1)(a_o + (a_1 z_{t-2}^2 + \beta_1)h_{t-2}) \\ &= a_o(1 + (a_1 z_{t-1}^2 + \beta_1)) + (a_1 z_{t-1}^2 + \beta_1)(a_1 z_{t-2}^2 + \beta_1)h_{t-2} \end{aligned} \tag{27}$$

and continuing the substitution for h_{t-i}, it follows that

$$h_t = a_o(1 + \sum_{i=1}^{t-1}\prod_{j=1}^{i}(a_1 z_{t-j}^2 + \beta_1)) + \prod_{i=1}^{t}(a_1 z_{t-i}^2 + \beta_1)h_o \tag{28}$$

as shown by Nelson (1990), a necessary condition for the strict stationary of the GARCH (1, 1) model is $E[\ln(a_1 z_{t-i}^2 + \beta_1)] < 0$. If this condition is satisfied, the impact of h_o disappears asymptotically. As expected, the autocorrelations of e_t^2 for an IGARCH model are not defined properly. However, Ding and Granger (1996) show that the approximate autocorrelations are given by

$$\rho_k = \frac{1}{3}(1+2a)(1+2a^2)^{-k/2} \tag{29}$$

Hence, the autocorrelations still decay exponentially. This is in sharp contrast with the autocorrelations for a random walk model, for which the autocorrelation are approximately equal to 1.

III. Stationarity In The GARCH (1, 1) Model

This section establishes necessary and sufficient conditions for the stationarity and of the GARCH (1, 1) model. We have already defined the GARCH (1, 1) model as:

$$\begin{aligned} &e_t = z_t\sqrt{h_t}, \\ &z_t \sim iid, non\deg enerate, P[-\infty < z_t < \infty] = 1 \\ &h_t = \alpha_o + \alpha_1 e_{t-1}^2 + \beta h_{t-1} \end{aligned} \tag{30}$$

where $\alpha_o \geq 0$, $\alpha_1 > 0$, $\beta \geq 0$. In most papers using GARCH (1, 1), a further restriction has been placed on $\{z_t\}$, namely that

$$E[z_t] = 0, E[z_t^2] = 1 \tag{31}$$

Under restriction (31), h_t is the conditional variance of e_t, given the history of the system. If we assume $E[z_t^2]=1$ but allow $E[z_t] \neq 0$, then h_t is the conditional moment of z_t. If we allow the second moment of z_t to be infinite or undefined, then h_t is the conditional scale parameter. Since the restrictions play no role in the main results of this paper, we adopt the less stringent condition,

$$z_t \sim iid, non\deg enerate, P[-\infty < z_t < \infty] = 1 \tag{32}$$

along with the requirement that

$$E[\ln(\beta+\alpha_1 z_t^2)] \tag{33}$$

exists. Note that (33) does not require that $E[\ln(\beta+\alpha_1 z_t^2)]$ be finite, only that the expectations of the positive and negative parts of $\ln(\beta+\alpha_1 z_t^2)$ are not both infinite. Relation (33) holds trivially for $\beta > 0$.

We also define as the conditional model the $\{h_t, e_t\}_{t=0,\infty}$ and as the unconditional model the process $\{{}_u h_t, {}_u e_t\}_{t=-\infty,\infty.}$ If we denote as B the Borel sets on $[0, \infty)$, we define as μ_t, the probability measure for h_t : $\mu_t(\Gamma)=P[h_t \in \Gamma]$, $\Gamma \in B$.
The results concerning the stationarity of the GARCH (1, 1) model rely on the relation between the coefficients α_1, β of the model. So different combinations of α_1, β provide different stationarity characteristics and different moment results for the unconditional process. The analysis below comes directly from Nelson (1990).

Theorem 2. Let $\alpha_o > 0$. If $E[\ln(\beta+\alpha_1 z_t^2)]<0$ then:

$$\alpha_o/(1-\beta) \leq {}_u h_t < \infty \text{ for all t a.s ,} \tag{34}$$

and ${}_u h_t$ is strictly stationary and ergodic with a well-defined probability measure

$$\mu_\infty \text{ on } [\alpha_o/(1-\beta), \infty]\ .\ t, \tag{35}$$

$$ {}_u h_t - h_t\ 0 \text{ a.s,} \tag{36}$$

$$\mu_t \rightarrow \mu_\infty \text{ and} \tag{37}$$

$$\mu_\infty \text{ is nondegenerate} \tag{38}$$

Corollary of theorem 3. Let $\alpha_o > 0$, $p>0$ and $E[\ln(\beta+\alpha_1 z_t^2)] <0$.

$$E[h_t^{-2p}] < \infty \quad t \geq 1 \tag{39}$$

$$E[{}_u h_t^{-2p}] < \infty \text{ for all t} \tag{40}$$

$$E[h_t^{2p}] < \infty \text{ iff } E[h_o^{2p}] < \infty \text{ and } E[(\beta+\alpha_1 z_t^2)^p] < \infty \tag{41}$$

$$E[{}_u h_t^{2p}] < \infty \quad \text{iff } E[(\beta+\alpha_1 z_t^2)^p] <1 \tag{42}$$

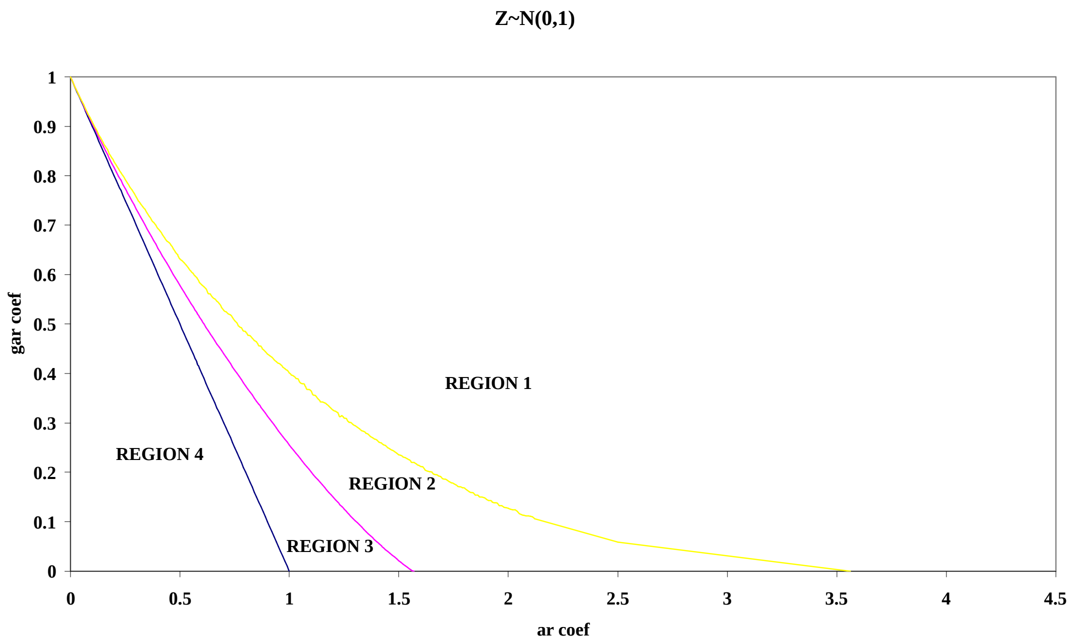
$$\limsup E[h_t^{2p}] < \infty \text{ iff } E[h_o^{2p}] < \infty \text{ and } E[(\beta+\alpha_1 z_t^2)^p] <1 \tag{43}$$

$$\lim E[h_t^{2p}] = E[{}_u h_t^{2p}] \text{ if } E[h_o^{2p}] < \infty \tag{44}$$

Theorem 4. (a) Let $\alpha_o > 0$ and $E[\ln(\beta+\alpha_1 z_t^2)] <0$. If $E[|z_t|^{2q}] < \infty$ for some $q>0$, then there exists a p, $0<p<q$, such that $E[(\beta+\alpha_1 z_t^2)^p] <1$.(b) If, in addition, $E[(\beta+\alpha_1 z_t^2)^r] <1$ for $0<r<q$, then exists a $\delta>0$ such that $E[(\beta+\alpha_1 z_t^2)^{r+\delta}] <1$.

Theorem 4(a) says that if ${}_u h_t$ is strictly stationary and z_t^2 has a finite moment of some (arbitrarily small, possibly fractional) order, then ${}_u h_t$ has a finite (possibly fractional) moment as well. The existence of such a finite fractional moment implies, for example, that $E[\ln({}_u h_t)]<\infty$. In addition we notice that in order for $E[(\beta+\alpha_1 z_t^2)^p] <1$ to hold for p=1, the iid innovation must have at least a fractional moment of order larger than 2. Part (b) gives a condition for $E[{}_u h_t^{2p+\delta}] < \infty$ for some $\delta>0$, given that $E[{}_u h_t^{2p}] < \infty$. It says, for

example, that if E $[(\beta+\alpha_1 z_t^2)^{1/2}] < 1$ and $E[|z_t|^{2p}] < \infty$ for some p> $\frac{1}{2}$, then not only is $E[|_u e_t|] < \infty$, but there is also a δ>0 with $E[|_u e_t|^{1+\delta}] < \infty$.
Summarizing the above results Nelson (1991) produced the following figure for the case that z_t~NIID(0,1) :



- Region 1: E $[\ln(\beta+\alpha_1 z_t^2)] > 0$ and h_t is explosively nonstationary.
- Region 2: E $[\ln(\beta+\alpha_1 z_t^2)] < 0$ and ${}_u h_t$, ${}_u e_t$ are strictly stationary and ergodic
- Region 3: E $[\ln(\beta+\alpha_1 z_t^2)] < 0$, E $[(\beta+\alpha_1 z_t^2)]^{1/2} < 1$ and ${}_u h_t$, ${}_u e_t$ are strictly stationariy and ergodic and $E[{}_u e_t] = 0$, $E[{}_u e_t^2] = \infty$
- Region 4: E $[\ln(\beta+\alpha_1 z_t^2)] < 0$, E $[(\beta+\alpha_1 z_t^2)]^{1/2} < 1$, E $[(\beta+\alpha_1 z_t^2)] < 1$ and ${}_u h_t$, ${}_u e_t$ are strictly stationary and ergodic and $E[{}_u e_t] = 0$, $E[{}_u e_t^2] =< \infty$.

We have to note that for region 4 the first and second moments exist, and for some combinations of α_1 and β there is also fourth moment. For region 3 there is only unconditional mean and for region 2, even though the unconditional process is strictly stationary, no moments exist.

IV. Survey Of Applications And Extensions

In a series of papers, the ARCH model has been analyzed, generalized, extended to the multivariate context, and used to test for time varying risk premia in the term structure of interest rates and in other financial markets. These papers include Engle (1983) and Engle and Kraft (1983) where a measure of the variance of inflation is given. The ARCH model is extended to a multivariate framework in Kraft and Engle (1982). In Engle, Granger and Kraft (1984) and Granger, Robins and Engle (1984) bivariate ARCH models of inflation with changing covariances as well as variances are constructed. Engle, Brown and Stern (1984) and granger and Engle (1984) examine the effectiveness of ARCH models for forecasting purposes. The power of ARCH tests and the finite sample properties of various estimators, are analyzed in Engle, Hendry and Trumble (1985) by means of Monte Carlo methods. In Bollerslev (1985a, 1985b) the Generalized ARCH model or GARCH model is developed, and the GARCH model with conditionally Student-t distributed errors is studied in Bollerslev (1985c). Engle, Lilien and Robins (1985) and Bollerslev, Engle and Wooldridge (1985) examine the term Structure of interest rates and a three asset Capital Asset Pricing Model (CAMP) to determine whether risk premia are varying over time. These papers introduce the ARCH in mean or ARCH-M model in a univariate and multivariate context respectively. Engle and Watson recast the GARCH-M model in a full information state space form.

In addition to this work, a variety of papers have begun to appear from different parts of the world. Particularly interesting are the papers by Milhøj (1985) who develops far more general moment conditions than those in the original Engle (1982) paper. Linnel-Nemec (1984a, 1984b) establishes stationarity and ergodicity of ARCH models. Pantula (1984) and Weiss (1982) derive the limiting distribution of ARCH estimators in more general contexts. Pagan, Hall and Trivedi (1983), Weiss (1984) and Coulson and Robins (1985) provide empirical time series examples of ARCH and related models of changing variances. Domowitz and Hakkio (1985), Diebold and Pauly (1985), Diebold and Nerlove (1985), Milhøj (1985b), Hsieh (1985) and McCurdy and Morgan (1985), apply the ARCH, the ARCH-M and the GARCH model to the foreign exchange market. Amsler (1984a, 1984b) investigated whether using the risk premia estimated by ARCH models will make long bonds satisfy the Shiller variance bounds. Poterba and Summers (1984) derive a pricing formula for stock market prices in the spirit of the asset pricing formulas presented in this paper. The price is related to its own variance, which is modeled as a simple AR (1) process. Similar ideas are employed in the paper by French, Schwert and Stambaugh (1985). Blanchard and Watson (1984) and Bodie, Kane and Mcdonald (1983, 1984) present evidence that macroeconomic and financial time series models can usefully be reformulated as a form of multivariate ARCH processes.

Engle and Bollerslev (1986) introduce a new class of models defined to be integrated in variance. This new class of models includes the variance analogue of a unit root in the mean as a special case. The models are argued to be both empirically important for the asset pricing models and empirically relevant. The conditional density is then generalized from a normal to a Student-t with unknown degrees of freedom. By estimating the degrees of freedom, implications about the conditional kurtosis of these models and time aggregated models can be drawn. By using a further generalization, they allow the conditional variance to be a non-linear function of the squared innovations. They conclude to the fact that the integrated GARCH models constitute an interesting development, judging from an empirical and theoretical point of view, and by utilizing Monte Carlo evidence they find that the knife edge properties while estimating and testing for unit root in the mean might not be as severe in the integrated variance models.

Nelson (1991) points some major drawbacks of the GARCH models and introduced a new form of ARCH that not suffer from these problems. He wishes the existence ARCH models that allow the same degree of simplicity and flexibility in representing conditional variances as ARIMA and related models have allowed in representing conditional mean. Bollerslev (1987) presents the correlation structure for the squares from the generalized autoregressive conditional heteroskedastic (GARCH) process. It is shown that the behaviour of the correlations for the squares mimics the usual correlations of an appropriately defined ARMA process, although the admissible regions for the correlations are somewhat more restrictive. Pagan and Schwert (1990) compare several statistical models for monthly stock return volatility. They use U.S data from 1834-1925, in an attempt to examine the ex-1926 data since the post–1926 have been analyzed in more detail. They point down that the Great Depression had levels of stock volatility that are inconsistent with stationary models for conditional heteroskedasticity. They demonstrate the importance of non-linearities in stock return behavior that are not captured by conventional ARCH or GARCH models. Ray Chou (1988) investigates issues of volatility and the changing risk premium in the stock market by using the GARCH estimation techniques. He claims that the persistence of shocks to the stock return volatility is so high that the data cannot distinguish whether the volatility process is stationary or not. Assuming stationarity the half-life of volatility shocks is about 1 year. The parameter estimates and the non-stationary test result are both robust to changes in the frequency of data measurements. By using monthly data the persistency result is also maintained for a longer sample period. He also claims that his findings are very different from that of Poterba and Summers (1986) who suggest that shocks to volatility are transitory and hence cannot have much impact on the market. He claims that the deviation stems from the difference in estimation methodologies and supports the fact that their methodology is limited in its nature and hence may give misleading results.

David Hsieh estimates ARCH and GARCH models for five foreign currencies, using 10 years of daily data. He utilizes a variety of ARCH and GARCH specifications, a number of nonlinear error densities, and a comprehensive set of diagnostic checks. He finds that ARCH and GARCH models can usually remove all heteroskedasticity in price changes in all five currencies. By using goodness-of-fit diagnostics indicate that exponential GARCH with certain nonnormal distributions fits the Canadian dollar extremely well and the Swiss franc and the Deutsche mark reasonably well. Baillie and Bollerslev (1989) use formal testing procedures to confirm the presence of a unit root in the autoregressive polynomial of the univariate time series representation of daily exchange-rate data. The first differences of the logarithms of daily spot rates are approximately uncorrelated through time, and a generalized autoregressive conditional heteroskedasticity model with dummy variables and conditionally t-distributed errors is found to provide good representation to the leptokurtosis and time-dependent conditional heteoskedasticity. From their close statistical examination on daily foreign exchange market data, many facts emerge; apart from similar day-of-the-week effects across currencies, the short run movements in daily log spot rates are well approximated by a martingale difference model with severe excess kurtosis and time- dependent heteroskedasticity. The conditional heteroskedasticity in daily spot rates is well represented by a GARCH (1,1) process with near unit roots. Distinctive daily seasonality and vacation effects are present in the conditional variance, which can be partly explained by differing information flows. ARCH effects are still very string in weekly data, less than on fortnight data, and minimal on monthly data. After taking account of any ARCH effects, the assumption of conditional normality is reasonable approximation on monthly and fortnight data, whereas for weekly data the validity of the assumption seems to vary across currencies.

Changli He and Timo Terasvirta (1999) present a complete characterization of the fourth moment structure of a general GARCH (p, q) process. By using these results, any investor is able to see what an estimated GARCH model implies about the second and the fourth moments, kurtosis, and the autocorrelation function of the centered and squared observations. Even though such considerations have previously been possible in the case of GARCH (1, 1), they manage to extend this possibility to other GARCH processes that are generalizations of the original GARCH (p, q) process. For example, some GARCH processes allowing for asymmetric effects to shocks belong to this category. Nelson (1990) establishes the necessary and sufficient conditions for the stationarity and ergodicity of the GARCH (1,1) process. As a special case, he shows that the IGARCH (1, 1) process with no drift converges almost surely to zero, while IGARCH (1, 1) with drift is strictly stationary and ergodic. He examines the persistence of shocks to conditional variance in the case of GARCH (1, 1) model, and shows that whether the shocks persistence or not depends on the definition of persistence. He also develops necessary and sufficient conditions for the finiteness of absolute moments of any order.

Robin Lumsdaine (1995) compares the GARCH (1, 1) and the IGARCH (1, 1) models via a Monte Carlo study of the finite sample properties of the maximum likelihood estimator and related test statistics. In all models considered, Lumsdaine finds that this estimator has a normal limiting distribution and constant covariance matrix. Although the asymptotic distribution is, for the most part, well approximated by the estimated t statistics, other commonly used statistics such as the Lagrange multiplier, likelihood ratio, and Wald do not behave as well in small samples. She also notices that the estimators themselves are skewed in small samples, particularly those of the ARCH parameters. Finally, a pileup effect at the boundary of the parameter values is also apparent; this pileup decreases as the sample size becomes larger. The tails of the small sample distributions are fatter than those of a normal distribution. Lumsdaine (1996) also provides the proof of the consistency and asymptotic normality of the quasi-maximum likelihood estimator in GARCH (1, 1) and IGARCH (1, 1) models. In contrast to the case of a unit root in the conditional mean, the presence of a “unit root” in the conditional variance does not affect the limiting distribution of the estimators; in both models, estimators are normally distributed. In addition, a consistent estimator of the covariance matrix is available, enabling the use of standard test statistics for inference.

V. An Empirical Application Of GARCH (1, 1)

a. Data and Methodology

Our focus is to compare the forecasting performance of two models for a univariate conditional variance, using bilateral daily data of stock market indices for the United States, Mexico, United Kingdom, France, Germany, Belgium, Switzerland, Australia, Singapore and Japan. We also use bilateral daily data for the dollar versus the currencies of Canada, France, Germany, Japan, United Kingdom, Spain, Singapore, Netherlands, Australia and Switzerland. This data are collected for ten years, from 1 January 1992 to 31 December 2001, and we consider the daily closing prices as the daily observations. This sample is divided into two different periods of five years each, and each period is examined separately. After an initial observation was lost due to differencing, the sample for each country in the case of the exchange rates includes 1258 observations for the first period, from 1 January 1992 to 31 December 1996, and 1256 observations for the second period, from 1 January 1997 to 31 December 2001. In the case of stock market indices the sample in each country is diversified (see appendix VII).

Prior to our formal analysis, we use logarithmic differences of the series. So, our stock return and exchange return series are defined as the natural logarithm of values relative.

$$R_t = \log(\frac{S_t}{S_{t-1}}) \tag{45}$$

For small values of R_t, such as in the case of daily data, this definition is very similar to the arithmetic rate of return.

Our empirical application, as well as our Monte Carlo experiment, is divided in two parts; the first part, in which we perform the estimation of the parameters of the GARCH (1, 1) model, selected for the description of the data, and the second one in which by utilizing the estimated parameters we perform forecasts of the conditional variance. It is important to state that before the estimation part, we choose an ARMA (p, q) model for the suitable representation of the conditional mean. This representation is thought as necessary in order to remove all the linear dependence of the original data. The selection of the appropriate order of the ARMA (p, q) model is done based on the serial correlation that the residuals present after the selection of an ARMA process upon the data. It is not surprising that for all our time series the appropriate ARMA (p, q) model that is chosen to remove the serial correlation of the data is the AR (1).

Vedat Akgiray in his article, “Conditional Heteroskedasticity In Time Series Of Stock Returns”, supports the idea that any realistic probability model of daily stock-price movements must be consistent with at least two empirical facts: (1) time series of returns exhibit significant first-lag autocorrelation, and (2) time series of absolute and squared returns are autocorrelated even at very long lags. A reasonable strategy to construct such a model may start with transforming the original return series so that the new series will no longer be correlated. Then the model to be fitted to this new series would be required to satisfy only the second property above.

One possible way of generating an uncorrelated sequence from the series $\{R_t\}$ is to obtain the ordinary least squares residuals of the following regression:

$$R_t = \phi_o + \phi_1 R_{t-1} + e_t \qquad (46)$$

The residuals series $\{e_t\}$ can be expected to be uncorrelated since second-order or higher-order autocorrelation is not observed in the return series.

By concluding to the fact that the daily return series R_t can be modeled as a AR (1)-GARCH (1, 1) process, we perform estimation upon the following model:

$$\begin{aligned} & R_t/F_{t-1} \sim D\,(m_t,h_t) \qquad (47) \\ & m_t = \varphi_o + \varphi_1 R_{t-1} \\ & h_t = \alpha_o + \alpha_1 e_t^2 + \beta_1 h_{t-1} \\ & e_t = R_t - \varphi_o - \varphi_1 R_{t-1} \end{aligned}$$

To estimate the parameters $\theta=(\varphi_o, \varphi_1, \alpha_o, \alpha_1, \beta_1)$ of the above model, it is necessary to specify the conditional distribution function of D (m_t, h_t). In all applications, a normal distribution function is assumed. For lack of a good reason for another distribution, this assumption is adopted in this study, although the model is flexible enough to admit other laws. Given a sample of daily returns $R_1, \ldots, R_T$ the log-likelihood function is then given by

$$L(\theta / p,q) = \sum_{t=r}^{T} \log f(m_t, h_t) \qquad (48)$$

where p,q = 1, r=max(p,q)=1. Numerical maximization gives the maximum likelihood estimates of the parameters for the AR (1)-GARCH (1, 1) model. We find that all of the parameters are statistically significant (except from very few cases), and the sum of the parameters of the GARCH (1, 1) model ($\alpha_1 + \beta_1$) is substantially smaller than unity and dominated by β, something that indicates that changes in market volatility tend to be persistent. The fact that the sum is smaller than unity ensures also the existence of second moment.

After the estimation part and since estimations of the parameters obtained for each time series (see appendix VI), we can move to the forecasting part where we will compare the forecasting performance of the AR (1)-GARCH (1, 1) model with the autoregressive conditional heteroskedasticity model of order 1, AR (1) (we have repeated the estimation part for this model too).

There are two alternative ways for forecasting with lagged dependent variables: the dynamic forecasting and the static forecasting. In the case of a dynamic forecasting, we have multi-step forecasts starting from the first period of the forecast sample. In the static forecast, a sequence of one-step-ahead forecasts is calculated, using actual, rather than forecasted values for lagged dependent variables.

The selection of the start of the forecast sample is very important for dynamic forecasting. The dynamic forecasts are true multi-step forecasts (from the start of the forecast sample), since they use the recursively computed forecast of the lagged values of the dependent variables. These forecasts may be interpreted as the forecasts for subsequent periods that would be computed using information available at the start of the forecast sample. Dynamic forecasting requires that data for the exogenous variables be available for every observation in the forecast sample, and that values for any lagged dependent variables be observed at the start of the forecast sample.

Static forecasting requires that the data for both the exogenous and any lagged endogenous variables be observed for every observation in the forecast sample. Both methods will always yield identical results in the first period of a multi-period forecast. Thus, two forecast series, one dynamic and the other static, should be identical for the first observation in the forecast sample. We decide to perform forecasts for the conditional variance by utilizing both methods and to compare the MSPE, computed as $MSPE(j) = E(\hat{h}_{t+j} - e^2_{t+j})^2$, with this obtained by the AR (1) model. Since the true volatility, h_{t+j}, is unobserved we choose to use as an unbiased estimator of it the squared shock, e^2_{t+j}, since E[e^2_{t+j}]=E[z^2_{t+j} h_{t+j}]=$1 \cdot h_{t+j}$. The $\hat{h}_{t+j}$ is the future conditional variance as it is forecasted through the AR (1)-GARCH (1, 1) model. As to the simple AR (1) process $R_t = f_o + f_1 R_{t-1} + \varepsilon_t$, the conditional variance of R_t is equal to the conditional variance of ε_t, $\text{var}(R_t / F_{t-1}) = \text{var}(\varepsilon_t / F_{t-1})$ =σ^2, and the statistical criterion that we utilize is the MSPE computed as $MSPE = E[\sigma^2 - \varepsilon^2_{t+j}]$. The right approximation for the true unobserved conditional variance is considered to be again the squared shock e^2_{t+j}.

The entire procedure was programmed in Eviews 4.0 software.

Tables for the group of stock indices, for the periods 1992-1996 and 1997-2001

USA1	MSE(1)	MSE(2)	MSE(3)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,008828	0,00491	0,006272	0,00652	0,011404	0,0296	0,049759
Static	0,008828	0,00506	0,006214	0,0066	0,011576	0,03033	0,04836
AR(1)	0,01191	0,00618	0,007879	0,00774	0,012796	0,02977	0,049229

USA2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,1086	0,05595	0,047995	0,06055	0,048603	0,11136	0,150934	0,15989
Static	0,1086	0,05455	0,043253	0,05	0,040399	0,11275	0,161557	0,17011
AR(1)	0,2261	0,12709	0,121093	0,13585	0,111246	0,15142	0,193034	0,19157

Ausrtalia1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,01558	0,0081	0,033856	0,0309	0,027025	0,02061	0,531789	0,41384
Static	0,01558	0,00832	0,034987	0,03283	0,02871	0,02	0,54437	0,42039
AR(1)	0,02193	0,011	0,030751	0,03	0,026992	0,02204	0,523631	0,40515

Australia2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,00874	0,02564	0,02221	0,02141	0,02457	0,03366	0,03854	0,03981
Static	0,00874	0,02362	0,01869	0,01657	0,01767	0,01884	0,01829	0,02029
AR(1)	0,02197	0,03874	0,03224	0,03213	0,0356	0,04159	0,04418	0,04367

UK1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,01318	0,00951	0,00731	0,00799	0,10101	0,05512	0,18204	0,15732
Static	0,01318	0,00902	0,00664	0,00688	0,10507	0,05816	0,18691	0,15836
AR(1)	0,0316	0,02515	0,02123	0,02224	0,09817	0,05843	0,17662	0,15043

UK2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,09728	0,13459	0,13636	0,18322	0,19688	0,18915	0,24055	0,23598
Static	0,09728	0,12539	0,118393	0,15022	0,15116	0,11765	0,19137	0,2044
AR(1)	0,02259	0,04395	0,049917	0,08311	0,09205	0,09642	0,17617	0,1849

France1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,02175	0,02821	0,03555	0,02746	0,02216	0,02734	0,20333	0,18483
Static	0,02175	0,02745	0,033703	0,02656	0,02163	0,02565	0,2045	0,18677
AR(1)	0,07279	0,08527	0,097047	0,07366	0,06033	0,06648	0,19542	0,18057

France2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	1,66	1,6295	1,188367	0,93235	0,94988	0,88045	1,06318	0,96243
Static	1,66	1,5175	1,060133	0,80634	0,77563	0,59085	0,84405	0,71924
AR(1)	0,4927	0,487	0,324927	0,245	0,2464	0,25087	0,62218	0,54532

Germany1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,04587	0,027489	0,038706	0,04091	0,03971	0,03915	0,02894	1,60487
Static	0,04587	0,026585	0,040433	0,04212	0,03989	0,03807	0,02786	1,55609
AR(1)	0,06034	0,037775	0,039077	0,04463	0,04486	0,04467	0,03335	1,54477

Germany2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	2,246	1,8425	1,685333	1,26838	1,02782	1,28075	1,38243	2,86809
Static	2,246	1,6665	1,3862	1,09785	0,8785	0,77112	0,92705	2,8896
AR(1)	0,6695	0,4588	0,387433	0,4615	0,39166	0,42334	0,62535	2,56251

Japan1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,05363	0,048285	0,032317	0,05961	0,07341	0,09627	0,625207	0,48335
Static	0,05363	0,044965	0,030797	0,05367	0,06188	0,07669	0,647679	0,50669
AR(1)	0,2227	0,1962	0,142513	0,19356	0,21743	0,23093	0,663869	0,52708

Japan2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	17,68	8,84846	6,000873	4,50748	3,67314	3,27015	2,663657	2,66189
Static	17,68	8,86288	6,156953	4,65582	3,84073	3,36984	2,724611	2,63066
AR(1)	18,1	9,072995	6,159663	4,63778	3,79264	3,3518	2,696975	2,61153

Mexico1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,1017	0,0783	0,094667	0,07214	0,06492	0,09542	0,160894	0,17033
Static	0,1017	0,070345	0,076563	0,06375	0,05386	0,06222	0,159475	0,1671
AR(1)	0,8206	0,7429	0,770767	0,64015	0,61386	0,62285	0,581527	0,5516

Mexico2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,6551	0,366295	0,497163	0,54387	0,6096	0,67107	0,774651	0,82734
Static	0,6551	0,337185	0,394323	0,37242	0,36489	0,2804	0,405102	0,3564
AR(1)	1,463	1,0102	1,173133	1,2201	1,27508	1,19231	1,20244	1,16646

Switzerland1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,01493	0,023145	0,02289	0,01819	0,01692	0,01823	0,116753	0,09327
Static	0,01493	0,020725	0,018463	0,01398	0,01219	0,01472	0,119466	0,09836
AR(1)	0,02872	0,03657	0,035417	0,02873	0,02623	0,02382	0,11682	0,09324

Switzerland2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,000207	0,012059	0,052806	0,05527	0,05563	0,06243	0,065354	0,07348
Static	0,000207	0,012369	0,057879	0,06363	0,06177	0,04958	0,052556	0,05719
AR(1)	0,07547	0,142685	0,096064	0,14322	0,16534	0,20231	0,189693	0,18826

Singapore1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,05751	0,03838	0,049123	0,0533	0,05023	0,04631	0,043823	0,04459
Static	0,05751	0,04723	0,055037	0,05104	0,04297	0,03218	0,032654	0,03067
AR(1)	0,06091	0,035388	0,050165	0,05634	0,05404	0,04935	0,045941	0,04669

Singapore2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	1,794	0,9609	1,5966	1,61045	1,68376	1,59857	1,547387	1,73224
Static	1,794	0,90116	1,27344	1,09388	0,9853	0,62278	0,980686	0,92449
AR(1)	0,159	0,22705	0,404233	0,35233	0,35488	0,32708	0,69524	0,68611

Belgium1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,00014	0,00678	0,00664	0,00614	0,00688	0,00656	0,007702	0,00965
Static	0,00014	0,006725	0,006366	0,0057	0,00619	0,00553	0,005913	0,00887
AR(1)	0,00099	0,009925	0,00998	0,00937	0,01029	0,0094	0,010481	0,01155

Belgium2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,2799	0,2469	0,185217	0,17236	0,20049	0,18861	0,184186	0,21982
Static	0,2799	0,2093	0,142615	0,11679	0,11927	0,09252	0,088927	0,07642
AR(1)	0,1885	0,16235	0,112627	0,10138	0,11835	0,1109	0,102926	0,11473

Tables for the group of foreign exchange rates, for the periods 1992-1996 and 1997-2001

Australia1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,001514	0,003546	0,097764	0,07437	0,06031	0,48944	0,32916	0,25586
Static	0,001514	0,003708	0,097138	0,07512	0,06161	0,48885	0,332583	0,25813
AR(1)	0,001981	0,003549	0,098866	0,07539	0,06129	0,4812	0,323791	0,25159

Australia2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,01329	0,009624	0,012096	0,01106	0,00938	0,00868	0,011275	0,0139
Static	0,01329	0,009159	0,010952	0,00955	0,00784	0,00644	0,007123	0,0102
AR(1)	0,01874	0,013898	0,017065	0,01579	0,01359	0,01218	0,014768	0,01677

Canada1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0008859	0,000515	0,000387	0,00029	0,00026	0,00018	0,000283	0,00048
Static	0,0008859	0,000532	0,000404	0,0003	0,00028	0,00019	0,000289	0,00045
AR(1)	0,0002063	0,000442	0,000503	0,00041	0,00047	0,00041	0,000375	0,00045

Canada2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,00348	0,002857	0,001906	0,00153	0,0013	0,00313	0,002209	0,00355
Static	0,00348	0,002729	0,001837	0,00153	0,00127	0,00298	0,002172	0,00347
AR(1)	0,002419	0,002058	0,001448	0,00132	0,00106	0,00269	0,001982	0,00323

Japan1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,01191	0,010422	0,009474	0,00972	0,01025	0,01706	0,012943	0,01152
Static	0,01191	0,00996	0,00867	0,00852	0,00863	0,01714	0,012536	0,01046
AR(1)	0,01821	0,016255	0,015017	0,01525	0,01584	0,01991	0,015668	0,01432

Japan2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0001905	0,001656	0,011681	0,01139	0,01208	0,00831	0,008975	0,00987
Static	0,0001905	0,00161	0,012057	0,01197	0,01256	0,00827	0,007909	0,00775
AR(1)	0,01292	0,01635	0,013818	0,01971	0,02471	0,02126	0,024053	0,02571

Singapore1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	7,168E-05	8,14E-05	5,82E-05	8,9E-05	0,00011	0,00018	0,000194	0,00023
Static	7,168E-05	5,94E-05	3,99E-05	5,6E-05	5,9E-05	5,1E-05	5,24E-05	5,2E-05
AR(1)	0,0002656	0,000254	0,000193	0,00021	0,00022	0,00023	0,000207	0,00021

Singapore2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	1,009E-05	0,00032	0,000452	0,00036	0,00044	0,00043	0,00047	0,00052
Static	1,009E-05	0,000314	0,000417	0,00032	0,00037	0,00033	0,000338	0,00032
AR(1)	0,002587	0,003904	0,004371	0,00404	0,00431	0,00405	0,004025	0,00415

Switzerland1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0361	0,020523	0,014328	0,01308	0,01205	0,02954	0,025346	0,02915
Static	0,0361	0,02128	0,015005	0,01372	0,01247	0,03073	0,026013	0,03002
AR(1)	0,009009	0,017645	0,017773	0,02183	0,02358	0,03134	0,026196	0,02988

Switzerland2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0134	0,008396	0,005788	0,00741	0,00898	0,00959	0,008133	0,00838
Static	0,0134	0,008517	0,005877	0,00747	0,00893	0,00967	0,008071	0,00822
AR(1)	0,01272	0,008111	0,00559	0,00733	0,00895	0,00955	0,008109	0,00837

U.K1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0009558	0,000737	0,001643	0,00219	0,00183	0,00238	0,261191	0,19665
Static	0,0009558	0,0007	0,001509	0,00228	0,00192	0,0024	0,262935	0,20095
AR(1)	0,007422	0,007083	0,009292	0,00698	0,00575	0,00719	0,247576	0,18843

U.K2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0001039	0,003395	0,002863	0,00354	0,00444	0,00531	0,004644	0,00443
Static	0,0001039	0,003423	0,002823	0,00339	0,00414	0,00476	0,004032	0,00381
AR(1)	0,0006591	0,002498	0,001905	0,00231	0,00297	0,00404	0,003568	0,00351

France1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0002711	0,000364	0,000805	0,0006	0,0005	0,01202	0,01232	0,01136
Static	0,0002711	0,000381	0,000825	0,00062	0,00052	0,01196	0,011664	0,01099
AR(1)	0,004361	0,007606	0,005593	0,0059	0,0063	0,01195	0,010744	0,01104

France2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0000248	0,000423	0,005719	0,00494	0,00729	0,00846	0,008691	0,0087
Static	0,0000248	0,00042	0,005647	0,00479	0,00693	0,00825	0,008076	0,00783
AR(1)	3,966E-06	0,000296	0,005281	0,00453	0,00677	0,00812	0,008317	0,00835

Germany1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,00098	0,001085	0,001185	0,00152	0,00123	0,01517	0,030262	0,02325
Static	0,00098	0,001147	0,001234	0,00161	0,00129	0,01571	0,029973	0,02372
AR(1)	0,003646	0,009603	0,007346	0,01027	0,00989	0,01808	0,028322	0,02468

Germany2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	7,877E-05	0,00057	0,00618	0,0054	0,00788	0,00886	0,009181	0,00922
Static	7,877E-05	0,000566	0,006114	0,00527	0,00757	0,00864	0,008575	0,00834
AR(1)	7,891E-06	0,000325	0,005384	0,00463	0,0069	0,00821	0,008425	0,00846

Netherlands1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0007321	0,001087	0,001416	0,00148	0,00118	0,01002	0,02296	0,02039
Static	0,0007321	0,001146	0,001479	0,00157	0,00125	0,01031	0,022474	0,02025
AR(1)	0,004389	0,0107	0,007869	0,01019	0,00965	0,01344	0,021718	0,02045

Netherlands2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	7,831E-05	0,000559	0,006142	0,00536	0,00784	0,00883	0,009142	0,00918
Static	7,831E-05	0,000554	0,006069	0,00522	0,0075	0,00859	0,008497	0,00825
AR(1)	1,148E-05	0,000347	0,005438	0,00468	0,00696	0,00825	0,008482	0,00852

Spain1	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,001405	0,00164	0,001504	0,00183	0,0015	0,00993	0,015003	0,01281
Static	0,001405	0,001782	0,001619	0,00198	0,00163	0,01027	0,014823	0,01315
AR(1)	0,004119	0,013025	0,009911	0,01334	0,01229	0,0137	0,015427	0,01542

Spain2	MSE(1)	MSE(2)	MSE(3)	MSE(4)	MSE(5)	MSE(10)	MSE(15)	MSE(20)
Dynamic	0,0002789	0,001007	0,007192	0,00638	0,0091	0,00974	0,010257	0,01034
Static	0,0002789	0,001	0,00713	0,00627	0,00887	0,00952	0,00972	0,00956
AR(1)	3,012E-07	0,000257	0,005121	0,00436	0,00655	0,00799	0,008154	0,00818

b. Results From The Empirical Application

The results obtained from our empirical application seem to confirm the skeptics about the forecasting performance of the GARCH models. In the majority of the time series that we examine it is hard to conclude surely which of the two models performs more accurately. There is a number of time series from the stock indices as well as from the foreign exchange rates in which the AR (1)- GARCH (1, 1) model provides forecasts of the conditional variance better than those of the simple AR (1) model. In particular, the AR (1)-GARCH (1, 1) model is appeared to have a slight edge over the other model at the one-week horizon, for some of the time series. This superiority of the GARCH model doesn't seem to hold at longer horizons, even though we don't observe a tendency for the MSPE to increase as we move forward in time. In most cases, where the GARCH model appears to have more accurate predictive ability, it ends to lose its superiority at the one-month horizon. There are also many cases in which the simple AR (1) model performs better for the entire forecast sample.

For the group of the stock indices, USA, United Kingdom, France, Germany and Switzerland for the first period, 1992-1997, configure the fact that the forecasts of the GARCH (1, 1) model are better than that of the AR (1) for the one-week horizon, but for longer periods its MSPE takes values larger than that of the second model. For the majority of time series of this group, and for the second period, 1997-2001, the GARCH model appears slight better behavior. But there are also significant exceptions as Germany, France, United Kingdom, Belgium, Singapore and Japan in which the superiority of the AR (1) process in describing and forecasting the conditional variance of the data seems to be unquestionable.

As for the group of the foreign exchange rates, for the currencies of Switzerland, United Kingdom, France and Spain versus dollar, for the first period, as well as for the currencies of Canada, Switzerland, United Kingdom, and France for the second period, the AR (1) model performs more accurately, for the short and also for the long-term horizons. In the case of Australia (first and second period), Canada (first period), Japan (first and second period) and Singapore (first and second period) the results are confusing; it is not obvious which of the models manage to conclude to more accurate forecasts for the entire forecast sample.

Motivated by these findings as well as by the existing literature which finds that the GARCH models explain little of the variability in ex-post squared returns and for this reason may be of limited practical value, we decide to perform a Monte Carlo experiment in order to examine if the utilization of GARCH models may be considered as inappropriate in comparison with other models. In or empirical experiment we compare an AR (1)-GARCH (1, 1) model with a simple AR (1). In the following section we evaluate a MD-GARCH (1, 1) in contrast with a homoskedastic one.

VI. Monte Carlo Design And Data Generating Process

Although estimated with growing frequency, autoregressive conditional heteroskedasticity (ARCH)- related models have experienced relatively few theoretical developments. Our goal in this part is to investigate the forecasting performance of a MD - GARCH (1, 1) model and to compare it with this of a homoskedastic one. We focus on three alternative models from the GARCH (1, 1) family. In particular, we examine the following pairs of values for (α, β):

	Model 1	**Model 2**	**Model 3**
α	0.20	0.40	0.15
β	0.55	0.50	0.85

The selected values in models 1-3 were chosen for their empirical and theoretical relevance. Models 1 and 2 are GARCH (1, 1) models, which differ in a very special point; model 1 is a typical GARCH (1, 1) model with finite second and fourth moment, while model 2 is a GARCH (1, 1) model, selected for its potential theoretical interest, since in its case the assumption for the existence of a fourth moment is violated. Bollerslev (1986, theorem 2) provided necessary and sufficient conditions on the parameters for the existence of higher moments of the e_t's. In particular, the condition for the existence of a finite fourth moment in the GARCH (1, 1) model is given by $k\alpha^2+2\alpha\beta+\beta^2 < 1$, where k is the kurtosis of the innovation process z_t. Even though the existence of the fourth moment is not necessary in order the estimators to be consistent and asymptotically normal (see Lumsdaine (1991)), we choose to consider this special case in order to examine whether or not the absence of the unconditional fourth moment has an effect in the forecasting performance of the GARCH (1, 1) model. Model 3 is considered in order to evaluate how accurate is the forecasts that we get from the GARCH (1, 1) model when $\alpha + \beta = 1$ (IGARCH (1, 1)). Keep in mind that under this restriction we are trying to forecast the future conditional variance under the absence of an unconditional one. For each of these three models, we perform a dynamic and a static way of forecasting and we consider alternative measures for the evaluation of the forecasting performance (see below).

The data are generating according to the following recursions:

$$y_t = e_{t,}\ e_t/F_{t-1} \sim D(m_t,h_t) \qquad (49)$$

$$\text{where } m_t = 0 \text{ and } h_t= \alpha_o + \alpha_1 e_{t-1}^2 + \beta h_{t-1}$$

$$\text{or } y_t = e_t = z_t h_t$$

$$h_t= \alpha_o + \alpha_1 e_{t-1}^2 + \beta h_{t-1}$$

$$\text{where } z_t \sim \text{iid } (0,1)$$

The innovation process z_t is drawn from a normal random-number generator, $h_{o,}$ as well as y_o, is supposed to be equal to 0.5, and the initial 50 observations are eliminated in order to minimize the effects of the initial values. The fact that the z_t is normally distributed is not required by the asymptotic theory but is often assumed in applied work. In addition, although the z_t's have finite unconditional moments of all orders, since they are drawn from a normal distribution, there is no guarantee that the compound term, $e_t = z_t h_t$, will have even finite fourth moments as discussed previously. The selection of 0.5 as the initial value of the h_t is considered to be asymptotically negligible. This asymptotic negligibility was shown by Lumsdaine (1993, lemma 6) for the likelihood, its first and second derivatives, and the components of the covariance matrices. Implicitly, all results reported here are conditioned on this initial choice. Engle and Bollerslev (1986) proposed using the sample mean of the squared estimated residuals, $\frac{1}{T}\sum_{t=1}^{T} e_t^2$. In the IGARCH (1, 1) model, however, the analogous population mean does not exist. Alternatively, an initial value for the h_t can be drawn from the asymptotic distribution of the filtering error of Nelson and Foster (1994), or the method of Foster and Nelson (1993) may be employed. Diebold and Schuermann (1992) considered finite-sample properties of ARCH estimators with respect to different choices of initial conditions.

We could say that our Monte Carlo experiment is divided into two parts; the first one has to do with the estimation of the parameters of the GARCH model, while the second one involves making forecasts for the future unconditional variance under the assumption of a GARCH (1, 1) structure. Utilizing quasi-maximum likelihood estimators, we make the estimation of the parameters of the GARCH process. Conventionally, the quasi-maximum likelihood estimators involve maximizing a normal likelihood function even though the true underlying distribution may not be normal. Maximum likelihood estimators refer to estimators computed under the assumption that the likelihood being maximized is indeed the true likelihood function; this is often taken to be normal. Thus quasi-maximum likelihood estimation is a generalization of maximum likelihood estimation to the case in which the true underlying distribution is unknown.

Results are reported for T=1000, using 1000 replications. In order to estimate the parameters of the GARCH (1, 1) models, we use a sample of 500 observations and the others 500 observations are used to carry out forecasts of the conditional variance. These samples were chosen to be representative of the sizes of daily data sets commonly used empirically. The size of the estimation sample is chosen with respect to our goal to report as good estimations as possible, while the large size of the forecasting sample is due to our aim to run three different regressions (see below).

We have already mentioned that for each model we consider three different expressions of a statistical criterion. In particular, we calculate the mean square prediction error under the availability or not of the true future conditional variance. The basic difference in the case of the Monte Carlo experiments is the fact that the future conditional variance is observed since it is generated, through the above data generating process, for the entire sample of 1000 observations. This offers us the capability to compare, under a different point of view, the GARCH (1, 1) (in all its alternatives forms) with the homoskedastic model, where we assume that the conditional variance is stable and indifferent of time t. The three alternative expressions of the MSPE for the case of the GARCH (1, 1) model are

$$MSPE(j) = E(\hat{h}_{t+j} - e_{t+j}^2)^2 \quad (50)$$

$$MSPE(j) = E(h_{t+j} - \hat{h}_{t+j})^2$$

$$MSPE(j) = E(h_{t+j} - e_{t+j}^2)^2$$

The first expression of the MSPE is the one that we use in our empirical application in order to evaluate the forecasting performance of a GARCH (1, 1) model. In the case of the empirical application, in order to make the forecast evaluation criteria operational, we use the squared shock $e_{t+j}^2 = z_{t+j}^2 h_{t+j}$ as a good approximation of the true unobserved volatility h_{t+j}. As $E[z_{t+j}^2] = 1$, e_{t+j}^2 is an unbiased estimate of h_{t+j}, $E[e_{t+j}^2] = E[z_{t+j}^2 h_{t+j}] = 1 \cdot h_{t+j} = h_{t+j}$. In the other two expressions of the statistical criterion MSPE, we observe the appearance of the true conditional volatility h_{t+j}. The MSPE, which is computed as $MSPE(j) = E(h_{t+j} - \hat{h}_{t+j})^2$, is the expression that we would prefer to have available for all our experiments. It offers us a direct and accurate measure of the forecasting performance of a GARCH (1, 1) model, since it compares the true conditional volatility h_{t+j} with the one that we get by assuming that the shock, e_t, is well represented by a GARCH (1, 1) process. The last expression of MSPE is used as a criterion for the comparison of the true volatility h_{t+j} with its approximation, e_{t+j}^2, in order to find out, if apart from the theoretical point of view, the use of the squared shock is well used for replacing h_{t+j}, when the true conditional variance is unobserved. For the homoskedastic model, where $var(e_t/F_{t-1}) = \sigma^2$ for every t, we use as a statistical criterion the mean squared prediction error which is computed as

$$MSPE(j) = E[h_{t+j} - \sigma^2] \text{ for j=1, 2, 3, 5, 10} \qquad (51)$$

Our goal, as in the case of the empirical application, is to compare this mean squared prediction error with those of the GARCH (1, 1) model and to examine if the predictions that we acquire when we suppose a homoskedastic structure for the conditional variance are better or not.

In our attempt to study the relation between the true volatility and each one of its approximations, $\hat{h}_{t+j}$ and e_{t+j}^2, as well as their relation between, we are running three different regressions:

$$\text{Regression \#1: } e_{t+j}^2 = a + b\hat{h}_{t+j} + n_{t+j} \qquad (52)$$

$$\text{Regression \#2: } h_{t+j} = a + b\hat{h}_{t+j} + k_{t+j}$$

$$\text{Regression \#3: } e_{t+j}^2 = a + bh_{t+j} + m_{t+j}$$

These three regressions provide us an alternative way to examine all these issues that are of great interest for us as in the case of the different expressions of mean squared prediction error. In particular, for the third regression, e_{t+j}^2 would be an unbiased estimator of the true volatility h_{t+j}, if $\alpha = 0$, $b = 1$ and $E(m_{t+j}) = 0$. For the same reason we would prefer for the second regression values 0, 1 and 0 for α, b and $E(k_{t+j})$ respectively, in order the future conditional variance $\hat{h}_{t+j}$ that we get from a GARCH process to be an unbiased estimator of h_{t+j}.

VII. Results From The Simulations

a. Model 1: Forecasting With GARCH (1,1) Under The Presence Of Finite Second And Fourth Unconditional Moments

DYNAMIC FORECASTING

i) $MSPE(j)=E(\hat{h}_{t+j} - e^2_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	3.6185895	4.3982382	2.7426020	3.3740240	2.7294505
std.dev.	39.107599	67.111446	20.001897	18.817564	15.135565

ii) $MSPE(j)=E(\hat{h}_{t+j} - h_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	0.0253691	0.1757629	0.2579264	0.1868534	0.2189938
std.dev.	0.1038183	1.6830446	2.4053255	0.8743073	0.8618254

iii) $MSPE=E(e^2_{t+j} - h_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	3.5758804	4.3471522	2.3778222	3.1369081	2.6622436
std.dev.	39.072997	68.110079	16.509856	16.312038	14.852504

- **Homoskedastic Model: $MSPE(j)=E(h_{t+j} - \hat{\sigma}^2)^2$ for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	0.2394041	0.3060065	0.3834270	0.2214123	0.2180635
std.dev.	1.2131117	2.7615450	5.1997577	1.2499291	0.8574242

Regression #1: $e^2_{t+j} = a + b\hat{h}_{t+j} + m_{t+j}$, j=1,2,…
Regression #2: $h_{t+j} = a + b\hat{h}_{t+j} + e_{t+j}$, j=1,2,…
Regression #3: $e^2_{t+j} = a + bh_{t+j} + n_{t+j}$, j=1,2,…

size=500,rep=1000	regression #2	regression #3	regression #1
a	8.5467578	0.1170924	35.671989
b	-6.9623708	0.8770900	-32.501735
R^2	0.0071891	0.0711726	0.0040225

STATIC FORECASTING

i) $MSPE(j)=E(\hat{h}_{t+j} - y^2_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	3.6185895	4.3133973	2.4385192	3.0537318	2.6793063
std.dev.	39.107599	65.748945	16.220567	15.221363	14.664600

ii) $MSPE(j)=E(\hat{h}_{t+j} - h_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	0.0253691	0.0330035	0.0308944	0.0303579	0.0263431
std.dev.	0.1038183	0.2841047	0.1658193	0.1633472	0.1439892

iii) $MSPE=E(y^2_{t+j} - h_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	3.5758804	4.3471522	2.3778222	3.1369081	2.6622436
std.dev.	39.072997	68.110079	16.509856	16.312038	14.852504

- **Homoskedastic Model: $MSPE(j)=E(h_{t+j} - \hat{\sigma}^2)^2$ for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	0.2394041	0.3060065	0.3834270	0.2214123	0.2180635
std.dev.	1.2131117	2.7615450	5.1997577	1.2499291	0.8574242

Regression #1: $e^2_{t+j} = a + b\hat{h}_{t+j} + m_{t+j}$, j=1,2,…
Regression #2: $h_{t+j} = a + b\hat{h}_{t+j} + e_{t+j}$, j=1,2,…
Regression #3: $e^2_{t+j} = a + bh_{t+j} + n_{t+j}$, j=1,2,…

size=500,rep=1000	regression #2	regression #3	regression #1
a	-0.0718274	0.1170924	0.0495819
b	1.0907753	0.8770900	0.9627832
R^2	0.9547408	0.0711726	0.0682077

In the first Model where $\alpha + \beta < 1$ and $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$, we observe that the quasi-maximum estimators of the parameters α_o, α_1, β, are very close to their theoretical values, since we examine a case where the second and the fourth order moment exist. So the estimators that we get are consistent and asymptotically normal (see appendix I.1).

After the estimation part, we examine the forecast results provided from the utilization of Model 1 when we perform dynamic and static forecast. Let us repeat that the first observation in the forecast sample will be identical for both methods and the difference between these two approaches will be obvious for subsequent periods since in a GARCH (1, 1) model there are lagged dependent variables. In the case of the dynamic forecast, it seems that among the three different expressions of the mean squared prediction error we manage to ensure better results when we utilize the true conditional variance and not an approximation of this and we compare it with the forecast that we get from a GARCH (1, 1) model. The mean squared prediction error, computed as $MSPE(j) = E(h_{t+j} - \hat{h}_{t+j})^2$, presents the smallest standard deviation and an apparent better behavior as it is kept in relatively small values even in the case of ten steps ahead forecast. When we use as an unbiased estimator of h_{t+j}, the squared shock, e_{t+j}, and calculate the MSPE as $MSPE(j) = E(h_{t+j} - e^2_{t+j})^2$, it is obvious that the approximation we use is not the best possible but it seems to be the only one when the true volatility is unobserved.

As for the MSPE that is denoted as $MSPE(j) = E(\hat{h}_{t+j} - e^2_{t+j})^2$, and which is also utilized when we perform our empirical application, it presents the largest values. The basic comparison that we have to state and which is of great interest is that between the above MSPE and that based on the homoskedastic model. This will give us the answer to the basic question if the utilization of a GARCH model can provide us with better forecasts of the conditional volatility than a simple homoskedastic model. It is obvious that the forecasts that we get by using the GARCH (1, 1) model are more accurate than that of the homoskedastic. But what are the results for the static forecasting?

When we use the static way in order to make forecasts, true values are used for the lagged dependent variables, and not the estimated ones as in the case of the dynamic forecasting, something that we expect to ameliorate the forecasting accuracy of the model as we move forward to time. The results are the same with those of the dynamic forecast for the MSPE in the cases where we don't utilize the forecast provided from the GARCH (1, 1) model, $\hat{h}_{t+j}$. But we observe that the first MSPE is slightly better in the case of the static forecasting, while the difference is more apparent in the case of the second MSPE. It is even more ameliorated and much smaller than that which we get when we suppose homoskedasticity, something that leads us to support the idea that the utilization of GARCH models for describing and forecasting the conditional variance of the process is correctly chosen.

Finally, we present the results of the regressions 1-3. We have already mentioned that the regression $e^2_{t+j} = a + bh_{t+j} + m_{t+j}$ is used as a way to test if the squared shock is an unbiased estimator of h_{t+j}. The results for this regression are the same for the dynamic and the static forecast, and give values for a and b that are close to 0 and 1 respectively, but with a low value for R^2 (in appendix II we present the results for this regression for samples of 250, 500, 1000 and 1500 observations). For the other two regressions the difference between the two ways of performing forecasting, static and dynamic, is obvious; while in the dynamic forecast the results from the regressions two and three are disappointing, since there is a large divergence from the desired values 0 and 1 for a and b respectively and apparently low value for R^2, the static forecasting seems to restore the problem. We could say that the utilization of true values for the lagged dependent variables manage the amelioration of the results and proves the actual superiority of the GARCH model in forecasting the future conditional variance, since we don't only get good values for a and b, but also a large value for $R^2 = 0.9547$.

b) Model 2: Forecasting With GARCH (1,1) When The Condition For Finite Fourth Unconditional Moment Is Violated

DYNAMIC FORECASTING

i) $MSPE(j)=E(\hat{h}_{t+j} - y^2_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	80.668638	111.86317	25.553032	60.340054	64.028742
std.dev.	1741.0094	2656.0445	176.97057	746.54478	608.66737

ii) $MSPE(j)=E(\hat{h}_{t+j} - h_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	0.8016463	10.552774	25.070144	13.208372	41.542258
std.dev.	15.202150	174.57808	564.21450	134.98120	563.15405

iii) $MSPE=E(y^2_{t+j} - h_{t+j})^2$ for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	88.612593	104.60611	43.293193	47.397305	35.243020
std.dev.	2057.8727	2223.0670	536.07211	656.88319	366.53120

- **Homoskedastic Model: $MSPE(j)=E(h_{t+j} - \hat{\sigma}^2)^2$ for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	27.560681	46.381565	42.679526	24.397172	37.970445
std.dev.	373.88486	907.72325	715.30982	344.63263	533.45504

Regression #1: e^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…
Regression #2: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…
Regression #3: e^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…

size=500,rep=1000	regression #2	regression #3	regression #1
a	1.9965772	0.6544921	3.7856035
b	0.5484549	0.7214938	-0.3008501
R^2	0.0130141	0.1940555	0.0075855

STATIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}$ -y^2_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	80.668638	123.91531	43.590545	48.404911	33.263833
std.dev.	1741.0094	2512.7331	529.36415	647.24492	316.50426

ii) MSPE(j)=E($\hat{h}_{t+j}$ -h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	0.8016463	2.5929391	0.8426444	0.5217397	0.4086858
std.dev.	15.202150	75.687200	20.689248	6.9038796	5.1702376

iii) MSPE=E(y^2_{t+j} - h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	88.612593	104.60611	43.293193	47.397305	35.243020
std.dev.	2057.8727	2223.0670	536.07211	656.88319	366.53120

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	27.560681	46.381565	42.679526	24.397172	37.970445
std.dev.	373.88486	907.72325	715.30982	344.63263	533.45504

Regression #1: e^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…
Regression #2: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…
Regression #3: e^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…

size=500,rep=1000	regression #2	regression #3	regression #1
a	-0.0432144	0.6544921	0.6216329
b	1.0468756	0.7214938	0.7545969
R^2	0.9905956	0.1940555	0.1924521

Model 2 is considered for its potential theoretical interest as it is a GARCH (1, 1) model in which the assumption for the existence of fourth-moment is violated. Even though the existence of the fourth moment is not required for the results obtain by Lumsdaine (1999) in order the estimators to be consistent and asymptotically normal, we decide to include this special case in order to examine what consequences, if any, would have this absence of finite fourth moment in the forecasting performance of GARCH (1, 1) model. The results from the estimation part (see appendix) verify the claims of Lumsdaine.

As to the dynamic way of forecasting, the forecasts which we get from GARCH (1, 1), when we use as a measure of the forecasting accuracy the $MSPE(j) = E(\hat{h}_{t+j} - e^2_{t+j})^2$, are again better than those of the homoskedastic one, and we observe that the difference between these two alternative models has become more apparent now. The utilization of the squared shock as an unbiased estimator of the true volatility is not again the best possible solution, especially in this case where the fourth moment does not exist, but seems to be the only one when the true volatility is unobserved. But what is of great importance is the fact that even though the absence of finite-fourth moment does not affect the performance of the quasi-maximum estimators, it seems to influence greatly the forecasts that we get by using a GARCH representation for the conditional variance as well as those of a homoskedastic model. If we compare the MSPE, $MSPE(j) = E(\hat{h}_{t+j} - e^2_{t+j})^2$, of model 1 and that of model 2 in the dynamic forecasting we conclude to the fact that the absence of fourth unconditional moment has an effect on the forecasting performance of the GARCH process.

In the case of the static forecast, the results are ameliorated, for the reasons we mention in section VII(a), since we obtain smaller values for the $MSPE(j) = E(h_{t+j} - \hat{h}_{t+j})^2$ and for the $MSPE(j) = E(\hat{h}_{t+j} - e^2_{t+j})^2$, but again inferior to those from the static forecasting when the values of a and b ensure the existence of fourth moment. The static forecasting makes the superiority of the GARCH model in comparison to the homoskedastic one to seem more apparent, and confirms for another time the fact the utilization of an autoregressive conditional heteroskedasticity process for the presentation of our data instead of an homoskedastic one has a logical basis.

We also present the results of the regressions 1-3. As to the previous model 1, the regression $e^2_{t+j} = a + bh_{t+j} + m_{t+j}$ is used as a way to test if the squared shock is an unbiased estimator of h_{t+j}. The results for this regression are the same for the dynamic and the static forecast, and give values for a and b that are close to 0 and 1 respectively but worse than that of model 1, and again with low value for R^2. For the other two regressions the difference between the two ways of performing forecasting, static and dynamic, is obvious; while in the dynamic forecast the results from the regressions two and three are disappointing, since there is a large divergence from the desired values 0 and 1 for a and b respectively and apparently low value for R^2, the static forecasting seems to restore the problem, since we acquire good values for a and b and also a large value for R^2.

c) Model 3: Estimation And Forecasting With IGARCH (1, 1)

DYNAMIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}$ -y^2_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	391422.92	1144521.9	102382.56	537467.46	621682.52
std.dev.	9380616.1	23299264.	1413662.8	11999283.	13979675.

ii) MSPE(j)=E($\hat{h}_{t+j}$ -h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	9436.6144	34616.334	80389.270	104901.07	367885.78
std.dev.	168339.32	760657.67	1642085.2	2289493.0	7406276.8

iii) MSPE=E(y^2_{t+j} - h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	309326.88	1034787.9	224269.15	246701.24	376322.23
std.dev.	7226838.7	23819070.	3616018.4	4129965.0	5910741.2

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	320089.48	258546.11	261948.97	237957.78	132427.63
std.dev.	5935422.1	4428775.4	3899712.9	3633540.4	1910334.1

Regression #1: e^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…
Regression #2: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…
Regression #3: e^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…

size=500,rep=1000	regression #2	regression #3	regression #1
a	9.3131519	19.028689	13.230903
b	4.6523193	0.7530529	4.4693883
R^2	0.1050917	0.1701713	0.0315440

STATIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}$ -y^2_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	391422.92	1066060.5	206819.72	267777.03	392185.66
std.dev.	9380616.1	23057798.	3166010.8	4839952.1	5965857.7

ii) MSPE(j)=E($\hat{h}_{t+j}$ -h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	9436.6144	7413.5839	3601.4398	7133.7798	3107.0126
std.dev.	168339.32	128722.43	53880.252	144930.69	39127.625

iii) MSPE=E(y^2_{t+j} - h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	309326.88	1034787.9	224269.15	246701.24	376322.23
std.dev.	7226838.7	23819070.	3616018.4	4129965.0	5910741.2

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	320089.48	258546.11	261948.97	237957.78	132427.63
std.dev.	5935422.1	4428775.4	3899712.9	3633540.4	1910334.1

Regression #1: e^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…

Regression #2: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…

Regression #3: e^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…

size=500,rep=1000	regression #2	regression #3	regression #4
a	-1.3454663	19.028689	18.064829
b	1.1145113	0.7530529	0.8383891
R^2	0.9895509	0.1701713	0.1686106

Model 3 is considered for its theoretical interest, since it belongs to the family of the Integrated GARCH model, where α + β = 1. This model founded on the boundary between regions 3 and 4, where the e_t is strongly stationary but not covariance stationary, since $E(e_t^2) = \infty$. It is well known that the IGARCH (1, 1) model has often been employed in empirical applications, and for this reason it is of great interest to see how good are the estimators that we acquire as well as the forecasting performance of the process. The results obtained from the estimation part are consistent with the argues of Lumsdaine (1999), since the quasi-maximum estimators are found to be consistent and asymptotically normal. As to the forecasting part, we observe that the absence of finite second unconditional moment influence the forecasting accuracy of the two utilized models. It is obvious that the presence of an infinite second moment prevents us from performing accurate forecasts of the conditional variance regardless of the model or the way we use to make the forecasts.

VIII. Conclusion

The empirical evidence of this study indicates that it is difficult to lead up to a unambiguous conclusion about the forecasting performance of a GARCH (1, 1) model. It seems to confirm the suggestions of many studies that present the GARCH models as of limited practical value. If we accept this idea, what is the reason of utilizing these volatility models in order to describe and forecast the progress of the conditional variance? It would be much more convenient for us, by using a simple homoskedastic model, to acquire more accurate forecasts for this measure of risk which considered so crucially important in modern financial theory.

The Monte Carlo experiment provides answers to the above question; the superiority of the GARCH model in forecasting the volatility of financial markets is unquestionable. Its seemingly poor performance is attributed to the way that we choose to compute the statistical criteria utilized for evaluating the forecasts. Since the true future volatility is unobserved, we are forced to replace it by an unbiased estimator. If we had available the future conditional variance, the evaluation of the GARCH forecasts would have been fairly straightforward. The suggestion of this study is that we must support, utilize and rely on the forecasts provided by a GARCH (1, 1) model even though the statistical criteria indicate the opposite.

APPENDIX

I. Estimation and forecasting with GARCH(1,1)

We assume $y_t=\varepsilon_t$, where ε_t/I_{t-1}~$D(0,h_t)$

$E(\varepsilon_t/I_{t-1})=0$ and $var(\varepsilon_t/I_{t-1})=h_t$

with $h_t=\omega+\alpha\varepsilon_{t-1}^2+\beta h_{t-1}$. We denote αo = ct, α = ar and β = gar

1. Region 4: αo =0.25, ar=0.20, gar=0.55, with $\alpha + \beta <1$ and $3(\alpha)^2+ (\beta)^2+2\alpha\beta <1$

size=500,rep=1000	mean	std.dev	kurtosis	skewness
ct	0.2675275	0.1223969	4.498484	1.013952
ar	0.1972171	0.0669083	2.967004	0.207329
gar	0.5341571	0.1574780	3.187160	-0.415760
bias(ct)	-0.0175275	0.1223969	4.498484	-1.013952
bias(ar)	0.0027829	0.0669083	2.967004	-0.207329
bias(gar)	0.0158429	0.1574780	3.187160	0.415760
ct-std	0.1110228	0.0534020	11.11731	2.067713
ar-std	0.0614420	0.0126273	3.285509	0.012769
gar-std	0.1430812	0.0583783	12.64773	2.285053
ct-tstat	2.4786209	0.6605983	3.802455	0.497911
ar-tstat	3.1682354	0.7536319	3.178489	-0.034082
gar-tstat	4.5186224	2.8684324	13.95792	2.397806

2. Estimation and forecasting with GARCH(1,1)

We assume $y_t=\varepsilon_t$, where ε_t/I_{t-1}~$D(0,h_t)$

$E(\varepsilon_t/I_{t-1})=0$ and $var(\varepsilon_t/I_{t-1})=h_t$

with $h_t=\omega+\alpha\varepsilon_{t-1}^2+\beta h_{t-1}$. We denote αo = ct, α = ar and β = gar

Region 4: αo =0.25, ar=0.40, gar=0.50, with $\alpha + \beta <1$ and $3(\alpha)^2+ (\beta)^2+2\alpha\beta >1$

size=500,rep=1000	mean	std.dev	kurtosis	skewness
ct	0.2607709	0.0790535	5.456581	0.919623
ar	0.3917348	0.0824681	3.028552	0.045670
gar	0.4968328	0.0824677	3.777029	-0.047017
bias(ct)	-0.0107709	0.0790535	5.456581	-0.919623
bias(ar)	0.0082652	0.0824681	3.028552	-0.045670
bias(gar)	0.0031672	0.0824677	3.777029	0.047017
ct-std	0.0755805	0.0201940	3.923819	0.698307
ar-std	0.0771361	0.0120335	3.246350	0.252237
gar-std	0.0773831	0.0169607	4.133701	0.851328
ct-tstat	3.4571809	0.5355649	3.052485	0.111826
ar-tstat	5.0899790	0.8206897	3.053394	-0.058311
gar-tstat	6.7765080	2.1304871	17.18134	1.859803

3. Estimation and Forecasting with IGARCH(1,1)

We assume $y_t=\varepsilon_t$, where $\varepsilon_t/I_{t-1}\sim D(0,h_t)$

$E(\varepsilon_t/I_{t-1})=0$ and $var(\varepsilon_t/I_{t-1})=h_t$

with $h_t=\omega+\alpha\varepsilon_{t-1}^2+\beta h_{t-1}$. We denote αo = ct, α = ar and β = gar

Boundary between region 4-3: αo =0.25, ar=0.15, gar=0.85, with ar+gar=1

size=500,rep=1000	mean	std.dev	kurtosis	skewness
ct	0.5072511	0.5290977	94.41957	7.486729
ar	0.1445523	0.0385753	4.359916	0.315433
gar	0.8416084	0.0391908	7.402882	-1.018063
bias(ct)	-0.2572511	0.5290977	94.41957	-7.486729
bias(ar)	0.0054477	0.0385753	4.359916	-0.315433
bias(gar)	0.0083916	0.0391908	7.402882	1.018063
ct-std	0.2963402	0.3142643	155.4544	10.22950
ar-std	0.0366587	0.0073464	3.914720	0.581948
gar-std	0.0387541	0.0131591	26.39838	3.230288
ct-tstat	1.6642779	0.4135621	3.256002	0.042451
ar-tstat	3.9434362	0.7438163	3.463494	-0.091024
gar-tstat	23.691012	6.7990543	3.497347	0.326659

4. Estimation and forecasting with IGARCH(1,1)

We assume $y_t=\varepsilon_t$, where $\varepsilon_t/I_{t-1}\sim D(0,h_t)$

$E(\varepsilon_t/I_{t-1})=0$ and $var(\varepsilon_t/I_{t-1})=h_t$

with $h_t=\omega+\alpha\varepsilon_{t-1}^2+\beta h_{t-1}$. We denote αo = ct, α = ar and β = gar

Boundary between region 4-3: αo =0.25, ar=0.05, gar=0.95, with ar+gar=1

size=500,rep=1000	mean	std.dev	kurtosis	skewness
ct	0.7411126	2.0809651	509.7877	20.34789
ar	0.0468959	0.0217278	13.69141	1.658200
gar	0.9412446	0.0580046	150.7336	-10.74200
bias(ct)	-0.4911126	2.0809651	509.7877	-20.34789
bias(ar)	0.0031041	0.0217278	13.69141	-1.658200
bias(gar)	0.0087554	0.0580046	150.7336	10.74200
ct-std	0.4471428	0.7910514	186.6749	11.52652
ar-std	0.0191397	0.0063578	21.92162	2.891867
gar-std	0.0240478	0.0236603	170.0274	10.92713
ct-tstat	1.6457272	0.7104495	5.018102	0.833293
ar-tstat	2.3921006	0.6799367	3.582909	-0.307886
gar-tstat	49.183341	18.858864	4.436135	0.554258

DYNAMIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}$ -y^2_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	40031.543	76026.108	66226.573	183166.96	101169.88
std.dev.	186615.42	894936.92	778712.01	2386574.4	1391835.7

ii) MSPE(j)=E($\hat{h}_{t+j}$ -h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	2267.4030	2439.5551	3399.3123	5887.1796	12603.057
std.dev.	52211.724	53452.228	78891.893	141856.91	311786.14

iii) MSPE=E(y^2_{t+j} - h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	41326.310	63654.609	52415.475	159316.71	65961.156
std.dev.	202047.44	547752.57	395644.21	2108058.4	511657.65

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	23242.794	23492.085	22152.731	19989.855	19657.401
std.dev.	362425.93	375613.63	334855.19	265664.13	182929.13

Regression 1: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…

Regression 2: y^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…

Regression 3: y^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…

size=500,rep=1000	regression #1	regression #2	regression #3
a	-165.86617	50.069687	-411.73364
b	9.0475207	0.7018961	14.847259
R^2	0.2481897	0.0623202	0.0301118

STATIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}$ -y^2_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	40031.543	76639.940	62428.055	168593.30	69293.852
std.dev.	186615.42	897635.17	662976.66	2198015.7	593438.61

ii) MSPE(j)=E($\hat{h}_{t+j}$ -h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	2267.4030	2278.1251	1902.7438	1452.8953	1067.5152
std.dev.	52211.724	53716.087	43551.397	28249.971	12950.933

iii) MSPE=E(y^2_{t+j} - h_{t+j})^2 for j=1,2,3,5,10

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	41326.310	63654.609	52415.475	159316.71	65961.156
std.dev.	202047.44	547752.57	395644.21	2108058.4	511657.65

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=500,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	23242.794	23492.085	22152.731	19989.855	19657.401
std.dev.	362425.93	375613.63	334855.19	265664.13	182929.13

obs=500,rep=1000	Q(0,025)	Q(0,975)
ctt-stat	0.4022022	3.1680457
art-stat	0.8160819	3.6357066
gart-stat	15.181642	89.472387

Regression 1: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…
Regression 2: y^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…
Regression 3: y^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…

size=500,rep=1000	regression #1	regression #2	regression #3
a	-10.665480	50.069687	42.746027
b	1.2992949	0.7018961	0.9142442
R^2	0.9571414	0.0623202	0.0603788

II. Running the regression y_{t+j}^2 =a + bh_{t+j} + n_{t+j} in order to examine if the squared shock e_t^2 is an unbiased estimator of true conditional variance h_{t+j}

a) GARCH (1,1) with z_t ~ N(0,1)

Regression 2:

$$y_{t+j}^2 = a + bh_{t+j} + n_{t+j} \text{ , } j=1,2,\ldots$$

size=500,rep=1000	regression #2
a	0.1339942
b	0.8572372
R^2	0.0679507

size=1000,rep=1000	regression #2
a	0.0853890
b	0.9092759
R^2	0.0731900

size=1500,rep=1000	regression #2
a	0.0656341
b	0.9294964
R^2	0.0757475

b) IGARCH(1,1) with z_t~N(0,1)

size=250,rep=1000	regression #2
a	8.0575266
b	0.6291477
R^2	0.1840895

size=500,rep=1000	regression #2
a	5.1440961
b	0.6886564
R^2	0.2320105

size=1000,rep=1000	regression #2
a	5.3656582
b	0.7148060
R^2	0.2575208

size=1500,rep=1000	regression #2
a	5.7160072
b	0.7256345
R^2	0.2687411

c) GARCH(1,1) with z_t~ t-student(5)

size=250,rep=1000	regression #2
a	0.8890817
b	0.9677933
R^2	0.0873648

size=500,rep=1000	regression #2
a	0.8333022
b	1.0830922
R^2	0.1045488

size=1000,rep=1000	regression #2
a	0.7729615
b	1.1175269
R^2	0.1084866

size=1500,rep=1000	regression #2
a	0.7431182
b	1.1621180
R^2	0.1173120

°

° The procedure was programmed in Eviews 4.0 software

III. Estimation and forecasting with GARCH(1,1)

We assume $y_t=\varepsilon_t$, where $\varepsilon_t/I_{t-1}\sim D(0,h_t)$

$E(\varepsilon_t/I_{t-1})=0$ and $var(\varepsilon_t/I_{t-1})=h_t$

with $h_t=\omega+\alpha^* \varepsilon_{t-1}^2+\beta^* h_{t-1}$. We denote αo = ct, α = ar and β = gar

1. Region 4:ct=0.25, ar=0.20, gar=0.55, with ar+gar<1

Obs=250,rep=1000	mean	std.dev	kurtosis	skewness
Ct	0.2571869	0.1346345	3.860012	0.883402
Ar	0.1971130	0.0881009	2.950798	0.383942
Gar	0.5421934	0.1833839	2.762591	-0.330645
bias(ct)	-0.0071869	0.1346345	3.860012	-0.883402
bias(ar)	0.0028870	0.0881009	2.950798	-0.383942
bias(gar)	0.0078066	0.1833839	2.762591	0.330645
Ctstd	0.1831536	0.2222933	361.9820	16.45374
Atstd	0.0950457	0.0243695	2.868855	-0.036091
garstd	0.2431243	0.3156860	608.1393	22.33379
cttstat	1.5732194	0.5568174	3.421150	0.556468
artstat	2.0248224	0.6758899	3.184052	0.167051
gartstat	3.1540871	2.6534869	54.74312	4.617296

i)**MSE(j)=E($\hat{h}_{t+j}$ - y_{t+j}^2)^2 for j=1,2,3,5,10**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
Mean	2.0465730	2.5901899	2.3979843	2.5588621	2.8044210
std.dev.	9.9343670	10.983790	10.281098	13.326115	14.729544

ii)**MSE(j)=E($\hat{h}_{t+j}$ - h_{t+j})^2 for j=1,2,3,5,10**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
Mean	0.0871942	0.1734569	0.2129174	0.2606958	0.2417653
std.dev.	1.1506904	1.4520915	0.9022505	1.1363094	1.3666689

iii)**MSE(j)=E($\hat{h}_{t+j}$ - σ^2)^2 for j=1,2,3,5,10 where σ^2=var(y_t)**

Obs=250,rep=1000	sSE(1)	SE(2)	SE(3)	SE(5)	SE(10)
Mean	0.4352040	0.2817778	0.1926262	0.1007509	0.0312036
std.dev.	4.1997542	2.8626751	1.9773807	0.9868167	0.2560636

Obs=250,rep=1000	Q(0,025)	Q(0,975)
ctt-stat	0.6124124	2.8030052
art-stat	0.7041731	3.4191523
gart-stat	0.4442922	9.9420232

2. Region 3: ct=0.20, ar=0.9, gar=0.3

Obs=250,rep=1000	mean	std.dev	kurtosis	skewness
ct	0.2468819	0.1274965	5.385927	1.661366
ar	0.8242672	0.2685130	2.552783	0.555143
gar	0.3165695	0.1216023	2.309952	-0.209870
bias(ct)	0.0031181	0.1274965	5.385927	-1.661366
bias(ar)	0.0757328	0.2685130	2.552783	-0.555143
bias(gar)	-0.0165695	0.1216023	2.309952	0.209870
ctstd	0.0995610	0.0442733	2.799231	0.102214
atstd	0.1823869	0.0399483	3.173673	0.239382
garstd	0.0880977	0.0297703	4.537377	0.857531
cttstat	4.560E+98	2.139E+99	4.143861	0.690305
artstat	4.4756256	0.8430296	1.744590	0.093090
gartstat	3.8000364	1.4413690	1.915304	-0.281739

i)**MSE(j)=E(** $\hat{h}_{t+j}$ **-** y^2_{t+j} **)^2 for j=1,2,3,5,10**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	938.28805	984.04592	4600.1326	23048.868	1984500.8
std.dev.	3212.5596	2873.7170	15567.235	76373.098	6214474.7

ii)**MSE(j)=E(** $\hat{h}_{t+j}$ **-** h_{t+j}**)^2 for j=1,2,3,5,10**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	64.979189	1424.3399	2769.3529	21300.342	1979194.1
std.dev.	240.38371	5211.3687	9006.0333	67224.917	6198692.2

iii)**MSE(j)=E(** $\hat{h}_{t+j}$ **-** σ^2 **)^2 for j=1,2,3,5,10 where σ^2=var(y_t)**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	11568.682	10193.151	9091.4947	15984.556	1794852.9
std.dev.	31618.170	30302.835	29831.016	35055.908	5595722.1

obs=250,rep=1000	Q(0,025)	Q(0,975)
Ctt-stat	1.1868372	9.530E+99
Art-stat	3.0868374	6.1282282
gart-stat	1.2213135	6.0359799

3. Region 2: ct=0.20, ar=1.6, gar=0.2

Obs=250,rep=1000	mean	std.dev	kurtosis	skewness
ct	3.5100588	8.0414661	4.198858	1.788218
ar	1.5948339	0.5891584	2.700197	0.568705
gar	0.2209408	0.1227167	2.065807	0.508506
bias(ct)	-3.2600588	8.0414661	4.198858	-1.788218
bias(ar)	0.0051661	0.5891584	2.700197	-0.568705
bias(gar)	-0.0209408	0.1227167	2.065807	-0.508506
ctstd	0.1694709	0.1316415	3.070260	0.756251
atstd	0.2867853	0.1200670	3.896286	1.609410
garstd	0.0548274	0.0237659	2.415298	0.445639
cttstat	3.061E+99	7.497E+99	1.688411	-0.171308
artstat	5.6208482	0.9367395	2.068602	-0.745793
gartstat	3.8519838	0.6251948	1.504515	0.436313

i)**MSE(j)=E($\hat{h}_{t+j}$ - y^2_{t+j})^2 for j=1,2,3,5,10**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	1.532E+11	1.257E+12	4.877E+10	3.724E+13	9.511E+15
std.dev.	3.418E+11	2.785E+12	1.041E+11	8.317E+13	2.066E+16

ii)**MSE(j)=E($\hat{h}_{t+j}$ - h_{t+j})^2 for j=1,2,3,5,10**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	805643700	3.350E+11	3.374E+12	5.771E+12	9.491E+15
std.dev.	1.730E+09	7.476E+11	7.478E+12	1.196E+13	2.060E+16

iii)**MSE(j)=E($\hat{h}_{t+j}$ - σ^2)^2 for j=1,2,3,5,10**

Obs=250,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	4.231E+11	1.314E+12	4.040E+12	3.751E+13	9.575E+15
std.dev.	9.448E+11	2.923E+12	8.947E+12	8.252E+13	2.078E+16

IV. Estimation and forecasting with GARCH(1,1)

We assume $y_t=\varepsilon_t$, where $\varepsilon_t/I_{t-1}\sim D(0,h_t)$

$E(\varepsilon_t/I_{t-1})=0$ and $var(\varepsilon_t/I_{t-1})=h_t$

with $h_t=\omega+\alpha\varepsilon_{t-1}^2+\beta h_{t-1}$. We denote αo = ct, α = ar and β = gar

Region 4: αo =0.25, ar=0.20, gar=0.55, with ar+gar<1

size=100,rep=1000	mean	std.dev	kurtosis	skewness
ct	0.2907009	0.1801373	7.084218	1.558221
ar	0.3930813	0.1786923	3.266706	0.518824
gar	0.4716926	0.1823276	2.674364	-0.171429
bias(ct)	-0.0407009	0.1801373	7.084218	-1.558221
bias(ar)	0.0069187	0.1786923	3.266706	-0.518824
bias(gar)	0.0283074	0.1823276	2.674364	0.171429
ct-std	0.1778979	0.1048984	7.203551	1.739290
ar-std	0.1651514	0.0532704	4.537564	0.791025
gar-std	0.1814675	0.0913725	10.03950	2.209989
ct-tstat	1.6628670	0.5184004	3.836052	0.622864
ar-tstat	2.3424549	0.7432170	3.062458	0.254599
gar-tstat	3.2245406	2.0744254	6.092882	1.371826

DYNAMIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}-y_{t+j}^2$)^2 for j=1,2,3,5,10

obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	79.743701	197.53269	220.22468	538.62784	3738.1506
std.dev.	1094.7634	4205.5551	5517.0447	14080.574	113316.57

ii) MSPE(j)=E($\hat{h}_{t+j}-h_{t+j}$)^2 for j=1,2,3,5,10

obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	31.879944	69.673552	170.98586	312.52266	3916.1930
std.dev.	906.00898	1728.4832	4550.3962	8549.7739	112179.81

iii) MSPE=E($y_{t+j}^2-h_{t+j}$)^2 for j=1,2,3,5,10

obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	60.560830	69.079768	31.842275	66.140612	219.87954
std.dev.	1047.5148	920.16474	222.15404	1064.1300	6240.1131

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

Obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	30.401096	36.968015	26.228826	44.846236	305.39274
std.dev.	449.73015	532.82261	334.97981	707.80932	7982.4615

Regression 1: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…

Regression 2: y^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…

Regression 3: y^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…

size=100,rep=1000	regression #1	regression #2	regression #3
a	-18.468704	2.9323788	-6.4077171
b	10.387543	-0.1602969	8.5373952
R^2	0.3143456	0.1018420	0.1513578

STATIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}$ -y^2_{t+j})^2 for j=1,2,3,5,10

obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	79.743701	128.21615	37.478151	153.19704	376.98617
std.dev.	1094.7634	1990.5099	233.58284	3575.4345	10680.751

ii) MSPE(j)=E($\hat{h}_{t+j}$ -h_{t+j})^2 for j=1,2,3,5,10

obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	31.879944	20.896053	4.4559801	39.024444	32.014842
std.dev.	906.00898	403.54340	51.327063	1100.0512	679.29214

iii) MSPE=E(y^2_{t+j} - h_{t+j})^2 for j=1,2,3,5,10

obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	60.560830	69.079768	31.842275	66.140612	219.87954
std.dev.	1047.5148	920.16474	222.15404	1064.1300	6240.1131

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=100,rep=1000	SPE(1)	SPE(2)	SPE(3)	SPE(5)	SPE(10)
mean	30.401096	36.968015	26.228826	44.846236	305.39274
std.dev.	449.73015	532.82261	334.97981	707.80932	7982.4615

obs=100,rep=1000	Q(0,025)	Q(0,975)
ctt-stat	0.7539936	2.8151181
art-stat	1.0138856	3.8688775
gart-stat	0.3476028	8.3467788

Regression 1: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…
Regression 2: y^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…
Regression 3: y^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…

size=100,rep=1000	regression #1	regression #2	regression #3
a	-0.2470597	2.9323788	3.1922213
b	1.2956111	-0.1602969	-0.4034980
R^2	0.9372535	0.1018420	0.1047704

V. Estimation and forecasting with GARCH(1,1)

We assume $y_t=\varepsilon_t$, where $\varepsilon_t/I_{t-1}\sim D(0,h_t)$

$E(\varepsilon_t/I_{t-1})=0$ and $var(\varepsilon_t/I_{t-1})=h_t$

with $h_t=\omega+\alpha\varepsilon_{t-1}^2+\beta h_{t-1}$. We denote αo = ct, α = ar and β = gar

Region 4: αo =0.25, ar=0.20, gar=0.55, with ar+gar<1

size=200,rep=1000	mean	std.dev	kurtosis	skewness
ct	0.2833091	0.1402242	5.590914	1.254727
ar	0.3892930	0.1269543	2.999718	0.221250
gar	0.4843046	0.1411477	3.686637	-0.280975
bias(ct)	-0.0333091	0.1402242	5.590914	-1.254727
bias(ar)	0.0107070	0.1269543	2.999718	-0.221250
bias(gar)	0.0156954	0.1411477	3.686637	0.280975
ct-std	0.1221314	0.0533597	4.847657	1.129881
ar-std	0.1158418	0.0282987	3.759125	0.479177
gar-std	0.1216340	0.0434987	6.962421	1.614211
ct-tstat	2.3234060	0.5638368	4.394722	0.605743
ar-tstat	3.3523982	0.7734613	3.319101	0.116745
gar-tstat	4.5503899	2.4397224	15.49686	2.202699

DYNAMIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}-y_{t+j}^2)^2$ for j=1,2,3,5,10

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	32.695700	31.881659	32.468159	65.677475	74.465816
std.dev.	274.33955	287.19066	281.50295	733.64902	1189.0803

ii) MSPE(j)=E($\hat{h}_{t+j}-h_{t+j})^2$ for j=1,2,3,5,10

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	1.9221152	8.8521407	13.648998	19.884119	35.052824
std.dev.	28.588099	108.73787	153.21482	225.58891	298.72846

iii) MSPE=E($y_{t+j}^2-h_{t+j})^2$ for j=1,2,3,5,10

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	24.637105	21.333108	26.051453	36.117133	46.913594
std.dev.	148.86740	189.99543	234.97502	370.09850	568.03660

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	16.848112	13.370022	14.738813	15.277296	25.710358
std.dev.	120.34677	65.681473	88.651022	136.60129	208.41977

Regression 1: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…

Regression 2: y^2_{t+j} =a + bh_{t+j} + n_{t+j} , j=1,2,…

Regression 3: y^2_{t+j} =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…

size=200,rep=1000	regression #1	regression #2	regression #3
a	3.0037112	2.5558005	6.8318703
b	1.2737369	-0.1321762	-0.2284874
R^2	0.3127950	0.0982598	0.1594447

STATIC FORECASTING

i)MSPE(j)=E($\hat{h}_{t+j}$ -y^2_{t+j})^2 for j=1,2,3,5,10

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	32.695700	24.926600	26.829581	34.250988	46.972709
std.dev.	274.33955	217.64286	226.18220	331.12045	653.08234

ii) MSPE(j)=E($\hat{h}_{t+j}$ -h_{t+j})^2 for j=1,2,3,5,10

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	1.9221152	1.0475305	0.6032094	0.7152167	1.8479474
Std.dev.	28.588099	11.695660	3.6903530	7.8569616	19.395147

iii) MSPE=E(y^2_{t+j} - h_{t+j})^2 for j=1,2,3,5,10

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	24.637105	21.333108	26.051453	36.117133	46.913594
std.dev.	148.86740	189.99543	234.97502	370.09850	568.03660

- **Homoskedastic Model: MSPE(j)=E(h_{t+j} - $\hat{\sigma}^2$)^2 for j=1,2,3,5,10**

obs=200,rep=1000	SE(1)	SE(2)	SE(3)	SE(5)	SE(10)
mean	16.848112	13.370022	14.738813	15.277296	25.710358
std.dev.	120.34677	65.681473	88.651022	136.60129	208.41977

Regression 1: h_{t+j}= a+ b$\hat{h}_{t+j}$ + e_{t+j} , j=1,2,…
Regression 2: y_{t+j}^2 =a + bh_{t+j} + n_{t+j} , j=1,2,…
Regression 3: y_{t+j}^2 =a + b$\hat{h}_{t+j}$ + m_{t+j} , j=1,2,…

size=200,rep=1000	regression #1	regression #2	regression #3
a	-0.1157436	2.5558005	2.6019644
b	1.1355933	-0.1321762	-0.1956672
R^2	0.9624140	0.0982598	0.1010572

obs=200,rep=1000	Q(0,025)	Q(0,975)
ctt-stat	1.3444653	3.5528893
art-stat	1.8898639	4.9710594
gart-stat	1.1567196	10.023322

VI. Results Of The Empirical Application As To The Estimation Part

A. Group of stock indices for the period 1992-1996

Period 1992-1996					
Parameter	**USA**	**Australia**	**UK**	**France**	**Germany**
φ_0	0,0005353 (3,18)	0,0002724 (1,38)	0,0004215 (2,20)	0,0003459 (1,23)	0,0005624 (2,42)
φ_1	0,0495651 (1,66)	0,1368127 (4,68)	0,0628113 (1,97)	0,0208194 (0,73)	0,0236227 (0,78)
α_o (thousands)	0,001867 (1,55)	0,0002417 (2,02)	0,00117 (2,27)	0,00009611 (1,43)	0,01083 (1,23)
α_1	0,0454314 (2,58)	0,0607002 (3,02)	0,0567073 (2,43)	0,0277248 (2,66)	0,0455234 (3,07)
β	0,9088276 (21,51)	0,8693317 (18,42)	0,9183651 (2,43)	0,9637298 (71,24)	0,9270001 (27,85)
Sum ($\alpha_1 + \beta$)	0,954259	0,9300319	0,9750724	0,9914546	0,9725235

Period 1992-1996					
Parameter	**Japan**	**Mexico**	**Switzerland**	**Singapore**	**Belgium**
φ_0	0,0001665 (0,51)	0,0010553 (2,64)	0,0007783 (3,79)	0,0003258 (1,50)	0,0003708 (2,11)
φ_1	-0,0270145 (-0,83)	0,2127678 (6,05)	0,0741209 (2,28)	0,1692701 (4,81)	0,1558048 (4,70)
α_o (thousands)	0,00113 (1,79)	0,001261 (2,44)	0,0006736 (2,80)	0,0007456 (4,24)	0,0002022 (0,99)
α_1	0,0719849 (3,26)	0,1160533 (4,37)	0,1175407 (3,86)	0,2206441 (3,24)	0,0369301 (2,40)
β	0,9005538 (32,32)	0,8597108 (28,21)	0,7430872 (11,06)	0,5848359 (9,44)	0,9056419 (13,17)
Sum ($\alpha_1 + \beta$)	0,9725387	0,9757641	0,8606279	0,80548	0,942572

B. Group of the foreign exchange rates for the period 1992-1996

Period 1992-1996					
Parameter	**Australia**	**Canada**	**Japan**	**Singapore**	**Switzerland**
φ_0	-0,00009651 (-0,72)	0,00006991 (0,97)	-0,00006691 (-0,39)	-0,0001846 (-3,14)	-0,000021 (-0,11)
φ_1	0,0649609 (2,01)	0,0534766 (1,72)	0,028723 (0,92)	-0,0863582 (-2,25)	0,0483358 (1,47)
α_o (thousands)	0,001427 (2,63)	0,00007624 (0,90)	0,001032 (1,75)	0,001054 (4,42)	0,001459 (1,90)
α_1	0,0564306 (3,30)	0,0443004 (3,18)	0,044946 (2,66)	0,3182965 (4,32)	0,0533131 (3,03)
β	0,8816753 (26,54)	0,9466472 (59,20)	0,9300238 (33,94)	0,5137332 (6,05)	0,9197578 (38,58)
Sum (α_1 +β)	0,9381059	0,9909476	0,9749698	0,8320297	0,9730709

Period 1992-1996					
Parameter	**UK**	**France**	**Germany**	**Netherlands**	**Spain**
φ_0	-0,0001258 (-0,92)	-0,00002463 (-0,16)	0,0000321 (0,20)	0,00002909 (0,18)	0,00007397 (0,47)
φ_1	-0,0110547 (-0,36)	0,0430463 (1,36)	0,0343821 (1,07)	0,0266658 (0,83)	-0,0201411 (-0,64)
α_o (thousands)	0,0003505 (2,61)	0,0005845 (2,26)	0,000751 (2,57)	0,0007178 (2,27)	0,001054 (3,27)
α_1	0,0530538 (3,74)	0,0505505 (4,00)	0,0569718 (3,84)	0,0576719 (3,92)	0,0864645 (3,21)
β	0,934803 (71,37)	0,9322209 (83,22)	0,9240472 (57,86)	0,9239626 (57,87)	0,8902507 (34,47)
Sum (α_1 +β)	0,9878568	0,9827714	0,981019	0,9816345	0,9767152

C. Group of stock indices for the period 1997-2001

Period 1997-2001					
Parameter	**USA**	**Australia**	**UK**	**France**	**Germany**
φ_0	0,0007089 (2,25)	0,0006069 (2,41)	0,0004292 (1,39)	0,0008342 (2,19)	0,001062 (2,86)
φ_1	0,023838 (0,79)	0,0024938 (0,06)	0,0669532 (2,22)	0,0523914 (1,75)	0,0374018 (1,33)
α_o (thousands)	0,01146 (2,39)	0,007999 (3,39)	0,003953 (2,69)	0,008318 (2,51)	0,008118 (3,27)
α_1	0,1068279 (2,81)	0,1073489 (1,74)	0,0765809 (4,47)	0,0715773 (3,14)	0,1054873 (4,38)
β	0,8252824 (15,16_	0,7988642 (11,17)	0,8962559 (42,58)	0,8908385 (28,09)	0,8647521 (34,38)
Sum (α_1 +β)	0,9321103	0,9062131	0,9728368	0,9624158	0,9702394

Period 1997-2001					
Parameter	**Japan**	**Mexico**	**Switzerland**	**Singapore**	**Belgium**
φ_0	-0,0002212 (-0,52)	0,001077 (2,47)	0,0007353 (2,51)	0,00009399 (0,21)	0,0005471 (2,08)
φ_1	-0,0345512 (-1,15)	0,1563182 (4,66)	0,0611619 (1,94)	0,1019504 (2,17)	0,1763467 (4,84)
α_o (thousands)	0,01173 (2,38)	0,02286 (2,74)	0,006833 (3,52)	0,02015 (1,22)	0,003456 (3,03)
α_1	0,0805748 (4,10)	0,1725854 (2,93)	0,1341482 (5,51)	0,1902965 (2,83)	0,1687616 (4,58)
β	0,8735663 (28,77)	0,7789435 (13,09)	0,8245226 (32,59)	0,7692546 (8,30)	0,8208543 (26,09)
Sum (α_1 +β)	0,9541411	0,9515289	0,9586708	0,9595511	0,9896159

D. Group of foreign exchange rates for the period 1997-2001

Period 1997-2001					
Parameter	**Australia**	**Canada**	**Japan**	**Singapore**	**Switzerland**
φ_0	0,0003721 (1,94)	0,0001561 (1,68)	0,0002403 (1,12)	0,0001561 (1,98)	0,0001878 (1,00)
φ_1	0,0286006 (0,94)	0,0594204 (2,08)	0,0145078 (0,44)	0,0052902 (0,17)	0,0141088 (0,48)
α_o (thousands)	0,001867 (1,53)	0,0002417 (1,52)	0,00117 (1,93)	0,00009611 (1,00)	0,01083 (0,50)
α_1	0,0570788 (2,52)	0,0475935 (3,23)	0,048168 (2,50)	0,0664324 (2,89)	0,0171447 (0,61)
β	0,9081525 (23,49)	0,9316964 (56,64)	0,933967 (40,16)	0,930774 (45,60)	0,7375541 (1,46)
Sum (α_1 +β)	0,9652313	0,9792899	0,982135	0,9972064	0,7546988

Period 1997-2001					
Parameter	**UK**	**France**	**Germany**	**Netherlands**	**Spain**
φ_0	0,0001061 (0,77)	0,0002953 (1,64)	0,0002939 (1,63)	0,00029 (1,60)	0,0003144 (1,83)
φ_1	0,0405397 (1,36)	0,0242735 (0,87)	0,0254094 (0,92)	0,0249137 (0,89)	0,0230327 (0,85)
α_o (thousands)	0,00113 (1,45)	0,001261 (2,17)	0,0006736 (1,81)	0,0007456 (1,58)	0,0002022 (0,83)
α_1	0,0292726 (2,00)	0,024092 (2,27)	0,0177131 (2,40)	0,0194922 (2,03)	0,0106421 (1,85)
β	0,9242562 (21,32)	0,9438958 (44,70)	0,9653695 (75,54)	0,9618378 (49,87)	0,9843453 (116,29)
Sum (α_1 +β)	0,9535288	0,9679878	0,9830826	0,98133	0,9949874

Note: Numbers in parentheses are t-statistics

VII.

a.

Group of stock indices					
Country	**Sample**	**Period**	**Country**	**Sample**	**Period**
usa	1265	1992-1996	Japan	1236	1992-1996
usa	1257	1997-2001	Japan	1229	1997-2001
Australia	1268	1992-1996	Mexico	1244	1992-1996
Australia	1260	1997-2001	Mexico	1247	1997-2001
England	1265	1992-1996	Switzerland	1259	1992-1996
England	1260	1997-2001	Switzerland	1320	1997-2001
France	1249	1992-1996	Singapore	1254	1992-1996
France	1258	1997-2001	Singapore	1248	1997-2001
Germany	1256	1992-1996	Belgium	1243	1992-1996
Germany	1254	1997-2001	Belgium	1184	1997-2001

b.

Group of foreign exchange rates					
Country	**Sample**	**Period**	**Country**	**Sample**	**Period**
Australia	1259	1992-1996	U.K	1259	1992-1996
Australia	1257	1997-2001	U.K	1257	1997-2001
Canada	1259	1992-1996	France	1259	1992-1996
Canada	1257	1997-2001	France	1257	1997-2001
Japan	1259	1992-1996	Germany	1259	1992-1996
Japan	1257	1997-2001	Germany	1257	1997-2001
Singapore	1259	1992-1996	Netherlands	1259	1992-1996
Singapore	1257	1997-2001	Netherlands	1257	1997-2001
Switzerland	1259	1992-1996	Spain	1259	1992-1996
Switzerland	1257	1997-2001	Spain	1257	1997-2001

References

Baillie, Richard T., and Tim Bollerslev. 1989. "The Message in Daily Exchange Rates: A Conditional Variance Tale." Journal of Business and Economic Statistics 7:297-305.

Baillie, Richard T., and Tim Bollerslev. 1992. "Prediction in Dynamic Models with Time-Dependent Conditional Variances." Journal of Econometrics 52:91-113.

Bollerslev, Tim.1986. "Generalized Autoregressive Conditional Heteroskedasticity." Journal of Econometrics 31:307-27

Bollerslev, Tim.1987. " A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return." Review of Economics and Statistics 69:542-47.

Bollerslev, Tim.1988. "On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process." Journal of Time Series Analysis 9:121-31.

Bollerslev Tim, Ray y. Chou, and Kenneth F. Kroner. 1992. "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence." Journal of Econometrics 52:5-59.

Bollerslev Tim, Robert F. Engle, and Jeffrey M. Wooldridge. 1988. "A Capital Asset Pricing Model with Time Varying Covariances." Journal of Political Economy 96:116-31.

Bollerslev Tim and Jeffrey M. Woodridge. 1992. " Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Varying Covariances." Econometrics Reviews 11:143-72.

Engle, Robert F. and Tim Bollerslev. 1986. "Modeling the Persistence of Conditional Variances." Econometric Reviews 5:1-50.

Engle Robert F. 1982. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." Econometrica 50:987-1007.

Engle Robert F., David M. Lilien, and Russell P. Robins. 1987. "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model." Econometrica 55:391-407.

Engle Robert F. and Victor K. Ng. 1993. "Measuring and Testing the Impact of News on Volatility." Journal of Finance 48:1749-78

Engle Robert F., Victor K. Ng, and Michael Rothschild. 1990. "Asset Pricing with a FACTOR-ARCH Covariance Structure: Empirical Estimates for Treasury Bills." Journal of Econometrics 45:213-37.

Geweke, John. 1986. "Modeling the Persistence Of Conditional Variances: A Comment ." Econometric Reviews 5:57-61.

Closten, Lawrence R., Ravi Jagannathan, and David Runkle. 1993. "Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." Journal of Finance 48:1779-1801.

Hamilton D. James. 1994. “Time Series Analysis.” Princeton University Press

Hsieh, David A. 1989. “Modeling Heteroskedasticity in Daily Foreign-Exchange Rates.” Journal of Business and Economic Statistics 7:307-17.

Mark, Nelson. 1988. “Time Varying Betas and Risk Premia in the Pricing of Foreign Exchange Contracts.” Journal of Financial Economics 22:335-54.

Milhoj, Anders. 1985. “The Moment Structure of ARCH Processes.” Scandinavian Journal of Statistics 12:281-92.

Nelson, Daniel B. 1990. “Stationarity and Persistence in the GARCH(1,1) Model.” Econometrica 59:347-70.

Nelson Daniel B., and Charles Q. Cao. 1992. “Inequality Constraints in the Univariate GARCH Model.” Journal of Business and Economic Statistics 10:229-35.

Pagan, Adrian R., and G. William Schwert. 1990. “Alternative Models for Conditional stock Volatility.” Journal of Econometrics 45:267-90.

Philip Hans Franses and Dick van Dijk. “ Nonlinear Time Series Models in Empirical Finance”

Choo Wei Chong. 1999. “Performance of GARCH Models in Forecasting Stock Market Volatility”. Journal of Forecasting. 18: 333-343.

Robin L. Lumsdaine. 1995. “Finite-Sample Properties of the Maximum Likelihood Estimator in GARCH(1, 1) and IGARCH(1, 1) Models: A Monte Carlo Investigation”. American Statistical Association. 13:1-10.

Vedat Akgiray. “ Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts”. Clarkson University.

Chahgli He and Timo Terasvirta. 1999. “Fourth Moment Structure of the GARCH(p, q) Process”. Econometric Theory. 15:824-846.

Torben G. Andersen and Tim Bollerslev. 1998. “Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts”. 39:885-905

Nelson C. Mark. 1998. “Time-Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts”. Journal of Financial Economics. 22:335-354

G. William Schwert. 1989.” Why Does Stock Market Volatility Change Over Time?”. Thw Journal of Finance. 5:1115-1151

Richard T. Baillie and Ramon P. DeGennaro.1990.”Stock Returns and Volatility”. 25:203-214

Ray Yeutien Chou. 1988.”Volatility Persistence and Stock Valuations: Some Empirical Evidence Using GARCH”. 1988. 3:279-294.

Kenneth D. West and Dongchul Cho. 1994.”The Predictive Ability of Several Models Of Exchange Rate Volatility”. Technical Working Paper Series.

Christopher G. Lamoureux and William D. Lastrapes. 1990. " Persistence in Variance, Structure Change, and the GARCH Model". American Statistical Association. 8:225-233

Robin L. Lumsdaine. 1996." Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1, 1) and Covariance Stationary GARCH(1, 1) Models. Econometrica. 3: 575-596

Daniel B. Nelson. 1991." Conditional Heteroskedasticity in Asset Returns: A New Approach". Econometrica. 59: 347-370.

Stephen J. Brown and Jerold B. Warner. 1985." Using Daily Stock Returns". Journal of Finacial Econometrics. 14:3-31