Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts

Vedat Akgiray

Vedat Akgiray

Clarkson University

Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts*

I. Introduction

This article presents new findings about the temporal behavior of stock market returns and summarizes the results of applying some new time-series models to daily return series. It is discovered that daily series exhibit much higher degrees of statistical dependence than has been reported in previous studies. This finding is the result of recognizing the possibility of nonlinear stochastic processes generating security prices. The dependence structure is then exploited to obtain forecasts of the conditional moments of return distributions. Forecasts of conditional variances in particular are shown to have reasonably high accuracy.

Empirical research on the statistical properties of stock returns dates back to the pioneering works of Mandelbrot (1963) and Fama (1965). These and many other studies since then have shown that time series of daily stock returns exhibit some autocorrelation for short lags. The magnitudes of the autocorrelations, however, are too small to form profitable trading rules. Moreover, since the empirical distributions of returns are significantly different from Gaussian distributions, the statistical significance of the usual

This article presents new evidence about the time-series behavior of stock prices. Daily return series exhibit significant levels of second-order dependence, and they cannot be modeled as linear white-noise processes. A reasonable return-generating process is empirically shown to be a first-order autoregressive process with conditionally heteroscedastic innovations. In particular, generalized autoregressive conditional heteroscedastic GARCH (1, 1) processes fit to data very satisfactorily. Various out-of-sample forecasts of monthly return variances are generated and compared statistically. Forecasts based on the GARCH model are found to be superior.

autocorrelation estimates may be even lower. Consequently, the assumption of serially uncorrelated stock returns has been widely accepted as a safe approximation.

Unless the underlying stochastic process is Gaussian, the lack of serial correlation does not imply statistical independence. However, many empirical studies of stock price behavior have made the stronger assumption of intertemporally independent returns. Based on this assumption, several probability distributions have been suggested to represent the empirical distribution of returns, ranging from infinite-variance stable laws to finite-variance mixture models (Cox and Rubinstein [1985] and Bookstaber and McDonald [1987] provide many references to studies in this area). Some of these models require the additional assumption of identically distributed returns while others allow for random variation in some or all of the parameters. The common characteristics of these models are two: (1) returns are independent, and (2) the return-generating process is a linear process with parameters that are independent of the past realizations of the process. The mean and the variance are the parameters that are of great interest because they are the key variables in theoretical and practical valuation models of finance.

The empirical evidence reported in this article challenges the common assumptions of independence and linearity. In fact, there seem to be no compelling theoretical reasons for assuming either (Neftçi 1984), and their necessities are questionable. On the contrary, the workings of speculative markets suggest that nonlinearities and intertemporal dependence in return series are to be expected. As the discussions and empirical evidence in the articles by Perry (1982), Pindyck (1984), and Poterba and Summers (1986) further imply, the assumption of constant conditional (conditional on the past values) means and variances is both unrealistic and dubious. All that is required by the theory of speculation and competitive markets is that the return-generating process must be representable as a martingale. This requirement is sufficient to guarantee market efficiency (Fama 1970).

This article is organized as follows. The following section investigates the statistical properties of return distributions and identifies the class of return-generating stochastic processes that are consistent with these properties. The third section includes statistical tests of the fit of two conditional heteroscedastic processes, autoregressive conditional heteroscedastic (ARCH) and generalized autoregressive conditional heteroscedastic (GARCH) processes. The fourth section presents various forecasts of market volatility, and compares their accuracies. Conclusions and suggestions for future research make up the fifth and final section of the article.

II. Statistical Analysis

This section includes a comprehensive analysis of the distributional and time-series properties of stock returns. Preliminary time-series analysis is conducted both in the frequency domain and in the time domain. The purpose is to determine whether stock price movements can be adequately represented by linear white-noise processes with independent increments, a common assumption in models of stock prices. The terms "white noise," "random walk," and "linear process" are sometimes applied outside of their precise meanings in mathematical statistics. Therefore, it may be useful expositionally to define these terms first.

Let $m = E[x_t]$ and $c_s = E[x_{t+s}x_t] - m^2$ denote the mean and the covariance function of a second-order stochastic process $\{x_t\}$. The process is said to be (weakly, or second-order) stationary if m and c_s do not depend on t. The mean can be set equal to zero without loss of generality. If $c_s = 0$ for all $s > 0$ (i.e., x_{t+s} and x_t are uncorrelated), then the process is "white noise." It is important to note that whiteness does not imply independence between x_t and x_{t+s} (unless it is a Gaussian white noise). In other words, zero autocorrelation does not necessarily mean that the probability distribution of x_{t+s} is independent of realized values of x_t. This is true even if the $\{x_t\}$ have identical (unconditional) distributions. If, in addition, x_{t+s} and x_t are statistically independent, then the process is called "strict white noise" or "purely random." If the process x_t is strict white noise, the processes $\{|x_t|\}$ and $\{x_t^2\}$ are also strict white noise. The term "random walk" as used in the finance literature is synonymous with strict white noise. Sometimes, however, it is used confusingly to mean white noise only. Detailed discussions of some of these issues and their implications for modeling financial time series can be found in Taylor (1986).

Finally, a stationary process $\{x_t\}$ is said to be a "linear process" if it can be expressed as a linear function of strict white noises $\{u_t\}$ as $x_t = \sum_{s=0}^{\infty} a_s u_{t-s}$. Stationary normal processes are always linear but other white-noise processes need not be linear (Priestley 1981). For finite-variance processes, linearity implies certain restrictions on the covariance structure. In general, significant sample estimates of $E[x_t x_{t+s} x_{t+s+r}]$, or high values for $E[x_t^2 x_{t+s}^2]$ for large s, would indicate a nonlinear process. A nonlinear process can be white noise but not strict white noise.

A. The Data

The data are obtained from the Center for Research in Security Prices (CRSP) tapes, and they contain 6,030 daily returns on the CRSP value-weighted and equal-weighted indices covering the period from

January 1963 to December 1986. To conform with the literature and to avoid some computational problems, return is defined as the natural logarithm of value relatives; $R_t = \log(I_T/I_{t-1})$, where dividends are part of total value. For small values of R_t, such as in daily data, this definition is very similar to the arithmetic rate of return.

The 24-year period between 1963 and 1986 involves a sample of 6,030 observations. This sample is divided into four different periods of 6 years each, and each period is analyzed separately. The partitioning of the data is motivated by the observation that the series do not exhibit homogeneous behaviour over the entire 24-year period. This particular partition is somewhat arbitrary and not necessarily the best scheme to obtain homogeneity. Nevertheless, plots of monthly sample medians and interquartile ranges during each 6-year period do not indicate apparent nonhomogeneities. Conversely, as the subsequent empirical evidence will show, there are significant statistical differences between the four periods, and the entire series may not be represented by a stationary process with constant (unconditional) parameters. Each period contains about 1,500 observations, which is a sufficiently large sample. Nevertheless, for purposes of completeness and comparison, statistical analyses are conducted and findings are reported for the entire sample as well.

The results are very similar for the two index series and therefore they are reported only for the value-weighted index. Whenever interesting and notable differences in the calculated values are found, reference to the equal-weighted index results is made. In addition to daily data, some of the analysis is conducted also for weekly and monthly data. They are summarized at the end of the paper.

B. Statistical Findings

In table 1, a wide range of descriptive statistics for the return series $\{R_t\}$, $t = 1, \ldots, T$, for the four periods are reported. These include the following distributional parameters: mean, variance, skewness, kurtosis, range, median, interquartile range (IQR), Kolomogorov-Smirnov D-statistic for the null hypothesis of normality, and the maximum log-likelihood function value when a normal distribution is fitted to data. Also included are statistics to test the null hypothesis of strict white noise both in the time domain (Fisher's kappa and Bartlett's Kolmogorov-Smirnov-type statistic) and in the frequency domain (Ljung-Box portmanteau test).

The sample moments in all four periods indicate that the empirical distributions have heavy tails and sharp peaks at the center compared to the normal distribution. The Kolmogorov-Smirnov test leads to the rejection of normality in every sample. The Kiefer-Salmon (1983) tests for normal kurtosis (zero-excess kurtosis) and normal skewness (zero) also reject the normality hypothesis. Zero skewness cannot be rejected

TABLE 1   Sample Statistics on Daily Return Series

Statistic	Period				
	1963–68	1969–74	1975–80	1981–86	1963–86
Sample size	1,485	1,513	1,516	1,516	6,030
Mean (thousands)	.5311	−.1692	.7496	.5792	.4224
t (mean = 0)	3.881	−.7317	3.6610	2.7453	4.229
Variance (thousands)	.0278	.0808	.0635	.0675	.0602
Skewness	−.1197	.3556	.0162	.1312	.1512
Kurtosis	7.9016	5.7217	4.4773	5.2156	5.874
Range	.0672	.0879	.0762	.0878	.0969
Median (thousands)	.8010	−.0271	.8560	.3985	.5505
IQR	.0053	.0104	.0079	.0076	.0082
D-statistic	.0752	.0647	.0296	.0442	.0540
Log-likelihood	5,683.4	4,981.9	5,174.4	5,128.8	20,745.2
Fisher's kappa	9.6710	8.2318	8.9878	7.4501	9.8389
Bartlett's test	.1292	.2078	.1390	.1004	.1437
LB(6)	66.98	164.03	67.32	32.10	312.51
LB(12)	70.90	173.39	71.71	39.74	318.40
LB(24)	80.39	193.53	86.80	48.38	341.84

NOTES.—The critical values of the D statistic are .021, .023, and .026 for significance levels of .10, .05, and .01, respectively. The critical values for Fisher's kappa are 8.83, 9.52, and 11.10 for significance levels of .10, .05, and .01, respectively. The critical values for Bartlett's test statistic are .0496 and .0595 for significance levels of .05 and .01, respectively. LB(n) is the Ljung-Box statistic at lag n, distributed as a chi-squared variate with n degrees of freedom.

in the first and third periods but normal kurtosis is safely rejected in all periods. That daily stock returns are not normally distributed is a well-known result. Several leptokurtic distributions have therefore been proposed as more descriptive models than the normal.

All of the findings for the four periods also hold for the entire 1963–86 period. However, the sample moments from different periods are generally significantly different. The differences between the means, medians, and the interquartile ranges may be interpreted as distribution-free indicators of nonhomogeneity in the entire series. Consequently, separate analyses of the four periods seem to be justified and this should not result in any loss of generality.

Before any probability distribution model is fitted to data, the underlying assumptions of the model have to be verified empirically. Almost all of the popular models of stock returns require that returns be independent random variables, and many also require that they be identically distributed. In order to test the hypothesis of independence, the periodogram of each series is estimated, and Fisher's kappa and Bartlett's test statistics are calculated (see Fuller 1976). Fisher's kappa is the largest periodogram ordinate divided by the average of these ordinates. Bartlett's procedure calculates the normalized cumulative periodogram, which is a uniform (0, 1) variate under the null hypothesis.[1] As the entries in table 1 indicate, the hypothesis of independence is rejected in periods 1 and 3 by Fisher's test, and in all four periods by Bartlett's test. Visual inspections of the periodograms support this conclusion. In the time domain, Ljung-Box tests also lead to the same result. The test statistic is calculated for lags up to 60 days, and those for lags 6, 12, and 24 are reported in table 1. The null hypothesis of strict white noise is rejected in all cases, even at lower significance levels than 1%. The conclusion must be that daily return series are not made up of independent variates.

In order to investigate the reasons for lack of independence, the sample autocorrelation functions may be analyzed. The estimated autocorrelations for the series $\{R_t\}$, $\{|R_t|\}$, and $\{R_t^2\}$ for the whole period from 1963 to 1986 are shown in figure 1. When the four periods are analyzed separately, similar results are obtained. The return series display high first-lag autocorrelations (ranging from 0.18 in period 4 to 0.31 in period 2) and apparently insignificant autocorrelations at longer lags. Using the usual approximation of $1/\sqrt{T}$ as the standard error of these estimates, all of the first-lag autocorrelations are greater than $7/\sqrt{T}$. Even though the $1/\sqrt{T}$ value may be an understatement of the standard error (due to the nonnormality of returns), seven times this

---

1. Strictly speaking, both procedures are tests of the hypothesis of normal (strict) white noise but in large samples they provide good approximations for general tests of independence. Bartlett's test is known to be more robust to distributional assumptions.
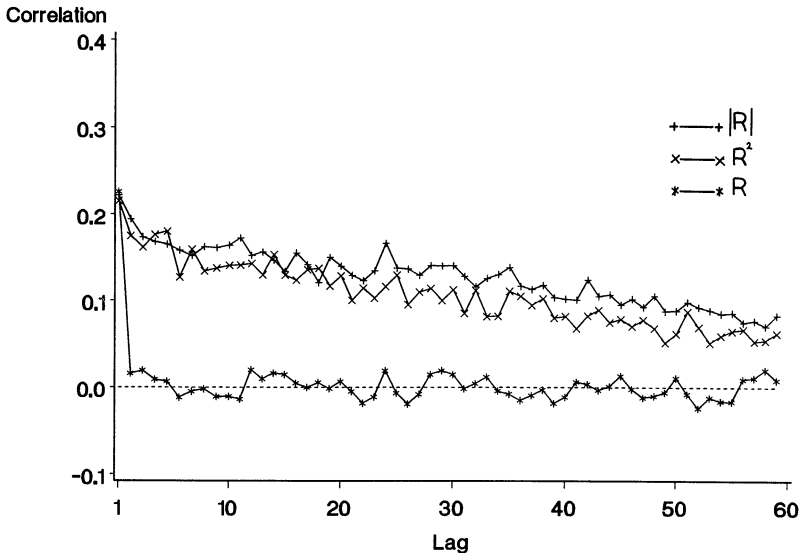
FIG. 1.—Autocorrelation functions for return series

value is a sufficiently large confidence bound. The autocorrelations in the absolute and squared return series are always much higher than those in the return series, and they are consistently significantly positive for lags up to 60 days. The autocorrelation between absolute returns, however, is generally higher than that in squared returns. As the lag increases, both autocorrelation functions slowly decay but never go below $3/\sqrt{T}$ at any lag up to 60. This finding agrees with those reported in the classic work of Fama (1965), that large price changes are followed by large changes, and small, by small, of either sign. More generally, the distribution of the next absolute or squared return depends not only on the current return but also on several previous returns. This is a conclusive rejection of the hypothesis that return series are strict white-noise processes. Indeed, the presence of significant first-lag correlations in $\{R_t\}$ implies the rejection of white noise too.

C. Implications for Model Building

The presence of linear dependence in daily return series of market indexes can be attributed to various market phenomena and anomalies. The presence of a common market factor, the problem of thin trading in some stocks, the speed of information processing by market participants, and day-of-the-week effects could contribute partially to the observed first-order autocorrelations. For example, the autocorrelations in the series of equal-weighted index returns are generally higher than those in the value-weighted series. However, as subsequent em-

pirical evidence will show, thin trading and day-of-the-week effects alone cannot account for the linear dependence structure in either series.

Nonlinear dependence, on the other hand, may be explained by the well-documented fact of changing variances (see, for example, Hsu, Miller, and Wichern 1974; Epps and Epps 1976; Perry 1982; and Tauchen and Pitts 1983). Changing variance can also explain the high levels of kurtosis in return distributions. Variance changes are often related to the rate of information arrivals, level of trading activity, and corporate financial and operating leverage decisions, which tend to affect the level of stock price. A natural way of modeling this phenomenon is to represent return distributions as mixtures of distributions, or as distributions with stochastic moments. Clark (1973), Blattberg and Gonedes (1974), Oldfield, Rogalski, and Jarrow (1977), Merton (1982), Kon (1984), and many others have proposed models of this type. While these models allow for changing variances and can explain the leptokurtosis (and, possibly, skewness) of empirical distributions, their theoretical statistical assumptions are not consistent with the empirical evidence reported here. Most important, all of these models assume that successive observations are independent random variables and hence the return series are strict white-noise processes.[2] These models are not compatible with the nonlinear dependence structure observed in the return series.

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 (as applied to the transformed series).

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

$$R_t = \phi_0 + \phi_1 R_{t-1} + e_t. \tag{1}$$

The residual series $\{e_t\}$ can be expected to be uncorrelated since second-order or higher-order autocorrelation is not observed in the return series. The OLS estimates of this regression model, and a number of

2. It has to be noted that an erroneous assumption of strict white noise makes questionable the validity of parameter-estimation methods requiring the calculation of likelihood-function values such as the maximum-likelihood method. This is because a likelihood function can usually be calculated only for samples of independent random variables.

statistics describing the distribution of the residuals, are reported in table 2.

The estimates of ϕ_1 are significantly greater than zero, confirming the presence of first-order autocorrelation in $\{R_t\}$. Applying the Dickey-Fuller (1979) test for unit roots shows that ϕ_1 is also significantly smaller than unity in all four periods. This implies that the return series in each period may be generated by a stationary random walk, which is more complex than a simple random walk. This is consistent with the previous findings. Furthermore, simple likelihood-ratio tests show that adding extra lagged variables to the model is not necessary.

The distributions of the residuals e_t are expectedly very similar to those of the returns. They are leptokurtic, slightly skewed, and non-normal as indicated by the Kolmogorov-Smirnov tests. The variance and range parameters are only slightly different. Durbin-Watson tests show that there is no first-order autocorrelation in the residual series. Thus, an AR(1) transformation of returns gives an uncorrelated series of residuals as desired.

In order to check the hypothesis of independence for the residual series, the periodograms and the correlograms are estimated, and the Fisher, Bartlett, and Ljung-Box test statistics are calculated. In almost all cases, these tests fail to reject the hypothesis that the $\{e_t\}$ is strict white noise. This result is surprisingly very different from what is observed in the return series. It is surprising because it is difficult to understand how a linear AR(1) transformation can eliminate the long autocorrelations in the absolute and squared return series. This paradox can be resolved by analyzing the correlograms of $\{|e_t|\}$ and $\{E_t^2\}$ in addition to that of the residuals because, if the residuals are strict white noise, so, too, are their absolute values and squares. All three autocorrelation functions are displayed in figure 2. It is seen that the correlograms of the absolute and squared residuals are very similar to their counterparts in the return series shown in figure 1. The autocorrelations in the squared and absolute residual series are significantly positive (many times $1/\sqrt{T}$) at even very long lags. Since the calculation of these autocorrelations is not based on any stringent distributional assumptions (with the possible exception of the existence of the fourth moment), they show clearly that the residuals exhibit high levels of intertemporal dependence. The residual series are not likely to be realizations of strict white-noise processes.

The Fisher, Bartlett, and Ljung-Box tests could not reject the hypothesis of independence in the residual series because all of these tests are based on the behavior of the autocorrelation function, or the periodogram. In other words, the full probability distribution of e_t is not taken into account. Therefore, when the series is not normal (which is the case here), failure to reject the independence hypothesis by these tests is nothing more than failure to reject the hypothesis of

TABLE 2 **The Regression Model and Residuals**

	Period				
	1963–68	1969–74	1975–80	1981–86	1963–86
A. Estimates of the model, $R_t = \phi_0 + \phi_1 R_{t-1} + e_t$:					
ϕ_0	.0041	−.00009	.00055	.00047	.00033
	(3.10)	(−.84)	(3.08)	(2.71)	(3.36)
ϕ_1	.1946	.3223	.2008	.1406	.2254
	(6.93)	(12.8)	(8.13)	(5.22)	(17.9)
F-statistic	58.57	173.0	66.32	30.52	322.6
R^2	.0374	.1022	.0420	.0191	.0506
D-W statistic	1.983	1.929	1.989	2.000	1.992
B. Sample statistics on the residual series e_t:					
Sample size	1,485	1,513	1,516	1,516	6,030
Mean (thousands)	.0006	−.0000	.0004	.0003	.0000
t (mean = 0)	.0045	−.0003	.0021	.0016	.0001
Variance (thousands)	.0310	.0726	.0607	.0662	.0571
Skewness	−.2381	.5317	.0847	.1297	.2564
Kurtosis	9.2401	6.9254	4.6296	5.1095	6.2468
Range	.0617	.0917	.0774	.0867	.0991
Median (thousands)	.2472	−.0375	.1698	−.1692	.0578
IQR	.0051	.0089	.0095	.0097	.0081
D-statistic	.0896	.0631	.0326	.0868	.0512
Log-likelihood	5,707.3	5,059.7	5,195.1	5,129.4	20,902.3
Fisher's kappa	7.3318	6.2596	8.6011	7.0568	8.5450
Bartlett's test	.0255	.0627	.0444	.0182	.0235
LB(6)	9.85	13.00	16.16	2.60	15.56
LB(12)	13.83	19.24	23.84	11.59	21.94
LB(24)	23.40	35.21	41.68	20.82	44.28

NOTES.—The critical values of the *D* statistic are .021, .023, and .026 for significance levels of .10, .05, and .01, respectively. The critical values for Fisher's kappa are 8.83, 9.52, and 11.10 for significance levels of .10, .05, and .01, respectively. The critical values for Bartlett's test statistic are .0496 and .0595 for significance levels of .05 and .01, respectively. LB(*n*) is the Ljung-Box statistic at lag *n*, distributed as a chi-squared variate with *n* degrees of freedom.
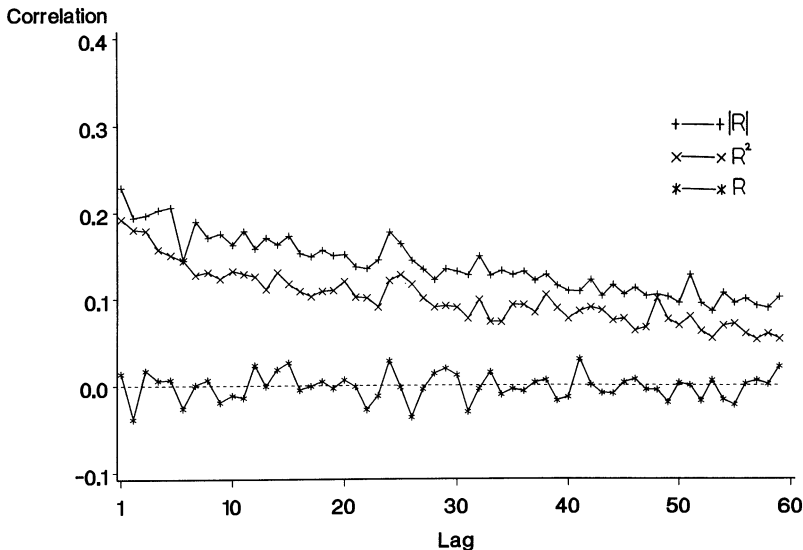
Correlation

Fig. 2.—Autocorrelation functions for residual series

white noise. The most that can be inferred from these test results is that the residual series is white noise (uncorrelated).[3]

The presence of significant autocorrelation in squared residual (and return) series explains the thick tails and peakedness of the empirical distributions. Therefore, the return-generating process may be empirically represented by a linear process of the form (which can be derived from the AR(1) specification by repeated substitution) $R_t = a + \sum_{s=0}^{\infty} b_s e_{t-s}$, where the innovations are nonnormal (heavy-tailed and peaked) random variables. Although this model could show a good empirical fit to data, it would have severe shortcomings. First, since the series $\{e_t\}$ is not independent (though uncorrelated), the traditional time series or regression estimation of the model would be theoretically erroneous. More important, any such linear specification would neglect valuable information for prediction purposes, namely, information about the dependence in the squared values of returns. Considering the critical role of ex ante parameters in many financial economic theories and algorithms, this type of information should hardly be overlooked.

---

3. The values of Bartlett's test statistic from the squared residual series are .2116, .1611, .1091, and .0873 for periods 1 through 4, respectively. Fisher's kappa values are 14.86, 28.47, 34.13, and 19.03. Therefore, the hypothesis of independence can alternatively be rejected by these tests too. The same conclusion applies to the absolute value series.

A nonlinear process, however, that includes functions of past values of e_t^2, would explicitly allow the probability distribution of R_t (at least its second-order properties) to depend on past realizations. This is what the presented empirical evidence demands. Hinich and Patterson (1985) report evidence supporting this point. As discussed by Priestley (1981), statistical estimation of general nonlinear processes is unfortunately often intractable. An alternative model, which closely approximates second-order nonlinear processes, has been developed by Engle (1982) under the name Autoregressive Conditional Heteroscedasticity (ARCH). The process allows the first and second moments of R_t to depend on its past values. This dependence is formulated as a linear function, yielding easy statistical estimation. In the next section, this process will be fitted to the return series.

It was mentioned previously that thin trading and day-of-the-week effects could not be the full causes of dependence in return series. Since thin trading is more prevalent in the equal-weighted index series than in the value-weighted series, higher degrees of autocorrelation may be found in the equal-weighted returns. Indeed, the first-order autocorrelation in this series for the 24-year period is .3625 while it is .2251 for the value-weighted series. However, autocorrelations at longer lags are similarly negligible. When the first-order autocorrelation for a value-weighted series of 30 blue-chip stocks is calculated, it is found to be equal to .1988. Therefore, the impact of thin trading, while present, seems to be small. In order to ascertain the role of the weekday effect, the following regression model is estimated by OLS:

$$R_t = \phi_0 + \phi_1 R_{t-1} + \phi_2 D + \eta_t, \tag{2}$$

where $D = 1$ if t is a Monday, and 0 otherwise. This is to be compared with the previous model where D is not included. It turns out that, for the purposes of this study, it is not necessary to include the Monday dummy variable. Although the estimates of ϕ_2 are marginally significant in all four periods (the largest t-statistic is -2.37), there are very small increases in the R^2 values (the R^2 values are .041, .114, .043, and .021 in the four periods, respectively). Estimates of the other parameters are also practically unchanged. Furthermore, the distributions and correlograms of η_t are almost identical to those of e_t. The weekday anomaly does not seem to have much effect on the temporal dependence in stock-return series. This is true for both indices.

III. Conditional Heteroscedastic Models

In this section, two closely related conditional heteroscedastic time-series models are fitted to data in order to represent the observed autocorrelation structure in daily-return and squared-return series. These are the ARCH model of Engle (1982) and the generalized ARCH

(GARCH) model developed by Bollerslev (1986). In addition to being approximations to more general nonlinear processes, these models have appealing economic and statistical implications.

A. Description of Models

It was shown in the previous section that the first-lag autocorrelation in daily-return series $\{R_t\}$ could be modeled as a simple AR(1) process. This specification is also included in the full conditional heteroscedastic process below. An ARCH process obtains as a special case of a GARCH process. A GARCH process of orders p and q, denoted as GARCH (p, q), can be described as follows:

$$R_t|\Omega_{t-1} \sim F(\mu_t, v_t), \tag{3}$$

$$\mu_t = \phi_0 + \phi_1 R_{t-1}, \tag{4}$$

$$v_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i e_{t-i}^2 + \sum_{j=1}^{q} \beta_j v_{t-j}, \tag{5}$$

and

$$e_t = R_t - \phi_0 - \phi_1 R_{t-1}, \tag{6}$$

where $p > 0$ and $q \geq 0$ are the orders of the process, and the parameters satisfy the conditions $\alpha_0 > 0$, $\alpha_i, \beta_j \geq 0$, $i = 1, \ldots, p, j = 1, \ldots, q$. $F(\mu_t, v_t)$ is the conditional distribution of the variable, with conditional mean μ_t and variance v_t. Ω_{t-1} is the set of all information available at time $t(R_{t-1}, R_{t-2}, \ldots)$. When $q = 0$, an ARCH(p) process results. The statistical properties of this class of processes has been studied by Weiss (1984), Milhoj (1987), and also Bollerslev (1986). The empirical distribution of variables generated by these processes are heavy tailed, compared to the normal distribution. However, no general expression for the distribution function is available.

The unconditional mean and variance of a GARCH process are constant, but the conditional mean and variance are time dependent as shown above. The fact that conditional variances are allowed to depend on past realized variances is particularly consistent with the actual volatility pattern of the stock market where there are both stable and unstable periods. The conditional variance v_t of R_t is large when $e_{t-1}^2, \ldots$, and $v_{t-1}, \ldots$, are large, and vice versa. Based on the results of Tsay (1987), GARCH processes can be seen as special cases of general random coefficient ARMA models. But they have the advantage that the conditional variance v_t is expressed as a simple linear function of past "forecast errors" $(e_{t-1}^2, \ldots,)$ and past conditional variances $(v_{t-1}^2, \ldots,)$. This greatly simplifies model estimation and prediction. ARCH models have been applied successfully on time series of foreign exchange rates by Domowitz and Hakkio (1985) and Milhoj (1987).

B.   Model Estimation and Results

To estimate the parameters $\theta \equiv (\phi_0, \phi_1, \alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)$ of a GARCH(p, q) process, it is necessary to specify the conditional distribution function $F(\mu_t, v_t)$. In all applications, a normal distribution function is assumed. For lack of a good reason for another distribution, this assumption is adopted here, although the model is flexible enough to admit other laws. Given a sample of daily returns $R_1, \ldots, R_T$ and initial values R_0, e_s, v_s for $s = 0, \ldots, r = \max(p, q)$, the log-likelihood function is then given by

$$L(\theta|p, q) = \sum_{t=r}^{T} \log f(\mu_t, v_t), \tag{7}$$

where $f(\mu_t, v_t)$ is the normal density function, and μ_t and v_t are calculated recursively by equations (4)–(6). For large samples, choice of initial values is not critical. Numerical maximization of $L(\theta|p, q)$ gives the maximum likelihood estimates of the parameters for the GARCH-(p, q) model. The values of p and q are to be prespecified. The likelihood function can be maximized for several combinations of p and q, and the maximum values can be compared statistically to obtain the optimal order of the process. Engle (1982) and Bollerslev (1986) have developed Lagrangean multiplier tests for ARCH and GARCH models. The Lagrangean multiplier test requires estimation under the null hypothesis only. Alternatively, χ^2 tests based on likelihood ratios can also be used. Since a normal process (for e_t) with time-dependent mean μ_t and constant variance $v_t = \alpha_0$, different ARCH(p) processes, and several GARCH(p, q) processes are all nested within some higher order GARCH model, likelihood ratio tests are readily available. If $L(\theta_n)$ and $L(\theta_a)$ are the maximum log-likelihood function values under the null and the alternative hypothesis, respectively, then the statistic $-2\{L(\theta_n) - L(\theta_a)\}$ is asymptotically χ^2 distributed with degrees of freedom equalling the difference in the number of parameters under the null and the alternative.

   Numerical maximization of the log-likelihood functions is carried out using the NPSOL package from Stanford University's Systems Optimization Laboratory. Numerical stability and rapid convergence to the optimum is obtained in all cases. The results for the four periods are presented in table 3 and table 4. The standard errors of the point estimates are calculated using the Hessian matrix at the optimum. In the tables, the numbers below the parameter estimates are the usual t-statistics based on these standard errors.

   Table 3 includes the results of fitting pure ARCH(p) processes to daily returns. The order p^* of the process is found by applying likelihood-ratio tests successively until the improvement in the log-likelihood function becomes insignificant. In most samples, if an

TABLE 3   **ARCH(p) Model Estimates**

Parameter	Period				
	1963–68	1969–74	1975–80	1981–86	1963–86
ϕ_0	.00041	−.00014	.00057	.00047	.00032
	(3.04)	(−0.63)	(2.83)	(2.24)	(5.672)
ϕ_1	.1945	.3212	.2045	.1406	.2307
	(7.65)	(13.2)	(8.14)	(5.53)	(24.6)
p^*	3	5	5	2	2
α_0(thousands)	.00961	.02042	.03279	.06352	.03628
	(129.8)	(2.16)	(11.31)	(23.52)	(34.55)
α_1	.24138	.06879	.11712	.03419	.18257
	(5.79)	(2.45)	(35.4)	(1.48)	(9.93)
α_2	.20540	.14984	.02974	.01773	.20149
	(5.40)	(4.38)	(1.21)	(.91)	(10.56)
α_3	.21927	.16835	.05432	...	...
	(54.7)	(4.81)	(2.11)		
α_4	...	.17641	.17413	...	...
		(5.03)	(5.67)		
α_5	...	.17572	.08584	...	...
		(7.86)	(2.83)		
$\sum_{i=1}^{p}\alpha_i$	.6661	.7391	.4612	.0519	.3841
σ_e^2(thousands)	.0288	.0783	.0681	.0670	.0589
σ_R^2(thousands)	.0297	.0872	.0710	.0684	.0620
Log-likelihood	5,892.3	5,228.9	5,236.5	5,133.3	21,138.9
LM statistic	162.01	176.11	78.53	40.48	258.68
χ^2 statistic	364.0	334.0	84.3	16.7	473.2

NOTE.—Numbers in parentheses are t-statistics.

ARCH process of an order higher than necessary is fitted, the estimates of the parameters corresponding to longer lags also tend to become insignificant. Maximum of five lags seems to give a satisfactory fit to daily series. All of the parameters are statistically significant (except α_2 in period 4). The estimates of ϕ_0 and ϕ_1 are very similar to those obtained by OLS in the AR(1) regression model of the previous section. This may show indirectly that if a process for conditional means alone is desired, daily series can be modeled as AR(1). The estimates of α_0 are all positive and considerably smaller than the sample variances shown in table 2. This is due to changing conditional variances over time and their eventual contribution to unconditional variance. The sum of the other ARCH parameters ($\alpha_1 + \ldots + \alpha_p$) is substantially smaller than unity. This indicates that the fitted models are second-order stationary and that at least the second moment exists (Bollerslev 1986). The unconditional variances of e_t and R_t are given by $\sigma_e^2 = \alpha_0/(1 - \sum_{i=1}^{p}\alpha_i)$ and $\sigma_R^2 = \sigma_e^2/(1 - \phi_1^2)$, and they are also reported. These are comparable to the sample variances reported in table 2 and table 1.

TABLE 4   GARCH(1, 1) Model Estimates

| | Period | | | | |
Parameter	1963–68	1969–74	1975–80	1981–86	1963–86
ϕ_0	.00045	.00007	.00055	.00049	.00034
	(3.04)	(.56)	(2.79)	(2.34)	(4.02)
ϕ_1	.1987	.3341	.2103	.1405	.2218
	(7.13)	(13.4)	(7.99)	(5.45)	(31.4)
α_0(thousands)	.00127	.00765	.00154	.00223	.00084
	(3.15)	(19.1)	(2.91)	(8.92)	(3.57)
α_1	.20018	.08706	.08506	.04659	.07906
	(3.70)	(4.35)	(3.27)	(4.65)	(6.14)
β	.76053	.90452	.89325	.92000	.90501
	(13.8)	(29.2)	(34.3)	(46.0)	(64.6)
$\alpha_1 + \beta$	.96071	.99158	.97831	.96659	.98407
σ_ε^2(thousands)	.0323	.9086	.0710	.0668	.0527
σ_R^2(thousands)	.0337	1.023	.0743	.0681	.0555
Log-likelihood	5,919.8	5,273.0	5,254.7	5,169.8	21,527.8

NOTE.—Numbers in parentheses are t-statistics.

The Lagrangean multiplier (LM) and the χ^2 test statistics in table 3 indicate the presence of significant ARCH effects. The null hypothesis of a homoscedastic normal process (that is, $\alpha_1 = \ldots = \alpha_p = 0$) is rejected in all four periods. This is true even for the period from 1981 to 1986, where the ARCH parameters are relatively smaller. The ARCH process describes stock-price fluctuations much better than a normal process with constant variance and with or without time-varying mean.

As a diagnostic check on the appropriateness of ARCH processes for daily return series, the autocorrelation function (ACF) of the squared residual series $\{e_t^2\}$ shown in figure 2 and also the partial autocorrelation function (PACF) of the same series are examined. Bollerslev (1986) shows that the ACF and PACF of an ARCH(p) process of $\{e_t^2\}$ are similar to those of an AR(p) process, where the ACF exhibits exponential and/or oscillatory decay and the PACF cuts off after lag p. Although the estimated ACF of the squared residual series seems to decay as the lag increases (the rate of decay may be slower than exponential), the PACF does not become zero after p or even longer lags. Therefore, as far as the ACF and PACF are concerned, the data do not seem to show full agreement with a pure ARCH process.[4]

Table 4 includes the results of fitting a GARCH(1, 1) process to daily return series. Within the class of GARCH processes, GARCH(1, 1) shows the best fit. Other models such as GARCH(p, q) for $p = 1, \ldots,$

---

4. The ACF and PACF of the squared series are calculated for all four periods (using the ARIMA procedure of the Statistical Analysis System). The same conclusion applies. In the fourth period, the disagreement is more pronounced.

5 and $q = 1, \ldots, 3$ were also tried, but there were no significant improvements in goodness-of-fit based on likelihood-ratio tests. The number of parameters in GARCH(1, 1) is smaller than that in ARCH(p) when $p > 2$. As reported in table 3, the smallest ARCH order for daily returns is 2, which corresponds to an equal number of parameters as in a GARCH(1, 1) process. When the log-likelihood function values for the GARCH(1, 1) estimates in table 4 are compared with those for the ARCH estimates in table 3, it can be seen that they are substantially greater. Therefore, without having to calculate any χ^2 or Lagrangean multiplier tests, it is concluded that GARCH processes fit to data much better than ARCH and normal processes.

The parameter estimates of the GARCH(1, 1) model in table 4 are all statistically significant (except ϕ_0 in period 2). The estimates of ϕ_0 and ϕ_1 are very similar to their counterparts in table 2 and table 3 for the ARCH case. The estimates of α_0 are much smaller than the sample variances of e_t in table 2, showing again that conditional variances are changing over time. The unconditional variances of e_t and R_t, calculated as $\sigma_e^2 = \alpha_0/(1 - \alpha_1 - \beta)$ and $\sigma_R^2 = \sigma_e^2/(1 - \phi_1^2)$, are expectedly of similar magnitudes as their sample variances in table 2 and table 1. The estimates of β are always markedly greater than those of α_1, and the sum $\alpha_1 + \beta$ is very close to but always smaller than unity. When the Dickey-Fuller test for unit roots is applied, the null hypothesis that $\alpha_1 + \beta \geq 1.0$ is rejected in all periods except the second. Therefore, the fitted process seems to be second-order stationary (admittedly, this is not a strong conclusion). The fact that $\alpha_1 + \beta$ is close to one, however, is useful for purposes of forecasting conditional variances. This will be explored at length in the next section.

The ACF and PACF for e_t^2 for a GARCH(p, q) process are similar to those for an ARMA(r, q) process ($r = \max(p, q)$), where both functions tail off after $r - q$ lags. Based on the results in Bollerslev (1986), it can be shown that, for the special case of GARCH(1, 1), $c_n = (\alpha_1 + \beta)c_{n-1}$ for $n \geq 2$, where c_n is the correlation between e_t^2 and e_{t-n}^2. The ACF for e_t^2 shown in figure 2 is unusually consistent with this difference equation. Of the 60 autocorrelations plotted in figure 2, 48 fall within plus/minus 10% of the values implied by the equation. Furthermore, the PACF (not shown) is also generally nonzero but decays at higher lags. It is concluded that the empirical correlation structure is consistent with a GARCH(1, 1) specification. This holds for the whole 24-year period as well as for all four periods, and for both index series.

The fit of the GARCH(1, 1) model is further evaluated by investigating whether the standardized residuals $(e_t - \mu_t)/\sqrt{v_t}$ have a standard normal distribution, as they should under the specification given by expressions (3)–(6) and the additional assumption of conditional normality in (7). Kolmogorov-Smirnov tests and the Kiefer-Salmon tests cannot reject the hypothesis of normality. The GARCH model reduces

the excess kurtosis to zero (the highest value observed is 0.37 in period 4). Skewness is also no longer present (the largest absolute value is 0.02 in period 2). These results are expected but still they provide supporting evidence of good fit to data. They show that large returns (of either sign) are more frequently observed in more volatile periods, and vice versa. This is a realistic description of stock-market behavior.

Finally, the results of fitting ARCH and GARCH models to the entire 1963–86 period are reported in the last columns of table 3 and table 4. Within their classes, ARCH(2) and GARCH(1, 1) specifications for this series provide satisfactory fit. GARCH(1, 1), however, is superior to any ARCH specification. These results have to be interpreted with consideration of the presence of nonhomogeneity in the series. As a case in point, the log-likelihood function of either model for the entire series is significantly smaller than the sum of the log-likelihood functions for the four periods. This indicates that there are gains in descriptive power when the models are fitted separately to the four periods, and it also provides further justification for dividing the data into four samples.

IV. Forecasts of Volatility

The parameter estimates in the previous section show that any realistic process for stock returns must allow for high degrees of dependence in the series of conditional variances and, to a lesser extent, in the series of conditional means. As mentioned before, any intertemporal dependence is also valuable information for forecasting purposes. In this section, several forecasts of return variances are calculated and their accuracies are compared. Given the set Ω_0 of all information about past and present returns $(R_0, R_{-1}, \ldots)$, forecasts of the variance of future returns (either $\text{var}(R_1|\Omega_0)$, or $\text{var}(R_1 + \ldots + R_N|\Omega_0)$ for some N) may be obtained. Forecasting the variance for a period that contains multiple observations is useful here because it is then possible to compare forecasts with actual realized values. Since return is defined as the continuously compounded return, the sum $R_1 + \ldots + R_N$ is the N-day return. In the following analysis, N is chosen as 20, which roughly corresponds to one month, and the sum is called monthly return. In other words, forecasts of monthly variances are estimated from samples of daily returns.

Forecasts of future variance are useful for several reasons. First of all, the predictive capabilities of ARCH and GARCH models constitute further evidence as to their overall usefulness as practical models of stock returns and also about their relative merits as such. Second, since risk is inherently related to volatility, expected future volatility is a major factor in the pricing of securities. Good forecasts of volatility can be used to investigate any relation between current prices and

expected risk. For example, future variance is an explicit argument in the popular option-pricing models.

A. Methodology

The time series of returns for each of the four periods is divided into two parts: $R_1, R_2, \ldots, R_{T-480}$ and $R_{T-479}, R_{T-478}, \ldots, R_T$, where T is the total sample size for the period. The forecasting models are estimated from the first $T - 480$ observations, and forecasts are generated for the variance of return for the following month (20 days). This gives the first forecast. For each subsequent forecast, the estimation sample is shifted forward by one month (by dropping the initial 20 observations and adding in the new 20 observations). Thus, the number of forecasts generated for each period is 24 (480 divided by 20). This scheme of successive out-of-sample forecasting allows for the model parameters to be modified over time.

The forecast of the variance of return in month s ($s = 1, \ldots, 24$) immediately following day z ($z = T - 480, T - 460, \ldots, T - 20$) is denoted by $V_{s,z} = \text{var}(R_{z+1} + \ldots + R_{z+20}|\Omega_z)$. The actual variance for the same month (taking into account the first-lag autocorrelation in daily returns) is calculated ex post as follows:

$$V_{s,z}^{(a)} = \sum_{i-1}^{20}(R_{z+i} - \bar{R})^2\left[1 + .10\sum_{j=1}^{19}(20 - j)\phi^j\right], \tag{8}$$

where $\bar{R}$ is the mean, and ϕ is the first-lag autocorrelation. This equation provides a more realistic measure of variance than the usual calculation $\sum_{i=1}^{20} R_{z+i}^2$ (Merton 1980, Perry 1982). These ex post values are to be compared with the forecasts. Four different sets of forecasts are obtained.

Benchmark forecast. This is the simple historical average:

$$V_{s,z} = \left(\frac{20}{T - 480}\right)\sum_{t=z-T+479}^{z}(R_t - \bar{R})^2.$$

This would be the best forecast if the time series of returns were strict white noise. It is an unbiased forecast over long periods of time.

Exponentially weighted moving average forecast (EWMA). This forecast is given by

$$V_{s,z} = (1 - w)\sum_{i=1}^{12} w^{i-1}V_{s-i,z-20i}.$$

The smoothing constant w is estimated by minimizing the sum of

squared forecast errors using the first $T - 480$ observations. The minimization is done by the Newton-Raphson algorithm, and the estimated values of w are .18, .24, .19, and .11 for the four periods. This simple exponential smoothing approach is consistent with the phenomena of infrequent changes in variances (unlike the GARCH models where variances are continuously chan$_\cup$ ng). Thus, notwithstanding the findings in the preceding section, the EWMA forecasting model should perform relatively better if the return-generating process is nonstationary.

ARCH forecast.   Since the ARCH model specifies the actual conditional variances as an explicit function of observed values, one-step-ahead forecasts are readily obtained. More distant forecasts are simply generated by repeated substitution until they are reduced to functions of only present and past values. After some suitable algebra and successive operations with conditional expectations, the final forecasting equation takes the following form:

$$V_{s,z} = \sum_{t=1}^{20} \left(\frac{1 - \phi_1^t}{1 - \phi_1}\right)^2 \left[A^{20-t}v_{z+1} + \sum_{j=0}^{19-t} \alpha_0 A^j\right],$$

where $A = \alpha_1 + \ldots + \alpha_{p*}$, and $v_{z+1} = \alpha_0 + \Sigma_{i=1}^{p*} \alpha_i e_{z+1-i}^2$ by equation (5). The second sum in the brackets is set equal to zero if $t = 20$.

GARCH forecast.   The GARCH forecast $V_{s,z}$ is given by the same expression as above, where now $A = \alpha_1 + \beta$ and v_{z+1} is given by equation (5) for $p = q = 1$. The calculation of each of these variance forecasts is computationally fast and easy. Unless the model parameters are to be revised as new data become available, forecast updates are also readily available. In this application, this approach is not taken and the parameters are reestimated as new observations come in. However, in most practical applications, less frequent revisions may be acceptable.

B.   Results

The forecasts of 24 monthly return variances by each of the four methods and the actual ex post variances are shown in figure 3. It is clearly seen that the ARCH and GARCH models can simulate the actual pattern of stock market volatility very closely. When these two related models are compared with each other, the GARCH specification is superior. For example, in the 1981–86 period, the ARCH model performs very poorly while GARCH does not. As expected, historical averages do not reflect short-term changes in volatility. They are virtually unchanged throughout the 24-month period. As for the exponentially weighted moving average representation, this is also unable to

TABLE 5 Forecasts of Monthly Variances

Period	Statistic	Historical Estimate	EWMA Forecast	ARCH Forecast	GARCH Forecast
1963–68	ME	.000341	.000108	.000071	.000069
	RMSE	.000748	.000722	.000643	.000587
	MAE	.000460	.000518	.000455	.000387
	MAPE	.559	.736	.653	.525
1969–74	ME	.004292	.003643	.000746	.000011
	RMSE	.005428	.004762	.002970	.002081
	MAE	.004292	.003643	.002033	.001527
	MAPE	.722	.603	.367	.338
1975–80	ME	.001333	.001072	.000471	.000150
	RMSE	.001940	.001769	.001326	.001052
	MAE	.001423	.001220	.001012	.000780
	MAPE	.487	.428	.502	.389
1981–86	ME	− .000018	.000431	− .000422	− .000251
	RMSE	.000772	.000873	.000854	.000693
	MAE	.000608	.000627	.000724	.000615
	MAPE	.557	.425	.786	.465

NOTE.—Letting $E_i = V_{s,z} - V_{s,z}^{(a)}$ denote the forecast error in the sth month, the statistics are calculated as following:

$$ME = (1/24) \sum_{s=1}^{24} E_s,$$

$$RMSE = \left[(1/24) \sum_{s=1}^{24} E_s^2 \right]^{1/2},$$

$$MAE = (1/24) \sum_{s=1}^{24} |E_s|,$$

$$MAPE = (1/24) \sum_{s=1}^{24} \left| \frac{E_s}{V_s} \right|.$$

model transitory changes in volatility. These findings show that the time-series behavior of market volatility can be realistically modeled by conditionally heteroscedastic processes. These models perform particularly well in periods of high overall volatility such as the 1970s. This can be seen in the second and third graphs in figure 3.

The forecasts are evaluated and compared through a number of statistics: mean error (ME), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percent error (MAPE). These are reported in table 5, along with their calculation methods. Based on the relative values of these statistics, the GARCH forecasts are far better than the other three. This difference is more pronounced in periods of high volatility (1969–74 and 1975–80). GARCH forecasts are generally less biased (if significant at all), as smaller ME values imply, and more accurate, as smaller values of the other three parameters imply. It seems safe to conclude that ex ante measures of variance can be satis-
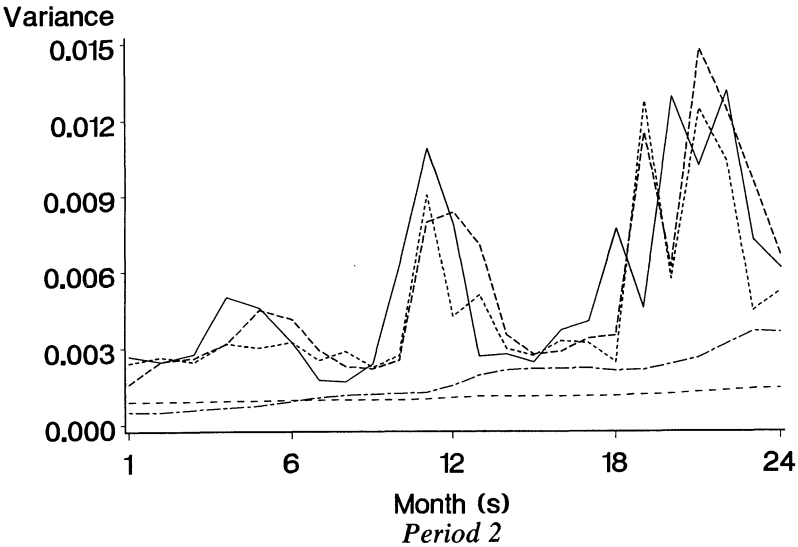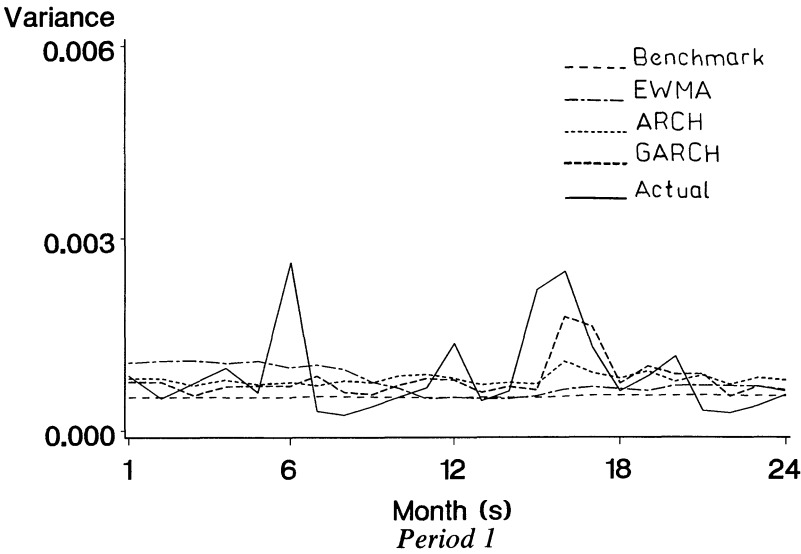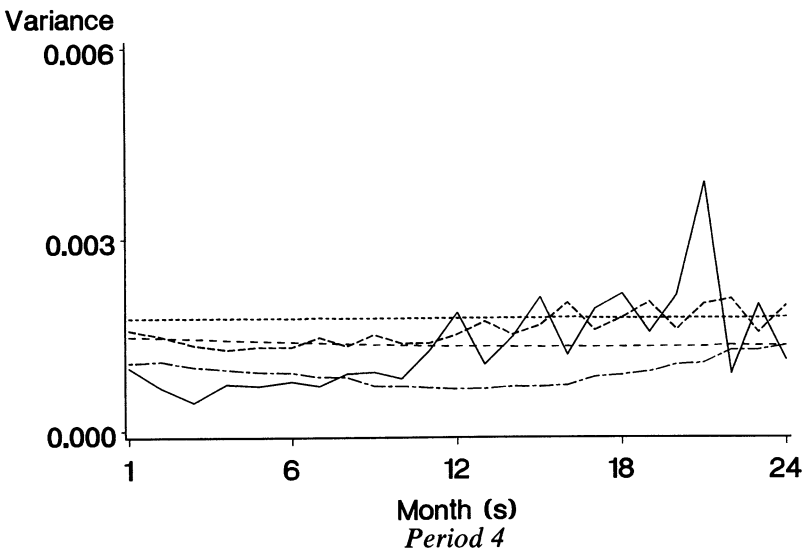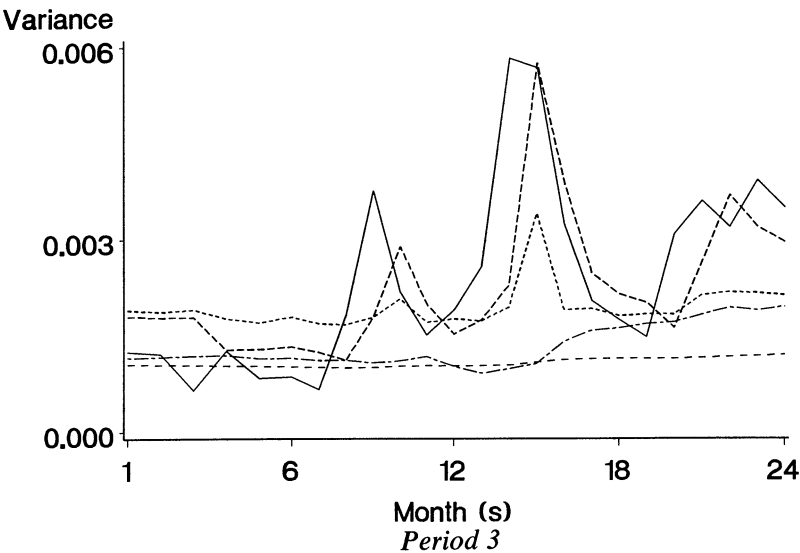
FIG. 3.—Forecasts of monthly variances, periods 1–4

Variance

Month (s)
Period 3

Variance

Month (s)
Period 4

factorily estimated by GARCH models of stock returns. Although none of the forecasts are as accurate as desirable (the smallest MAPE is greater than 30%), GARCH forecasts constitute substantial improvement over the traditional forecasts such as the historical sample averages.

An interesting observation from the plots in figure 3 is that GARCH (and, to a lesser extent, ARCH) forecasts are more accurate when actual changes in volatility are not one-time events but persist for at least a few months. This is due in part to the particular model specification, where next period's variance is a function of this period's realized variance. Alternatively, the finding that $\alpha_1 + \beta$ is very close to unity and that this sum is dominated by β indicates that changes in market volatility tend to be persistent. This is probably why GARCH forecasts are better than the others. To further investigate the pattern of volatility implied by the estimated model, daily conditional variances v_t are calculated for $t = 3, \ldots, T$, using equation (5), and the distribution of these statistics is analyzed. It is found that these conditional variances have a right-skewed distribution. This means that periods of above-average volatility are less frequent than periods of average or below-average volatility. This is consistent with the observed patterns in the stock markets. When the distributions of 12 subsamples containing roughly $T/12$ observations each (spanning a period of about 6 months) are analyzed, both the medians and the means of these distributions are found to be substantially different. This is interpreted as evidence that changes in market volatility in either direction are usually persistent changes. This would seem to be in agreement with the relative magnitudes of the estimated parameters.

V.   Discussion and Conclusions

The empirical evidence presented in this paper indicates that time series of daily stock returns exhibit significant levels of dependence. The probability distribution of R_t is not independent of R_{t+s} for several values of s. The conditional heteroscedastic processes allow for autocorrelation between the first and second moments of return distributions over time, and consequently they fit to data very satisfactorily. More important, they provide improved forecasts of volatility. Within the class of such models, GARCH(1, 1) processes show the best fit and forecast accuracy.

Several extensions to this study may be suggested. Currently, research is under way to investigate the plausibility of bilinear processes for stock prices, where conditional means evolve nonlinearly over time. More general nonlinear models, which allow for dependence in higher order moments, may also be used. The difficulty, however, is that there is not much statistical estimation theory for nonlinear processes. This is despite the fact that it is relatively easy to identify

nonlinearities. As immediate competitors to the (G)ARCH models, conditional heteroscedastic ARMA (Tsay 1987), ARMA-GARCH (Weiss 1984), and general random coefficient AR processes may be investigated. All of these models have similar financial theoretical backgrounds, and all can represent the empirical behavior of prices.

GARCH models may be used to further understand the relationship between volatility and expected returns. The fundamental valuation theories in finance, such as the capital asset pricing models and the option pricing models, are based on some hypothesized risk-return relationship. Most of these models hold for the "average" security but they do not explain the full valuation mechanism for "nonaverage" securities. Empirical evidence about the size effect and the deviation of out-of-the-money option premia from implied theoretical values is abundant. The apparent failure of the models for such securities may be largely due to an erroneous choice of values for the model parameters. Consequently, improved parameter estimates may explain the discrepancies between theory and reality. In this regard, GARCH models can be very useful. For example, since ex ante (rather than ex post) measures of variance are what the traders use in forming expectations of return, GARCH forecasts of variance are better choices than the usual historical estimates. Future research in this area is well warranted.

Finally, most of the findings in this study hold only for daily data. The preliminary statistical analysis conducted for daily return series in Section II is also conducted for weekly (Wednesday–Wednesday) and monthly series. Notable differences are found. First of all, the distributions of weekly and monthly returns are not as leptokurtic as those of daily returns. In fact, Kolmogorov-Smirnov tests cannot reject the hypothesis that monthly returns are normally distributed. Secondly, there is no significant autocorrelation in either return series. When the squared and absolute series are analyzed, weekly series exhibit some autocorrelation up to a maximum of four lags, but monthly series have no significant autocorrelation. Therefore, it is concluded that monthly returns are independently normally distributed (strict white noise). For weekly series, the log-likelihood functions corresponding to GARCH-(1, 1) processes are not significantly greater than those for normal process (except in period 2, where a χ^2 value of .103 is found). A central limit theorem for sums of *dependent* (daily) returns may be manifesting itself in these findings about weekly and monthly series.

References

Blattberg, R., and Gonedes, N. 1974. A comparison of the stable and student distributions as statistical models of stock prices. *Journal of Business* 47:244–80.
Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31:307–27.

Bookstaber, R., and McDonald, J. 1987. A general distribution for describing security price returns. *Journal of Business* 60:401–24.

Clark, P. 1973. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 41:135–55.

Cox, J., and Rubinstein, M. 1985. *Option Markets*. Englewood Cliffs, N.J.: Prentice-Hall.

Dickey, D., and Fuller, W. 1979. Distribution of estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 77:427–31.

Domowitz, I., and Hakkio, C. 1985. Conditional variance and the risk premium in the foreign exchange market. *Journal of International Economics* 19:47–66.

Engle, R. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. *Econometrica* 50:987–1008.

Epps, T., and Epps, M. 1976. The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distribution hypothesis. *Econometrica* 44:305–21.

Fama, E. 1965. The behavior of stock market prices. *Journal of Business* 38:34–105.

Fama, E. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25:383–417.

Fuller, W. 1976. *Introduction to Statistical Time Series*. New York: Wiley.

Hinich, M., and Patterson, D. 1985. Evidence of nonlinearity in daily stock returns. *Journal of Business and Economic Statistics* 3:69–77.

Hsu, D.; Miller, R.; and Wichern, D. 1974. On the stable Paretian behaviour of stock market prices. *Journal of the American Statistical Association* 69:108–13.

Kiefer, N., and Salmon, M. 1983. Testing normality in econometric models. *Economics Letters* 11:123–27.

Kon, S. 1984. Models of stock returns: A comparison. *Journal of Finance* 39:147–65.

Mandelbrot, B. 1963. The variation of certain speculative prices. *Journal of Business* 36:394–419.

Merton, R. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8:323–61.

Merton, R. 1982. On the mathematics and economics assumptions of continuous-time models. In W. Sharpe and C. Cootner (eds). *Financial Economics*. Englewood Cliffs, N.J.: Prentice-Hall.

Milhoj, A. 1987. A conditional variance model for daily deviations of an exchange rate. *Journal of Business and Economic Statistics* 5:99–103.

Neftçi, S. 1984. Are economic time series asymmetric over the business cycle? *Journal of Political Economy* 92:307–28.

Oldfield, S.; Rogalski, R.; and Jarrow, R. 1977. An autoregressive jump process for common stock returns. *Journal of Financial Economics* 5:389–418.

Perry, P. 1982. The time-variance relationship of security returns: Implications for the return-generating stochastic process. *Journal of Finance* 37:857–70.

Pindyck, R. 1984. Risk, inflation, and the stock market. *American Economic Review* 74:335–51.

Poterba, J., and Summers, L. 1986. The persistence of volatility and stock market fluctuations. *American Economic Review* 76:1142–51.

Priestley, M. 1981. *Spectral Analysis and Time Series*. Vol. 2. New York: Academic Press.

Tauchen, G., and Pitts, M. 1983. The price variability-volume relationship in speculative markets. *Econometrica* 51:485–505.

Taylor, S. 1986. *Modelling Financial Time Series*. Chichester: Wiley.

Tsay, R. 1987. Conditional heteroscedastic time series models. *Journal of the American Statistical Association* 82:590–604.

Weiss, A. 1984. ARMA models with ARCH errors. *Journal of Time Series Analysis* 5:129–43.