

## Chapter 2

# Realized Stock Volatility

### 2.1 Introduction

Financial market volatility is indispensable for asset and derivative pricing, asset allocation, and risk management. As volatility is not a directly observable variable, large research areas have emerged that attempt to best address this problem. By far the most popular approach is to obtain volatility estimates using the statistical models that have been proposed in the ARCH and Stochastic Volatility literature. Another method of extracting information about volatility is to formulate and apply economic models that link the information contained in options to the volatility of the underlying asset. All these approaches have in common that the resulting volatility measures are only valid under the specific assumptions of the models used and it is generally uncertain which or whether any of these specifications provide a good description of actual volatility.

A model-free measure of volatility is the sample variance of returns. Using daily data, for instance, it may be freely estimated using returns spanning over any number of days and, as such, one can construct a time series of model-free variance estimates. When one chooses the observation frequency of this series, an important trade-off has to be made, however. When the variances are calculated using a large number of observations (e.g. the returns over an entire year), many interesting properties of volatility tend to disappear (the

volatility clustering and leverage effect, for instance). On the other hand, if only very few observations are used, the measures are subject to great error. At the extreme, only one return observation is used for each daily variance estimate.

The approach taken in this dissertation is to calculate the daily volatility from the sample variance of intraday returns, the ‘realized’ volatility. Specifically, we use the transaction record of the Dow Jones Industrials Average (DJIA) portfolio over the period extending from January 1993 to May 1998, to obtain a time series of 1366 daily realized variances. These are free of the assumptions necessary when the statistical or economic approaches are employed and, as we have an (almost) continuous record of returns for each day, we can calculate the interdaily variances with little or perhaps negligible error.

In this chapter, we shall first give a through account of the theoretical properties that underlie the concept of realized volatility measurement. Using our data for the DJIA, we next document the empirical regularities of this volatility variable and then capture these using a parametric model. Finally, we compare the predictive ability of the realized volatility model to various ARCH formulations.

Almost all of the work on daily volatility is within the confines of ARCH and Stochastic Volatility models or derivative pricing formulas. There are exceptions, however. Schwert (1990) and Hsieh (1991) have computed sample standard deviations from intradaily returns on the S&P 500 index. However, the modeling and investigation of the properties of volatility have not been the major focus and consequently these two papers do not present a thorough analysis of the constructed series. More recently, Andersen and Bollerslev (1998) have calculated a time series of realized exchange rate variances to evaluate one-day-ahead GARCH model forecasts while Andersen, Bollerslev, Diebold and Labys (1999) use realized variance estimates to document the properties of daily exchange rate volatility. Our study is in spirit close to the latter paper, but distinct in two key aspects. Firstly, our analysis is on stock return volatility and as a result we characterize important empirical regularities not

found for exchange rates. Secondly, we not only examine but also model realized volatility and determine whether this new approach is of practical relevance.

Following this introduction, Section 2.2 gives an account of the theoretical underpinnings of the realized volatility measure. Section 2.3 details the construction of the data that provide the basis for our subsequent empirical analysis. In Section 2.4 we investigate the properties of stock return volatility and, in Section 2.5, we fit parametric models to our volatility series. We compare the performance of these models to four ARCH formulations in Section 2.6. We finish in Section 2.7 with concluding remarks.

## 2.2 Realized Volatility Measurement

A common model-free indicator of volatility is the daily squared return. In this paper we measure interdaily volatility using intradaily high-frequency returns. We highlight in this section the relation between these two measures and discuss their individual properties.

To set forth the notation, let  $p_{n,t}$  denote the time  $n \geq 0$  logarithmic price at day  $t$ . The discretely observed time series of continuously compounded returns with  $N$  observations per day is then defined by:

$$r_{n,t} = p_{n,t} - p_{n-1,t}$$

where  $n = 1, \dots, N$  and  $t = 1, \dots, T$ . If  $N = 1$ , for any series we ignore the first subscript  $n$  and thus  $r_t$  denotes the time series of daily return.

We shall assume that:

$$\text{A.1:} \quad E[r_{n,t}] = 0$$

$$\text{A.2:} \quad E[r_{n,t} r_{m,s}] = 0 \quad \forall n, m, s, t \quad \text{but not } n = m \text{ and } s = t$$

$$\text{A.3:} \quad E[r_{n,t}^2 r_{m,s}^2] < \infty \quad \forall n, m, s, t$$

Hence, returns are assumed to have mean zero and to be uncorrelated and it is assumed that the variance and covariances of squared returns exist and are finite.

The continuously compounded daily squared returns may be decomposed as:

$$r_t^2 = \left( \sum_{n=1}^N r_{n,t} \right)^2 = \sum_{n=1}^N r_{n,t}^2 + \sum_{n=1}^N \sum_{\substack{m=1 \\ n \neq m}}^N r_{n,t} r_{m,t} = \sum_{n=1}^N r_{n,t}^2 + 2 \sum_{n=1}^N \sum_{m=n+1}^N r_{n,t} r_{m-n,t} \quad (2.2.1)$$

Assuming that A.1 holds, the squared daily return is therefore the sum of two components: the sample variance (at the daily unit) and twice the sum of  $N - 1$  sample autocovariances (at the  $1/N$ th day interval unit). In this decomposition it is the sample variance that is of interest – the sample autocovariances are measurement error and induce noise in the daily squared return measure.

From (2.2.1) and A.1 and A.2 it therefore follows that an unbiased estimator of the daily return volatility is the sum of intraday squared returns, the realized volatility:

$$s_t^2 = \sum_{i=1}^N r_{n,t}^2$$

as:

$$E[s_t^2] = \sigma_t^2$$

where  $\sigma_t^2$  is daily population variance.

Because the realized volatility  $s_t^2$  is an estimator, it has itself a variance which can be interpreted as measurement error. From now on we shall assume that A.1 to A.3 hold, and then the variance of  $s_t^2$  is given by:

$$\begin{aligned} V(s_t^2) &= E \left[ \left( \sum_{n=1}^N r_{n,t}^2 - \sigma_t^2 \right)^2 \right] \\ &= E \left[ \sum_{n=1}^N \sum_{m=1}^N \left( r_{n,t}^2 - \frac{\sigma_t^2}{N} \right) \left( r_{m,t}^2 - \frac{\sigma_t^2}{N} \right) \right] \end{aligned}$$

Thus the variance of  $s_t^2$  depends on the sum of all covariances of the squared return process. Upon separating the double sum for all  $n \neq m$ , taking expectations and rearranging terms it follows:

$$= E \left[ \sum_{n=1}^N \left( r_{n,t}^2 - \frac{\sigma_t^2}{N} \right)^2 \right] + 2 E \left[ \sum_{n=1}^N \sum_{m=n+1}^N \left( r_{n,t}^2 - \frac{\sigma_t^2}{N} \right) \left( r_{m,t}^2 - \frac{\sigma_t^2}{N} \right) \right]$$

The first term is the variance of the intraday squared returns process (at the daily unit) and the second term is the sum of all squared return autocovariances (at the  $1/N$ th day interval unit). Upon dividing the term on the right by 1 over  $N$  times the expression on the left and taking expectations one obtains:

$$= E \left[ \sum_{n=1}^N \left( r_{n,t}^2 - \frac{\sigma_t^2}{N} \right)^2 \right] \left( 1 + 2 \sum_{n=1}^N \frac{N-n}{N} \rho_{N,n,t} \right)$$

where  $\rho_{N,n,t}$  the  $n$ th autocorrelation of  $\{r_{n,t}^2\}_1^N$ . Finally, after expanding the factor on the left and taking expectations it follows:

$$V(s_t^2) = \frac{\sigma_t^4}{N} \left( K_{N,t} - 1 \right) \left( 1 + 2 \sum_{n=1}^{N-1} \frac{N-n}{N} \rho_{N,n,t} \right) \quad (2.2.2)$$

where  $K_{N,t}$  denotes the kurtosis of  $\{r_{n,t}\}_1^N$ . Note that the kurtosis and autocorrelations have subscript  $N$  as these may change with the number of intraday returns. From (2.2.2) follows that for any particular value of  $N$ , measurement error increases with the daily population variance, with the kurtosis of intraday returns and with the autocorrelations of intraday squared returns.

Special cases of equation 2.2.2 reduce to familiar expressions. For instance, if  $r_{n,t}$  is *i.i.d.* normal with  $E[r_{n,t}^2] = \sigma_t^2/N$  (variances are constant within the day), equation (2.2.2) becomes:  $V[s_t^2] = 2 \sigma_t^4 / N$ . This result can be found in Kendall and Stuart (1963, p. 243), for instance. Note that under these assumptions the variance of the realized volatility decreases at rate  $N$ . However, for various assets it is well documented that returns have kurtosis in

excess of three and that the squares of returns are correlated (the ARCH effect). Under these circumstances, this expression will therefore give the lower bound of measurement error.

To establish consistency of  $s_t^2$ , we require the two additional assumptions that:

$$\text{A.4:} \quad K_{N,t} < \infty \quad \forall N$$

$$\text{A.5:} \quad \exists \rho_{N,n,t} \text{ s.t. } \rho_{N,n,t} < 1$$

Boundedness of  $K_{N,t}$  rules out jump-diffusions (Drost, Nijman and Werker 1998) and implies continuity of the sample paths of  $\sigma_{n,t}^2$  by the Kolmogorov criterion (Revuz and Yor 1991, Theorem I.1.8). Assumption A.5 states that the squared return process has at least one autocorrelation that is less than unity.

Suppose  $\rho_{N,n,t} = 1$  for  $n = 1, \dots, N$ , then the last factor in (2.2.2) becomes:  $(1 + 2(N-1) - 2N^{-1} \sum_{n=1}^{N-1} n) = N$ , since  $\sum_{n=1}^{N-1} n = 0.5(N-1)N$ . Therefore,  $V(s_t^2) = \sigma_t^4(K_{N,t} - 1)$ . By A.5, however,  $V(s_t^2)$  will decrease in  $N$  and by A.4 it follows therefore that:

$$\lim_{N \rightarrow \infty} V[s_t^2] = 0$$

Thus, the realized volatility measure converges in mean-square and is consistent.<sup>1</sup> The daily variance may therefore be estimated to any desired degree of accuracy by the realized volatility.

Recall that the results reported thus far are derived under the assumption that returns are uncorrelated. This assumption is questionable when  $N$  is large, as serial correlation in returns is a common symptom of market micro-structure effects such as price discreteness,

---

<sup>1</sup>Consistency may alternatively be established under the assumption that the price process  $p_{n,t}$  follows  $dp_{n,t} = \sigma_{n,t} dW_{n,t}$ , where  $W_{n,t}$  denotes a Wiener process. Under the assumption that  $\sigma_{n,t}$  is continuous, it follows from the results in Karatzas and Shreve (1988, Chapter 1.5) or Barndorff-Nielsen and Shephard (1999) that  $\text{plim}_{N \rightarrow \infty} \sum_{i=1}^N r_{n,t}^2 = \int_1^\infty \sigma_{n,t}^2 dn = \sigma_t^2$ . See also Andersen *et al.* (1999) for a thorough treatment along these lines in the context of special semi-martingales.

bid-ask bounces and non-synchronous trading (see for instance the textbook treatment by Campbell, Lo and MacKinlay 1997; Chapter 3). Any violation of this assumption can easily be studied when considering the MA (q) (moving average) representation of  $r_{n,t}$ :

$$r_{n,t} = \epsilon_{n,t} + \sum_{i=1}^q \psi_{i,t} \epsilon_{n-i,t} \quad (2.2.3)$$

where the innovations  $\epsilon_{n,t}$  are assumed to be uncorrelated across all leads and lags. Note that we allow the moving average representation to change across  $t$ . This simply reflects that our realized volatility measure does not require processes to remain constant over time. Upon squaring (2.2.3), taking expectations, and summing over  $n = 1, \dots, N$ , it follows that:

$$E \left[ \sum_{n=1}^N r_{n,t}^2 \right] = E[s_t^2] = (1 + \sum_{i=1}^q \psi_{i,t}^2) E \left[ \sum_{n=1}^N \epsilon_{n,t}^2 \right] \quad (2.2.4)$$

where  $E[\sum_{n=1}^N \epsilon_{n,t}^2] = \sigma_t^2$ . At day  $t$  therefore, the cumulative squared returns measure has a multiplicative bias that is given by the squared dynamic coefficients of the moving average representation. Under conditions of serial correlation, the realized variance will therefore unambiguously overestimate actual volatility. One may, of course, test for the statistical significance of the parameters that are used to capture any temporal dependence in returns and use (2.2.4) to determine whether any bias is economically important.

## 2.3 Data Source and Construction

Our empirical analysis is based on data from the NYSE Transaction and Quote (TAQ) database which records all trades and quotations for the securities listed on the NYSE, AMEX, NASDAQ, and the regional exchanges (Boston, Cincinnati, Midwest, Pacific, Philadelphia, Instinet, and CBOE). Our sample consists of the Dow Jones Industrial Average (DJIA) index constructed from the transaction prices of the 30 stocks that are contained in it.<sup>2</sup> The data span from January 4, 1993 to May 29, 1998 (1366 observations). Within

---

<sup>2</sup>The reconstruction of the index is straightforward. The DJIA is the sum of all prices adjusted by a divisor that changes once an index stock splits, pays a stock dividend of more than 10% or when one

each day, we consider the transaction record extending from 9:30 to 16:05, the time when the number of trades noticeably dropped. Next to transaction prices, volume and time (rounded to the nearest second) TAQ records various codes describing each trade. We used this information to filter trades that were recorded in error and out of time sequence.<sup>3</sup>

Taking all 30 stocks together, we observe a trade about every one second. Naturally, the trading frequency of the index components is lower. It also varies greatly across stocks: the median time between trades in a single stock ranges from a low of 7 seconds to a high of 54 seconds. This suggests that one should worry that non-synchronous trading induces serial correlation in the returns process which, in turn, would render the cumulative squared returns measure biased. Since we are focusing on an index, the market micro-structure effects that are due to price discreteness and bid-ask bounces are of less concern as these tend to wash out in the aggregate (see Gottlieb and Kalay 1985, Ball 1988 and Harris 1990 for instance).<sup>4</sup>

To mitigate the problem of bias, following Andersen and Bollerslev (1998) and Andersen *et al.* (1999), we shall rely on five-minute returns to obtain daily variance estimates. These are constructed from the logarithmic difference between the prices recorded at or immediately before the five-minute marks. When considering the transactions record extending from 9:30 to 16:05, this provides us with  $N = 79$  returns for each of the  $T = 1366$  days.

---

company in the group of thirty is replaced by another. Over our sample, the composition of the DJIA index changed March 17, 1997, when four stocks were replaced. Naturally, we accounted for this.

<sup>3</sup>Specifically, we omitted all trades that carry the ‘correction indicators’ 2 (symbol correction; out of time sequence), 7 (trade canceled due to error), 8 (trade cancelled) and 9 (trade canceled due to symbol correction). Moreover, we filtered all trades with the ‘condition indicator’ G (bunched sold; a bunched trade not reported within 90 seconds of execution time), L (sold last; a transaction that occur in sequence but is reported to the tape at a later time) and Z (sold sale; a transaction that is reported to the tape at a time later than it occurred and when other trades occurred between the time of the transaction and its report time). We refer to corresponding data manual for a more complete description of these and other codes.

<sup>4</sup>The theoretical literature on price discreteness suggests that the upward bias of the volatility estimate decreases with the price level. This suggests that one can detect the presence of bias by examining whether a time series of volatility estimates is negatively correlated with the corresponding prices. For our data we find a slight positive correlation – suggesting therefore that the discreteness of prices should not be of concern.



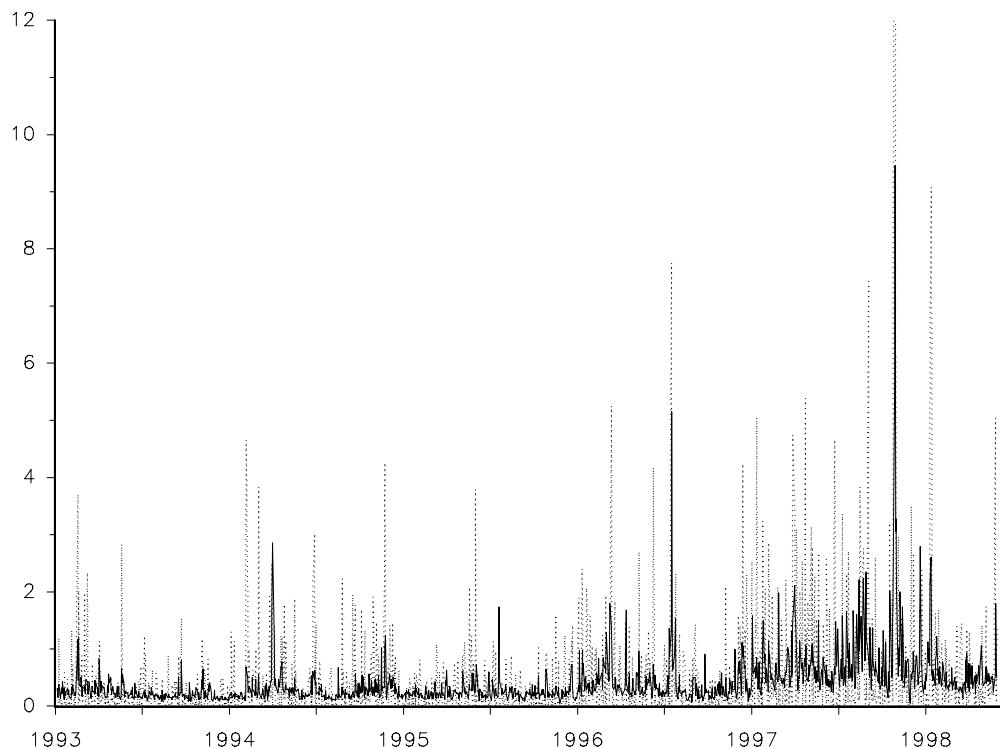
It remains an empirical question whether the five-minute cut-off is sufficient large enough so that the problem of bias due to market micro-structure effects is of no practical concern. For our data we find that the first two sample autocorrelations are 0.080 and  $-0.018$  and these are significant judged by the  $\pm 1.96\sqrt{(1/NT)}$  5% confidence interval. Consistent with the spurious dependencies that would be induced in an index by non-synchronous trading, the first order autocorrelation is positive. The consequences of serial correlation are minimal, however. Upon estimating the MA (2) model defined by equation (2.2.3), we obtain  $\hat{\psi}_1 = 0.0431$  and  $\hat{\psi}_2 = -0.0317$ . From (2.2.4) it follows that the bias resulting from serial correlation scales the volatility estimates upward by a factor of only 1.0029. Considering that the mean realized variance equals 0.4166, we view this bias too small to be of any economic significance – no correction is therefore made.

The reduction in measurement error when using intradaily data to calculate volatility measures becomes readily apparent in Figure 2.1. The solid line plots realized variances, *i.e.* cumulative 5-minute squared returns, while the dotted line displays squared daily returns. For the latter series the two highest values are 39.4 (October 27, 1997) and 48.5 (October 28, 1997) and these two observations are outside the plot region. Although both series are correlated, the variability in the realized volatility series is small compared to the variability in the squared returns.

## 2.4 Properties of Realized Volatility

The main subject of this section is to investigate the properties of realized stock return volatility. Using intradaily returns on the DJIA portfolio over the period January 4, 1993 to May 29, 1998 (1366 observations), we focus on three volatility measures: variances ( $s_t^2$ ) standard deviations ( $s_t$ ) and logarithmic variances ( $\ln(s_t)^2$ ). For each of these three measures, we investigate the distribution, persistency and relation to current and lagged returns. Our findings shall set the stage for the development of realized volatility models in our next section. In the literature on ARCH and Stochastic Volatility models considerable

**Figure 2.1:** Squared Daily Returns and Realized Variances



The graph displays realized variances (solid line) and daily squared returns (dashed line) for the Dow Jones Industrials Average over the period January 4, 1993 to May 29, 1998 (1366 observations). Variances are obtained using cumulative 5-minute squared returns. For the squared return series the two highest values are outside the plot region (39.4 and 48.5 recorded for October 27 and 28, 1997).

interest is in the distribution of daily returns divided by their daily standard deviations. Therefore, we characterize this distribution as well using our measures of volatility. Finally, as some of our analysis overlaps with the work by Andersen *et al.* (1999) on exchange rates, we will compare our results to theirs at the end of this section.

### 2.4.1 Distribution of Volatility

We graph the distributions of the variance, standard deviation and logarithmic variance series in Figure 2.2.<sup>5</sup> The skewness ( $\hat{S}$ ) and kurtosis ( $\hat{K}$ ) coefficients are displayed in the top-right corner of each plot. The distributions of variances and standard deviations are

<sup>5</sup>Density estimates throughout this paper are based on the Gaussian kernel. The bandwidths are calculated according to equation 3.31 of Silverman (1986).

clearly non-normal – both are skewed right and leptokurtic. The square root transformation of variances to standard deviations, however, reduces the skewness estimate from 8.19 to 2.57 and the kurtosis estimate from 121.59 to 16.78. The distribution of logarithmic variances appears to be approximately normal (the normal density is displayed by dashed lines). Nonetheless, standard tests reject normality. For instance, under the null hypothesis of *i.i.d.* normality,  $\hat{S}$  and  $\hat{K}$  are distributed normal as well with standard errors equal to  $\sqrt{6/T} = 0.066$  and  $\sqrt{24/T} = 0.133$ ; the skewness and kurtosis estimates for logarithmic variances are several standard errors away from their hypothesized values.<sup>6</sup>

### 2.4.2 Persistency of Volatility

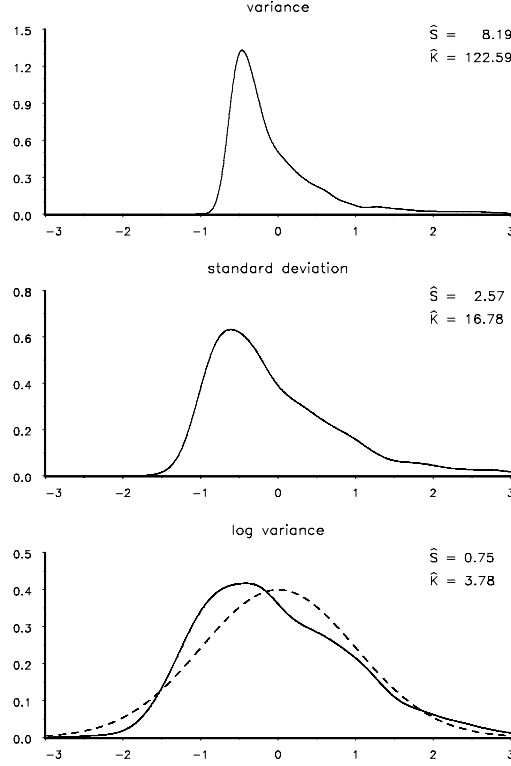
Dating back to Mandelbrot (1963) and Fama (1965), volatility clustering has a long history as a salient empirical regularity characterizing speculative returns. The literature on volatility modeling has almost universally documented that any such temporal dependency is highly persistent. The time series plot of the realized volatility in Figure 2.1 – displaying identifiable periods of high and low variances – seems to already intimate that view. The temporal dependency of volatility is reinforced by Figure 2.3, where we plot the sample autocorrelation function for each series (boxes). For all three volatility measures, the autocorrelations begin around 0.65 and decay very slowly. The 100 day correlation is 0.15, 0.28 and 0.32 for the variance, standard deviation and logarithmic variance series, respectively. At the 200 day offset, the functions take a value of 0.08, 0.15 and 0.18.

The low first-order autocorrelations already suggests that the three volatility measures do not contain a unit root. The Augmented Dickey-Fuller test, allowing for a constant and 22 lags, yields test statistics of -4.692, -3.666 and -3.096 – rejecting therefore the unit root hypothesis,  $I(1)$ , at the 5% level. The slow hyperbolic decay, however, indicates the presence

---

<sup>6</sup>It is often noted that tests for normality can be grossly incorrect in finite samples and/or when the observations are dependent (see for instance Beran 1994, Chapter 10 and the references therein). Upon simulating the fractionally integrated process  $(1-L)^{0.4}y_t = -0.04 + \varepsilon_t$ ,  $\varepsilon_t \sim i.i.d. N(0, 0.2)$  and  $t = 1, \dots, 1366$  (see Table 2.1 later in this chapter), we find for  $y_t$ , using 100,000 trials, that the 95% confidence interval for  $\hat{S}$  and  $\hat{K}$  is given by  $[-0.195, 0.194]$  and  $[2.707, 3.305]$ , respectively. For  $\varepsilon_t$ , we obtain  $[-0.130, 0.130]$  and  $[2.760, 3.278]$ . The asymptotic intervals under *i.i.d.* normality are  $[-0.130, 0.130]$  and  $[2.740, 3.260]$ . Under these conditions, the degree of bias is therefore not severe.

**Figure 2.2:** Distribution of Realized Volatility



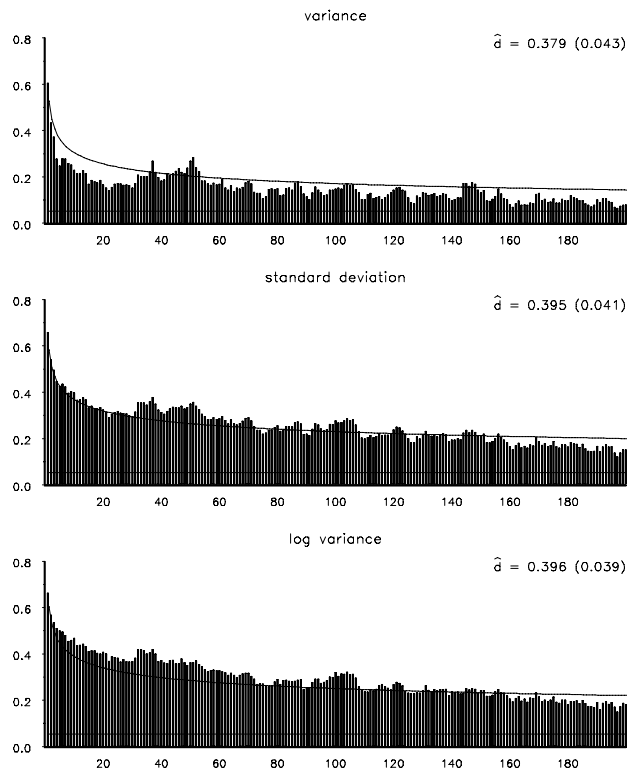
The graphs display the density estimates of variances (top panel), standard deviations (middle panel) and logarithmic variances (bottom panel). All series are standardized to mean zero and variance one. The bottom panel graphs along with the density estimates the standard normal probability distribution function (dashed line). Skewness ( $\hat{S}$ ) and kurtosis ( $\hat{K}$ ) coefficients are displayed in the top-right corner of each plot.

of long-memory. Using squared returns (or some transform thereof), this phenomenon has been documented by Ding, Granger and Engle (1993) and Crato and de Lima (1994), among others.

A covariance stationary fractionally integrated process,  $I(d)$ , has the property of long-memory in the sense that the autocorrelations decay at a slow hyperbolic rate when  $0 < d < 0.5$ . In the top-right corner of each plot in Figure 2.2 we display the Geweke and Porter-Hudak (1993) log-periodogram estimate for the fractional integration parameter  $d$ ; standard errors are given in parentheses.<sup>7</sup> The theoretical autocorrelation functions implied

<sup>7</sup>The estimates are obtained using the first to  $m = T^{4/5} = 322$  spectral ordinates and this choice is optimal according to Hurvich, Deo and Brodsky (1998). Standard errors are obtained using the usual OLS regression formula and are slightly higher than the asymptotic standard error of the estimator,  $\pi/\sqrt{24m} = 0.036$ .

**Figure 2.3:** Realized Volatility Sample Autocorrelation Functions



The graphs display the first 200 sample autocorrelations of variances (top panel), standard deviations (middle panel) and logarithmic variances (bottom panel). The horizontal lines are the upper limit of the 95% confidence interval,  $1.96/\sqrt{T}$ . Geweke and Porter-Hudak estimates for the fractional integration parameter are given in the top-right corner of each plot; standard errors are reported in parentheses. The lines are the theoretical autocorrelations implied by these estimates.

by these estimates for  $d$  match rather well the sample autocorrelations and the parameter estimates for  $d$  are more than two standard errors away from 0.5 and several standard errors away from zero.<sup>8</sup> This suggests that realized volatilities are covariance stationary and fractionally integrated.

---

<sup>8</sup>The unweighted minimum distance estimator proposed by Tieslau, Schmidt and Baillie (1996) minimizes the sum of squares between the theoretical autocorrelation function of an I(d) process and the sample autocorrelation function. Upon applying this estimator to the first 200 autocorrelations, we obtain for the fractional integration parameter estimates of 0.346, 0.389 and 0.401 for variances, logarithmic variances and standard deviations, respectively.

### 2.4.3 Volatility and Returns

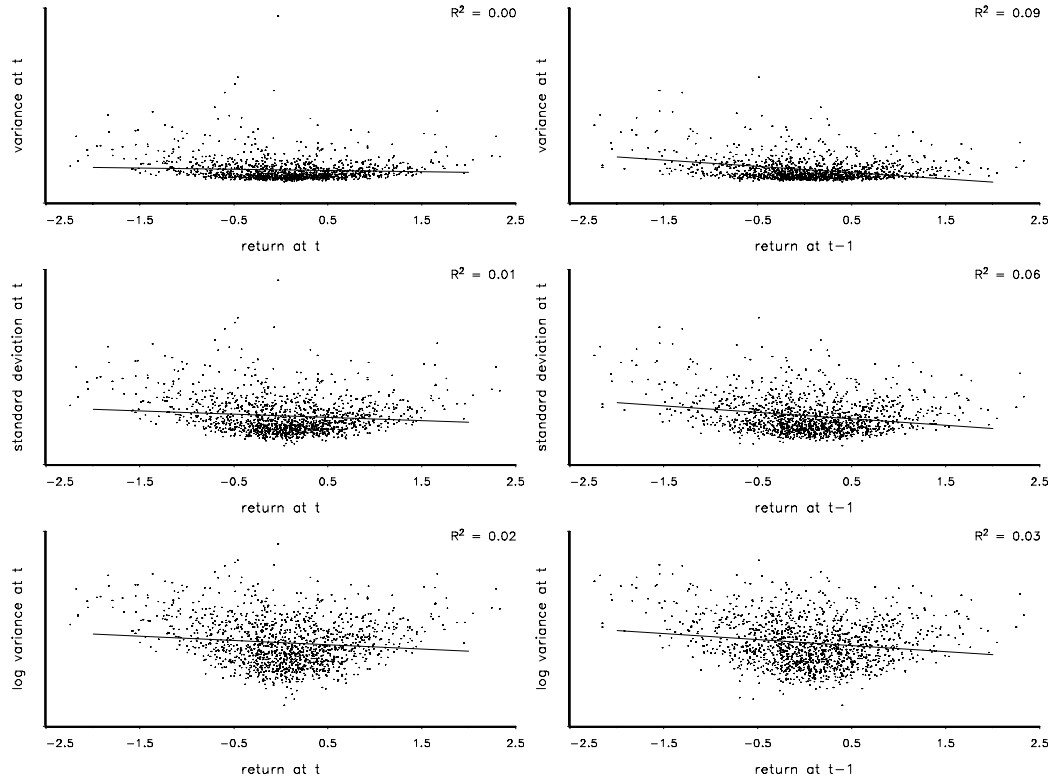
The relation between volatility and returns is of interest for two reasons. First, theoretical models such as Merton's (1973) intertemporal CAPM model relate current expected excess returns to volatility. This has motivated the ARCH in mean, or ARCH-M, model introduced by Engle, Lilien and Robins (1987). In the ARCH-M specification the conditional mean of returns (or excess returns) is a linear function of the conditional variance. Second, Black (1976), Pagan and Schwert (1990) and Engle and Ng (1993), among others, have documented asymmetries in the relation between news (as measured by lagged returns) and volatility – suggesting that good and bad news have different predictive power for future volatility. Generally it is found that a negative return tends to increase subsequent volatility by more than would a positive return of the same magnitude. This phenomenon is known as the 'leverage' or 'news' effect.

In Figure 2.4 the relation between our three volatility measures and current returns is displayed on the left whereas the relation with lagged returns is given on the right. Through the scatters, the graphs display an ordinary least squares regression line which is based on the displayed variables and a constant term. The  $R^2$  statistic from each regression is given in the top-right corner of the plots.

Focusing on the three graphs on the left, we can see that there is no 'important' linear relation between the three volatility measures and current returns – suggesting that the ARCH-M effect is negligible for our data. It becomes however obvious that volatilities are non-linear in returns; all three volatility measures increase with positive and negative returns. Note also how in each of the three plots a convex frontier seems to shape out. This implies that a particular daily price change generates some minimum level of volatility.

The plots on the right of Figure 2.4 suggest the presence of the leverage effect: lagged negative returns yield high volatility more frequently than lagged positive returns. This phenomenon is most pronounced for variances and least obvious for logarithmic variances.

**Figure 2.4:** Realized Volatilities and Current and Lagged Returns



The graphs display returns (left panels) and lagged returns (right panels) against variances (top panel), standard deviations (middle panel) and logarithmic variances (bottom panel). The lines are OLS regression lines which are based on the displayed variable and a constant term. The regression  $R^2$  measures are given in top-right corner of the plots. In all graphs we omit four observations that are to the left and three observations that are to the right of the plot region.

It is quite surprising that such asymmetry is less evident when looking at the graphs using current returns. If the news-effect is indeed the source of asymmetry, one would expect that current news, rather than past news, yield the suggested effect. Possibly it takes time for some market participants to react.

To investigate further the asymmetric response of volatility to past returns, we fit via ordinary least squares the following regression models to our data:

$$\begin{aligned}
s_t^2 &= \omega_1 + \omega_2 \mathbf{I}^- + \omega_3 r_{t-1}^2 + \omega_4 r_{t-1}^2 \mathbf{I}^- + \varepsilon_t \\
s_t &= \omega_1 + \omega_2 \mathbf{I}^- + \omega_3 r_{t-1} + \omega_4 r_{t-1} \mathbf{I}^- + \varepsilon_t \\
\ln(s_t^2) &= \omega_1 + \omega_2 \mathbf{I}^- + \omega_3 r_{t-1} + \omega_4 r_{t-1} \mathbf{I}^- + \varepsilon_t
\end{aligned} \tag{2.4.1}$$

where  $s_t^2$  denotes realized variances; the indicator  $\mathbf{I}^-$  takes value one when  $r_{t-1} < 0$  and is zero otherwise. Note that we allow for asymmetry in intercepts as well as slopes and that we consider for variances a quadratic relation between lagged returns and volatility.<sup>9</sup>

In Figure 2.5 we plot the regression lines implied by estimates of equation 2.4.1 (solid line) along with the regression lines implied by the nonparametric models of lagged returns on each of the three volatility measures (dashed line).<sup>10</sup> The  $R^2$  statistics from the parametric and nonparametric regressions (in parentheses) are displayed in the top-right corner of each plot.

Both the parametric and nonparametric regressions confirm the asymmetric news-effect – volatility increases more steeply with negative than with positive returns. The news-impact functions are centered around  $r_{t-1} = 0$ ; this suggests that asymmetry is only in slopes and not in intercepts. The close correspondence between the parametric and nonparametric regression lines indicates that the models given by equation 2.4.1 characterize well the news-impact functions for the DJIA portfolio. There are no obvious discrepancies that would suggest any other parametric specification to capture the lagged return volatility

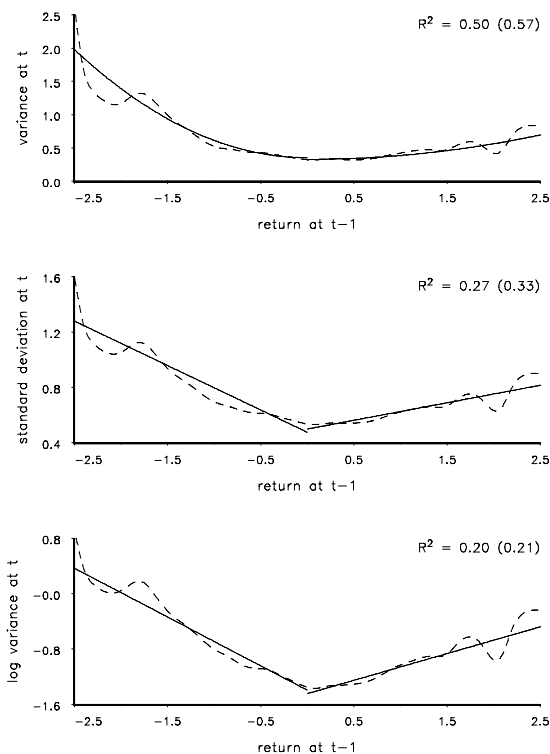
---

<sup>9</sup>In our estimations we find, as one would expect from the results in Section 2.4.2, that the residual innovations  $\hat{\varepsilon}_t$  are serially correlated and non-normal (see also Figure 2.6 which we shall discuss shortly). Nonetheless, the least squares estimator yields under these circumstances still unbiased, albeit not efficient, coefficient estimates.

<sup>10</sup>The nonparametric regression estimates are obtained using the Nadaraya-Watson estimator with a Gaussian kernel. The bandwidth parameters are determined using cross-validation scores. Estimation was done over the entire sample, yet the plot regions are restricted to returns in the -2.5 to 2.5 interval. Four observations are smaller than -2.5 and three observations are greater than 2.5. Note that the kernel estimator is consistent despite non-normal and correlated residuals. However bandwidth selection by cross-validation gives under-smoothed estimates (see Härdle and Linton 1994).



**Figure 2.5:** Parametric and Nonparametric News Impact Functions

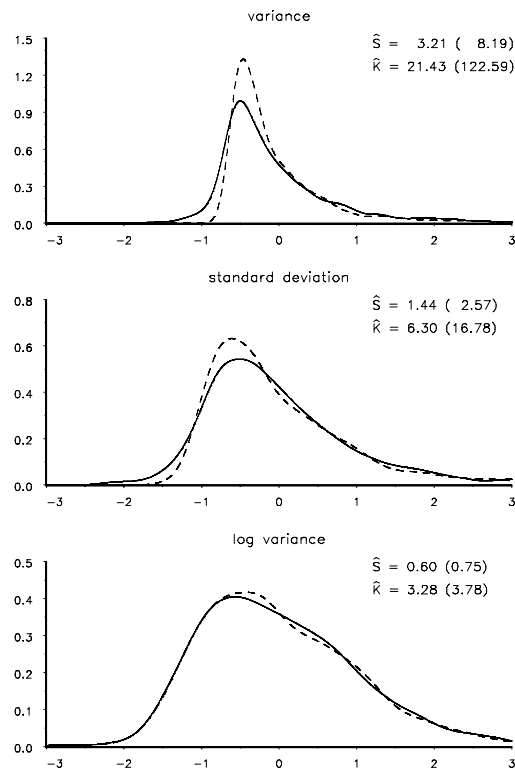


The graphs display the regression lines implied by estimates of equation 2.4.1 and nonparametric regression estimates of lagged returns on variances (top panel), standard deviations (middle panel) and logarithmic variances (bottom panel). The  $R^2$  of both the parametric and nonparametric regressions (in parentheses) are given in top-right corner of each plot.

relation.

For the modeling of volatility it becomes of interest whether the news-effect can account for the asymmetry and excess kurtosis we observe in the distribution of our volatility series (Figure 2.2). In Figure 2.6, we graph the distribution of the variance, standard deviation and logarithmic variance series after (using solid lines) and before (dashed lines) accounting for the news effect. For the ‘after news-effect’ distributions we use the residuals from the models defined by equation 2.4.1. The skewness and kurtosis coefficients are displayed in the top-right corner of each plot. The estimates before accounting for news-effects are reported in parentheses.

**Figure 2.6:** Distribution of Realized Volatility and News Impact



The graphs display the density estimates of variances, standard deviations and logarithmic variances before and after accounting for news-effects. The density estimates after accounting for news-effects (solid line) are obtained from the OLS residuals of the models defined by equation 2.4.1. The density estimates before accounting for news-effects (dashed lines) are identical to the ones displayed in Figure 2.2. Skewness ( $\hat{S}$ ) and kurtosis ( $\hat{K}$ ) coefficients are displayed in the top-right corner of each plot. The estimates before accounting for news-effects are reported in parentheses.

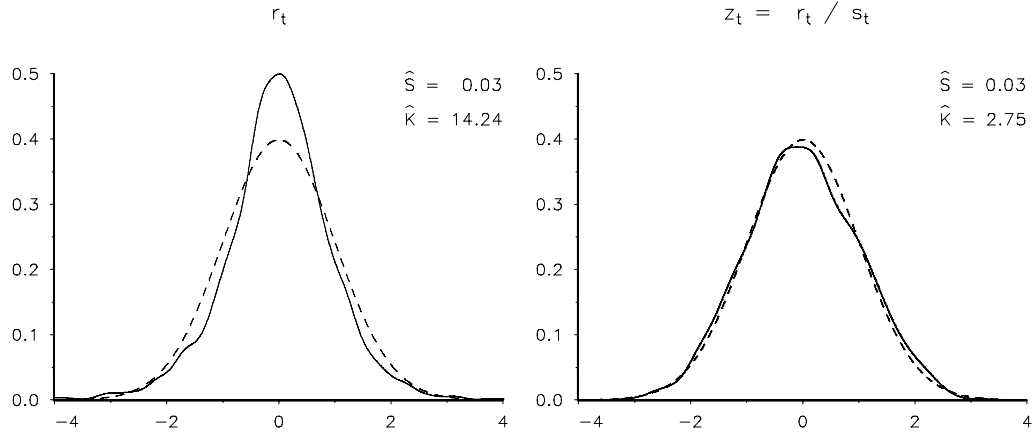
Even after accounting for the asymmetric response of volatility to lagged returns, the distribution of variances and standard deviations remains clearly non-normal. News-effects can however remove some of the asymmetry and flatness in the distribution of the volatility measures. For variances, standard deviations and logarithmic variances, the skewness coefficient is reduced from 8.19 to 3.21, 2.57 to 1.44 and 0.75 and 0.60, respectively. The kurtosis coefficient decreases from 122.59 to 21.43, 16.78 to 6.30 and 3.78 to 3.28 for the respective volatility measures. As there is little asymmetry in the distribution of logarithmic variances, the corresponding reduction in the skewness and kurtosis estimates is only modest, however. Judged by the standard errors of these estimates (see Section 2.4.1), normality of logarithmic variances is again rejected.

#### 2.4.4 Distribution of Returns and Standardized Returns

An empirical regularity found almost universally across all assets is that high frequency returns are leptokurtic. Early evidence for this dates back to Mandelbrot (1963) and Fama (1965). Clark (1973) established that a stochastic process is thick tailed if it is conditionally normal with changing conditional variance. ARCH and Stochastic Volatility models have this property, but it is often found that these models do not adequately account for leptokurtosis. Specifically, returns divided by the estimated standard deviations ( $z_t = r_t/\hat{\sigma}_t$ ) display frequently excess kurtosis. As a result, several other conditional distributions have been employed to fully capture the degree of tail fatness (see for instance Hsieh 1989 and Nelson 1991).

Realized standard deviations allow us to characterize the distribution of standardized returns without modeling changing variances. In Figure 2.7, we plot the density of the daily return series ( $r_t$ ) on the left whereas we depict the density of this series scaled by daily standard deviations ( $z_t = r_t/s_t$ ) on the right. In each graph we also plot the normal density. Skewness ( $\hat{S}$ ) and kurtosis ( $\hat{K}$ ) estimates are given in the top-right corner of the plots.

**Figure 2.7:** Distribution of Returns and Standardized Returns



The graphs display the density estimates of daily returns  $r_t$  (left panel) and scaled returns  $z_t = r_t/s_t$  (right panel), where  $s_t$  are daily realized standard deviations. The displayed series are standardized to mean zero and variance one (the mean and standard deviation of  $z_t$  equal 0.132 and 1.041, respectively). The standard normal density is plotted with dashed lines.

Returns are hardly skewed, but leptokurtic as expected. From the kurtosis estimate of scaled returns it becomes evident that changing variances can fully account for fat tails in returns – the estimate even suggests that this distribution is platykurtic. The density of  $z_t$  is very close to the one implied by the normal distribution. Based on the standard errors of the skewness and kurtosis estimates (see Section 2.4.1), normality cannot not be rejected at the 5% level –  $\hat{S}$  and  $\hat{K}$  are within two standard errors of their hypothesized values.

Recall from Section 2.4.1 that we found that logarithmic variances are distributed nearly normal – implying that standard deviations and variances are distributed approximately lognormal. Combined with the normality of  $z_t$ , this suggests that returns are approximately a normal-lognormal mixture which has been proposed by Clark (1973). In Clark’s model, however, the volatility process is assumed *i.i.d.* whereas we find that it is serially correlated (see Section 2.4.2).<sup>11</sup>

## 2.4.5 Comparison to Exchange Rates

Our results regarding the distribution and persistency of realized stock volatility are remarkably similar to the ones obtained by Andersen *et al.* (1999) in the setting of exchange rates. They also found that the distribution of variances and standard deviations is skewed right and leptokurtic, but that logarithmic variances are distributed approximately normal.<sup>12</sup> Exchange rate return variances, standard deviations and logarithmic variances display a high degree of persistency as well. Depending on the volatility measure and exchange rate series used, Andersen *et al.* report Geweke and Porter-Hudak (1993) log-periodogram estimates ranging between 0.346 and 0.421. Results on news impact however differ. Contrary to our results, they did not find much evidence for the asymmetric volatility effect. This is

---

<sup>11</sup>In Clark it is also assumed that the series  $z_t$  is independent. The BDS test (Brock, Dechert, LeBaron and Scheinkman 1987) yields test statistics of  $W_2 = -2.721$ ,  $W_3 = -2.839$  and  $W_4 = -2.089$ . As these are distributed standard normal, we therefore have to reject independence of  $z_t$  at the 5% level. However, Ljung-Box portmanteau statistics for up to  $\{10, 20, 100\}$  th-order serial correlation in  $z_t$  and  $z_t^2$  are insignificant at the 5% level.

<sup>12</sup>Without accounting for the leverage effect, our skewness and kurtosis estimates are higher than the ones reported for exchange rates by Andersen *et al.* After adjusting for the effect (see Section 2.4.3), our estimates become quite close to theirs, however.

to be expected, however, as this phenomenon is generally observed for equities only.

## 2.5 Realized Volatility Modeling and Predictions

In this section we first build models aimed to capture the temporal dependency of realized volatility. Treating volatility as observed instead of latent allows us to utilize the time series techniques employed when modeling the conditional mean. We thus can sidestep the relatively more complicated ARCH and Stochastic Volatility formulations that model and measure volatility simultaneously. Later in this section we investigate how well the developed models predict volatility *ex ante* one-step-ahead. We shall compare these predictions to the ones obtained by ARCH models in our next section.

### 2.5.1 Realized Volatility Modeling

As far as the modeling of our three volatility measures is concerned, the main findings of our previous section are: (1) the distributions of variances and standard deviations are asymmetric and leptokurtic, but logarithmic variances are distributed approximately normal; (2) realized volatilities appear covariance stationary and fractionally integrated; and (3) volatility is correlated with lagged negative and positive returns. Before detailing the specific models we shall employ to account for (2) and (3), we discuss first the implications of the distributional characteristics of our volatility measures for the modeling of these series.

Assuming that the only deterministic component of a covariance stationary process  $y_t$  is a (possibly non-zero) mean  $\omega$ , then it is well known that the Wold representation of  $y_t$  is a (possibly infinite) moving average, *i.e.*  $y_t = \omega + \varepsilon_t + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}$ ,  $\varepsilon_t \sim WN(0, \sigma^2)$  where WN denotes serially uncorrelated white noise. Estimation and inference generally require the stronger assumption that  $\varepsilon_t \sim i.i.d. WN(0, \sigma^2)$  and using this premise it is

straightforward to show that:

$$K_y - 3 = \frac{1 + \sum_i^q \alpha_i^4}{(1 + \sum_i^q \alpha_i^2)^2} (K_\varepsilon - 3)$$

$$S_y = \frac{1 + \sum_i^q \alpha_i^3}{(1 + \sum_i^q \alpha_i^2)^{3/2}} S_\varepsilon$$

where  $K_y, S_y$  ( $K_\varepsilon, S_\varepsilon$ ) denotes the kurtosis and skewness of  $y_t$  ( $\varepsilon_t$ ). Note, as  $S_y > 0$ ,  $K_y > 3$ ,  $\exists i$  s.t.  $\alpha_i \neq 0$  and  $0 \leq \alpha_i \leq 1 \forall i$  then  $S_\varepsilon > S_y$  and  $K_\varepsilon > K_y$ . Because we found in our previous section that variances as well as standard deviations – even after accounting for the news effect – are highly skewed and leptokurtic and that the sample autocorrelation functions of these volatility measures are positive and slowly decaying (suggesting  $0 \leq \alpha_i$ ), a model with a moving average representation would leave these distributional characteristics unexplained and even amplified in the residuals. When estimation is done by maximum likelihood, as it is commonly the case, this in turn would require one to either rely on quasi-maximum likelihood estimates or to condition the residuals on a density that allows for skewness and excess kurtosis. However, the former approach may not yield consistent estimates of the parameters and variance-covariance matrix whereas the latter would complicate analysis as it requires additional coefficients.<sup>13</sup> We found, however, that the distribution of logarithmic variances is almost symmetric and subject to little excess kurtosis. For these reasons we restrict our attention to modeling this series only. Of course, logarithmic variances are rarely of interest. We address this issue by investigating in our next subsection whether logarithmic predictions transformed into variances and standard deviations provide useful descriptions of these two volatility measures.

---

<sup>13</sup>One density that allows for skewness and kurtosis is the exponential generalized beta (McDonald and Xu 1995). We used this density to estimate the ARFIMAX specification discussed below to model variances and standard deviations directly. Any improvements, as measured by *ex ante* one-day-ahead prediction criteria, were minor only. Alternatively, one may consider estimation in the frequency domain, which can allow one to relax the normality assumption. However, such models do not easily allow for the type of exogenous variables we consider.

To account for long-memory and the correlation of volatility with lagged negative and positive returns we model logarithmic variances using the following ARFIMAX (p,d,q) model:

$$(1 - L)^d (1 - \beta(L_p)) \ln(s_t^2) = w_0 + w_1 r_{t-1} I^- + w_2 r_{t-1} I^+ + (1 + \alpha(L_q)) \varepsilon_t \quad (2.5.1)$$

where  $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$ ,  $\alpha(L_q) = \sum_{i=1}^q \alpha_i L^i$  and  $\beta(L_p) = \sum_{i=1}^p \beta_i L^i$ . Realized variances are denoted by  $s_t^2$ , the indicator  $I^-$  ( $I^+$ ) takes value one when  $r_{t-1} < 0$  ( $r_{t-1} \geq 0$ ) and is zero otherwise. Next to the standard ARMA (p,q) coefficients ( $w_0, \beta(L_p), \alpha(L_q)$ ) the above specification contains the following three coefficients: a fractional integration parameter ( $d$ ) to capture the slow hyperbolic decay in the sample autocorrelation function; lagged negative ( $w_1$ ) and positive ( $w_2$ ) returns to allow for the leverage effect as well as to account for the slight asymmetry and tail fatness in the distribution of  $\ln(s_t^2)$ . We estimate the above model using the conditional sum-of-squares maximum likelihood estimator suggested by Hosking (1984). The finite sample properties of this estimator have been investigated by Chung and Baillie (1993).

Parameter estimates of three specifications nested within the above model – an ARFIMA (0,d,0) labeled FI, an ARFIMAX (0,d,0) with label FIX and an ARFIMAX (0,d,1) labeled FIMAX – are given in Table 2.1. Standard errors are reported in parentheses under the coefficient estimates. All of the parameters are statistically significant at the 5% level on the basis of either Wald or likelihood ratio tests. The table reports in addition the Schwartz Bayesian Information Criterion (SBC) and Ljung-Box portmanteau statistics for up to  $K$ th-order serial correlation in the residuals ( $Q_K$ ). The numbers in parentheses below these statistics report the probability that the  $K$  autocorrelations are not significant.

Paying attention to the estimates of the fractional integration parameter  $d$  first, we can see that our estimation results confirm our earlier suspicion that the logarithmic variance process is stationary and fractionally integrated. Estimates for  $d$  range between 0.324 and 0.392 and are several standard errors away from both zero and 0.5. The FI model estimate

**Table 2.1:** Realized Volatility Model Estimates

	$\hat{\omega}_0$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{d}$	$\hat{\alpha}_1$	$\hat{\sigma}^2$	$SBC$	$Q_{10}$	$Q_{20}$	$Q_{100}$
<i>FI</i>	-0.043 (0.156)			0.392 (0.020)		0.221 (0.009)	-918.8	8.540 (0.287)	12.969 (0.738)	97.410 (0.469)
<i>FIX</i>	-0.153 (0.020)	-0.316 (0.030)		0.324 (0.017)		0.205 (0.008)	-870.6	16.064 (0.013)	21.866 (0.148)	102.520 (0.306)
<i>FIMAX</i>	-0.170 (0.025)	-0.336 (0.034)	0.067 (0.031)	0.344 (0.023)	-0.100 (0.037)	0.203 (0.008)	-871.9	12.880 (0.012)	18.043 (0.205)	94.291 (0.472)

Coefficients of the ARFIMAX model defined by equation (2.5.1) are obtained by conditional sum-of-squares maximum likelihood estimation using analytical gradients. The  $(1 - L)^d$  polynomial is truncated at lag 1000. Standard errors, based on the second derivatives of the log-likelihood function, are reported in parentheses under the coefficient estimates.  $SBC$  reports the Schwarz Bayesian Information Criterion ( $SBC = \mathfrak{L}^* - 0.5k \ln(1366)$ ), where  $\mathfrak{L}^*$  denotes the maximized log likelihood and  $k$  the number of estimated coefficients).  $Q_K$  refers to the Ljung-Box portmanteau tests for up to  $K$ th-order serial correlation in the residuals. The numbers in parentheses below these statistics report the probability that the  $K$  autocorrelations are not significant.

of  $d = 0.392$  corresponds closely to the Geweke and Porter-Hudak (1993) log-periodogram estimate of  $d = 0.396$  obtained in our previous section. The FI model estimate is also in accordance with Breidt, Crato and de Lima (1998) who on estimating a ARFIMA(1,d,0) Stochastic Volatility process (without allowing for the asymmetric volatility effect) report  $d = 0.444$  for the CRSP index. Upon fitting a FIEGARCH model to daily returns on the S&P 500 composite stock index, Bollerslev and Mikkelsen (1996) however found  $d = 0.633$ . This estimate is much higher than the ones we report and suggests, contrary to our results, that the logarithmic variance process is not covariance-stationary.

Looking next at the estimates for  $\omega_1$  and  $\omega_2$ , we find support for the asymmetric news-effect. It becomes evident, however, that it is mostly negative and not positive returns that are important for the modeling of logarithmic variances; the estimate of  $\omega_2$  in the FIMAX specification is small and only marginally significant at the 5% level.

The addition of lagged negative and/or positive returns to the FI model induces some low-order serial correlation in the residuals. While for the FI model all reported Ljung-Box  $Q$  statistics are insignificant at the conventional levels, for the FIX and FIMAX model we cannot – at the 5% significance level – reject the null of no 10th-order serial correlation in



the residuals. For the FIMAX model we mitigate this problem by allowing next to the two news-parameters a first-order moving average component; the coefficient on  $\alpha_1$  is however only small and accompanied by a relatively large standard error.

Judged by the Schwarz Bayesian Criterion the asymmetric return-volatility effect is important for the modeling of logarithmic variances. Among the two models that allow for lagged returns, the criterion favors the parsimonious FIX specification which does not include negative returns and a first-order moving average component.

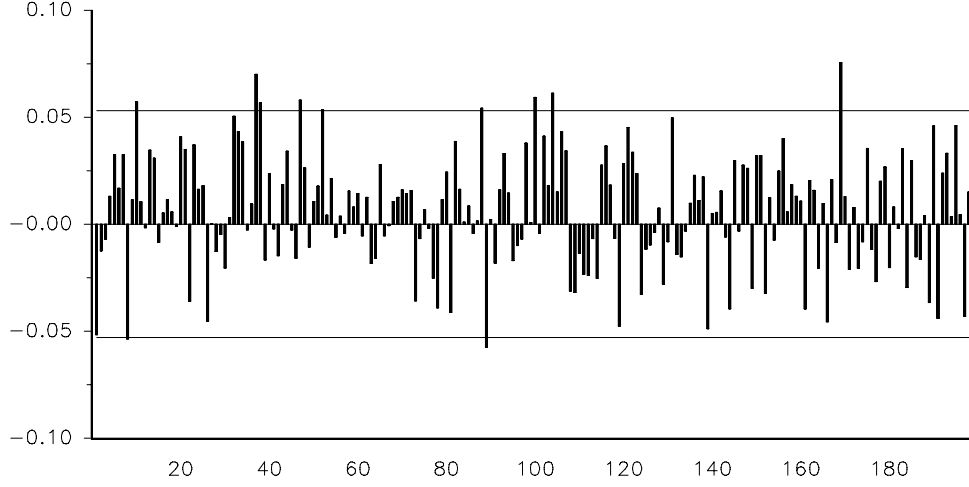
To investigate further the possibility that the FIX model leaves some time-dependency of volatility unexplained, we plot in Figure 2.8 its residual autocorrelation function. The significant  $Q_{10}$  statistic for this model is likely driven by the size of the first, eighth and tenth-order residual autocorrelation. Judged by the 95 percent confidence interval,  $\pm 1.96/\sqrt{T}$ , only the eight and tenth order autocorrelations are significant – however only marginally so. When considering all 200 autocorrelations it becomes evident that the FIX model captures logarithmic variance dynamics rather well. The eleven significant autocorrelations may be attributed to type II error of the test. Above all, the FIX model accounts fully for the slow hyperbolic decay found in the logarithmic variance autocorrelation function (see Figure 2.3). In Figure 2.8, no pattern of decay remains.

### 2.5.2 Realized Volatility Model Predictions

In this subsection we investigate how well the realized logarithmic volatility models set out above predict our three volatility series *ex ante* one-step-ahead. To determine the next-period predictions, it is convenient to rewrite the ARFIMAX model given by equation 2.5.1 more compactly as:

$$\ln(s_t^2) = f(\mathfrak{F}_{t-1}) + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. N(0, \sigma^2) \quad (2.5.2)$$

**Figure 2.8:** Realized Volatility Model Residual Autocorrelations



The graph displays the first 200 residual autocorrelations for the FIX model reported in Table 2.1. The parallel lines are the 95% confidence interval,  $\pm 1.96/\sqrt{T}$ .

where  $\mathfrak{F}_{t-1}$  denotes the information set available at time  $t-1$ . The one-step-ahead variance, standard deviation and logarithmic variance predictions of (2.5.2) evaluated at the estimates given in Table 2.1 are given by:

$$\begin{aligned}
 \hat{s}_t^2 &= E[s_t^2 | \mathfrak{F}_{t-1}] &= e^{\hat{f}(\cdot)} E[e^{\hat{\varepsilon}_t} | \mathfrak{F}_{t-1}] &= e^{\hat{f}(\cdot) + \frac{1}{2} \hat{\sigma}^2} \\
 \hat{s}_t &= E[s_t | \mathfrak{F}_{t-1}] &= e^{\frac{1}{2} \hat{f}(\cdot)} E[e^{\frac{1}{2} \hat{\varepsilon}_t} | \mathfrak{F}_{t-1}] &= e^{\frac{1}{2} \hat{f}(\cdot) + \frac{1}{8} \hat{\sigma}^2} \\
 \ln(\hat{s}_t^2) &= E[\ln(\hat{s}_t^2) | \mathfrak{F}_{t-1}] &= \hat{f}(\cdot) + E[\hat{\varepsilon}_t | \mathfrak{F}_{t-1}] &= \hat{f}(\cdot)
 \end{aligned} \tag{2.5.3}$$

Since in (2.5.2) it is assumed that  $\varepsilon_t \sim N(0, \sigma^2)$  it follows that  $\exp(\varepsilon_t) \sim LN(0, \sigma^2)$  and  $\exp(\frac{1}{2} \varepsilon_t) \sim LN(0, \frac{1}{4} \sigma^2)$ , where  $LN$  denotes the lognormal density.<sup>14</sup> Let  $y_t$  denote one of our three volatility series, *i.e.*  $s_t^2$ ,  $s_t$  and  $\ln(s_t^2)$ , then we evaluate its predictions  $\hat{y}_t$ , *i.e.*  $\hat{s}_t^2$ ,

<sup>14</sup>As one would expect from the discussion at the beginning of this section, the residual innovations coming from our logarithmic variance models display slight asymmetry and excess kurtosis. In particular, we obtain skewness and kurtosis estimates of  $\{0.560, 4.304\}$ ,  $\{0.380, 3.984\}$  and  $\{0.362, 4.134\}$  for the FI, FIX and FIMAX model respectively. However, when we instead compute the expectations of  $\exp(\hat{\varepsilon}_t)$  and  $\exp(\frac{1}{2} \hat{\varepsilon}_t)$  using the mean of these two measures in order to obtain variance and standard deviation predictions, our subsequent results change only little. Furthermore, when we condition the residual innovations on non-normal densities, results hardly change.

$\hat{s}_t$  and  $\ln(\hat{s}_t^2)$ , using the OLS regression:

$$y_t = \alpha + \beta \hat{y}_t + \varepsilon_t \quad (2.5.4)$$

If a prediction is unbiased,  $\alpha = 0$  and  $\beta = 1$ . Table 2.2 reports the ordinary least squares estimates of (2.5.4) and the associated  $R^2$  statistic when applied to variances, standard deviations and logarithmic variances. Standard errors using White's (1980) heteroskedasticity correction are in parentheses.

**Table 2.2:** Realized Volatility Model *Ex Ante* Predictions

	variances			standard deviations			log variances		
	$\hat{\alpha}$	$\hat{\beta}$	$R^2$	$\hat{\alpha}$	$\hat{\beta}$	$R^2$	$\hat{\alpha}$	$\hat{\beta}$	$R^2$
<i>FI</i>	-0.079 (0.058)	1.238 (0.162)	0.379	-0.048 (0.031)	1.086 (0.057)	0.486	0.028 (0.039)	1.024 (0.030)	0.515
<i>FIX</i>	0.000 (0.031)	1.026 (0.085)	0.627	-0.030 (0.023)	1.055 (0.041)	0.576	0.026 (0.035)	1.022 (0.027)	0.551
<i>FIMAX</i>	0.071 (0.040)	0.843 (0.103)	0.607	0.000 (0.028)	1.003 (0.049)	0.576	0.007 (0.035)	1.006 (0.026)	0.554

The table reports ordinary least squares coefficient estimates for the model defined by equation 2.5.4 using the variance, standard deviation and logarithmic variance predictions given by (2.5.3), *i.e.* the *ex ante* one-step-ahead volatility predictions coming from the FI, FIX and FIMAX models reported in Table 2.1. Standard errors using White's (1980) heteroskedasticity correction are in parentheses.

For all three models, the estimates of  $\alpha$  and  $\beta$  are within two standard errors of their hypothesized values. From the  $R^2$  statistics it becomes evident that the realized volatility specifications can explain much that is observed in volatility over our sampling period. The  $R^2$  statistics for logarithmic variances range between 51.5% and 55.4%, for standard deviations between 48.6% and 57.6% and for variances between 37.9% and 62.7%. The addition of lagged negative returns to the FI model (therefore yielding the FIX model) improves only slightly the predictions for logarithmic variances, but has important consequences for the predictions of standard deviations and most notably for variances; the  $R^2$  measure for this latter volatility measure increases by 24.8 percentage points to 62.7%. Little or nothing is however gained by adding positive lagged returns and a moving average component to

the FIX model (therefore yielding the FIMAX model). The  $R^2$  for logarithmic variances increases by only 0.3 percentage points. For standard deviations the  $R^2$  measures are identical and for variances the parsimonious FIX specification yields an even higher  $R^2$  measure than the FIMAX specification that requires two additional parameters.

We plot in Figure 2.9 the one-step-ahead *ex ante* variance predictions implied by the FIX model (solid line) along with the realized variance series (dotted line). Clearly, the one-day ahead predictions do a remarkable job of tracking realized variances over our sample period. Major discrepancies between the two depicted series are generally only noticeable when the realized volatility is unusually high (for instance March 31, 1994 and July 16, 1997). However, for the highest realized variance observation in the sample (October 28, 1997) the FIMAX model predicts a variance of 10.43 while the corresponding realized volatility measure takes value of 9.45 for that day.

The question remains whether employment of realized volatility measures to model volatility leads to improvements or whether perhaps one of the standard techniques yields similar or even better result. We tackle this issue in our next section.

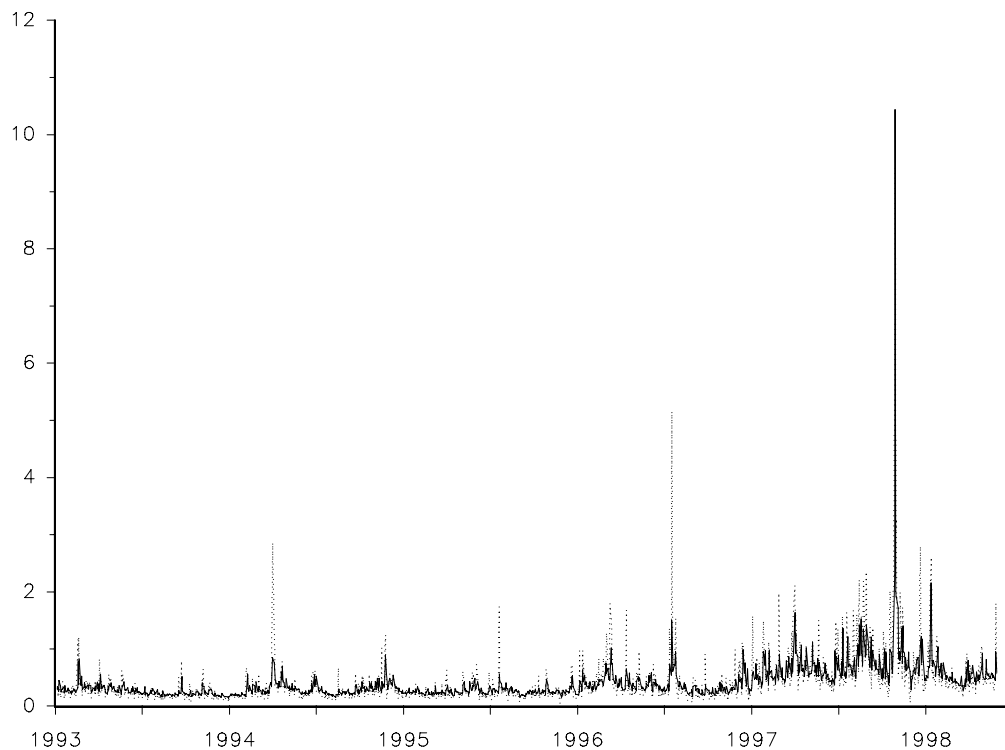
## 2.6 ARCH Volatility Modeling and Predictions

The most common tool for characterizing changing variances is to fit ARCH-type models to daily returns. The performance of some of these models relative to the ones just developed is the subject of this section. We detail next the exact formulations we shall be using. Later in this section we evaluate the volatility predictions implied by these models and this will allow us to directly compare the ARCH models to the realized volatility formulations employed before.

### 2.6.1 ARCH Volatility Modeling

For the parameterization of ARCH models the main findings of our previous sections are: (1) volatilities are covariance stationary and fractionally integrated, (2) volatilities are non-

**Figure 2.9:** Realized Volatility Model *Ex Ante* Variance Predictions



The graph displays the *ex ante* variance predictions implied by the FIX model (solid line) along with the realized variances (dashed line). The FIX model is defined by equation 2.5.1 and its estimates are reported in Table 2.1.

symmetric in lagged returns, (3) returns are not (at best only weakly) correlated with volatilities and (4) the distribution of returns divided by standard deviations is normal. In many applications with daily data it is however found that the distribution of standardized returns is leptokurtic. We shall therefore investigate whether our finding of normality is particular to the data underlying our study or whether perhaps non-normality is confined to the ARCH approach to modeling volatility.

Since the introduction of the ARCH model by Engle (1982) numerous extensions have been proposed.<sup>15</sup> However, only the FIGARCH model developed by Baillie, Bollerslev and Mikkelsen (1996) and the FIEGARCH model formulated Bollerslev and Mikkelsen (1996)

---

<sup>15</sup>Recent studies surveying the various ARCH models include Pagan (1996), Palm (1996), Bollerslev, Engle and Nelson (1994), Bera and Higgins (1993) and Bollerslev, Chou and Kroner (1992).

explicitly allow for the long-memory property of volatility. We shall focus on these two specifications although only the FIEGARCH model allows for the news-effect and can be covariance stationary while allowing for long-memory.

When modeling the conditional variance processes discussed below, we did not find any evidence for temporal dependencies in the conditional mean of returns ( $r_t$ ) other than a constant term ( $\mu$ ). Since we in Section 2.4.3 found hardly any evidence for the ARCH-M effect, we consider return representations of the form:<sup>16</sup>

$$\begin{aligned} r_t &= \mu + \varepsilon_t \\ \varepsilon_t &= \sigma_t z_t \end{aligned} \tag{2.6.1}$$

where  $E[z_t] = 0$  and  $E[z_t^2] = 1$ .

The conditional variance process in the FIGARCH (q,d,p) model is defined as:

$$\sigma_t^2 = \frac{\omega + \left[ (1 - \beta(L_p)) - (1 - \alpha(L_q) - \beta(L_p)) (1 - L)^d \right] \varepsilon_t^2}{(1 - \beta(L_p))} \tag{2.6.2}$$

where  $\alpha(L_q) = \sum_{i=1}^q \alpha_i L^i$ ,  $\beta(L_p) = \sum_{i=1}^p \beta_i L^i$ . The FIGARCH model is covariance stationary only in the special case where  $d = 0$  and then it reduces to Bollerslev's (1986) GARCH specification. The FIGARCH model displays however the important property of having a bounded cumulative impulse-response function for any  $d < 1$ . As in Bollerslev (1987), we condition the innovations  $z_t$  in (2.6.1) on the Student  $t$  distribution, *i.e.*  $z_t \sim T(0, 1, \eta_1)$ . This density has thicker tails than the normal when  $\eta_1 < \infty$ .

Although the FIGARCH model is consistent with our finding of long-memory, for  $d > 0$  the FIGARCH process is, contrary to our findings, not covariance stationary. Furthermore, variances are symmetric in lagged returns and therefore the FIGARCH model does not

---

<sup>16</sup>Nonetheless, for the GARCH (1,1), EGARCH (1,2) and FIEGARCH (0,d,1) models discussed below we still tested whether the ARCH-M specification is appropriate, *i.e.*  $r_t = \mu_1 + \mu_2 \sigma_t^2 + \varepsilon_t$ . As expected, the estimates for  $\mu_2$  were positive, yet insignificant.

permit the leverage effect. These two deficiencies are addressed by the FIEGARCH(p,d,q) model which is defined as:

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(L_q) (\gamma z_t + |z_t| - E[|z_t|])}{(1-L)^d (1-\beta(L_p))} \quad (2.6.3)$$

with all polynomials defined as before. If the leverage effect holds, we expect to find  $\gamma < 0$ . This formulation nests Nelson's (1991) EGARCH model when  $d = 0$ . We condition – as in the original formulation of the EGARCH model – the innovations  $z_t$  on the generalized error distribution, i.e.  $z_t \sim GED(0, 1, \eta_2)$ . The density is normal when  $\eta_2 = 2$  while it displays heavy tails for  $\eta_2 < 2$ . The fractional integration parameter in (2.6.3) has the same interpretation as in the models of our previous section, i.e., the logarithmic variance process is covariance stationary if  $d < 0.5$ . For  $d < 1$  the process is mean-reverting and shocks to volatility decay.

The FIEGARCH model is similar to our realized volatility model in that it seeks long-memory in the logarithmic variance process and allows for the asymmetric news-effect. Whereas our analysis of news impact in Section 2.4 and 2.5 suggests that logarithmic variances are linear in lagged positive and negative returns, the FIEGARCH model conjectures that logarithmic variances increase linearly with negative and positive standardized returns  $(r_{t-1} - \mu)/\sigma_{t-1}$ .<sup>17</sup> The main difference however is that our earlier specifications are in the spirit of Stochastic Volatility models and not of ARCH models.

Maximum likelihood estimates of some formulations nested within (2.6.2) and (2.6.3) – a GARCH(1,1), FIGARCH(1,d,1), EGARCH(1,2) and FIEGARCH(0,1,1) model – are reported in Table 2.3. Coefficient estimates for  $\eta$  carry suffix 1 when the ARCH innovations  $z_t$  are conditioned on the Student  $t$  density and suffix 2 when the generalized error density is used instead. Standard errors, based on the matrix of second derivatives of the log-likelihood function, are in parentheses. With the exception of the fractional integration parameter  $d$

---

<sup>17</sup>Upon holding constant the information dated  $t-2$  and earlier (as in the definition by Engle and Ng 1993), logarithmic variances are however linear in positive and negative  $r_{t-1}$ .

in the FIGARCH model, all reported estimates are significant at the 5% level on the basis of either Wald or log-likelihood ratio tests.  $\mathfrak{L}^*$  reports the maximized log-likelihood.

**Table 2.3:** ARCH Model Estimates

	$\hat{\mu}$	$\hat{\omega}$	$\hat{\beta}_1$	$\hat{d}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\gamma}$	$\hat{\eta}_{1,2}$	$\mathfrak{L}^*$
<i>GARCH</i>	0.063 (0.016)	0.008 (0.005)	0.930 (0.024)		0.054 (0.018)			6.250 <sub>1</sub> (1.043)	-1340.0
<i>FIGARCH</i>	0.063 (0.016)	0.021 (0.012)	0.652 (0.105)	0.375 (0.108)	-0.285 (0.108)			6.536 <sub>1</sub> (1.150)	-1338.6
<i>EGARCH</i>	0.050 (0.015)	-0.884 (0.125)	0.972 (0.014)		0.231 (0.043)	-0.117 (0.047)	-0.596 (0.158)	1.425 <sub>2</sub> (0.075)	-1329.1
<i>FIEGARCH</i>	0.065 (0.014)	-1.245 (0.274)		0.585 (0.056)	0.227 (0.041)		-0.668 (0.039)	1.418 <sub>2</sub> (0.072)	-1326.7

Coefficients of the models defined by equations 2.6.1 and either 2.6.2 or 2.6.3 are obtained by conditional sum-of-squares maximum likelihood estimation using analytical gradients. Coefficient estimates for  $\eta$  carry suffix 1 when the ARCH innovations  $z_t$  are conditioned on the Student  $t$  density and suffix 2 when the generalized error density is used instead. Standard errors, based on the second derivatives of the log-likelihood function, are reported in parentheses under the coefficient estimates.  $\mathfrak{L}^*$  reports the maximized log-likelihood. The  $(1 - L)^d$  polynomial in the FIGARCH and FIEGARCH model is truncated at lag 1000. The data are daily percentage returns for the Dow Jones Industrial Average from January 1993 to May 1998.

Consistent with prior literature on ARCH models, the innovations  $z_t$  are heavy tailed, the implied volatility processes are highly persistent and, when we allow for asymmetry in returns, coefficients on the news parameters suggest the leverage effect. The FIGARCH and FIEGARCH models indicate that shocks to volatility decay (eventually) at a slow hyperbolic rate. Our FIEGARCH estimate of  $d = 0.585$  is in line with the one reported by Bollerslev *et al.* (1996), who found  $d = 0.633$  for the S&P 500 composite stock index. Both estimates are however much higher than the ones we obtained in the context of our realized volatility models and imply, contrary to the findings of our previous sections, that the logarithmic variance process is not covariance stationary.

Judged by the maximized log-likelihood ( $\mathfrak{L}^*$ ), the FIEGARCH model is the most promising ARCH specification for characterizing changing variances. We shall next investigate whether this or any other of the above formulations provide useful volatility predictions.



### 2.6.2 ARCH Volatility Model Predictions

ARCH model predictions are generally evaluated by means of criteria that match squared returns with the volatility predictions implied by a particular model (or some transform of these two series). As we made clear in Section 2.2, the daily squared return is a very noisy indicator of volatility. Following Andersen and Bollerslev (1998), we therefore use realized volatilities to evaluate the ARCH model predictions. Specifically, let  $y_t$  denote one of our three volatility series, i.e.  $s_t^2$ ,  $s_t$  and  $\ln(s_t^2)$ , then we evaluate the corresponding ARCH predictions  $\hat{y}_t$ , i.e.  $\hat{\sigma}_t^2$ ,  $\hat{\sigma}_t$  and  $\ln(\hat{\sigma}_t^2)$ , using the regression:

$$y_t = \alpha + \beta \hat{y}_t + \varepsilon_t \quad (2.6.4)$$

If the predictions are unbiased,  $\alpha = 0$  and  $\beta = 1$ . Table 2.4 reports the ordinary least squares estimates of (2.6.4) and the associated  $R^2$  statistics when applied to variances, standard deviations and logarithmic variances. Standard errors using White's (1980) heteroskedasticity correction are in parentheses.

**Table 2.4:** ARCH Model *Ex Ante* Predictions

	variances			standard deviations			log variances		
	$\hat{\alpha}$	$\hat{\beta}$	$R^2$	$\hat{\alpha}$	$\hat{\beta}$	$R^2$	$\hat{\alpha}$	$\hat{\beta}$	$R^2$
<i>GARCH</i>	0.146 (0.044)	0.526 (0.098)	0.228	0.130 (0.036)	0.682 (0.056)	0.334	-0.443 (0.037)	0.866 (0.036)	0.388
<i>FIGARCH</i>	0.128 (0.057)	0.562 (0.123)	0.283	0.110 (0.040)	0.713 (0.062)	0.379	-0.421 (0.035)	0.888 (0.035)	0.424
<i>EGARCH</i>	-0.165 (0.079)	1.225 (0.175)	0.518	-0.037 (0.037)	0.953 (0.059)	0.457	-0.366 (0.034)	0.904 (0.033)	0.410
<i>FIEGARCH</i>	-0.121 (0.059)	1.163 (0.136)	0.572	-0.008 (0.032)	0.926 (0.051)	0.495	-0.347 (0.032)	0.877 (0.030)	0.444

The table reports ordinary least squares coefficient estimates for the model defined by equation 2.6.4 using the ARCH model variance, standard deviation and logarithmic variance predictions. Standard errors using White's (1980) heteroskedasticity correction are in parentheses.

Turning to the results, we can see that the ARCH model volatility predictions are not always unbiased, but all models can capture much of the variation we observe for our three volatility measures. The  $R^2$  statistics range between 22.8% and 57.2%. The FIEGARCH model clearly performs best. For variances and standard deviations the estimates for  $\alpha$  and  $\beta$  are roughly within two standard errors of their hypothesized values and, compared to all the other ARCH specifications, the  $R^2$  statistics are highest for all three volatility measures.

Recall that the FIX model employed in our previous section gave unbiased volatility predictions and that we obtained for this specification  $R^2$  statistics of 62.7%, 57.6% and 55.1% for variances, standard deviations and logarithmic variances, respectively. Judged by these measures, this realized volatility model clearly improves upon the four ARCH specifications. Yet, the extent of enhancement depends greatly on which formulation is employed. Compared to the standard GARCH model, the FIX model  $R^2$  measures are higher by 39.9, 24.2 and 16.3 percentage points for the respective volatility measure. Compared to the FIEGARCH model, the gains are more modest. The  $R^2$  measures are higher by only 5.5, 8.1 and 10.7 percentage points.

Our result that the realized volatility model performs better is of course only suggestive. There may exist other ARCH models that outperform the models used in this section. Nonetheless, only the FIEGARCH formulation is – in principle – consistent with all the empirical regularities we document. It is therefore doubtful that any other univariate model of the ARCH class could disinter anything more from returns that would be relevant for the prediction of stock return volatility. Moreover, the FIEGARCH model estimates suggest that scaled returns are non-normal and that the volatility process is not covariance stationary – two implications we did not observe using realized volatilities. This perhaps suggests some mis-specification that is confined to the ARCH approach to modeling volatility. An open question remains whether Stochastic Volatility models would perform better. The results throughout this paper suggest that any such formulation would need to account for the long-memory property of volatility. Although it is possible to obtain parameter estimates of

fractionally integrated Stochastic Volatility models (e.g. Breidt *et al.* 1998), for these type of models one cannot extract volatility predictions from the data.<sup>18</sup> Any comparison along the lines we have pursued is therefore not possible.

## 2.7 Conclusions

Using 5-minute squared returns on the Dow Jones Industrials Average portfolio over the January 1993 to May 1998 period, we documented the properties of daily stock return volatility. We found that the distributions of variances and standard deviations are skewed-right and leptokurtic, but that logarithmic variances are distributed approximately normal. All three volatility measures are (a) covariance stationary, (b) highly persistent, (c) very little correlated with current returns (no ARCH-M effect) and (d) correlated more strongly with lagged negative than lagged positive returns (news-effect). The news effect can explain some of the asymmetry and flatness of tails in the distribution of the three volatility series – most notably for variances and standard deviations.

We fitted a fractionally integrated model that accounts for the news-effect directly to logarithmic variances. Using *ex ante* one-day-ahead prediction criteria we found that this model yields unbiased and accurate variance, standard deviation and logarithmic variance predictions and that these predictions are better than the ones obtained by the GARCH, FIGARCH, EGARCH and FIEGARCH models. Among these four ARCH specifications, the FIEGARCH formulation performed best. However, the estimate of the fractional integration parameter given by this specification implies that the logarithmic variance process is not covariance stationary. For all ARCH models we found that the distribution of returns divided by the implied standard deviations is leptokurtic. When using realized standard deviations instead, normality of this distribution cannot be rejected.

---

<sup>18</sup>A survey of Stochastic Volatility models can be found in Ghysels, Harvey and Renault (1996).