

NBER WORKING PAPER SERIES

ANSWERING THE CRITICS: YES, ARCH
MODELS DO PROVIDE GOOD
VOLATILITY FORECASTS

Torben G. Andersen
Tim Bollerslev

Working Paper 6023

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1997

We are grateful to Olsen and Associates for making the intradaily exchange rate quotations underlying portions of the empirical analysis in the paper available. We also thank seminar participants at Wharton, USC, Georgetown and York Universities, Academia Sinica Taipei, the December 1996 Triangle Econometrics Workshop, and the Time Series Analysis of High Frequency Financial Data conference in San Diego, April 1997, for valuable comments. Needless to say, we remain fully responsible for the content. This paper is part of NBER's research program in Asset Pricing. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1997 by Torben G. Andersen and Tim Bollerslev. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Answering the Critics: Yes, ARCH Models Do
Provide Good Volatility Forecasts
Torben G. Andersen and Tim Bollerslev
NBER Working Paper No. 6023
April 1997
JEL Nos. C22, C52, C53, G12, G13
Asset Pricing

ABSTRACT

Volatility permeates modern financial theories and decision making processes. As such, accurate measures and good forecasts of future volatility are critical for the implementation and evaluation of asset pricing theories. In response to this, a voluminous literature has emerged for modeling the temporal dependencies in financial market volatility at the daily and lower frequencies using ARCH and stochastic volatility type models. Most of these studies find highly significant in-sample parameter estimates and pronounced intertemporal volatility persistence. Meanwhile, when judged by standard forecast evaluation criteria, based on the squared or absolute returns over daily or longer forecast horizons, ARCH models provide seemingly poor volatility forecasts. The present paper demonstrates that ARCH models, contrary to the above contention, produce strikingly accurate interdaily forecasts for the latent volatility factor that is relevant for most financial applications.

Torben G. Andersen
Department of Finance
J.L. Kellogg Graduate School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
t-andersen@nwu.edu

Tim Bollerslev
Department of Economics
Rouss Hall
University of Virginia
Charlottesville, VA 22901
and NBER
bollerslev@virginia.edu

Volatility is a key ingredient of the risk-return tradeoff that permeates modern financial theories. Moreover, it is a well-established fact, dating back to Mandelbrot (1963) and Fama (1965), that financial markets display pronounced volatility clustering. As such, accurate measures and a thorough understanding of the determinants of the volatility process are critical for issues related to the functioning of markets and the implementation and evaluation of asset pricing theories. However, it is only over the last decade that financial economists have begun to seriously model the temporal dependencies in return volatility. The majority of these studies rely on formulations within the Autoregressive Conditional Heteroskedastic (ARCH) class of models pioneered by Engle (1982). Almost universally, the reported parameter estimates point towards a very high degree of intertemporal volatility persistence, supposedly underscoring the importance of capturing this dimension of the return generating process; see, e.g., Bollerslev, Chou and Kroner (1992) and Bollerslev, Engle and Nelson (1994) for recent surveys. Yet, a number of studies find that, although the ARCH parameters are highly significant in-sample, the models explain little of the variability in ex-post volatility as measured by the squared or absolute returns over the relevant forecast horizon; see, e.g., Cumby, Figlewski and Hasbrouck (1993), Figlewski (1994), and Jorion (1995, 1996) among many others. Predictably, these rather striking findings have led to the perception that ARCH volatility forecasts may be of limited practical use.

In contrast, we demonstrate that well-specified ARCH models yield surprisingly accurate volatility forecasts. While normalized squared or absolute returns over the appropriate horizon provide unbiased estimates for volatility, the signal-to-noise ratio is diminutive, when evaluations are conducted over daily time-spans. Consequently, we demonstrate that the apparent poor predictive power of ARCH models, when judged by standard forecast evaluation criteria, is a natural, indeed inevitable, consequence of the inherent noise in the return generating process. Building on the continuous-time stochastic volatility framework developed by Nelson (1990) and Drost and Werker (1996), we furthermore demonstrate how high-frequency intraday returns allow for the construction of increasingly more accurate ex-post volatility measurements. When evaluated in this more appropriate setting, we find that daily ARCH models readily explain about half of the variability in the latent volatility factor. These findings thus document the importance and practical relevance of interdaily volatility modeling and forecasting.

Although the general conclusions apply equally well across most financial markets and instruments, our empirical analysis will, for concreteness, focus on the foreign exchange market and

daily volatility forecasts for the Deutschemark - U.S. Dollar (DM-\$) and Japanese Yen - U.S. Dollar (¥-\$) spot exchange rates. Similarly, all our model estimates and theoretical developments are based on the popular GARCH(1,1) specification of Bollerslev (1986), but the qualitative results carry over to other empirically relevant ARCH and stochastic volatility models.

The plan for the remainder of the paper is as follows. The notation and data sources are set out in section I. Section II provides a brief empirical illustration of the highly significant ARCH parameter estimates typically obtained in-sample, coupled with the apparent poor out-of-sample ARCH volatility forecasting performance. These illustrations are based on a sample of daily DM-\$ and ¥-\$ exchange rates. Formal theoretical justifications for the empirical findings are developed in section III within the context of a continuous-time stochastic volatility model. This section also demonstrates how the use of high-frequency data effectively reduces the measurement error involved in quantifying the ex-post latent volatility. Utilizing a one-year sample of five-minute returns, the empirical analysis in section IV highlights how the resulting improved daily volatility measures give rise to very different conclusions regarding the accuracy of the identical daily ARCH volatility forecasts for the two exchange rates discussed in section I. Section V concludes with suggestions for future research.

I. Notation and Data

We let p_t denote the time $t \geq 0$ logarithmic price for some financial asset, where the unit time interval corresponds to one day. The discretely observed time series process of continuously compounded returns with m observations per day, or a return horizon of $1/m$, is then defined by,

$$r_{(m),t} \equiv p_t - p_{t-1/m}, \tag{1}$$

where $t = 1/m, 2/m, \dots$. In line with this convention, conditional and unconditional expectations are indexed by the observation frequency of the variables in the information set, and are denoted by $E_{(m),t}(\cdot)$ and $E_{(m)}(\cdot)$, respectively, while the corresponding variance operators are given by $Var_{(m),t}(\cdot)$ and $Var_{(m)}(\cdot)$. We further refer to the continuous-time instantaneous returns process by $r_t \equiv r_{(\infty),t} \equiv dp_t$, while the instantaneous variance is denoted σ_t^2 . Likewise, the conditional expectation adapted to the continuous-time sample-path filtration, $\sigma(p_\tau, \tau \leq t)$, is referred to by $E_t(\cdot)$, whereas the corresponding unconditional expectation is denoted $E(\cdot)$. To facilitate comparison, all reported

population figures and model estimates are scaled to reflect daily percentage returns.

The model estimates underlying the continuous-time simulations are based on daily returns, or $r_{(1),t}$, for the DM-\$ and the ¥-\$ spot exchange rates from October 1, 1987 through September 30, 1992. Meanwhile, the empirical out-of-sample forecast analysis is based on temporal aggregates of the five-minute returns, or $r_{(288),t}$, for the same two exchange rates from October 1, 1992, through September 30, 1993. These intraday returns are constructed from the linearly interpolated logarithmic midpoint of the continuously-recorded bid and ask quotes that appeared on the interbank Reuters network over the one-year sample. Due to the extremely low market activity over the weekends, the returns from Friday 21:00 Greenwich Mean Time (GMT) through Sunday 21:00 GMT are excluded, resulting in a total of 74,880 5-minute returns spanning 260 days. For a more detailed discussion of the data construction we refer to Andersen and Bollerslev (1997a, 1997b), where the identical five-minute DM-\$ return series is analyzed from a different perspective.

II. Interdaily Volatility Modeling and Forecast Evaluation

The existence of volatility clustering in daily, weekly, or monthly speculative returns has been extensively documented in the literature. This feature is also evident in figure 1, which plots the daily returns on the DM-\$ and ¥-\$ spot exchange rates over the five-year sample. Even though the returns appear to be serially uncorrelated, the evidence for volatility clustering is ubiquitous, with both exchange rates exhibiting well-defined periods of relative tranquility and turbulence. This visual impression is confirmed by the Ljung and Box (1978) portmanteau tests for up to thirtieth-order serial correlation in the squared returns. The test statistics equal 89.0 and 206.1, respectively, which are highly significant in the asymptotic chi-square distribution with thirty degrees of freedom.

A. Daily GARCH(1,1) Volatility Modeling

Econometric modeling of the volatility clustering has been an active research area in recent years. Many studies find that the simple GARCH(1,1) model provides a good first approximation to the observed temporal dependencies; see, e.g., Baillie and Bollerslev (1989), Bollerslev (1987), Engle and Bollerslev (1986), Hsieh (1989) and McCurdy and Morgan (1988) for some early evidence. In order to formally define the model, let $\sigma_{(m),t}^2$ denote the conditional variance of $r_{(m),t}$ based on information up through time $t - 1/m$. With a sampling frequency of m observations per day, the GARCH(1,1) model for $r_{(m),t}$ is then given by the following system,

$$r_{(m),t} = \sigma_{(m),t} \cdot z_{(m),t} \quad (2)$$

and

$$\sigma_{(m),t}^2 = \psi_{(m)} + \alpha_{(m)} \cdot (\sigma_{(m),t-1/m} \cdot z_{(m),t-1/m})^2 + \beta_{(m)} \cdot \sigma_{(m),t-1/m}^2 \quad (3)$$

where $\psi_{(m)} > 0$, $\alpha_{(m)} \geq 0$, $\beta_{(m)} \geq 0$, and $z_{(m),t}$ is i.i.d. with mean zero and variance one.

The parameter estimates for the two daily exchange rates, corresponding to $m = 1$, are reported in table I.¹ For none of the rates do we find any evidence of important serial dependence in the mean. For instance, the portmanteau tests for up to thirtieth-order serial correlation in the standardized residuals, or $\hat{z}_{(1),t}$, equal 33.9 and 39.6, respectively, which are insignificant relative to the conventional ninety-five percent chi-square critical value of 43.8. In contrast, the estimates for the conditional variance parameters are all highly significant, and the robust Wald tests for no ARCH effects, $\alpha_{(1)} = \beta_{(1)} = 0$, overwhelmingly reject for both rates. The portmanteau tests for up to thirtieth-order serial dependence in the standardized squared residuals, $\hat{z}_{(1),t}^2$, equal 29.1 and 24.7, indicating that the GARCH(1,1) model does a good job of tracking the short-run interdaily volatility dependencies. Consistent with the prior literature, the estimates for $\hat{\alpha}_{(1)} + \hat{\beta}_{(1)}$ are close to unity, thus approaching the IGARCH(1,1) model of Engle and Bollerslev (1986).²

The high degree of volatility persistence, coupled with the significant parameter estimates, observed almost universally across different speculative returns, suggest that financial market volatility is highly predictable. Specifically, consider the GARCH(1,1) model in equations (2) and (3). Assuming that $\alpha_{(m)} + \beta_{(m)} < 1$, so that the model is covariance stationary, it follows that the minimum Mean Square Error (MSE) forecast for the conditional variance h-steps ahead is given as,

$$E_{(m),t}(\sigma_{(m),t+h/m}^2) = \sigma_{(m)}^2 + (\alpha_{(m)} + \beta_{(m)})^h \cdot (\sigma_{(m),t}^2 - \sigma_{(m)}^2),$$

¹ The estimates are quasi maximum likelihood (QMLE) under the assumption that $z_{(1),t}$ is normally distributed, with robust standard errors in parentheses; see Bollerslev and Wooldridge (1992). The models also allow for intercepts in the conditional mean equations, but these estimates are indistinguishably different from zero and consequently not reported.

² Recent evidence suggest that the long-run dependencies in financial market volatility may be better characterized by a fractionally integrated, or FIGARCH, model; see e.g. Andersen and Bollerslev (1996), Baillie, Bollerslev and Mikkelsen (1996) and Bollerslev and Mikkelsen (1996). Since the present analysis is focused exclusively on short-term volatility forecasting, we shall not pursue these more complicated specifications any further here.

where $\sigma_{(m)}^2 = \psi_{(m)} \cdot (1 - \alpha_{(m)} - \beta_{(m)})^{-1}$ denotes the unconditional one-period return variance. The half-life of a volatility shock, defined as the time it takes for half of the expected reversion towards the unconditional variance to occur, therefore equals $-\log(2) \cdot \log(\alpha_{(m)} + \beta_{(m)})^{-1}$. For the parameter estimates in table I, this translates into half-lives of 19.8 and 12.9 trading days, respectively.³ Nonetheless, as we confirm in the next section, when judged by standard criteria, the model appears to provide poor volatility forecasts, even over the immediate one-day-ahead horizon.

B. Daily GARCH(1,1) Volatility Forecast Evaluation

The majority of the volatility forecast evaluations reported in the literature rely on some MSE criteria involving the ex-post squared or absolute returns over the relevant forecast horizon.⁴ One particularly popular metric is obtained via the ex-post squared return - volatility regression,

$$r_{(m),t+1/m}^2 = a_{(m)} + b_{(m)} \cdot \sigma_{(m),t+1/m}^2 + u_{(m),t+1/m}, \quad (4)$$

where $t = 0, 1/m, 2/m, \dots$. This regression equation provides an analogue to the commonly used procedure for evaluating forecasts for the conditional mean.⁵ If the model for the conditional variance is correctly specified and $E_{(m),t}(r_{(m),t+1/m}^2) = \sigma_{(m),t+1/m}^2$, it follows that, in population, $a_{(m)}$ and $b_{(m)}$ equals zero and unity, respectively.^{6,7} Of course, in practice the values for $\sigma_{(m),t+1/m}^2$ are

³ These half-lives are actually lower than many previous estimates reported in the literature. This may in part reflect our reliance on a relatively short five year sample and the associated downward bias in the parameter estimates; see e.g. the developments in Linton (1997) and Lumsdaine (1995). For instance, on using a longer fourteen year time-span of DM-\$ exchange rates and a GARCH(1,1) specification, Andersen and Bollerslev (1997a) report a half-life of 31.2 trading days.

⁴ Although the MSE may be a natural choice when evaluating traditional model forecasts for the conditional mean, it is less obvious in a heteroskedastic environment; see, e.g., Bollerslev, Engle and Nelson (1994), Engle et al. (1993), Diebold and Mariano (1995), Lopez (1995), and West, Edison and Cho (1993). However, for simplicity we do not pursue any of these more complex non-linear forecast evaluation criteria here.

⁵ Following Mincer and Zarnowitz (1969), the corresponding procedure for evaluating forecasts for the conditional mean is frequently referred to as a Mincer-Zarnowitz regression.

⁶ This assumes that the conditional mean of $r_{(m),t}$ is zero. Otherwise, replace $r_{(m),t}^2$ in equation (4) by $(r_{(m),t} - \mu_{(m),t})^2$, where $\mu_{(m),t}$ denotes the conditional mean; Pagan and Ullah (1988) and Pagan and Sabau (1992) analyze the complications that arise when the conditional mean depends on the conditional variance. However, at the daily horizon, the predictability in the mean is of second order importance, as exemplified by the results for the DM-\$ and ¥-\$ exchange rates discussed above.

subject to estimation error, resulting in a standard errors-in-variables problem and a downward bias in the regression estimate for $b_{(m)}$.⁸ Nonetheless, the coefficient of multiple determination, or $R^2_{(m)}$, from the regression in (4) provides a direct assessment of the variability in the ex-post volatility, as measured by $r^2_{(m),t+1/m}$, that is explained by the particular estimates of $\sigma^2_{(m),t+1/m}$. The $R^2_{(m)}$ is therefore often interpreted as a simple gauge of the degree of predictability in the volatility process, and hence of the potential economic significance of the volatility forecasts.

The use of R^2 as a guide to the accuracy of the volatility forecasts is, however, problematic. Financial applications focus on the future volatility and not the subsequent realized squared returns. Under the null hypothesis that the estimated GARCH(1,1) model constitutes the correct specification, the true return variance is, by definition, identical to the GARCH volatility forecast. Thus, under this scenario the R^2 simply measures the extent of idiosyncratic noise in squared returns relative to the mean which is given by the (true) conditional return variance. On the other hand, if the regression is used as a diagnostic for potential misspecification, then an alternative measure of the realized return volatility is required. Implicitly, the observed squared returns are used in this capacity. This is justified to the extent that the squared returns provide an unbiased estimator of the underlying latent volatility. However, the realized squared returns are poor estimators of the day-by-day movements in volatility, as the idiosyncratic component of daily returns is large. In other words, it is unclear how to interpret the resulting R^2 , unless we establish a benchmark for the value expected under the null hypothesis of correct model specification.⁹

To illustrate these points, consider the GARCH(1,1) estimates for the daily DM-\$ and ¥-\$ exchange rates. The $R^2_{(t)}$'s from the one-step-ahead return volatility regressions in (4) for the 260 weekday returns over the subsequent year from October 1, 1992 through September 30, 1993, equal

⁷ A closely related regression, $|r_{(m),t+1/m}| = c_{(m)} + d_{(m)} \cdot \sigma_{(m),t+1/m} + v_{(m),t+1/m}$, has been employed in a number of studies; see, e.g., Jorion (1995). However, unlike $b_{(m)}$, the population value of $d_{(m)}$ hinges on distributional assumptions. For simplicity we therefore concentrate exclusively on the squared return - volatility regression in (4) in the present analysis.

⁸ If the forecasts are unbiased in population, the downward bias in the estimate for $b_{(m)}$ is given as $-\text{Var}_{(m)}(\nu_{(m),t}) \cdot [\text{Var}_{(m)}(\nu_{(m),t}) + \text{Var}_{(m)}(\sigma^2_{(m),t})]^{-1}$, where $\nu_{(m),t}$ denotes the measurement error in $\sigma^2_{(m),t}$; see, e.g., Chow (1983). Christensen and Prabhala (1997) explicitly recognize this bias within the context of evaluating variance forecasts based on implied volatilities from options prices.

⁹ The predication on R^2 as a convenient measure for summarizing predictable changes in returns, and stock prices in particular, is highlighted by Roll (1988) in his 1987 Presidential Address to the American Finance Association succinctly entitled "R²".

0.047 and 0.026, respectively.¹⁰ These "disappointingly" low $R^2_{(l)}$'s are in line with the evidence in the extant literature for other speculative returns and sample periods.¹¹ For instance, on evaluating the predictive power of a GARCH(1,1) model for weekly returns on the S&P100 stock index from 1983-1989, Day and Lewis (1992) report $R^2_{(l/5)} = 0.039$, while Pagan and Schwert (1990) find $R^2_{(l/22)} = 0.067$ with a GARCH(1,2) model for monthly aggregate U.S. stock market returns from 1835-1925. Jorion (1996) uses the same GARCH(1,1) specification as here, but a longer seven-year sample of daily DM-\$ returns from 1985-1992, to obtain $R^2_{(l)} = 0.024$. Modeling weekly stock and bond market volatility in the U.S. and Japan from 1977-1990 by an EGARCH model, Cumby, Figlewski and Hasbrouck (1993) report $R^2_{(l/5)}$'s ranging from 0.003 to 0.106, while West and Cho (1995) find $R^2_{(l/5)}$'s ranging from 0.001 to 0.045 with a GARCH(1,1) model for five different weekly U.S. dollar exchange rates from 1973-1989. Closely related results have been reported by, e.g., Akgiray (1989), Brailsford and Faff (1996), Canina and Figlewski (1993), Dimson and Marsh (1990), Frennberg and Hansson (1995), Figlewski (1994), Heynen and Kat (1994), Jorion (1995), Lamoureux and Lastrapes (1993), Schwert (1989, 1990a) and Schwert and Seguin (1990). Predictably, these systematically low $R^2_{(m)}$'s reported throughout the literature have led to the perception that standard ARCH models may be seriously misspecified and provide poor volatility forecasts, and consequently are of limited, if any, practical use.

To highlight the fallacy of such an inference, we derive the population R^2 under the null hypothesis that the returns are generated by a GARCH(1,1) model as in equations (2) and (3). Letting $\kappa_{(m)} \equiv E_{(m),t}(z^4_{(m),t})$ denote the conditional kurtosis of the standardized innovations, it is straightforward, following Bollerslev (1986), to show that, provided the unconditional kurtosis for $r_{(m),t}$ is finite, or $\kappa_{(m)} \cdot \alpha^2_{(m)} + \beta^2_{(m)} + 2 \cdot \alpha_{(m)} \cdot \beta_{(m)} < 1$, we have

$$\text{Var}_{(m)}(r^2_{(m),t}) = \psi^2_{(m)} \cdot (\kappa_{(m)} - 1) \cdot (1 - \beta^2_{(m)} - 2 \cdot \alpha_{(m)} \cdot \beta_{(m)}) \cdot (1 - \kappa_{(m)} \cdot \alpha^2_{(m)} - \beta^2_{(m)} - 2 \cdot \alpha_{(m)} \cdot \beta_{(m)})^{-1} \cdot (1 - \alpha_{(m)} - \beta_{(m)})^{-2},$$

¹⁰ Consistent with the results in table I, all our out-of-sample predictions use daily returns measured at 12:00 GMT. While the reported figures do reflect the actual definition of the daily return interval, the qualitative conclusions are robust. For instance, on measuring the daily returns at 0:00 GMT, the one-year out-of-sample $R^2_{(1)}$'s equal 0.021 and 0.012, respectively.

¹¹ Whereas the results for the DM-\$ and ¥-\$ exchange rates reported here are truly out-of-sample, most of the results reported in the literature rely on in-sample parameter estimates. If anything, this is likely to bias the $R^2_{(m)}$'s upward.

and,

$$\text{Var}_{(m)}(\sigma_{(m),t}^2) = \psi_{(m)}^2 \cdot (\kappa_{(m)} - 1) \cdot \alpha_{(m)}^2 \cdot (1 - \kappa_{(m)} \cdot \alpha_{(m)}^2 - \beta_{(m)}^2 - 2 \cdot \alpha_{(m)} \cdot \beta_{(m)})^{-1} \cdot (1 - \alpha_{(m)} - \beta_{(m)})^2.$$

Thus, the (true) population $R_{(m)}^2$ from the regression in equation (4) takes the simple form,

$$R_{(m)}^2 = \text{Var}_{(m)}(\sigma_{(m),t}^2) \cdot \text{Var}_{(m)}(r_{(m),t}^2)^{-1} = \alpha_{(m)}^2 \cdot (1 - \beta_{(m)}^2 - 2 \cdot \alpha_{(m)} \cdot \beta_{(m)})^{-1}. \quad (5)$$

By the implicit assumption of a finite unconditional fourth order moment underlying the squared return - volatility regression, the coefficient of multiple determination will never exceed $\kappa_{(m)}^{-1}$. In particular, with conditional Gaussian errors the $R_{(m)}^2$ from a correctly specified GARCH(1,1) model is bounded from above by $1/3$, while with conditional fat-tailed errors the upper bound is even lower. Moreover, with realistic parameter values for $\alpha_{(m)}$ and $\beta_{(m)}$, the population value for the $R_{(m)}^2$ statistic is significantly below this upper bound. In other words, low R^2 's are not an anomaly, but rather a direct implication of standard volatility models.

Consider again the daily DM-\$ and ¥-\$ GARCH(1,1) parameter estimates for $\alpha_{(t)}$ and $\beta_{(t)}$ in table I. The population $R_{(t)}^2$'s implied by the ex-post volatility regression in (4) equal 0.064 and 0.096, respectively. While these $R_{(t)}^2$'s are slightly higher than the actual one-year out-of-sample statistics calculated above, the values are in close accordance with the in-sample $R_{(m)}^2$'s reported in the extant literature.¹² Thus, even for a correctly specified model with $E_{(m),t}(r_{(m),t+1/m}^2) = \sigma_{(m),t+1/m}^2$ it is naive to expect a "high" $R_{(m)}^2$ from the squared return - volatility regression in (4).

The fact that the daily GARCH(1,1) models for the DM-\$ and ¥-\$ exchange rates do not explain much of the variability in the squared returns is also evident from figure 2, which graph the 260 one-day-ahead volatility forecasts from October 1, 1992 through September 30, 1993, along with the corresponding realized daily squared returns. The variability in $\sigma_{(t),t}^2$ is diminutive compared to the variability in $r_{(t),t}^2$. It is evident that the low $R_{(t)}^2$'s largely reflect the inherent noise in the daily squared returns as a measure for the underlying latent volatility factor. The next section further

¹² By ignoring the higher volatility following market closures, the GARCH(1,1) models reported in table I systematically over-estimate volatility on regular trading days, possibly explaining part of the discrepancy between the actual and population $R_{(t)}^2$'s; see Baillie and Bollerslev (1989) and Andersen and Bollerslev (1997b) for a detailed analysis of day-of-the-week and holiday effects in the foreign exchange market. For simplicity, we do not pursue this additional complication here.

explores this fundamental issue within the context of a continuous-time stochastic volatility model.

III. Continuous-Time Volatility Modeling and Forecast Evaluation

The results of the previous section pose some important challenges. First, while the low R^2 measures are consistent with standard volatility models, they do not answer the underlying question of interest, namely whether these models actually provide valuable volatility forecasts. Second, the null hypothesis that a GARCH(1,1) model constitutes the true data generating process at the daily frequency does not provide a convenient analytical framework for an exploration of this issue. Instead, we adopt a continuous-time diffusion setting in which so-called weak-form GARCH models apply to all time series obtained by sampling at a fixed frequency from this diffusion. The approach is particularly useful because it allows for a straightforward evaluation of interdaily volatility forecasts. Furthermore, many theoretical asset pricing models and most derivatives pricing theories are cast in a similar framework.

Specifically, we assume that the instantaneous returns are generated by the continuous-time martingale,

$$dp_t = \sigma_t \cdot dW_{p,t}, \quad (6)$$

where $W_{p,t}$ denotes a standard Wiener process.¹³ By Ito's Lemma, the minimum MSE forecast for the conditional variance for the one-day returns, or $r_{(1),t+1} \equiv p_{t+1} - p_t$, is then readily expressed as,

$$E_t(r_{(1),t+1}^2) = E_t\left(\int_0^1 r_{t+\tau}^2 d\tau\right) = E_t\left(\int_0^1 \sigma_{t+\tau}^2 d\tau\right) = \int_0^1 E_t(\sigma_{t+\tau}^2) d\tau.$$

Of course, with time-varying volatility it is generally the case that $E_{(1),t}(r_{(1),t+1}^2) \neq E_t(r_{(1),t+1}^2)$. Thus, any discrete-time daily ARCH forecast is necessarily inefficient in a MSE sense relative to the optimal forecast based on the continuous sample path. Furthermore, the earlier discussion of the $R_{(1)}^2$ statistics suggests that this inefficiency may be substantial, seriously questioning the usefulness of ARCH volatility modeling. Meanwhile, it is evident that the proper assessment of daily volatility

¹³ Any mean predictability could easily be incorporated into the subsequent analysis, but the assumption of serially uncorrelated mean-zero returns in (6) greatly simplifies the notation. This assumption is also consistent with the empirical evidence for the two exchange rates analyzed throughout.

in this setting should be based on $\int_0^t \sigma_{t+\tau}^2 d\tau$ as opposed to $r_{(1),t+1}^2$. This is explicitly acknowledged in the literature on derivative security pricing under stochastic volatility; see, e.g., Hull and White (1987), Melino (1994), Scott (1987) and Wiggins (1987). The relevant gauge on the precision of daily ARCH volatility forecasts is therefore also given by the former statistic.

A. Continuous Time GARCH(1,1) Volatility Modeling

In our setting, the natural continuous-time model for the volatility process is given by the diffusion limit of the GARCH(1,1) process, as developed in Nelson (1990).¹⁴ It takes the form,

$$d\sigma_t^2 = \theta(\omega - \sigma_t^2) \cdot dt + (2\lambda\theta)^{1/2} \sigma_t \cdot dW_{\sigma,t}, \quad (7)$$

where $\omega > 0$, $\theta > 0$, $0 < \lambda < 1$, and the Wiener processes, $W_{p,t}$ and $W_{\sigma,t}$, are independent.

While the exact discretization for stochastic volatility models typically are not available in closed form, it follows from Drost and Nijman (1993) and Drost and Werker (1996) that returns obtained by discrete sampling from the system defined by the equations (6) and (7), $r_{(m),t} \equiv p_t - p_{t-1/m}$, satisfy the weak GARCH(1,1) model restrictions,

$$\zeta_{(m),t}^2 = \psi_{(m)} + \alpha_{(m)} \cdot r_{(m),t-1/m}^2 + \beta_{(m)} \cdot \zeta_{(m),t-1/m}^2, \quad (8)$$

where $\zeta_{(m),t}^2$ refers to the linear projection of $r_{(m),t}^2$ on the Hilbert space spanned by 1 , $r_{(m),t-1/m}$, $r_{(m),t-2/m}$, ... and $r_{(m),t-1/m}^2$, $r_{(m),t-2/m}^2$, Although the formal interpretations differ, the recursions for the weak GARCH model defined by (8) and the conditional variance in (3) obviously result in identical numerical values for $\zeta_{(m),t}^2$ and $\sigma_{(m),t}^2$. We therefore refer to the one-day-ahead weak GARCH(1,1) projections as $\sigma_{(1),t}^2$, instead of $\zeta_{(1),t}^2$, in the sequel. However, insofar as $E_{(1),t}(r_{(1),t}^2) \neq \sigma_{(1),t}^2$, the results for the daily GARCH(1,1) forecasts provide only a lower bound on the predictability afforded by higher order discrete-time ARCH approximations. Nonetheless, given the weak GARCH(1,1) interpretation of the diffusion approximation in (8), more complicated stochastic differential equations should at best result in minor improvements relative to the findings below.

¹⁴ Note, however, that many other properly designed ARCH filters will yield consistent estimates for the same σ_t process as the sampling frequency increases; see Nelson (1996) and Nelson and Foster (1994).

The exact one-to-one relationship between the discrete-time weak GARCH(1,1) parameters and the continuous-time stochastic volatility parameters in equation (7) is conveniently expressed by,

$$\theta = -m \cdot \log(\alpha_{(m)} + \beta_{(m)}), \quad (9)$$

$$\omega = m \cdot \psi_{(m)} \cdot (1 - \alpha_{(m)} - \beta_{(m)})^{-1}, \quad (10)$$

and

$$\begin{aligned} \lambda = & 2 \cdot \alpha_{(m)} \cdot \log^2(\alpha_{(m)} + \beta_{(m)}) \cdot [1 - \beta_{(m)} \cdot (\alpha_{(m)} + \beta_{(m)})] \cdot \\ & \{ [1 - (\alpha_{(m)} + \beta_{(m)})^2] \cdot (1 - \beta_{(m)})^2 - \alpha_{(m)} \cdot [1 - \beta_{(m)} \cdot (\alpha_{(m)} + \beta_{(m)})] \cdot \\ & [6 \cdot \log(\alpha_{(m)} + \beta_{(m)}) + 2 \cdot \log^2(\alpha_{(m)} + \beta_{(m)}) + 4 \cdot (1 - \alpha_{(m)} - \beta_{(m)})] \}^{-1}. \end{aligned} \quad (11)$$

Equation (9) implies that $\lim_{m \rightarrow \infty} (\alpha_{(m)} + \beta_{(m)}) = 1$, so the weak GARCH(1,1) model converges to the IGARCH(1,1) case of Engle and Bollerslev (1986) as the sampling frequency increases. This diffusion approximation therefore provides a possible rationale for the widespread empirical findings of apparent IGARCH behavior, as originally argued in Nelson (1990).

The continuous-time parameters implied by the daily, or $m = 1$, GARCH(1,1) estimates for the exchange rates reported in table I above, are listed in table II. These parameters correspond quite closely to those implied by the daily GARCH(1,1) estimates in Baillie and Bollerslev (1989) over the earlier 1980-85 sample period for the same exchange rates, as reported in Drost and Werker (1996). The parameters in table II are also in line with the results reported elsewhere for other stochastic volatility models and alternative estimation procedures; see, e.g., Andersen (1994), Jacquier, Polson and Rossi (1994), Shephard (1996) and the collection of papers in Rossi (1996). As such, the findings based on the particular diffusion parameterizations in table II serve as a realistic gauge on the degree of predictability afforded by daily discrete-time ARCH approximations to the continuous-time specifications typically employed in the theoretical asset pricing literature.¹⁵

¹⁵ While continuous-time diffusions provide a convenient framework for asset pricing, the specifications in (6) and (7) ignore pertinent market microstructure features. For instance, non-synchronous trading induces negative serial correlation in individual returns, whereas index returns become positively correlated. Similarly, the bid-ask spread on organized exchanges, as well as the systematic positioning of quotes in dealer markets, cause the observed returns to be negatively serially correlated. Moreover, the return variances differ over trading versus non-trading periods, and there are pronounced intraday volatility patterns in financial markets. Finally, several studies argue for the simultaneous incorporation of jumps and time-varying volatility in continuous time models. While it is of interest to pursue the specification of richer continuous-

B. Temporal Aggregation and Continuous Time GARCH(1,1) Volatility Measurement

We have established that the relevant gauge for the performance of daily volatility forecasts is given by $\int_0^1 \sigma_{t+\tau}^2 d\tau$. Although the corresponding daily squared returns, $r_{(1),t+1}^2$, constitute an unbiased estimator of this quantity, it is also an extremely noisy estimator. Specifically, for the diffusions in table II, the population values of $E[(\int_0^1 \sigma_{t+\tau}^2 d\tau - r_{(1),t+1}^2)^2]$ equal 1.138 and 0.842, respectively, while the variances for the one-day-ahead latent volatility factor, $Var(\int_0^1 \sigma_{t+\tau}^2 d\tau)$, equal 0.166 and 0.191. The latter are thus orders of magnitude less than the corresponding MSE's for the daily squared returns.¹⁶

To further illustrate the pitfalls in using the squared daily returns for ex-post volatility forecast evaluation, consider the following decomposition of the ideal one-day-ahead latent volatility forecast error for the GARCH(1,1) model,

$$\begin{aligned} E[(\sigma_{(1),t}^2 - \int_0^1 \sigma_{t-1+\tau}^2 d\tau)^2] &= E[(\sigma_{(1),t}^2 - r_{(1),t}^2)^2] + E[(r_{(1),t}^2 - \int_0^1 \sigma_{t-1+\tau}^2 d\tau)^2] \\ &+ 2 \cdot E[(\sigma_{(1),t}^2 - r_{(1),t}^2) \cdot (r_{(1),t}^2 - \int_0^1 \sigma_{t-1+\tau}^2 d\tau)]. \end{aligned} \quad (12)$$

The prediction error calculated in practice using squared daily returns is given by the first term on the right-hand-side of equation (12). For the diffusions in table II, this term equals 1.221 and 0.944, respectively. In contrast, the ideal MSE for each of the daily weak-form GARCH(1,1) models, given by the left-hand-side of equation (12), equal 0.084 and 0.097. This glaring discrepancy reflects the impact of the measurement error, comprised of the second and third term on the right-hand-side of equation (12). Thus, whereas the population $R_{(1)}^2$'s from the daily squared return-volatility regressions in equation (4) suggest that the true GARCH(1,1) model only explains between five and ten percent of the daily variability, when measured by the more appropriate statistic

time stochastic volatility models that accommodate all of the above features, this is well beyond the scope of the present analysis. See Goodhart and O'Hara (1997) for a recent survey of the relevant empirical literature.

¹⁶ The numbers reported here, and throughout the remainder of the paper, are based on numerical simulations of the continuous-time model in equations (6) and (7) using a standard Euler discretization scheme; i.e., $p_{t+\Delta} = p_t + \sigma_t \Delta^{1/2} \cdot w_{p,t}$ and $\sigma_{t+\Delta}^2 = \theta \cdot \omega \cdot \Delta + \sigma_t^2 \cdot (1 - \theta \cdot \Delta + [2 \cdot \lambda \cdot \theta \cdot \Delta]^2 \cdot w_{\sigma,t})$, where $w_{p,t}$ and $w_{\sigma,t}$ denote independent standard normal variables. In the actual implementation we took $\Delta = 1/2,880$, corresponding to 10 observations per five-minute interval, while the $N(0,1)$ random variables were generated by the RNDNS routine in the GAUSS computer language. The sample size was fixed at 1,000,000 "daily" observations, which along with the use of antithetic variates based on $-w_{p,t}$ and $-w_{\sigma,t}$, was deemed sufficient to reduce the sampling variation beyond the reported decimal points for all relevant summary statistics; see Geweke (1995) for a recent discussion of simulation-based methods in econometrics.

$$R_{(1)\infty}^2 \equiv 1 - E[(\sigma_{(1),t}^2 - \int_0^1 \sigma_{t+\tau}^2 d\tau)^2] \cdot \text{Var}(\int_0^1 \sigma_{t+\tau}^2 d\tau)^{-1}, \quad (13)$$

both of the weak GARCH(1,1) models account for close to fifty percent ($1 - 0.084/0.166 \approx 1/2$ and $1 - 0.097/0.191 \approx 1/2$) of the variance in the one-day-ahead volatility factors. These findings underscore the importance of proper ex-post evaluation criteria when judging the quality of volatility forecasts.

Of course, in practice the sample path realization for the volatility process is inherently unobservable, rendering the ex-post sample equivalent of the $R_{(1),\infty}^2$ statistic in equation (13) infeasible. However, if the discretely sampled returns are serially uncorrelated, and the sample path for σ_t is continuous, it follows by Ito's Lemma that,

$$\lim_{m \rightarrow \infty} E([\int_0^1 \sigma_{t+\tau}^2 d\tau - \sum_{j=1, \dots, m} r_{(m), t+j/m}^2]^2) = 0. \quad (14)$$

This observation suggests that the use of high-frequency intraday returns may greatly improve the ex-post volatility measurement, in turn resulting in more meaningful volatility forecast evaluations.

To illustrate, consider again the measurement errors for the two continuous-time diffusions, or $E([\int_0^1 \sigma_{t+\tau}^2 d\tau - \sum_{j=1, \dots, m} r_{(m), t+j/m}^2]^2)$, reported in table III. As previously noted, with daily returns, or $m = 1$, the measurement errors equal 1.138 and 0.842. However, increasing the sampling frequency to eight hours, or $m = 3$, lowers the measurement errors to 0.381 and 0.289. Further reducing the length of the return interval to one hour, or $m = 24$, yield 0.048 and 0.036, respectively. For the five-minute returns, or $m = 288$, the measurement errors of 0.004 and 0.003, are both less than 2.5 percent of the daily variability in the latent volatility factor.^{17,18}

¹⁷ Note that the measurement errors are almost perfectly inversely related to m . Hence, the findings effectively extend the theoretical developments in Merton (1980), which show that the variance of the sample variance of a homoskedastic diffusion is inversely related to the sampling frequency, whereas the accuracy of the estimate for the drift in the logarithmic price process only depends on the span of the data. A similar idea for more efficiently estimating the daily volatility of a homoskedastic diffusion allowing for measurement noise in the observed high-frequency price process has been explored by Zhou (1994). The results in table III may also be seen as a practical guide to the applicability of the continuous-record asymptotics for rolling regressions formally developed by Foster and Nelson (1996).

¹⁸ While high-frequency intraday data have only recently become readily available, intraday high-low prices - *the intraday range* - have long been recorded daily for some equity markets. Given the availability of these statistics Garman and Klass (1980), Parkinson (1980), Ball and Torous (1984), and Kunitomo (1992), among others, argued for the use of the intraday range in order to develop more accurate daily volatility estimates for homoskedastic diffusions, while the properties of extreme value estimators in continuous-time models allowing for jumps are analyzed by Rogers and Satchell (1991) and

Motivated by these findings, consider the one-day-ahead squared return - volatility regression obtained by replacing the squared daily returns on the left-hand-side of equation (4) with the sum of the corresponding squared intraday returns,

$$\sum_{j=1, \dots, m} r_{(m), t+j/m}^2 = a_{(1)m} + b_{(1)m} \cdot \sigma_{(1), t+1}^2 + u_{(1)m, t+1}, \quad (15)$$

where $t = 0, 1, 2, \dots$, and by definition $a_{(1)l} \equiv a_{(l)}$, $b_{(1)l} \equiv b_{(l)}$, and $u_{(1)l, t+1} \equiv u_{(l), t+1}$. Irrespective of the sampling frequency m , if the model for the conditional variance is correctly specified, the population values of $a_{(1)m}$ and $b_{(1)m}$ equal zero and unity, respectively.¹⁹ However, the more precise ex-post volatility measurements afforded by the high-frequency intraday returns allow for more meaningful qualitative assessments of the daily volatility forecasts, $\sigma_{(1)l, t+1}^2$, when judged by the resulting coefficient of multiple determination, say $R_{(1)m}^2$.

The numerical results for the two continuous-time diffusions are reported in table IV. The $R_{(1)l}^2$'s again indicate that the weak GARCH(1,1) forecasts explain little of the ex-post variability.²⁰ Meanwhile, the population $R_{(1)m}^2$'s increase monotonically with the sampling frequency to the much larger $R_{(1)\infty}^2$'s defined in equation (13). For instance, using the sum of the hourly squared returns on the left-hand-side of equation (15), the $R_{(1)24}^2$'s equal 0.383 and 0.419. Reducing the return interval to five-minutes result in $R_{(1)288}^2$'s of 0.483 and 0.488, both of which are extremely close to the ideal $R_{(1)\infty}^2$'s of 0.495 for each of the rates. These findings highlight the advantage of using high-frequency intraday returns in the construction of interdaily volatility forecast evaluation criteria.

Maheswaran (1996). Meanwhile, the time-series models estimated in Hsieh (1993) show that the empirical distribution of the intraday range is strongly time-varying. Although the high-low range is not an unbiased estimator for the latent volatility over the day, it follows by numerical simulation, that the MSE for the correspondingly scaled unbiased estimator, $E[(\gamma [\max_{0 \leq \tau \leq 1} p_{t+\tau} - \min_{0 \leq \tau \leq 1} p_{t+\tau}] - \int_0^1 \sigma_{\tau, t}^2 d\tau)^2]$, equal 0.103 and 0.114 for the two stochastic volatility models in table II. Thus, compared to the measurement errors reported in table III, this puts the accuracy of the high-low estimator around that afforded by the intraday sample variance based on two- or three-hour returns.

¹⁹ The same errors-in-variables problem that plagues the estimation of $b_{(m)}$ in equation (4) will result in a downward bias in the estimate for $b_{(1)m}$ formally given by $-\text{Var}_{(m)}(v_{(1), t}) \cdot [\text{Var}_{(m)}(v_{(1), t}) + \text{Var}_{(m)}(\sigma_{(1), t}^2)]^{-1}$.

²⁰ Note that the numerical values for the $R_{(1)l}^2$'s from the daily weak GARCH(1,1) approximations in table IV are slightly lower than the implied $R_{(1)l}^2$'s from the daily strong GARCH(1,1) model with the same conditional variance parameters which, by equation (5), equal 0.064 and 0.096, respectively.

IV. Intraday Returns and Interdaily Volatility Forecast Evaluation

The computation of daily return variances from high-frequency intraday returns parallels the use of daily returns in calculating monthly ex-post volatility, as exemplified by Schwert (1989, 1990a) and Schwert and Seguin (1990). The idea has also been applied by, among others, Hsieh (1991) and Schwert (1990b) for measurement of daily equity market volatility from sample standard deviations of intraday returns.²¹ However, the connection between ARCH volatility modeling and forecasting on the one hand and the ex-post volatility measurements on the other has hitherto not been explored. We now turn to the practical application of this relation in the context of evaluating the out-of-sample volatility forecasts from the daily GARCH(1,1) models discussed in section II.

A. Improved Daily GARCH(1,1) Volatility Forecast Evaluation

Direct interpretation of the low $R_{(1)}^2$'s for the one-day-ahead GARCH(1,1) DM-\$ and ¥-\$ volatility forecasts suggests that the models perform very poorly, explaining less than five percent of the ex-post variability in either rate. However, increasing the sampling frequency of the ex-post squared returns on the left-hand-side of equation (15), the fallacy of this conclusion becomes evident. For instance, with an hourly sampling frequency the two $R_{(1)24}^2$'s reported in table V equal 0.331 and 0.237, respectively. The reduction in the measurement error is also apparent in figure 3, which graphs the one-step-ahead volatility forecasts, $\sigma_{(1)1,t+1}^2$, along with the ex-post volatility based on the hourly returns, $\sum_{j=1,\dots,24} r_{(24),t+j/24}^2$. Clearly, the sum of the hourly squared returns correlates much more closely with the daily GARCH(1,1) predictions than do the squared daily returns in figure 2. Further increasing the sampling frequency results in still higher correlations, with $R_{(1)288}^2$'s at the five-minute level of 0.479 and 0.392, respectively. The latter statistics signify more than a ten-fold increase in the explanatory power of the GARCH(1,1) models relative to the inference based on the conventional $R_{(1)1}^2$'s reported in the literature. Figure 4 underscores the point.²² Except for a few isolated episodes, the one-day-ahead GARCH(1,1) predictions do a remarkable job of tracking the ex-post volatility in both markets.

²¹ The estimation of standard time series models for these ex-post volatility measures generally confirm the very high degree of intertemporal volatility dependencies documented in the ARCH literature.

²² A similar figure for the scaled absolute DM-\$ returns and standard deviations is contained in Andersen and Bollerslev (1997b), where the volatility measurement technique formally justified here has been used in the analysis of news and calendar effects in the foreign exchange market.

The close correspondence between the implied continuous-time GARCH(1,1) predictive $R^2_{(1)m}$'s in table IV and the actual empirical results for the DM-\$ rate in table V is particularly noteworthy. It suggests that the market microstructure rigidities and pronounced intraday volatility patterns not accommodated by the continuous-time process in equations (6) and (7) effectively are annihilated at the daily level. Moreover, it indicates that the simple GARCH(1,1) model does a good job of characterizing the volatility clustering for the DM-\$ rate over the ex-post sample period. Meanwhile, the out-of-sample $R^2_{(1)m}$'s for the ¥-\$ rate in table V are all slightly below the corresponding theoretical values for the ¥-\$ diffusion in table IV. However, the discrepancy between the empirical and theoretical ¥-\$ results is in part attributable to a few pronounced appreciations that occurred during the out-of-sample period.²³ Eliminating the two largest ex-post volatility measures, that are readily recognized in figure 4.B, the value of $R^2_{(1)288}$ for the ¥-\$ rate increases from 0.392 to 0.456.

V. Concluding Remarks

Numerous studies in the empirical finance literature have suggested that ARCH models provide poor volatility forecasts. Contrary to this widespread perception, both the theoretical and empirical analysis in this paper demonstrate that for empirically relevant ARCH specifications the forecasts correlate closely with the future latent volatility factor that is relevant for most applications, typically accounting for close to fifty percent of the variability in ex-post volatility. Yes, ARCH models do provide good volatility forecasts!

Several important questions remain. First, it is of interest to further explore the role of model misspecification. The formal conditions developed by Nelson (1992) and Nelson and Foster (1997) pertaining to the use of misspecified ARCH models in forecasting, along with the robustness results in Nelson and Foster (1994), should provide a useful guide for future work along these lines. Furthermore, when extending the forecast horizon beyond one day, issues related to the proper modeling of the long-term volatility dependencies become especially important; see, e.g., Baillie, Bollerslev and Mikkelsen (1996).

Our main results hinge on the effective use of frequently sampled data in constructing more accurate ex-post volatility measurements. A closely related question pertains to the precision of the volatility forecasts as a function of the sampling frequency. Do the additional costs and

²³ An analysis of the economic determinants behind these large rate changes is beyond the scope of the present paper.

complications in model construction and data gathering warrant the use of intraday data for volatility forecasting as well? The decomposition of the volatility into short- and long-lived volatility components along with distinct calendar effects proposed by Andersen and Bollerslev (1997b) may be helpful in quantifying the relevant tradeoff inherent in this important practical problem.

The volatility forecasts analyzed above are based solely on ad-hoc time-series models. There is a voluminous literature on alternative ways in which to extract information about the latent volatility factor from sources other than, or in addition to, the corresponding squared or absolute returns. They include implied volatilities extracted from options prices, as in the recent work of Canina and Figlewski (1993), Jorion (1995) and Lamoureux and Lastrapes (1993), along with information provided by the joint distribution of return and trading volume, as in the work by Andersen (1996) and Gallant, Rossi and Tauchen (1992). The evaluation criteria proposed here should allow for more meaningful comparisons of these structural methods for estimating volatility.

Most of the volatility forecast comparisons in the literature rely on some variant of the squared return - volatility regression utilized here. While such evaluation criteria may be natural when evaluating forecasts for the conditional mean, it is less obvious when evaluating volatility forecasts; see, e.g., the discussion in Engle et al. (1993), West, Edison and Cho (1993) and Lopez (1995). Our results suggest that further analysis along these lines may benefit from the use of high-frequency data. We leave all of these issues for future research.

REFERENCES

- Akgriray, V. (1989), "Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts," **Journal of Business**, 62, 55-80.
- Andersen, T.G. (1994), "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling," **Mathematical Finance**, 4, 75-102.
- Andersen, T.G. (1996), "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility," **Journal of Finance**, 51, 169-204.
- Andersen, T.G. and T. Bollerslev (1996), "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," NBER Working Paper No.5752.
- Andersen, T.G. and T. Bollerslev (1997a), "Intraday Periodicity and Volatility Persistence in Financial Markets," **Journal of Empirical Finance**, forthcoming.
- Andersen, T.G. and T. Bollerslev (1997b), "DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies" **Journal of Finance**, forthcoming.
- Baillie, R.T. and T. Bollerslev (1989), "The Message in Daily Exchange Rates: A Conditional Variance Tale," **Journal of Business and Economic Statistics**, 7, 297-305.
- Baillie, R.T., T. Bollerslev and H.O. Mikkelsen (1996) "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," **Journal of Econometrics**, 74, 3-30.
- Ball, C.A. and W.N. Torous (1984), "The Maximum Likelihood Estimation of Security Price Volatility: Theory, Evidence and Applications to Option Pricing," **Journal of Business**, 57, 97-112.
- Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," **Journal of Econometrics**, 31, 307-327.
- Bollerslev, T. (1987), "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," **Review of Economics and Statistics**, 69, 542-547.
- Bollerslev, T., R.Y. Chou and K.F. Kroner (1992) "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," **Journal of Econometrics**, 52, 5-59.
- Bollerslev, T., R.F. Engle and D.B. Nelson (1994), "ARCH Models," in **Handbook of Econometrics Vol.IV** (R.F. Engle and D. McFadden, eds.). Amsterdam: North Holland Press.
- Bollerslev, T. and H.O. Mikkelsen (1996), "Modeling and Pricing Long-Memory in Stock Market Volatility," **Journal of Econometrics**, 73, 151-184.
- Bollerslev, T. and J.M. Wooldridge (1992), "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances," **Econometric Reviews**, 11, 143-172.
- Brailsford, T.J. and R.W. Faff (1996), "An Evaluation of Volatility Forecasting Techniques," **Journal of Banking and Finance**, 20, 419-438.

- Canina, L. and S. Figlewski (1993), "The Informational Content of Implied Volatility," **Review of Financial Studies**, 6, 659-681.
- Chow, G.C. (1983). **Econometrics**. New York: McGraw-Hill, Inc.
- Christensen, B.J. and N.R. Prabhala (1997), "The Dynamic Relationship Between Implied and Realized Volatility," **Journal of Financial Economics**, forthcoming.
- Cumby, R., S. Figlewski and J. Hasbrouck (1993), "Forecasting Volatility and Correlations with EGARCH Models," **Journal of Derivatives**, Winter, 51-63.
- Day, T.E. and C.M. Lewis (1992), "Stock Market Volatility and the Information Content of Stock Index Options," **Journal of Econometrics**, 52, 267-287.
- Diebold, F.X. and R.S. Mariano (1995), "Comparing Predictive Accuracy," **Journal of Business and Economic Statistics**, 13, 253-263.
- Dimson, E. and P. Marsh (1990), "Volatility Forecast without Data-Snooping," **Journal of Banking and Finance**, 14, 399-421.
- Drost, F.C. and T.E. Nijman (1993), "Temporal Aggregation of GARCH Processes," **Econometrica**, 61, 909-927.
- Drost, F.C. and B.J.M. Werker (1996), "Closing the GARCH Gap: Continuous Time GARCH Modeling," **Journal of Econometrics**, 74, 31-57.
- Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," **Econometrica**, 50, 987-1008.
- Engle, R.F. and T. Bollerslev (1986), "Modelling the Persistence of Conditional Variances," **Econometric Reviews**, 5, 1-50.
- Engle, R.F., C-H. Hong, A. Kane, and J. Noh (1993), "Arbitrage Valuation of Variance Forecasts with Simulated Options," in **Advances in Futures and Options Research** (D.M. Chance and R.R. Trippi, eds.). Greenwich Connecticut: JAI Press.
- Fama, E. (1965), "The Behavior of Stock Market Prices," **Journal of Business**, 38, 34-105.
- Figlewski, S. (1994), "Forecasting Volatility Using Historical Data," manuscript, Department of Finance, New York University.
- Foster, D.P. and D.B. Nelson (1996), "Continuous Record Asymptotics for Rolling Sample Variance Estimators," **Econometrica**, 64, 139-174.
- Frennberg, P. and B. Hansson (1996), "An Evaluation of Alternative Models for Predicting Stock Volatility: Evidence from a Small Stock Market," **Journal of International Financial Markets, Institutions and Money**, 5, 117-134.
- Gallant, A.R., P.E. Rossi and G.E. Tauchen (1992), "Stock Prices and Volume," **Review of Financial Studies**, 5, 199-242.

- Garman, M.B. and M.J. Klass (1980), "On the Estimation of Security Price Volatilities from Historical Data," **Journal of Business**, 53, 67-78.
- Geweke, J. (1995), "Monte Carlo Simulation and Numerical Integration," Staff Report No.192, Federal Reserve Bank of Minneapolis.
- Goodhart, Charles A.E. and Maureen O'Hara (1997), "High frequency data in financial markets: Issues and applications," **Journal of Empirical Finance**, forthcoming.
- Heynen, R.C. and H.M. Kat (1994), "Volatility Prediction: A Comparison of the Stochastic Volatility, GARCH(1,1), and EGARCH(1,1) Models," **Journal of Derivatives**, Winter, 50-65.
- Hsieh, D.A. (1989), "Modeling Heteroskedasticity in Daily Foreign Exchange Rates," **Journal of Business and Economic Statistics**, 7, 307-317.
- Hsieh, D.A. (1991), "Chaos and Nonlinear Dynamics: Application to Financial Markets," **Journal of Finance**, 46, 1839-1877.
- Hsieh, D.A. (1993), "Implications of Nonlinear Dynamics for Financial Risk Management," **Journal of Financial and Quantitative Analysis**, 28, 41-64.
- Hull, J. and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatilities," **Journal of Finance**, 42, 381-400.
- Jacquier, E., N.G. Polson and P.E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models," **Journal of Business and Economic Statistics**, 12, 371-417.
- Jorion, P. (1995), "Predicting Volatility in the Foreign Exchange Market," **Journal of Finance**, 50, 507-528.
- Jorion, P. (1996), "Risk and Turnover in the Foreign Exchange Market," in **The Microstructure of Foreign Exchange Markets** (J.A. Frankel, G. Galli and A. Giovannini, eds.). Chicago: The University of Chicago Press.
- Kunitomo, N. (1992), "Improving the Parkinson Method of Estimating Security Price Volatilities," **Journal of Business**, 65, 295-302.
- Lamoureux, C.G. and W.D. Lastrapes (1993), "Forecasting Stock Returns Variances: Towards Understanding Stochastic Implied Volatility," **Review of Financial Studies**, 6, 293-326.
- Linton, O. (1997), "The Asymptotic Bias of the QMLE in the GARCH(1,1) Model," **Econometric Theory**, forthcoming.
- Ljung, G.M. and G.E.P. Box (1978), "On a Measure of Lack of Fit in Time Series Models," **Biometrika**, 65, 297-303.
- Lopez, J.A. (1995), "Evaluating the Predictive Accuracy of Volatility Models," Manuscript, Research and Market Analysis Group, Federal Reserve Bank of New York.
- Lumsdaine, R.L. (1995), "Finite Sample Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models: A Monte Carlo Investigation," **Journal of Business and Economic Statistics**, 13, 1-10.

- Maheeswaran, S. (1996), "Volatility Estimation Using Extreme Values of Asset Prices," manuscript, Olin School of Business, Washington University.
- Mandelbrot, B.B. (1963), "The Variation of Certain Speculative Prices," **Journal of Business**, 36, 394-419.
- McCurdy, T. and I. Morgan (1988), "Testing the Martingale Hypothesis in Deutschemark Futures with Models Specifying the Form of Heteroskedasticity," **Journal of Applied Econometrics**, 3, 187-202.
- Melino, A. (1994), "The Estimation of Continuous-Time Models in Finance," in **Advances in Econometrics, Sixth World Congress Vol.II** (ed. C. A. Sims), Cambridge University Press: Cambridge, Great Britain.
- Merton, R.C. (1980), "On Estimating the Expected Return on the Market," **Journal of Financial Economics**, 8, 323-361.
- Mincer, J. and V. Zarnowitz (1969), "The Evaluation of Economic Forecasts," in **Economic Forecasts and Expectations** (J. Mincer, ed.). New York: National Bureau of Economic Research.
- Nelson, D.B. (1990), "ARCH Models as Diffusion Approximations," **Journal of Econometrics**, 45, 7-38.
- Nelson, D.B. (1992), "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model," **Journal of Econometrics**, 52, 61-90.
- Nelson, D.B. (1996), "Asymptotic Filtering Theory for Multivariate ARCH Models," **Journal of Econometrics**, 71, 1-47.
- Nelson, D.B. and D.P. Foster (1994), "Asymptotic Filtering Theory for Univariate ARCH Models," **Econometrica**, 62, 1-41.
- Nelson, D.B. and D.P. Foster (1997), "Filtering and Forecasting with Misspecified ARCH Models II: Making the Right Forecast with the Wrong Model," **Journal of Econometrics**, forthcoming.
- Pagan, A.R. and H. Sabau (1992), "Consistency Tests for Heteroskedasticity and Risk Models," **Estudios Económicos**, 7, 3-30.
- Pagan, A.R. and G.W. Schwert (1990), "Alternative Models for Conditional Stock Volatility," **Journal of Econometrics**, 45, 267-290.
- Pagan, A.R. and A. Ullah (1988), "The Econometric Analysis of Models with Risk Terms," **Journal of Applied Econometrics**, 3, 87-105.
- Parkinson, M. (1980) "Extreme Value Method for Estimating the Variance of the Rate of Return," **Journal of Business**, 53, 61-65.
- Rogers, L.C.G. and S.E. Satchell (1991), "Estimating Variance from High, Low and Closing Prices," **Annals of Applied Probability**, 1, 504-512.
- Roll, R. (1988), "R²," **Journal of Finance**, 43, 541-566.
- Rossi, P. (1996). **Modelling Stock Market Volatility: Bridging the Gap to Continuous Time**. San Diego: Academic Press.

- Schwert, G.W. (1989), "Why Does Stock Market Volatility Change Over Time?" **Journal of Finance**, 44, 1115-1153.
- Schwert, G.W. (1990a), "Stock Volatility and the Crash of '87," **Review of Financial Studies**, 3, 77-102.
- Schwert, G.W. (1990b), "Stock Market Volatility," **Financial Analysts Journal**, May-June, 23-34.
- Schwert, G.W. and P.J. Seguin (1990), "Heteroskedasticity in Stock Returns," **Journal of Finance**, 45, 1129-1155.
- Scott, L. (1987), "Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application," **Journal of Financial and Quantitative Analysis**, 22, 419-436.
- Shephard, N. (1996), "Statistical Aspects of ARCH and Stochastic Volatility," in **Likelihood, Time Series with Econometric and Other Applications** (D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen, eds.). London: Chapman and Hall.
- West, K.D. and D. Cho (1995), "The Predictive Ability of Several Models of Exchange Rate Volatility," **Journal of Econometrics**, 69, 367-391.
- West, K.D., H.J. Edison, and D. Cho (1993), "A Utility Based Comparison of Some Models of Exchange Rate Volatility," **Journal of International Economics**, 35, 23-45.
- Wiggins, M. (1987), "Option Values under Stochastic Volatility: Theory and Empirical Estimates," **Journal of Financial Economics**, 19, 351-372.
- Zhou, B. (1994), "Estimating the Variance Parameter from Noisy High Frequency Financial Data," Manuscript, Sloan School of Management, MIT.

Table I
Daily GARCH(1,1) Model Estimates

	DM-\$	¥-\$
$\psi_{(1)}$	0.022 (0.009)	0.026 (0.008)
$\alpha_{(1)}$	0.068 (0.016)	0.104 (0.026)
$\beta_{(1)}$	0.898 (0.023)	0.844 (0.032)

The table reports Quasi Maximum Likelihood estimates (QMLE) for the GARCH(1,1) model defined in equation (3) based on the daily percentage returns for the Deutsche mark - dollar and yen - dollar spot exchange rates from October 1, 1987, through September 30, 1992. The QMLE are obtained under the assumption of conditional normality. Robust standard errors, as described in Bollerslev and Wooldridge (1992), are reported in parentheses.

Table II
Implied Continuous Time GARCH(1,1):
Model Estimates

	DM-\$	¥-\$
θ	0.035	0.054
ω	0.636	0.476
λ	0.296	0.480

The table gives the parameters for the continuous-time stochastic volatility model defined by equations (6) and (7) implied by the discrete-time daily GARCH(1,1) model estimates reported in table I. The formal relationship between the continuous- and discrete-time parameters is detailed in equations (9), (10) and (11).

Table III
Implied Continuous Time GARCH(1,1):
Volatility Measurement Errors

m	DM-\$	¥-\$
1	1.138	0.842
3	0.381	0.289
24	0.048	0.036
288	0.004	0.003

The table reports the measurement errors using the sum of squared intraday returns as a measure for the true daily latent volatility; i.e., $E([\int_0^1 \sigma_{t+\tau}^2 d\tau - \sum_{j=1, \dots, m} R_{(m), t+j/m}^2]^2)$. The returns are generated by the stochastic volatility model in equations (6) and (7) at the parameter values in table II. The aggregation frequencies for $m = 1, 3, 24, 288$, correspond to daily, 8-hours, hourly, and 5-minute returns, respectively. The numbers are computed by simulation using antithetic variates and 1,000,000 "daily" observations.

Table IV
 Implied Continuous Time GARCH(1,1):
 Predictive R²'s

m	DM-\$	Y-\$
1	0.063	0.089
3	0.151	0.198
24	0.383	0.419
288	0.483	0.488
∞	0.495	0.495

The table reports the population R² from the squared return - volatility regression in equation (15); i.e., $R_{(1)m}^2 \equiv 1 - \text{Var}_{(m)}(\sum_{j=1, \dots, m} r_{(m), t+j/m}^2 - \sigma_{(1), t+1}^2) \cdot \text{Var}_{(m)}(\sum_{j=1, \dots, m} r_{(m), t+j/m}^2)^{-1}$. The returns are generated by the stochastic volatility model in equations (6) and (7) at the parameter values in table II. The daily GARCH(1,1) forecasts, $\sigma_{(1), t+1}^2$, are based on equation (3) with $m \equiv 1$ and the parameter values in table I. The rows labelled $m = 1, 3, 24, 288$ correspond to daily, 8-hours, hourly, and five-minute returns, respectively. The last row gives the population $R_{(1)\infty}^2 \equiv 1 - \text{Var}(\int_0^1 \sigma_{t+\tau}^2 d\tau - \sigma_{(1), t+1}^2) \cdot \text{Var}(\int_0^1 \sigma_{t+\tau}^2 d\tau)^{-1}$. The figures are calculated by simulation using antithetic variates and 1,000,000 "daily" observations.

Table V
Daily GARCH(1,1) Predictive R^2 's

m	DM-\$	¥-\$
1	0.047	0.025
3	0.133	0.095
24	0.331	0.237
288	0.479	0.392

The table reports the R^2 , or $R^2_{(1)m}$, from the squared return-volatility regression in equation (15). The returns are calculated from continuously recorded quotations for the Deutsche mark - dollar and yen - dollar spot exchange rates from October 1, 1992, through September 29, 1993. Quotes from Friday 21:00 GMT through Sunday 21:00 GMT are excluded, resulting in a total of 260 weekdays, or 74,880 observations for the five-minute return interval. The daily GARCH(1,1) volatility forecasts, $\sigma^2_{(1)t+1}$, are based on equation (3) with $m = 1$ evaluated at the parameters in table I. The rows labelled $m = 1, 3, 24, 288$ give the results for daily, 8-hours, hourly, and five-minute sampling frequencies, respectively.

FIGURE LEGENDS

Figure I.

The figure plots the daily percentage returns on an open position in the two spot foreign exchange markets from October 1, 1987 through September 30, 1992. Panel A graphs the returns for the Deutsche mark - dollar rate, whereas panel B graphs the daily yen - dollar rate.

Figure II.

The solid line in the figure graphs the daily one-step-ahead GARCH(1,1) volatility forecasts, $\sigma_{(1),t+1}^2$, given by equation (3) with $m \equiv 1$ and the parameter values in table 1. The dotted line gives the corresponding realized daily squared returns, $r_{(1),t}^2$. The results for the Deutsche mark - dollar and yen - dollar rates are given in panels A and B, respectively.

Figure III.

The figure is identical to figure II, except that the dotted line graphs the daily sample variance based on hourly returns; i.e., $\sum_{j=1, \dots, 24} r_{(24),t+j/24}^2$.

Figure IV.

The figure is identical to figure II, except that the dotted line graphs the daily sample variance based on five-minute returns; i.e., $\sum_{j=1, \dots, 288} r_{(288),t+j/288}^2$.

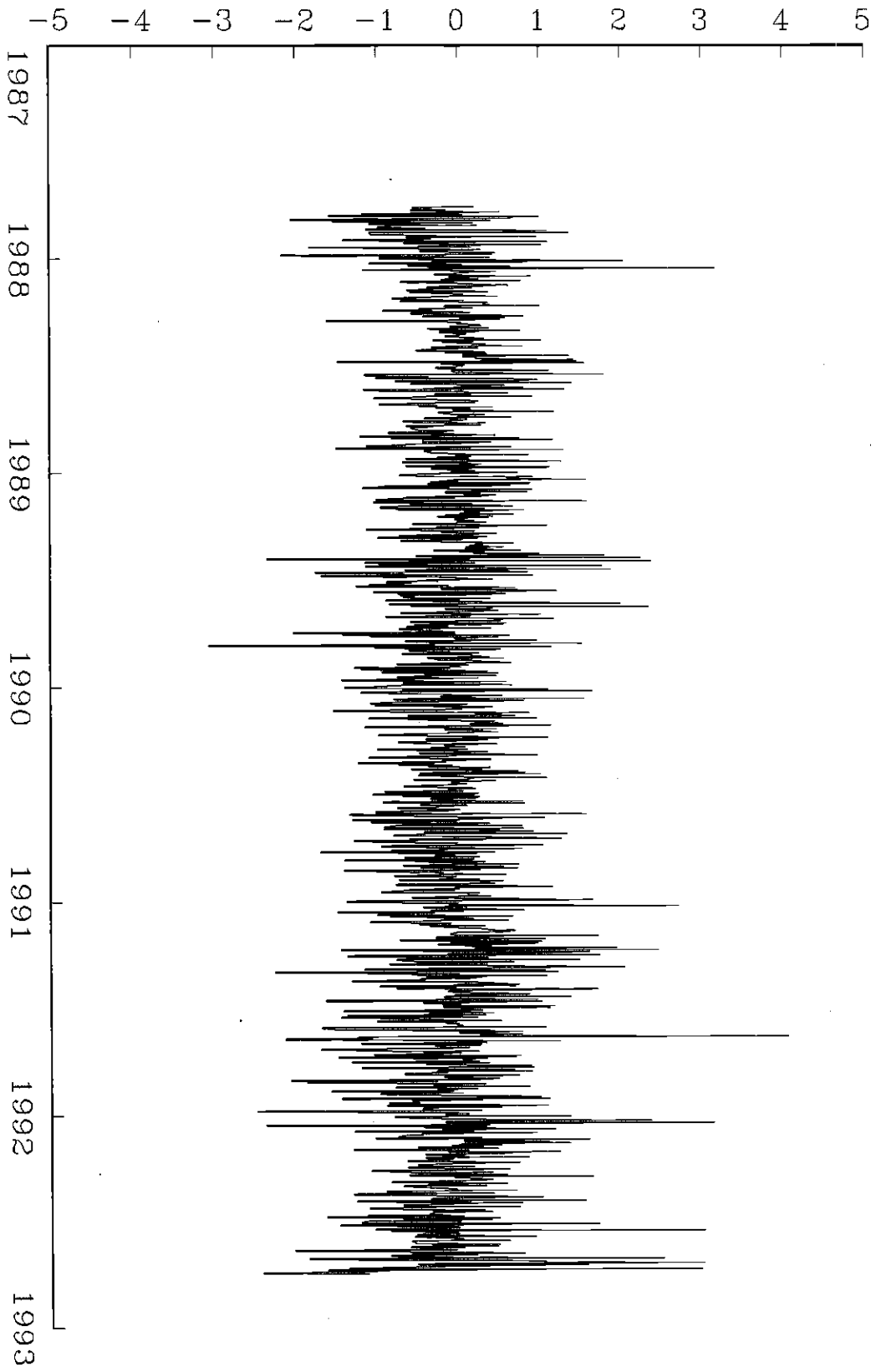


Figure 1.A
Daily DM-\$ Returns

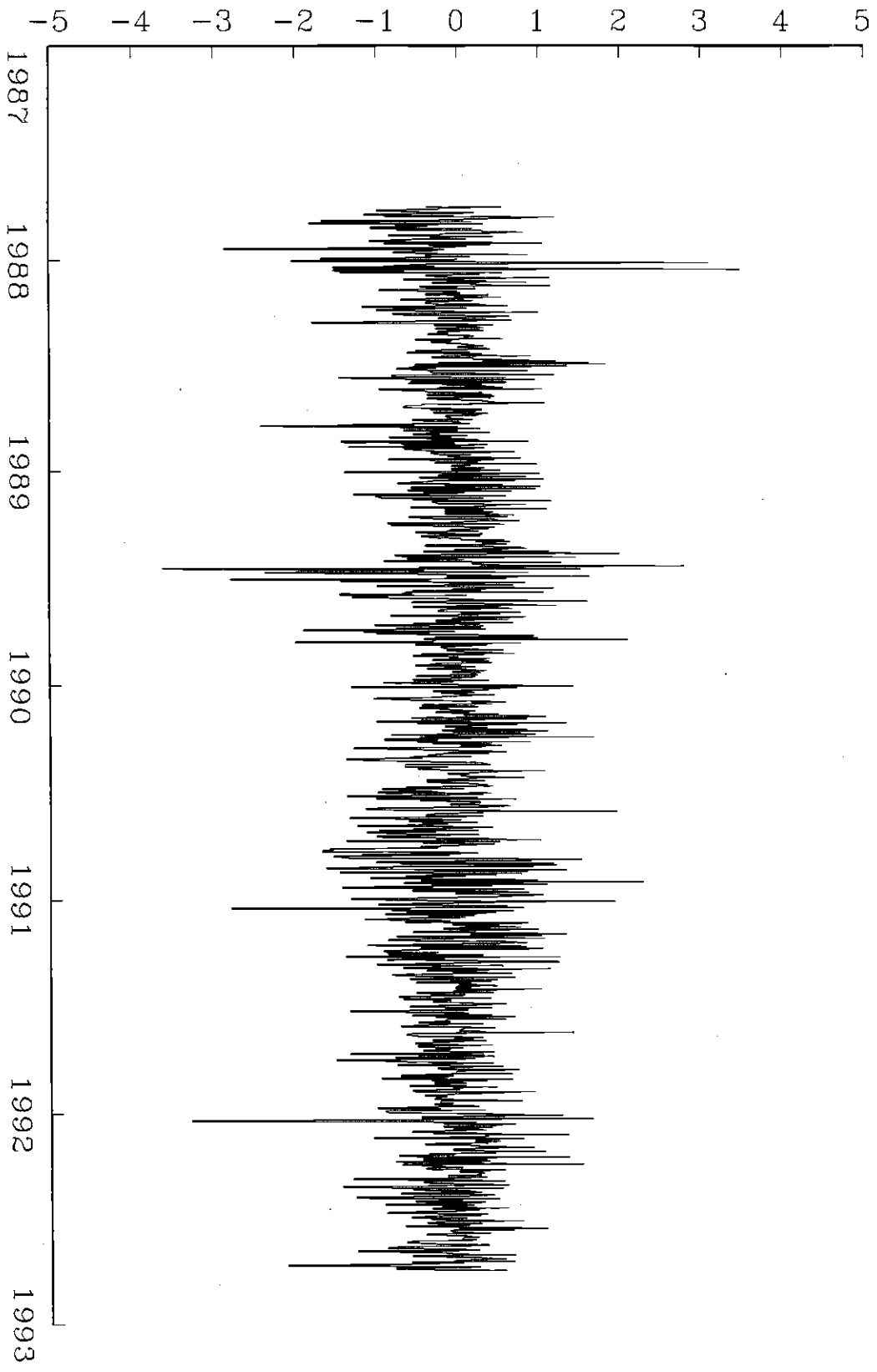


Figure 1.B
Daily Yen-\$ Returns

Figure 2.A
Daily DM-\$ Volatility

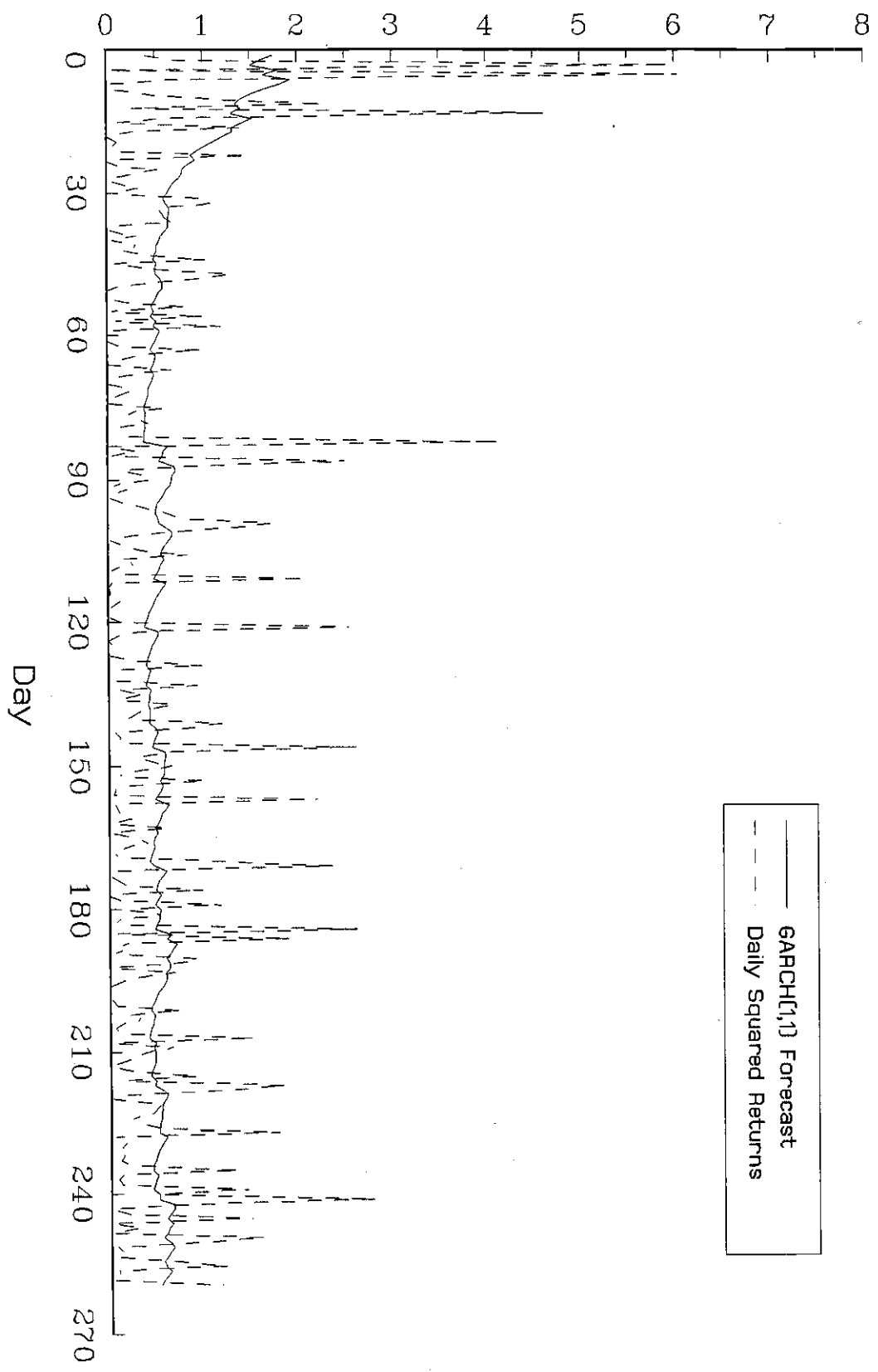


Figure 2.B
Daily Yen-\$ Volatility

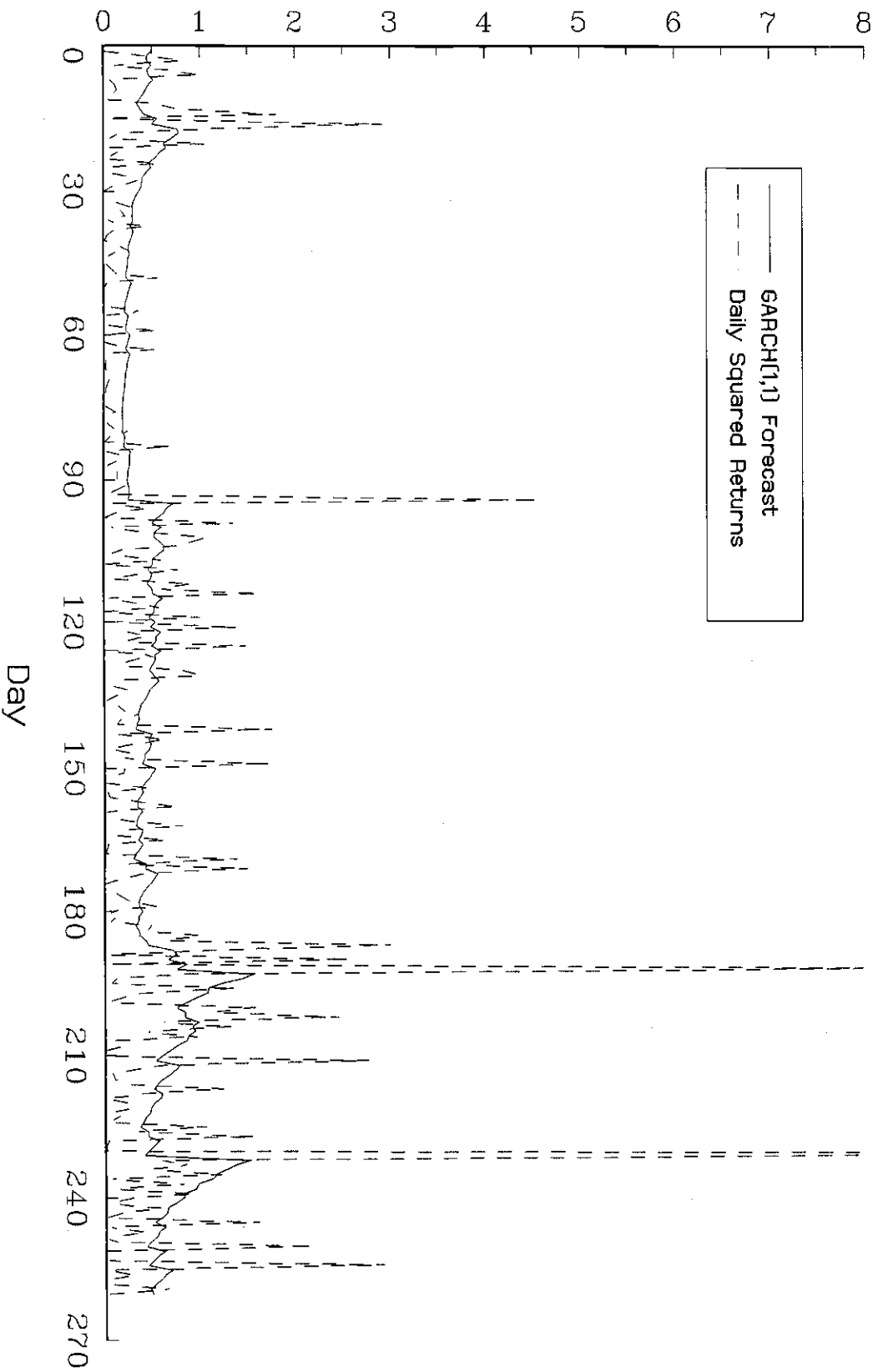


Figure 3.A
Daily DM-\$ Volatility

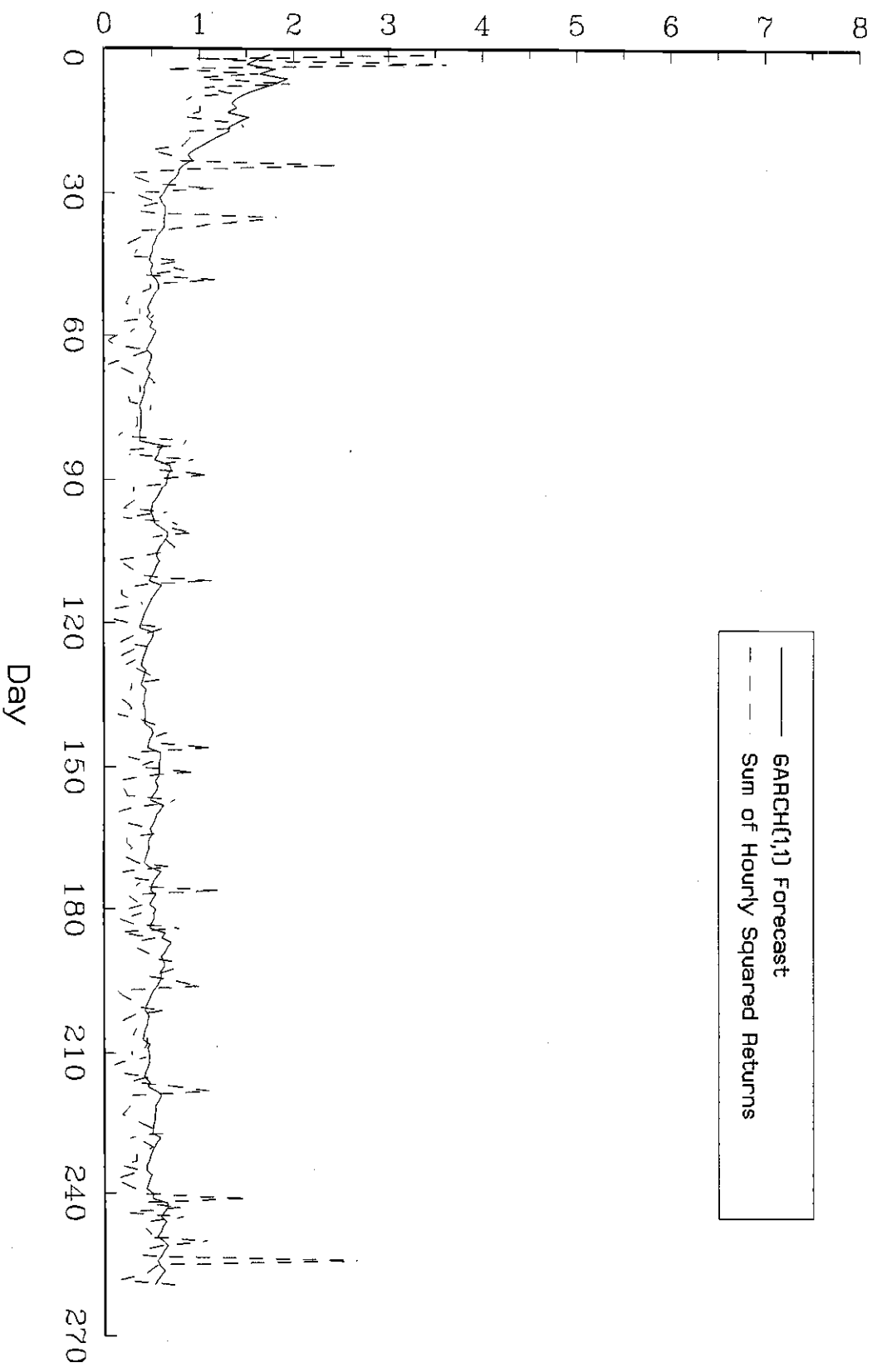


Figure 3.B
Daily Yen-\$ Volatility

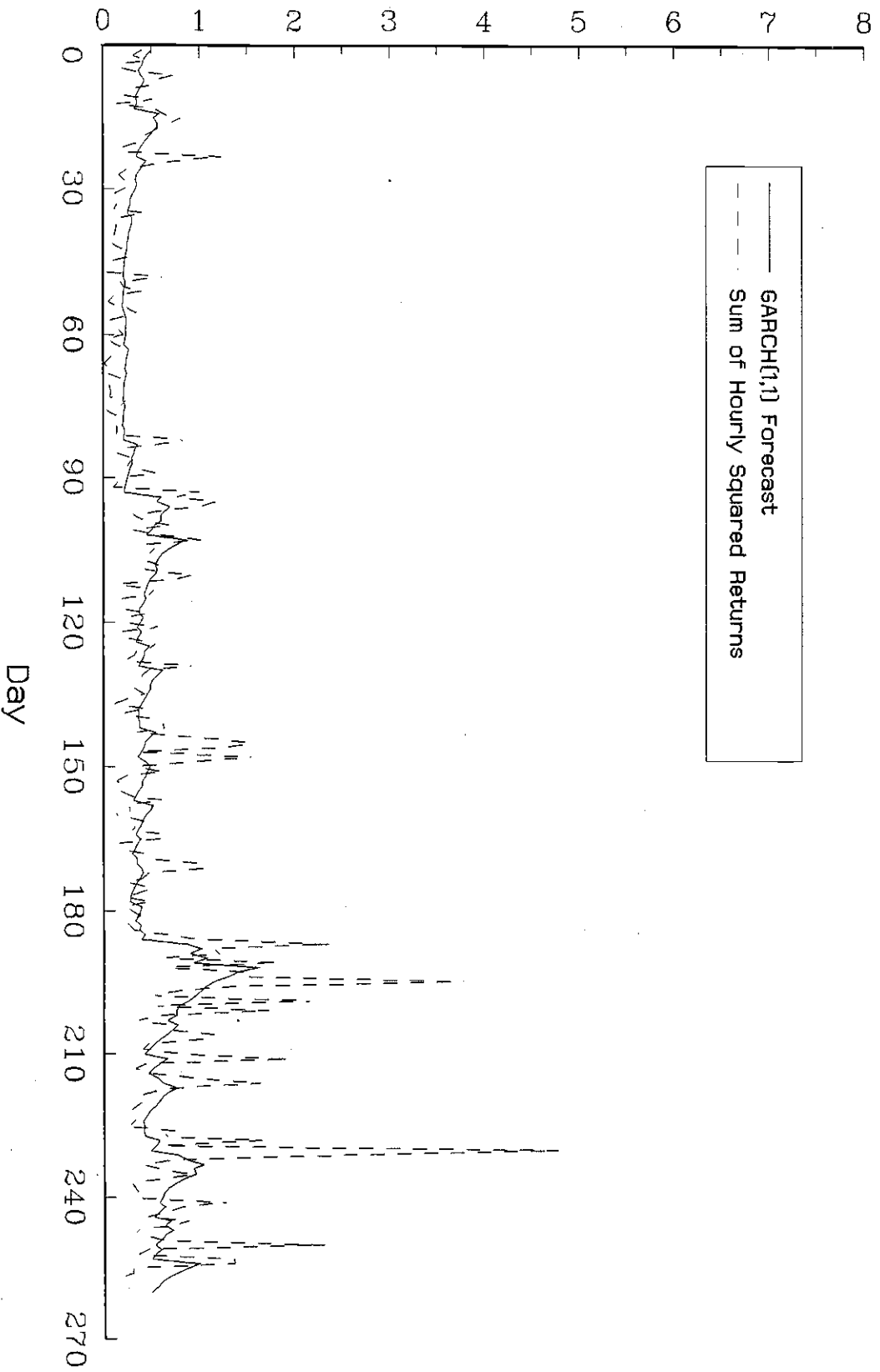


Figure 4.A
Daily DM-\$ Volatility

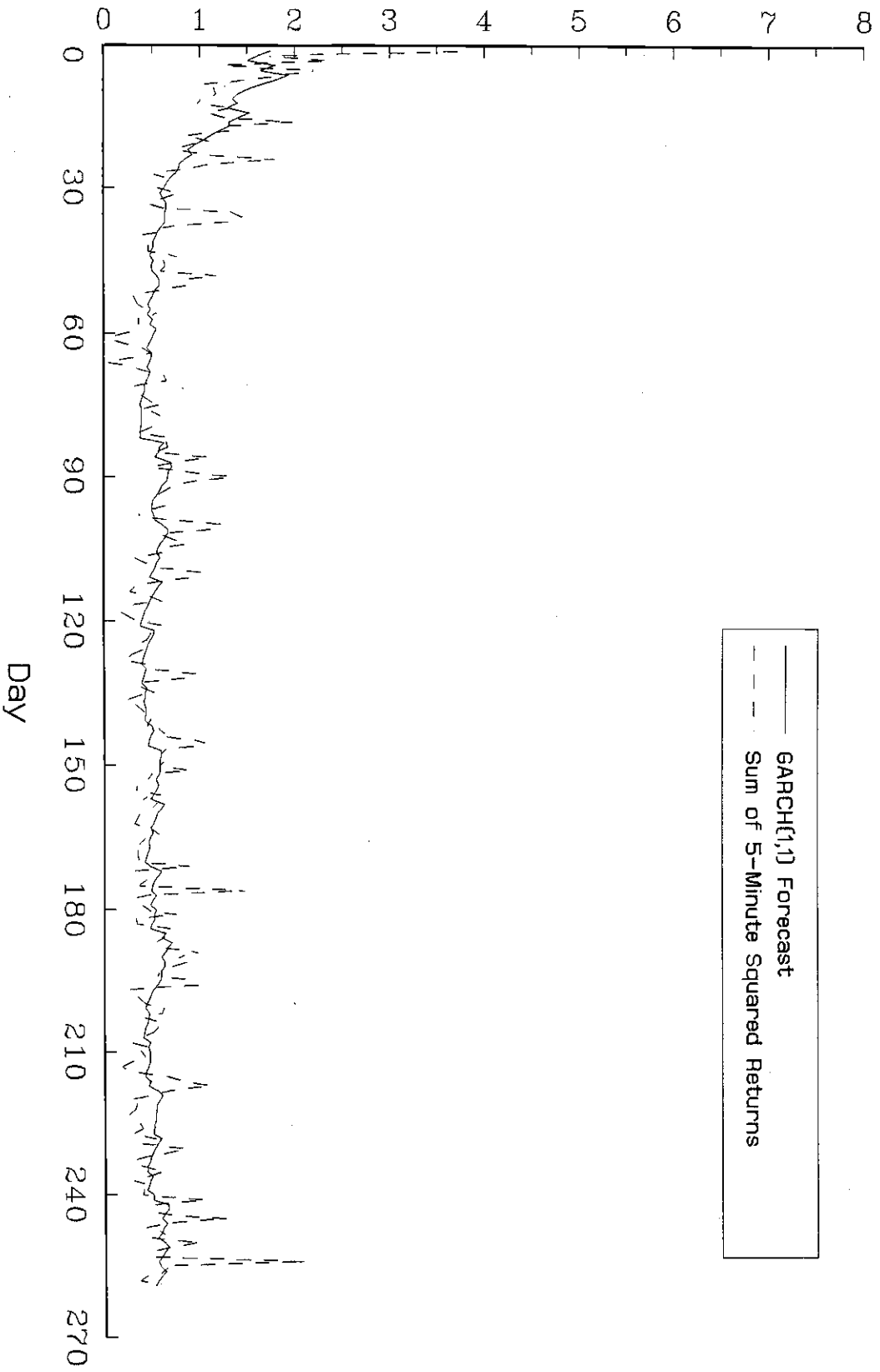


Figure 4.B
Daily Yen-\$ Volatility

