
MEASURING AND FORECASTING S&P 500 INDEX-FUTURES VOLATILITY USING HIGH-FREQUENCY DATA

MARTIN MARTENS

In the 24-hr foreign exchange market, Andersen and Bollerslev measure and forecast volatility using intraday returns rather than daily returns. Trading in equity markets only occurs during part of the day, and volatility during nontrading hours may differ from the volatility during trading hours. This paper compares various measures and forecasts of volatility in equity markets. In the absence of overnight trading it is shown that the daily volatility is best *measured* by the sum of intraday squared 5-min returns, excluding the overnight return. In the absence of overnight trading, the best daily *forecast* of volatility is produced by modeling overnight volatility differently from intraday volatility. © 2002 Wiley Periodicals, Inc. Jrl Fut Mark 22:497–518, 2002

The author thanks Thomas Henker, David Michayluk, Ian Sharpe, two anonymous referees, and seminar participants at the First World Congress of the Bachelier Finance Society in Paris (2000) for providing useful comments. The author is grateful to the Futures Industry for making the intraday futures data available. Remaining errors are, of course, the author's responsibility. For correspondence, Martin Martens, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000DR Rotterdam, The Netherlands; e-mail: mmartens@few.eur.nl

Received May 2000; Accepted September 2001

■ *Martin Martens is an Associate Professor in the Econometric Institute at Erasmus University Rotterdam in Rotterdam, The Netherlands.*

INTRODUCTION

Forecasting financial market volatility is important in the short and the long run. For example, Riskmetrics™ uses short-term volatility forecasts to produce Value-at-Risk (VaR) measures while longer term volatility forecasts are needed for option pricing and asset allocation. Many studies, however, report disappointing out-of-sample forecast performances using standard volatility models (see, e.g., Figlewski, 1997, and references herein). Recent work by Andersen and Bollerslev (1998) addressed the problem. Andersen and Bollerslev argued that standard volatility models provide good forecasts, but the ex-post volatility measures used to evaluate the forecasts need to be improved. For the 24-hr foreign exchange market, Andersen and Bollerslev showed that although the daily squared return is an unbiased estimator of the true volatility, it is also an extremely noisy estimator. Both theoretically and empirically the sum of the intraday squared returns¹ provides the best measure of realized volatility. Using this improved volatility measure it is shown that standard volatility models provide good volatility forecasts. Andersen, Bollerslev, and Lange (1999) and Martens (2001) showed that intraday returns can be used to improve not only the measuring of volatility, but also the forecasting of volatility.

The objective of this study is to improve the measuring and forecasting of stock market volatility. Andersen and Bollerslev's results apply to the 24-hr foreign exchange market. Stock markets do not trade on a 24-hr basis and this implies that the observation frequency can only be increased during trading hours.² Furthermore, stock prices potentially have different dynamics during daytime trading hours compared to overnight nontrading hours. International news replaces domestic news during nontrading hours. This differs from foreign exchange markets where most news is released during the business hours of two countries and there is substantial trading activity even outside those business hours.

Various *measures* of stock market volatility are considered that combine the "overnight" (close-to-open) return and the intraday returns that make up the open-to-close return. The best measure of daily stock market volatility is given by a rescaled sum of squared intraday returns, which is better than the sum of the squared "overnight" return and the

¹The highest intraday frequency Andersen and Bollerslev used is a 5-min interval, and this frequency provides the best measure for realized volatility.

²Although some international markets (e.g., United States, France, and Australia) have overnight futures trading, many do not.

squared intraday returns. If overnight futures trading is available,³ the best measure of stock market volatility is given by the sum of squared "intranight" and intraday returns. The relative usefulness of the various measures of stock market volatility is demonstrated by comparison with the daily volatility forecasts that are provided by a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model for S&P 500 index futures returns. Regressing the commonly used daily squared return on the daily GARCH forecasts provides an R^2 of 0.035. However, when regressing the sum of squared intranight and intraday returns on the daily forecasts the R^2 is 0.150. This implies that using the sum of squared intranight and intraday returns is better than using the daily squared return to measure stock market volatility.

Three *forecasting* models are proposed that use intraday returns. An intraday GARCH model that explicitly models the volatility of overnight and intraday returns provides the best daily volatility forecasts. The regression R^2 rises from 0.150 for the daily GARCH(1,1) model to 0.420 for the intraday GARCH model. A longer (weekly or monthly) forecast horizon reduces the importance of explicitly modeling the overnight return volatility. For the weekly and monthly horizon it is sufficient to expand the daily GARCH(1,1) model with the lagged sum of intraday squared returns.

The results show the importance of using intraday returns to the measuring and forecasting of volatility. The outcomes and conclusions of existing studies of volatility forecasting should be reviewed using the measurement of volatility that is based on intraday returns. In addition, the improved GARCH-type forecasts may change the outcome between the competing approaches of using either the historical or the implied volatility. For example, Day and Lewis (1992), studying S&P 100 index returns and options on the S&P 100 index from 1983 to 1989, reported a regression R^2 of only 0.039 for weekly GARCH forecasts and 0.037 for option-implied volatility forecasts. Canina and Figlewski (1993), studying the same index and options from 1983–1987, reported a regression R^2 of 0.053 for implied volatility forecasts and 0.151 for historical volatility averaging over 32 equations and forecast horizons of 7 to 127 calendar days. These regression R^2 s are relatively small vis-à-vis those obtained when using the measurement of volatility that is based on intraday returns.

The remainder of this study is organized as follows. The next section describes the data. The third section discusses various measures of stock

³S&P 500 index-futures started trading on GLOBEX, the electronic overnight trading system, in 1994.

market volatility and compares these measures both theoretically and empirically. The fourth section proposes three forecasting models that use intraday returns and provides an out-of-sample forecast evaluation of the competing models. The final section concludes.

DATA

S&P 500 index-futures transaction prices were obtained for the period from January 1990 through July 1998, a total of 2,140 trading days. Futures data are preferred over the index data for two reasons. First, the S&P 500 index is calculated based on the last transaction price of each of the (500) stocks comprising the index, and not every stock trades every minute. This results in an infrequent trading problem,⁴ meaning the index lags actual developments, especially at the opening of trading since it takes some time before each of the 500 stocks begins trading.⁵ For the measuring and forecasting of stock market volatility it is of crucial importance to measure correctly the close-to-open return. Second, the U.S. S&P 500 index-futures have traded on GLOBEX during the night since 1994, whereas the stock market is closed overnight.

Futures floor trading at the Chicago Mercantile Exchange (CME) is open from 8:30 a.m. to 3:15 p.m., Chicago time, for daytime trading. GLOBEX⁶ overnight trading is from 3:30 p.m. to 8:00 a.m. (8:15 a.m. from February 1995 onward). To avoid market microstructure problems such as bid-ask bounce, the last price in each 5-min interval is selected for floor trading, and the last price in each 30-min interval is selected for overnight trading. All intraday returns are based on prices of the futures contract closest to maturity. For example, prices on the March 1997 futures contract are gathered until March 14, 1997. At this point, prices on the June 1997 futures contract are gathered until that contract is within one week of expiration. When rolling over from the March to the June contract on March 14, the 3:15 p.m. prices for both the March and June contracts are used to compute the correct last 5-min (3:10–3:15 p.m.) return from the March contract and the correct overnight (3:15 p.m.–8:30 a.m.) return for the June contract. This procedure creates a continuous sequence of futures returns. The first 1,000

⁴For a more detailed explanation of the infrequent trading problem, see Stoll and Whaley (1990).

⁵Lee and Linn (1994) demonstrate that for 1983–1985 the close-to-open standard deviation for futures returns is 0.377%, whereas it is only 0.055% for index returns due to the stale opening price.

⁶See Coppejans and Domowitz (1996) for a detailed description of the GLOBEX trading system.

trading days (January 1990–January 1994) are used to estimate the parameters of the various models. The next 1,140 trading days (for which intraday returns are available) are used to test the out-of-sample forecasting performance of 1,140 daily, 228 weekly, and 57 monthly volatility forecasts.

MEASURING STOCK MARKET VOLATILITY

Andersen and Bollerslev (1998) argue that, in most financial applications, the asset price is assumed to follow a continuous time diffusion process, and hence the correct measure for daily volatility is

$$\sigma_{t,1}^2 = \int_0^1 \sigma_{t+\tau}^2 d\tau \quad (1)$$

Unfortunately, volatility is not directly observable. A popular approach is to use the daily squared return to measure daily volatility. Andersen and Bollerslev (1998) showed that although the daily squared return is an unbiased estimator of the expression in equation (1), it is also an extremely noisy estimator. One extreme example would be a very volatile day with wildly fluctuating prices, but where the closing price is the same as the previous closing price. The daily (close-to-close) squared return would then equal zero, whereas the actual volatility has been nonzero.

For the 24-hr foreign exchange market, Andersen and Bollerslev proposed the use of the sum of 288 squared 5-min returns to approximate equation (1) accurately. For stock markets, however, intraday returns are only available during trading hours. In the absence of overnight trading we can only observe the close-to-open return for the overnight period. A logical candidate to measure stock market volatility is then

$$(r_t^N)^2 + \sum_{d=1}^D (r_{t,d}^D)^2 \quad (2)$$

where r_t^N is the overnight return on day t , and $r_{t,d}^D$ is the intraday return on day t for intraday period d ($d = 1 \dots D$). The squared overnight return is a noisy estimate for overnight volatility. An alternative volatility measure, therefore, uses only intraday returns,

$$(1 + c) \sum_{d=1}^D (r_{t,d}^D)^2 \quad (3)$$

where c is a positive constant such that the volatility measure in equation (3) corresponds to daily volatility. When overnight (futures) trading is available, the overnight volatility can be approximated by the sum of intranight squared returns,

$$\sum_{n=1}^N (r_{t,n}^N)^2 + \sum_{d=1}^D (r_{t,d}^D)^2 \quad (4)$$

where $r_{t,n}^N$ is the intranight return on day t and intranight period n ($n = 1 \dots N$).

A Monte Carlo Simulation

The same continuous time GARCH framework as in Andersen and Bollerslev (1998) is used to provide a theoretical framework to compare the various measures of stock market volatility. The model has the property that all discretely sampled time series follow a GARCH(1,1) model. The main motivation for conducting a simulation experiment is that, since the true volatility is known, the candidate volatility measures can be compared with certainty. For details on the Monte Carlo simulation experiment and the relationship between the continuous time GARCH parameters and the parameters of a daily GARCH(1,1) model, see Andersen and Bollerslev (1998, pp. 892–899).⁷

Table I reports the daily GARCH(1,1) parameter estimates using daily closing prices for the S&P 500 futures contract from January 3, 1990 through January 10, 1994 (1,000 return observations). The combined value of the volatility parameters, $\alpha + \beta$, is 0.9986. This is similar to the 0.9985 reported by Randolph and Najand (1991) for S&P 500 futures from 1986 to 1988.⁸ Hence the volatility is highly persistent.

⁷The simulation here differs in one important aspect from the one in Andersen and Bollerslev. They use actual clock time, but in our case this would overestimate the importance of the overnight period. The open-to-close variance equals 0.585, whereas the close-to-open variance averages only 0.120, similar to findings in Lee and Linn (1994) for an older sample of S&P 500 futures. It is therefore assumed that one day consists of ninety-eight “5-min” intervals, of which 81 intervals are attributed to floor trading.

⁸Randolph and Najand (1991) use several models to forecast the monthly standard deviation of S&P 500 futures returns for the period 1986–1988. The competing models are the daily GARCH(1,1) model, implied volatility from the nearest-the-money call option on the S&P 500 futures contract, the previous month's standard deviation, and a mean-reversion model. GARCH performs marginally better than implied volatility and the previous month's a standard deviation, with the mean-reversion model providing the best forecasts.

TABLE I
GARCH(1,1) Parameter Estimates

	Daily Frequency		Continuous Time
μ	0.0186 (0.025)		
ω	0.0002 (0.002)	θ	0.0014
α	0.0175 (0.009)	ϕ	0.1798
β	0.9811 (0.011)	λ	0.2598

Note. The daily GARCH(1,1) model is estimated for 1,000 close-to-close S&P 500 futures returns, r_t , from January 4, 1990 through January 10, 1994:

$$\begin{aligned} r_t &= \mu + \varepsilon_t \\ \varepsilon_t | \Phi_{t-1} &\sim N(0, h_t) \\ h_t &= \omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot h_{t-1} \end{aligned}$$

where r_t is the daily (close-to-close) futures return, Φ_{t-1} contains all information up to day $t-1$, and the residuals ε_t follow a conditional normal distribution with mean zero and variance h_t .

The parameter estimates of the daily GARCH(1,1) model are used to compute the parameters of the continuous time GARCH diffusion model following Drost and Werker (1996):

$$\begin{aligned} dS_t &= \sigma_t \cdot dW_t^{(1)} \\ d\sigma_t^2 &= \theta \cdot (\phi - \sigma_t^2) \cdot dt + \sqrt{2 \cdot \lambda \cdot \theta} \cdot \sigma_t^2 \cdot dW_t^{(2)} \end{aligned}$$

where S_t is the log futures value, $\theta > 0$, $\phi > 0$, $0 < \lambda < 1$, and $W_t^{(1)}$ and $W_t^{(2)}$ are independent Brownian motions. Heteroskedasticity consistent errors are inside parentheses.

In Andersen and Bollerslev's experiment the various measures of volatility are compared by using each of them to evaluate the forecast performance of the daily GARCH(1,1) model. The best measure of volatility is that which most closely estimates the true forecast performance, which is known as we have the true volatility as well.

There is a wide range of evaluation criteria used in the literature. One of the most popular evaluation criteria is the regression R^2 obtained from the regression of realized volatility on forecasted volatility. In addition, the heteroskedasticity adjusted root mean squared error (HRMSE) and mean absolute error (HMAE) are computed, as follows:

$$HRMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(1 - \frac{\text{Realized}_t}{\text{Forecast}_t} \right)^2} \quad (5)$$

$$HMAE = \frac{1}{T} \sum_{t=1}^T \left| \left(1 - \frac{\text{Realized}_t}{\text{Forecast}_t} \right) \right| \quad (6)$$

These two forecast criteria follow the statistical tradition of reporting statistics based directly on the deviation between forecasts and realizations, while adjusting for heteroskedasticity in the forecast errors.⁹

The results of the Monte Carlo simulation are provided in Table II. If overnight trading is available, the trading process becomes similar to a 24-hr foreign exchange market. In that case volatility is best measured by the sum of squared intranight and intraday returns, as defined in equation (4). For example, the regression R^2 for the daily horizon is 0.415.

In the absence of overnight trading, the best *available* measure for realized volatility is the rescaled sum of squared 5-min returns from the trading day, excluding the overnight return (as defined in equation (3) with $1 + c$ equal to 98/81). To illustrate the effect of excluding or including the overnight return, the regression R^2 is 0.400 when excluding the squared overnight return, whereas the regression R^2 is only 0.267 when including the squared overnight return.

Similar to the daily squared return, the squared overnight return is a very noisy estimate of the true volatility for the nontrading period. In this case, it is even better to assume that volatility measured for part of the 24-hr day (81 out of 98 intervals) applies to the entire day, than to use the aggregated information including the remaining 17 intervals. There is a potential concern with this procedure. This procedure may overestimate the forecasting performance of the models. After all, the most difficult (if not impossible!) part in forecasting is to foresee unexpected events that cause major jumps in volatility.¹⁰ The GARCH model, however, is providing forecasts conditional upon past information, and as such should be interpreted as *expected* volatility. In any case, it is an interesting finding that, if the data-generating process is continuous time GARCH, the overnight squared return should be excluded when measuring stock market volatility.¹¹

The relative differences between the measures of stock market volatility using intraday returns become smaller when measuring volatility over multiple days. A large gap still remains between using the sum of daily squared returns and the sum of intraday squared returns when measuring monthly volatility. The regression R^2 is 0.213 for the sum of daily squared returns, compared to at least 0.466 for the intraday measures.

Finally, it is important to compare the different measures of stock market volatility with the true volatility. Regressing the true volatility on

⁹The HRMSE and HMAE are also used in Andersen et al. (1999).

¹⁰Some extreme examples would be the start of the Gulf War, and the recent Russian and Asian crises.

¹¹Note that given the general framework these conclusions may well extend to non-U.S. markets and other assets for which only the close-to-open return and intraday trading returns can be observed.

TABLE II
Theoretical Forecast Evaluation for Daily GARCH(1,1) Forecasts

Realized	Forecast Horizon		
	1 Day [1,140 Obs]	5 Days (1 Week) [228 Obs]	20 Days (1 Month) [57 Obs]
<i>Panel A: Regression R²</i>			
r_m^2			0.0370
r_w^2		0.0256	0.0851
r_i^2	0.0223	0.0900	0.213
$(r_t^N)^2 + \sum (r_{t,d}^D)^2$	0.267	0.429	0.466
$\frac{98}{81} \sum (r_{t,d}^D)^2$	0.400	0.488	0.484
$\sum (r_{t,n}^N)^2 + \sum (r_{t,d}^D)^2$	0.415	0.494	0.486
$\int \sigma^2(u) du$	0.552	0.543	0.512
<i>Panel B: HRMSE</i>			
r_m^2			1.41
r_w^2		1.43	0.734
r_i^2	1.45	0.671	0.377
$(r_t^N)^2 + \sum (r_{t,d}^D)^2$	0.344	0.233	0.211
$\frac{98}{81} \sum (r_{t,d}^D)^2$	0.252	0.208	0.204
$\sum (r_{t,n}^N)^2 + \sum (r_{t,d}^D)^2$	0.243	0.206	0.204
$\int \sigma^2(u) du$	0.188	0.189	0.196
<i>Panel C: HMAE</i>			
r_m^2			0.972
r_w^2		0.974	0.560
r_i^2	0.976	0.510	0.296
$(r_t^N)^2 + \sum (r_{t,d}^D)^2$	0.252	0.183	0.168
$\frac{98}{81} \sum (r_{t,d}^D)^2$	0.199	0.165	0.163
$\sum (r_{t,n}^N)^2 + \sum (r_{t,d}^D)^2$	0.192	0.164	0.163
$\int \sigma^2(u) du$	0.150	0.151	0.157

Note. The continuous time GARCH(1,1) model is simulated for the parameters reported in Table I. Every time the equivalent of 2,140 days is simulated; this is repeated 1,000 times. Each time the antithetic variable technique is used. The first 1,000 daily returns are used to load the conditional variances of the daily GARCH(1,1) model using the parameters reported in Table I. Next, 1,140 daily, 228 weekly, and 57 monthly forecasts are produced. These forecasts are evaluated vis-à-vis seven realized volatility measures: (a) The monthly squared return (r_m^2); (b) The weekly squared return (r_w^2); (c) The daily squared return (r_i^2); (d) The sum of the squared overnight return and squared intraday 5-min returns ($(r_t^N)^2 + \sum (r_{t,d}^D)^2$); (e) The rescaled sum of squared intraday 5-min returns ($\frac{98}{81} \sum (r_{t,d}^D)^2$); (f) The sum of squared overnight and intraday 5-min returns ($\sum (r_{t,n}^N)^2 + \sum (r_{t,d}^D)^2$); and (g) The "true" volatility using the simulation frequency of 1-min returns. The table shows the average regression R^2 , HRMSE, and HMAE taken over the 2,000 simulation runs.

the various proposed measures of volatility, the regression R^2 for the daily horizon is only 0.0376 (not reported in tables) for the daily squared return, whereas the regression R^2 is 0.721 for the sum of squared intranight and intraday returns [equation (4)]. The HRMSE for the daily squared return is 1.412 whereas for the sum of squared intranight and intraday returns, the HRMSE is only 0.151. These results confirm the large gains that are obtained from using intraday returns when measuring stock market volatility.

Measuring Stock Market Volatility: Empirical Results

The simulation results depend on the continuous time GARCH model assumption. It is therefore important to verify the conclusions from the simulation experiment by repeating the analysis for the actual data, even though this only provides one sample path and the true volatility is unknown. Again we follow Andersen and Bollerslev's work on foreign exchange markets in that GARCH(1,1) forecasts are determined for the 1,140 days, 228 weeks, and 57 months from January 1994 through July 1998. S&P 500 index futures have traded on GLOBEX since January 1994, and hence all realized volatility measures examined in the simulation are available. The results are reported in Table III.

The conclusions are similar to those of the simulation experiment. The sum of squared intranight and intraday returns almost always result in the highest regression R^2 , and lowest HRMSE and HMAE. For example, when forecasting one day ahead and measuring stock market volatility using the daily squared return, the regression R^2 is only 0.035. This low R^2 is in line with existing studies.¹² However, when the sum of squared intranight and intraday returns is used, the regression R^2 increases to 0.150.¹³ The HRMSE drops from 2.62 for the daily squared return to 1.15 for the sum of intranight and intraday squared returns.

The differences between the sum of the squared overnight return and squared intraday returns [equation (2)] and the rescaled sum of intraday squared returns [equation (3) with $1 + c$ equal to 98/81] are smaller for the empirical data than in the simulation. As the simulation predicts, the HRMSE and HMAE are lower for all horizons (except for HMAE at the monthly horizon) when the overnight return is excluded. For the daily

¹²See, e.g., Day and Lewis (1992) and Pagan and Schwert (1990).

¹³In the absence of intraday the daily high and low could be used for measuring volatility (see, e.g., Garman & Bliss, 1980), although similar to findings in Andersen and Bollerslev (1998) it is not as good as the use of intraday returns. The high-low range results in a regression R^2 of 0.124.

TABLE III
Empirical Forecast Evaluation for Daily GARCH(1,1) Forecasts

Realized	Forecast Horizon		
	1 Day [1,140 Obs]	5 Days (1 Week) [228 Obs]	20 Days (1 Month) [57 Obs]
<i>Panel A: Regression R²</i>			
r_m^2			0.056
r_w^2		0.057	0.167
r_i^2	0.035	0.068	0.200
$(r_i^N)^2 + \Sigma(r_{i,d}^D)^2$	0.132	0.198	0.361
$\frac{98}{81}\Sigma(r_{i,d}^D)^2$	0.130	0.183	0.316
$\Sigma(r_{i,n}^N)^2 + \Sigma(r_{i,d}^D)^2$	0.150	0.226	0.391
<i>Panel B: HRMSE</i>			
r_m^2			1.07
r_w^2		1.57	0.740
r_i^2	2.62	1.52	0.824
$(r_i^N)^2 + \Sigma(r_{i,d}^D)^2$	1.23	0.910	0.649
$\frac{98}{81}\Sigma(r_{i,d}^D)^2$	0.956	0.791	0.646
$\Sigma(r_{i,n}^N)^2 + \Sigma(r_{i,d}^D)^2$	1.15	0.870	0.631
<i>Panel C: HMAE</i>			
r_m^2			0.846
r_w^2		1.00	0.593
r_i^2	1.22	0.803	0.573
$(r_i^N)^2 + \Sigma(r_{i,d}^D)^2$	0.633	0.511	0.459
$\frac{98}{81}\Sigma(r_{i,d}^D)^2$	0.527	0.478	0.485
$\Sigma(r_{i,n}^N)^2 + \Sigma(r_{i,d}^D)^2$	0.593	0.484	0.441

Note. The daily GARCH(1,1) model is estimated for S&P 500 futures returns from January 4, 1990 through January 10, 1994 (1,000 observations); see Table I. Without re-estimating the model parameters, forecasts are produced for the next 1,140 days, 228 weeks, or 57 months from January 1994 through July 1998. These forecasts are evaluated vis-à-vis six realized volatility measures: (a) The monthly squared return (r_m^2); (b) The weekly squared return (r_w^2); (c) The daily squared return (r_i^2); (d) The sum of the squared overnight return and squared intraday 5-min returns ($(r_i^N)^2 + \Sigma(r_{i,d}^D)^2$); (e) The rescaled sum of squared intraday 5-min returns ($\frac{98}{81}\Sigma(r_{i,d}^D)^2$); and (f) The sum of squared overnight and intraday 5-min returns ($\Sigma(r_{i,n}^N)^2 + \Sigma(r_{i,d}^D)^2$).

horizon, for example, the HRMSE drops from 1.23 to 0.956 and the HMAE drops from 0.633 to 0.527. Although the difference in regression R^2 (0.132 vs. 0.130) is not significant, the HRMSE, HMAE, and simulation results suggest that for the daily horizon the overnight return should be excluded. In contrast to the simulation results the regression R^2 is higher for the weekly and monthly horizons when the squared overnight return is included. The higher R^2 indicates that the continuous time GARCH diffusion model may not be the Data Generating Process (DGP),

and that the overnight (squared) return may play a different role than is assumed under the continuous time GARCH model.

Given the mixed evidence for the weekly and monthly horizons, more empirical research is needed to test which measure of stock market volatility is superior when no overnight trading is available. GLOBEX trading data is available for the S&P 500 index-futures and therefore the sum of intranight and intraday squared returns is the best measure of stock market volatility in the present study.^{14,15}

FORECASTING STOCK MARKET VOLATILITY

This section addresses the question of whether intraday returns can also be used to improve the forecasting of stock market volatility. Three different models are proposed; each model treats the overnight return volatility in a different way. The forecasting performance of each model is reported at the end of the section.

EXTENDING THE DAILY GARCH MODEL

One possible approach to the forecasting of stock market volatility is to extend the daily GARCH(1,1) model by including intraday formation. A popular specification¹⁶ is

$$r_t = \mu + \varepsilon_t$$

$$\varepsilon_t | \Phi_{t-1} \sim N(0, h_t) \quad (7)$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma I_{t-1}$$

¹⁴Results are also obtained for other frequencies, splitting up the floor trading session in twenty-seven 15-min intervals, nine 4-min intervals, three 135-min intervals, and one 405-min interval. This leads to similar patterns, in that the higher the frequency, the higher the regression R^2 , and the lower the HRMSE and HMAE. For the daily horizon and the volatility measure in equation (2), for example, the regression R^2 is 0.132, 0.090, 0.108, 0.065, and 0.028 for the 5-min, 15-min, 45-min, 135-min, and 405-min frequencies, respectively. Similarly, for the overnight period the results are replicated when combining 60-min and 90-min intranight returns with the 5-min floor returns in equation (4). For the daily horizon this results in a regression R^2 (HRMSE; HMAE) of 0.150 (1.15; 0.593), 0.152 (1.14; 0.596), and 0.154 (1.16; 0.601) for the 30-min, 60-min, and 90-min intranight frequencies, respectively.

¹⁵All results are reproduced for standard deviations instead of variances, to mitigate the impact of outliers. The results and conclusions are similar, except that regression R^2 s are larger and the HRMSEs and HMAEs are smaller for standard deviations.

¹⁶See, e.g., Bessembinder and Seguin (1993) and Laux and Ng (1993). These studies use volume and the number of trades, respectively, to extend the daily GARCH model. Taylor and Xu (1997) used implied volatility (from currency options) and the sum of intraday squared returns for the 24-hr foreign exchange market.

where r_t is the daily (close-to-close) return, Φ_{t-1} contains all information up to day $t - 1$, and the residuals ε_t follow a conditional normal distribution with mean zero and variance h_t . I_{t-1} is the intraday information from the previous trading day. Here we use either equation (2) or (3) for I_{t-1} . The sum of intranight and intraday squared returns is not available for forecasting, since during the in-sample period (January 1990 through January 1994) overnight trading did not exist.

Table IV presents the parameter estimates. The results for the daily GARCH(1,1) model are repeated from Table I for comparison, and the log-likelihood value is included. The log-likelihood value increases from $-1,155.1$ to $-1,136.4$ when including the sum of the squared overnight returns and the squared intraday returns as defined in equation (2). The log-likelihood value increases from $-1,155.1$ to $-1,138.6$ when including the sum of squared intraday returns as defined in equation (3). In both cases a log-likelihood ratio test rejects the null hypothesis of

TABLE IV
Estimates for Daily GARCH Model Extended with Intraday Returns

	Daily GARCH(1,1)	Extended with Equation (2)	Extended with Equation (3)
μ	0.0186 (0.025)	0.0023 (0.040)	0.0003 (0.025)
ω	0.0002 (0.002)	0.0224 (0.031)	0.0169 (0.025)
α	0.0175 (0.009)	-0.0363 (0.044)	-0.0270 (0.036)
β	0.9811 ^a (0.011)	0.7567 ^a (0.201)	0.8051 ^a (0.178)
γ	—	0.2405 (0.202)	0.2267 (0.205)
Log-likelihood	-1,155.1	-1,136.4	-1,138.6

Note. The daily GARCH(1,1) model extended with intraday returns is estimated for S&P 500 futures returns from January 4, 1990 through January 10, 1994 (1,000 days).

$$r_t = \mu + \varepsilon_t$$

$$\varepsilon_t | \Phi_{t-1} \sim N(0, h_t)$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma I_{t-1}$$

where r_t is the 24-h close-to-close return, and I_{t-1} is either the sum of the squared overnight return and squared intraday returns [equation (2)] or the sum of squared intraday returns (equation (3) with c equal to zero). Heteroskedasticity consistent errors are reported inside parentheses. Superscripts "a" and "b" denote significance at the 1 and 5% level, respectively.

$\gamma = 0$.¹⁷ Hence the intraday measures significantly contribute to the in-sample fit of the data. Although the models including the two intraday measures are non-nested, the model including the squared overnight return provides a slightly better in-sample fit.

Modeling Overnight and Intraday Returns

Alternatively, the overnight returns and the returns during trading hours can be modeled directly. Chan, Chan, and Karolyi (1991) proposed a bivariate GARCH model for the overnight and intraday 5-min returns for the S&P 500 index and index-futures. Their model is used to investigate the intraday return and volatility spillovers between the spot and futures market. Chan et al. stacked the overnight and intraday 5-min returns and included a dummy to scale the overnight return for the mean and the volatility of returns.

In this study an alternative model for intraday returns is proposed specifically taking into account the potentially different dynamics of overnight returns and returns during trading hours:

$$\begin{aligned}
 r_t^N &= \mu_N + \varepsilon_t^N \\
 r_{t,d}^D &= \mu_D + \varepsilon_{t,d}^D \quad (d = 1, 2, \dots, D) \\
 \varepsilon_t^N | \Psi_{t-1} &\sim N(0, h_t^N) \\
 \varepsilon_{t,d}^D | \Psi_{t,d-1} &\sim N(0, h_{t,d}^D) \\
 h_t^N &= \omega_1 + \phi_1(\varepsilon_{t-1}^N)^2 + \phi_2 \sum_{d=1}^D (\varepsilon_{t-1,d}^D)^2 + \phi_3 h_{t-1}^N \\
 h_{t,1}^D &= \omega_2 + \omega_3 + \theta_1(\varepsilon_t^N)^2 + \theta_2(\varepsilon_{t-1,D}^D)^2 + \theta_3 h_{t-1,D}^D \\
 h_{t,d}^D &= \omega_2 + \theta_2(\varepsilon_{t,d-1}^D)^2 + \theta_3 h_{t,d-1}^D
 \end{aligned} \tag{8}$$

Implicitly it is assumed that the overnight return, r_t^N , and the intraday returns, $r_{t,d}^D$, each follow a GARCH(1,1) process. The conditional volatility

¹⁷For example, for the model using the intraday measure in equation (2), the LR test statistic for $\gamma = 0$ is equal to $-2(-1,155.1 \text{ to } -1,136.4) = 37.4$. The LR test statistic follows a chi-squared distribution with one degree of freedom, and hence the 1% critical value is 6.63. A standard t test on γ would not reject the null hypothesis. This is caused by the relatively large standard errors of the GARCH(1,1) model extended with intraday returns. A closer investigation reveals this is caused by a relatively flat log-likelihood function and several outliers (not necessarily on the same days) for the squared return and the sum of intraday squared returns. For example, imposing $\alpha = 0$ more than halves the standard error of γ and makes it significant at the 5% level using a t test. Winsorizing the intraday measure (if the intraday measure exceeds 5%, then make the intraday measure equal to 5%) also approximately halves the standard error making γ significant at the 1% level using a t test (even with $\alpha \neq 0$).

of the overnight return, h_t^N , is affected by the realized volatility during the last trading session through ϕ_2 , as well as the previous overnight volatility (shock) through ϕ_1 and ϕ_3 . The conditional volatility in the first intraday interval ($d = 1$), $h_{t,1}^D$, is affected by the overnight return through θ_1 . For out-of-sample forecasting the overnight squared residual is replaced by its expectation, that is, the forecasted overnight conditional variance.

This model is only one of numerous possibilities to model the overnight interval differently from the daytime trading intervals. The underlying idea is that the overnight interval is different due to the closure of the stock market and the absence of domestic news. Another intuition for this model is obtained by viewing it as a discrete version of a continuous time jump diffusion process. The difference here is that the time of the jump is known—that is, the overnight interval—but the magnitude of the jump is unknown. The magnitude of the jump can depend on the previous jump as well as the daytime volatility. Henceforth the model in equation (8) will be referred to as the “intraday model.”

Table V presents the estimation results for the intraday model¹⁸ for various frequencies for the in-sample period from January 4, 1990 through January 10, 1994. The volatility of the last trading session, measured by the sum of squared intraday (residual) returns, has a significant impact on the overnight volatility through the parameter ϕ_2 . Interestingly, ϕ_2 gradually increases with an increase of the frequency of observations, indicating that the open-to-close volatility becomes more important when it is more accurately measured. The estimate for ϕ_2 rises from 0.0160 for the squared open-to-close return to 0.0517 for the sum of intraday squared 5-min returns. As a consequence, less weight is placed on the lagged overnight volatility through ϕ_3 , which is the smallest at 0.6755 for the 5-min frequency. Although the parameter θ_1 is always positive, the overnight squared return only has a statistically significant impact on the volatility during the subsequent trading session for the 15-min and 5-min frequencies.

Continuous Time GARCH Diffusion

If the underlying process is continuous time GARCH, as used in the simulation, the returns measured over discrete intervals follow a GARCH process. Andersen et al. (1999) used this property to produce

¹⁸The Chan et al. model was also estimated. It is found that the more complex model in equation (8) provides better forecasts at all horizons.

TABLE V
Parameter Estimates for the Intraday Model

	405 Min (2,000 Obs)	135 Min (4,000 Obs)	45 Min (10,000 Obs)	15 Min (28,000 Obs)	5 Min (82,000 Obs)
μ_1	-0.0055 (0.0070)	-0.0033 (0.0077)	-0.0062 (0.0086)	-0.0031 (0.0610)	0.0024 (0.0365)
μ_2	0.0363 (0.0204)	0.0100 (0.0074)	0.0042 (0.0022)	0.0024 ^a (0.0008)	0.0009 ^a (0.0002)
ω_1	0.0002 (0.0008)	-0.0002 (0.0010)	-0.0007 (0.0012)	-0.0012 (0.0014)	-0.0013 (0.0023)
ω_2	0.0145 (0.0147)	0.0130 ^b (0.0052)	0.0012 ^a (0.0004)	0.0006 ^a (0.0002)	0.0001 ^a (0.0000)
ω_3	—	-0.0332 ^b (0.0147)	-0.0073 ^a (0.0026)	0.0008 (0.0009)	0.0012 (0.0006)
ϕ_1	0.0723 ^b (0.0304)	0.0749 ^b (0.0309)	0.0684 ^b (0.0282)	0.0721 ^a (0.0057)	0.0706 ^a (0.0004)
ϕ_2	0.0160 ^b (0.0063)	0.0220 ^b (0.0102)	0.0348 ^b (0.0156)	0.0453 ^a (0.0060)	0.0517 ^a (0.0036)
ϕ_3	0.8482 ^a (0.0472)	0.8231 ^a (0.0676)	0.7718 ^a (0.0847)	0.7163 ^a (0.0045)	0.6755 ^a (0.0005)
θ_1	0.2455 (0.1777)	0.0707 (0.0407)	0.0410 (0.0223)	0.0540 ^a (0.0097)	0.0270 ^a (0.0043)
θ_2	0.0305 (0.0200)	0.0301 ^b (0.0134)	0.0329 ^a (0.0099)	0.1214 ^a (0.0150)	0.1197 ^a (0.0036)
θ_3	0.8949 ^a (0.0702)	0.9448 ^a (0.0253)	0.9527 ^a (0.0158)	0.8455 ^a (0.0231)	0.8677 ^a (0.0031)

Note. The model in equation (8) is estimated for S&P 500 futures returns from January 4, 1990 through January 10, 1994 (1,000 days). The trading session of 405 minutes (8:30 a.m. through 3:15 p.m.) is taken as it is, or split up in $D = 3$ (135-min), 9 (45-min), 27 (15-min) or 81 (5-min) intervals. Heteroskedasticity consistent errors are reported inside parentheses. Superscripts "a" and "b" denote significance at the 1 and 5% level, respectively.

$$\begin{aligned} r_t^N &= \mu_N + \varepsilon_t^N \\ r_{t,d}^D &= \mu_D + \varepsilon_{t,d}^D \quad (d = 1, 2, \dots, D) \\ \varepsilon_t^N | \Psi_{t-1} &\sim N(0, h_t^N) \\ \varepsilon_{t,d}^D | \Psi_{t,d-1} &\sim N(0, h_{t,d}^D) \\ h_t^N &= \omega_1 + \phi_1(\varepsilon_{t-1}^N)^2 + \phi_2 \sum_{d=1}^D (\varepsilon_{t-1,d}^D)^2 + \phi_3 h_{t-1}^N \\ h_{t,1}^D &= \omega_2 + \omega_3 + \theta_1(\varepsilon_t^N)^2 + \theta_2(\varepsilon_{t-1,D}^D)^2 + \theta_3 h_{t-1,D}^D \\ h_{t,d}^D &= \omega_2 + \theta_2(\varepsilon_{t,d-1}^D)^2 + \theta_3 h_{t,d-1}^D \end{aligned}$$

foreign exchange volatility forecasts from GARCH(1,1) models for various high frequencies. For the stock market, however, only the aggregated return is observable for the overnight period. To solve this problem, intranight returns are simulated in the overnight period such that, on average, the total overnight return is correct and the draws come from

the conditional distribution with the conditional variance updated every period. The average path for the conditional variance is subsequently used to fill the overnight period. This model is applied in two different ways. First, for 5-min intervals it is assumed there are 288 five-min intervals in a day, and the time allocation to daytime trading and the overnight return is according to actual time. Second, the number of periods during daytime trading (81) and during the night return (17) is set such that it reflects the relative contribution of daytime trading and the overnight return to daily volatility.

Forecasting Stock Market Volatility: Empirical Results

The forecasting models presented previously are compared to each other, and also to the forecasts from the daily, weekly, and monthly GARCH(1,1) models in order to show the relative importance of intraday returns. All forecasts are compared to realized volatility using the three forecast criteria introduced in the previous section— R^2 , HRMSE, and HMAE. Based on the previous findings, the sum of squared intranight and intraday returns [see equation (4)] is used to measure stock market volatility.¹⁹

The results for the out-of-sample (January 1994 through January 1998) forecasting performance of the various models are presented in Table VI. The first three models are standard GARCH(1,1) models applied to monthly, weekly, or daily returns. The parameters for the monthly and weekly GARCH(1,1) models are inferred from the daily GARCH(1,1) parameters presented in Table I using the formulae derived in Drost and Nijman (1993). The next two models are the two versions of the GARCH(1,1) model extended with intraday returns, as presented in equation (7). The sixth model is the intraday model presented in equation (8), using 5-min returns. Finally, there are two versions of the continuous GARCH model. Columns 2 to 4 present the forecast performance for the daily, weekly, and monthly forecasting horizon.

For the *daily horizon* the results show that the daily GARCH model is inferior to the three models: the daily GARCH model extended with intraday information, the intraday model, and the continuous time GARCH model. The regression R^2 increases from 0.150 for the daily GARCH model to 0.226 for the daily GARCH(1,1) model extended with

¹⁹The conclusions are unaffected when measuring stock market volatility using the sum of the squared overnight return and squared intraday returns [equation (2)] or the rescaled sum of squared intraday returns [equation (3)]. These would be the best candidates for measuring volatility when there is no overnight trading.

TABLE VI
Out-of-Sample Forecast Performance

Forecast	Forecast Horizon		
	1 Day [1140 Obs]	5 Days (1 Week) [228 Obs]	20 Days (1 Month) [57 Obs]
<i>Panel A: Regression R²</i>			
Monthly GARCH(1,1)			0.342
Weekly GARCH(1,1)		0.139	0.227
Daily GARCH(1,1)	0.150	0.226	0.391
Daily GARCH(1,1) extended			
$l_{1t} = (r_t^N)^2 + \sum (r_{t,d}^D)^2$	0.226	0.288	0.317
$l_{2t} = \sum (r_{t,d}^D)^2$	0.229	0.301	0.376
Intraday model (5-min)	0.420	0.221	0.259
Continuous GARCH real time	0.224	0.309	0.447
Continuous GARCH adjusted	0.255	0.307	0.375
<i>Panel B: HRMSE</i>			
Monthly GARCH			1.70
Weekly GARCH		1.54	1.21
Daily GARCH(1,1)	1.15	0.870	0.631
Daily GARCH(1,1) extended			
$l_{1t} = (r_t^N)^2 + \sum (r_{t,d}^D)^2$	0.915	0.778	0.643
$l_{2t} = \sum (r_{t,d}^D)^2$	0.953	0.794	0.460
Intraday model (5-min)	0.638	0.654	0.607
Continuous GARCH real time	0.703	0.672	0.627
Continuous GARCH adjusted	0.881	0.800	0.644
<i>Panel C: HMAE</i>			
Monthly GARCH			1.15
Weekly GARCH		0.820	0.814
Daily GARCH(1,1)	0.593	0.484	0.441
Daily GARCH(1,1) extended			
$l_{1t} = (r_t^N)^2 + \sum (r_{t,d}^D)^2$	0.518	0.427	0.401
$l_{2t} = \sum (r_{t,d}^D)^2$	0.522	0.409	0.358
Intraday model (5-min)	0.509	0.536	0.568
Continuous GARCH real time	0.673	0.648	0.603
Continuous GARCH adjusted	0.507	0.425	0.439

Note. Forecast performance of standard GARCH(1,1) and intraday models. The models are estimated for S&P 500 futures returns from January 4, 1990 through January 10, 1994 (1,000 observations), and forecasts are produced for January 1994 through July 1998 (1,140 observations). The results in Drost and Nijman (1993) are used to infer the parameters for the weekly and monthly GARCH(1,1) model. The models using intraday returns are (a) The daily GARCH(1,1) extended with intraday volatility measures l_{1t} or l_{2t} , see equation (7); (b) The intraday model defined in equation (8) for 5-min returns; and (c) "Continuous GARCH" for 5-min returns. This is a GARCH(1,1) model for 5-min returns, with parameters again inferred using Drost and Nijman. For the close-to-open interval the 5-min returns are nonobservable, and the conditional variance is determined by simulation. The close-to-open interval is either assumed to consist of 207 five-min returns (labeled "real time") or 17 five-min returns (labeled "adjusted") to reflect the relative contribution to volatility of close-to-open and open-to-close returns. Realized variance is the sum of squared 5-min returns from floor trading and the sum of squared 30-min returns from overnight trading.

the sum of the squared overnight return and squared intraday returns. The daily GARCH model extended with the sum of squared intraday returns has an R^2 of 0.229. The intraday model has the highest R^2 of 0.420, whereas the continuous time GARCH forecasts have R^2 s of 0.224 and 0.255 using real time and adjusted time, respectively. Hence, for forecasting stock market volatility it is important to include the intraday returns.²⁰

For the daily horizon the intraday model provides the best forecasts using the three performance measures. The intraday model has the highest regression R^2 and has the smallest or second smallest HRMSE and HMAE. The HRMSE is the lowest at 0.638—well below, for example, 1.15 for the daily GARCH(1,1) model. These results show the useful potential for modeling the overnight return in a different way from the daytime intraday returns.

For the *weekly horizon*, the performance of the weekly GARCH model is the worst of all models considered. The use of returns at a higher frequency than the weekly period leads to a superior forecasting performance. Interestingly, the intraday model is not the best model for the weekly horizon. Seemingly, for longer horizons, the modeling of the overnight return separately is no longer important.²¹ In this case, the daily GARCH model extended with the sum of intraday returns (but not the squared overnight return) provides the best results, closely followed by the continuous time GARCH models. For the continuous time GARCH models there is not much difference between the use of 207 (“real time”) and 19 (“adjusted”) 5-min returns for the overnight period.

For the *monthly horizon* the monthly GARCH model is the worst of all models. Similar to the daily and weekly frequencies, the use of returns of a higher frequency than the monthly period leads to a superior forecasting performance. Against the daily GARCH model extended with intraday information and the continuous time GARCH models are the best models, with the simple daily GARCH model also performing well. The HRMSE, for example, is 1.70 for the monthly GARCH model, whereas it is only 0.631 for the daily GARCH model. The daily GARCH model extended with intraday returns has the lowest HRMSE of 0.460.

²⁰Extending the daily GARCH model with the high-low range produces slightly inferior forecasts (R^2 is 0.211) compared to those using intraday returns.

²¹One explanation is that U.S. market volatility is affected the day immediately after large foreign volatility shocks, but not in the longer run. For other (non-U.S.) markets, where U.S. market movements are closely monitored, this result may well be different.

For the foreign exchange market, Andersen et al. (1999) reported that, for very high frequencies, the forecast performance actually deteriorates. Using the continuous GARCH model they obtain the best (daily, weekly, and monthly) forecasts for the 1-hr return frequency, which provides better results than, for example, the 5-min return frequency. They attributed this finding to market microstructure effects. To check whether this is also the case for the S&P 500 index-futures, the open-to-close return is divided into twenty-seven 15-min intervals, nine 45-min intervals, three 135-min intervals, or one 405-min interval. Note that the estimation results for the corresponding intraday models are presented in Table IV. The results (not reported here) show that the 5-min frequency provides the best forecasts, contrary to the findings of Andersen et al. for the foreign exchange market. For the daily horizon, for example, the regression R^2 for the intraday model is 0.420, 0.420, 0.309, 0.238, and 0.196 for the 5-min, 10-min, 45-min, 135-min, and 405-min frequencies, respectively.

To summarize, it appears that for the daily horizon, it is important to have a good model for the overnight returns. At the weekly frequency, this is no longer important, but the use of intraday returns still substantially improves forecasting power. At the monthly frequency, intraday returns become relatively less important, but they still improve results vis-à-vis using the daily GARCH model. The monthly GARCH(1,1) model performs very poorly compared to the daily and higher frequency models.

CONCLUSION

This study extends the work of Andersen and Bollerslev (1998) on the 24-hr foreign exchange market to the stock market. For *measuring* stock market volatility, the rescaled sum of the squared intraday trading returns provides the best approximation of the true (unobservable) volatility in the absence of overnight trading. The squared overnight return is a noisy estimate of the overnight volatility, and it is better to exclude it from the volatility measure. When there is overnight (futures) trading it is best to measure volatility by the sum of squared intranight and intraday returns.

For *forecasting* daily stock market volatility, explicitly modeling the volatility of the overnight return in a different way from the volatility of the intraday returns leads to the optimal forecasting performance. In the proposed "intraday" model, the overnight returns follow a separate GARCH(1,1) model, where the lagged sum of squared intraday returns

has a significant impact on overnight volatility. The intraday returns are also modeled as GARCH(1,1), with the first intraday volatility depending on the overnight squared return. This model performs better than the daily GARCH model extended with intraday information and continuous time GARCH. For the weekly and monthly horizons it is much less important to model overnight returns and intraday returns differently.

For further research it will be interesting to compare the new GARCH-type forecasts using intraday returns with alternative volatility forecasts, such as historical volatility or implied volatility from options. Such a comparison should be based on the volatility measures using intraday returns. In addition, with more and more GLOBEX data becoming available, it will be interesting to investigate whether the GARCH forecasts can be further improved using intranight returns.

BIBLIOGRAPHY

- Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39, 885–905.
- Andersen, T. G., Bollerslev, T., & Lange, S. (1999). Forecasting financial market volatility: Sampling frequency vis-à-vis forecast horizon. *Journal of Empirical Finance*, 6, 457–477.
- Bessembinder, H., & Seguin, P. J. (1993). Price volatility, trading volume, and market depth: Evidence from futures markets. *Journal of Financial and Quantitative Analysis*, 28, 21–39.
- Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *Review of Financial Studies*, 6, 659–681.
- Chan, K., Chan, K. C., & Karolyi, A. G. (1991). Intraday volatility in the stock index and stock index futures markets. *Review of Financial Studies*, 4, 637–684.
- Coppejans, M., & Domowitz, I. (1996). Automated trade execution and open outcry trading: A first look at the GLOBEX trading system. Working paper, Northwestern University.
- Day, T. E., & Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics*, 52, 267–287.
- Drost, F. C., & Nijman, T. E. (1993). Temporal aggregation of GARCH processes. *Econometrica*, 61, 909–927.
- Drost, F. C., & Werker, B.J.M. (1996). Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics*, 74, 31–57.
- Figlewski, S. (1997). Forecasting volatility. *Financial Markets, Institutions and Instruments*, 6, 1–88.
- Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53, 67–78.

- Laux, P. A., & Ng, L. K. (1993). The sources of GARCH: Empirical evidence from an intraday returns model incorporating systematic and unique risks. *Journal of International Money and Finance*, 12, 543–560.
- Lee, J. H., & Linn, S. C. (1994). Intraday and overnight volatility of stock index and stock index futures returns. *Review of Futures Markets*, 13, 1–30.
- Martens, M. (2001). Forecasting daily exchange rate volatility using intraday returns. *Journal of International Money and Finance*, 20, 1–23.
- Nelson, D. B. (1990). ARCH models as diffusion approximations. *Journal of Econometrics*, 45, 7–38.
- Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, 45, 267–290.
- Randolph, W. L., & Najand, M. (1991). A test of two models in forecasting stock index futures price volatility. *Journal of Futures Markets*, 11, 179–190.
- Stoll, H. R., & Whaley, R. E. (1990). The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative Analysis*, 25, 441–468.
- Taylor, S. J., & Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance*, 4, 317–340.

Copyright of Journal of Futures Markets is the property of John Wiley & Sons, Inc. / Business and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.