Realized Volatility of Stocks in Hong Kong

Chow, Y.F. [1], T.K. Lam [1] and H.S. Yeung [1]

[1] Department of Finance, The Chinese University of Hong Kong, Hong Kong
Email: yfchow@baf.msmail.cuhk.edu.hk

Keywords: *Equity markets, high-frequency data, realized volatility, correlation*

EXTENDED ABSTRACT

The joint distributional characteristics of asset returns, especially second moment structure (i.e., volatility), are essential to the analysis of financial issues such as asset pricing, portfolio allocation and risk management. Many traditional financial models have assumed constant volatilities and correlations among returns, but empirical evidence have shown them to vary significantly over time, and much effort has been put into the modeling and forecasting of their properties and structure.

Over time, the availability of data for increasingly shorter return horizons has enabled practitioners to narrow the frequencies of their volatility models. However, progress has stalled over the modeling of daily return volatilities, partly due to standard volatility models' general inability to handle intra-day data, while few high-dimensional models have been suggested for practical use as well. As for several notable multivariate autoregressive conditional heteroskedasticity (ARCH) and stochastic volatility models being proposed, namely Bollerslev, Engle and Nelson (1994) and Ghysels, Harvey and Renault (1996), have also been challenged for being unable to deal with more than a few assets simultaneously.

At the same time, others have relied on ad hoc, model-free approaches like simple exponential smoothing based on the counterfactual assumption of conditionally normally distributed returns, or through squared returns over the relevant return horizon, which in turn does not account for the presence of noise.

The limitations of the above approaches have led researchers to seek an alternative framework. Andersen, Bollerslev, Diebold and Ebens (2001) (henceforth ABDE) and Barndorff-Nielsen and Shephard (2002) both turned to a new "realized volatility" approach. In particular, ABDE proposed using continuously recorded transaction prices and by summing squares and cross-products of intra-day high-frequency returns, they constructed estimates known as ex post realized daily volatilities. Under the theory of quadratic variation, the realized daily volatility obtained is an unbiased and highly efficient estimator of return volatility. The advantages of such a measure are that the volatility measures are model-free, and are free from measurement error as the sampling frequency of the returns approaches infinity.

Under the realized volatility approach, ABDE (2001) focused on 30 stocks in the Dow Jones Industrial Average (DJIA). They also settled down on 5-minute intervals as the effective "continuous time record" so as to tackle microstructure frictions such as bid-ask bounce effects and price discreteness. Meanwhile, Andersen, Bollerslev, Diebold and Labys (2003) again adopted a similar approach to estimate exchange rate volatilities, while choosing 30-minute intervals.

Nevertheless, we observed that similar studies on high-frequency financial data have mainly concentrated on developed markets while relatively little has been done on emerging markets like Hong Kong. This paper is a preliminary attempt to study the volatility structure of shares listed on the Hong Kong Exchanges and Clearing Limited (HKEx) using transaction data. Here, we will follow ABDE (2001) approach on high-frequency financial data to observe the volatility structure of Hong Kong shares. Meanwhile, due to the need for high-frequency price observations, we would focus on the constituents stocks of HSI and HSCEI for all transactions between January 2004 and December 2005.

Using the sample of constituent stocks of Hang Seng Index (HSI) and Hang Seng China Enterprises Index (HSCEI or "H-shares Index"), we found that the mean daily realized volatilities of HSCEI to be significantly higher than its HSI counterpart, while the correlations between H-shares stay relatively lower than that of HSI stocks. A long-memory effect is also reported for the logarithmic standard deviations of all shares, with most of them showing slow decay over the series.

1. INTRODUCTION

Financial market volatility is central to the theory and practice of asset pricing, asset allocation, and risk management. The need for an appropriate framework of (conditional) variance of financial asset returns has led to the analysis of "realized volatility" using high frequency intraday data. See McAleer and Medeiros (2007) for an excellent review of the rapidly expanding literature on this.

To set out the basic idea and intuition of realized volatility, let us first consider the case of no microstructure frictions such as price discreteness, infrequent trading, and bid-ask bounce effects. Assume that the logarithmic $N \times 1$ vector price process, p_t, follows a multivariate continuous time stochastic volatility diffusion,

$$\mathrm{d}\,p_t = \mu_t\,\mathrm{d}t + \Omega_t\,\mathrm{d}W_t, \tag{1}$$

where W_t denotes a standard N-dimensional Brownian motion, the process for the $N \times N$ positive definite diffusion matrix, Ω_t, is strictly stationary, and we normalize the unit time interval, or $h = 1$, to represent one trading day. Conditional on the sample path realization of μ_t and Ω_t, the distribution of the continuously compounded h-period returns, $r_{t+h,h} = p_{t+h} - p_t$; is then

$$r_{t+h,h} \mid \sigma\{\mu_{t+\tau}, \Omega_{t+\tau}\}_{\tau=0}^{h} \sim N\left(\int_0^h \mu_{t+\tau}\,\mathrm{d}\tau, \int_0^h \Omega_{t+\tau}\,\mathrm{d}\tau\right), \tag{2}$$

where $\sigma\{\mu_{t+\tau}, \Omega_{t+\tau}\}_{\tau=0}^{h}$ denotes the σ-field generated by the sample paths of $\mu_{t+\tau}$ and $\Omega_{t+\tau}$ for $0 \leq \tau \leq h$. The integrated diffusion matrix thus provides a natural measure of the true latent h-period volatility. By the theory of quadratic variation, we have that under weak regularity conditions,

$$\sum_{j=1,\ldots,[h/\Delta]} r_{t+j\Delta,\Delta} r'_{t+j\Delta,\Delta} - \int_0^h \Omega_{t+\tau}\,\mathrm{d}\tau \to 0, \tag{3}$$

almost surely for all t as the sampling frequency of the returns increases, or $\Delta \to 0$. Thus, by summing sufficiently finely sampled high-frequency returns, it is possible to construct ex post realized volatility measures for the integrated latent volatilities that are asymptotically free of measurement error. ABDE (2001) obtained the realized daily covariance matrix as

$$Cov_t \equiv r_{t+j\Delta,\Delta} r'_{t+j\Delta,\Delta}, \tag{4}$$

with Δ referring to each incremental time interval, which is small enough to be considered continuous in the model.

2. DATA

Our empirical analysis is based on selected stocks in HSI and HKCEI with data from HKEx database. Specifically, we focus on 25 firms with highest 90-day average trading volume in each index to ensure a reasonable degree of liquidity.

It is known that intraday quotes and transaction prices are subject to various regulatory measures and bid-ask bounce effects. Such market microstructure characteristics potentially distort the distributional properties of intraday returns thereby distorting the results of statistical inference. Following ABDE (2001), we resort to artificially constructed five-minute returns which are believed to be a balance between reducing the effect of various market microstructures and finding appropriate discrete approximations of quadratic variations. In case there are no transactions in our predetermined time intervals, the price of the one immediately before the mark will be selected to replace the missing data.

With daily transactions from 10:00 HKT to 12:30 HKT and 14:30 HKT to 16:00 HKT, a total of 48 records can be collected each day, i.e., $\Delta = 1/48$. Our sampling starts from 2 January 2004 to 31 December 2005, with a total of 496 trading days and hence we have a total of 23,808 samples for each stock.

It should be noted, however, microstructure effects can introduce a severe bias on the daily volatility estimation rendering the approach along the lines of ABDE (2001) sub-optimal in comparison with the more recent approaches as in Barndorff-Nielsen, Hansen, Lund and Shephard (2006a, 2006b), Aït-Sahalia, Mykland and Zhang (2005, 2006) and Zhang, Mykland and Aït-Sahalia (2005).

3. UNIVARIATE UNCONDITIONAL RETURN AND VOLATILITY DISTRIBUTIONS

3.1. Returns

The summary statistics in Table 1 show that the daily return of HSI and HKCEI stocks, $r_{i,t}$, have fatter tails than the normal distribution and for majority of stocks, they are skewed to the right. In addition, they are in line with previous empirical evidence about the fat-tail characteristics of returns. The average values of skewness coefficients are 0.14/0.01 (HSI/HSCEI) for raw returns and 0.10/0.07 (HSI/HSCEI) for standardized returns, while the average values of excess kurtosis are 2.33/2.28 (HSI/HSCEI) for raw

Table 1. Unconditional daily return distributions.

Panel A: $r_{i,t}$				
HSI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.026	1.463	0.135	2.329
S.D.	0.106	0.446	0.313	1.273
Min.	–0.121	0.533	–0.596	0.232
Max.	0.412	2.411	0.885	4.906
HSCEI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	–0.004	2.213	0.005	2.275
S.D.	0.111	0.339	0.246	1.285
Min.	–0.242	1.439	–0.564	0.627
Max.	0.174	2.748	0.461	5.875
Panel B: $r_{i,t} / v_{i,t}$				
HSI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.005	0.702	0.104	0.488
S.D.	0.049	0.137	0.182	0.576
Min.	–0.137	0.492	–0.321	–0.369
Max.	0.121	0.920	0.365	1.857
HSCEI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.000	0.644	0.073	0.552
S.D.	0.042	0.122	0.132	0.494
Min.	–0.116	0.445	–0.270	–0.350
Max.	0.081	0.994	0.434	1.468

returns and 0.49/0.55 (HSI/HSCEI) for standardized returns.

Figure 1 plots the standardized returns for Esprit (0330.HK) as an example. It can be seen that normal distribution may be reasonable but not a very good description of the data. This result is in line with the leptokurtic distributions for standardized high-frequency returns implicitly assumed in ARCH and stochastic volatility model.

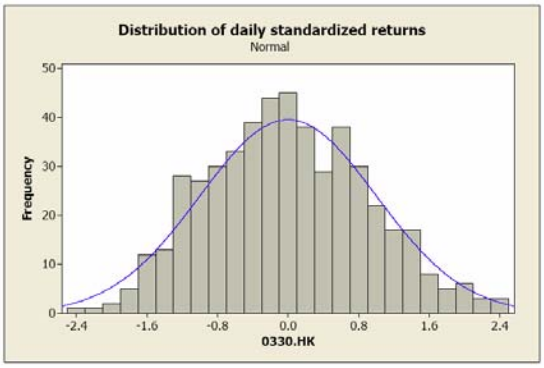

Figure 1. Unconditional distribution of the standardized daily returns on Esprit (0330.HK). The sample period extends from January 2, 2004 and December 30, 2005, for a total of 496 daily observations. The curve refers to the normal density.

3.2. Variances and Log Standard Deviations

Table 2 provides the same set of summary statistics for the unconditional realized daily return variances and logarithmic standard deviations. In Panel A, the averages of daily variance (column 1) are 4.81 and 11.65 for HSI and HSCEI stocks, which translate to an annualized volatility of 34.5% and 53.7%. The standard deviation of the realized volatility (column 2) indicates that the realized volatility of different assets vary significantly over time. The skewness (column 3) and excess kurtosis (column 4) of daily realized variances show that the distributions are highly skewed to the right and leptokurtic with average skewness of 3.33/3.05 (HSI/HSCEI) and average excess kurtosis of 25.6/28.47 (HSI/HSCEI) respectively. Such results show that normality assumption is not appropriate for return variances.

Panel B reports the summary statistics of logarithmic realized volatility, $lv_{i,t}$. It can be seen that the average values of the sample skewness are close to zero for the realized variances albeit being slightly negative. The excess kurtosis coefficients are significantly lower and hence the assumption

Table 2. Unconditional volatility distributions.

Panel A: $v_{i,t}^2$				
HSI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	4.810	3.329	3.334	25.647
S.D.	3.323	2.045	2.125	34.962
Min.	1.078	0.677	0.859	0.308
Max.	12.632	7.860	10.108	158.828
HSCEI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	11.655	7.506	3.053	28.470
S.D.	2.712	2.839	3.457	73.735
Min.	4.912	4.007	0.702	0.521
Max.	17.237	18.674	18.010	370.773
Panel B: $lv_{i,t}$				
HSI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.582	0.318	–0.043	0.936
S.D.	0.331	0.080	0.467	1.089
Min.	–0.055	0.199	–0.980	–0.869
Max.	1.227	0.586	0.702	3.353
HSCEI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	1.125	0.290	–0.299	1.276
S.D.	0.170	0.060	0.456	0.933
Min.	0.659	0.189	–1.076	–0.134
Max.	1.374	0.471	0.559	3.353

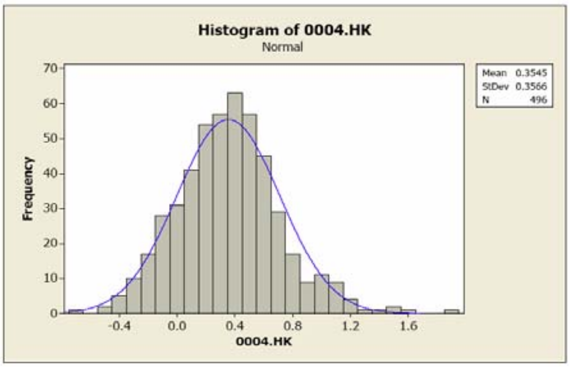

Figure 2. Unconditional distribution of the standardized daily realized logarithmic standard deviations for Wharf Holdings (0004.HK). The realized volatilities are calculated from five-minute intraday returns. The curve refers to the normal density.

of normality may be more appropriate in this case.

3.3. Covariances and Correlations

Table 3 reports the summary statistics of the covariances and correlations between log realized volatilities. The averages of covariances and correlations across all stocks are 0.41/0.81 (HSI/HSCEI) and 0.09/0.06 (HSI/HSCEI) with substantial variations over time indicated by their standard deviations (column 2). The skewness and excess kurtosis of realized covariances and correlations are reported in columns 3 and 4 respectively. The average skewness and excess kurtosis for realized covariances are 2.2/1.4 (HSI/HSCEI) and 15.0/9.0 (HSI/HSCEI), compared to almost zeros for realized correlations. It seems that the distributions of the realized covariances are right skewed, but the correlations

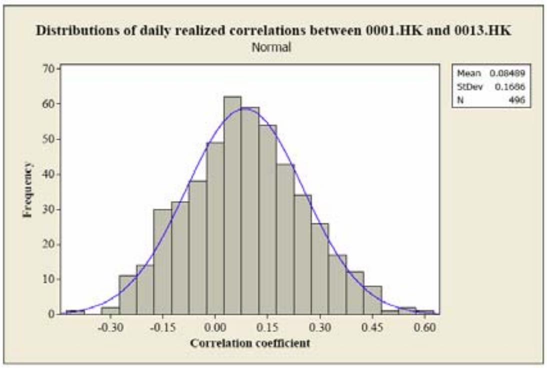

Figure 3. Unconditional distribution of the standardized realized daily correlations between Cheung Kong Holdings (0001.HK) and Hutchison Whampoa (0013.HK). The realized correlations are calculated from five-minute intraday returns. The curve refers to the normal density.

Table 3. Unconditional covariances and correlations distributions.

Panel A: $Cov_{i,j,t}$				
HSI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.411	0.946	2.197	15.030
S.D.	0.193	0.379	1.437	18.445
Min.	0.049	0.275	–0.286	0.756
Max.	1.091	2.380	8.595	123.234
HSCEI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.807	2.342	1.413	9.045
S.D.	0.354	0.530	1.119	10.137
Min.	0.210	1.204	–4.165	0.822
Max.	2.288	4.863	5.894	80.562
Panel B: $Corr_{i,j,t}$				
HSI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.093	0.177	–0.003	–0.025
S.D.	0.045	0.010	0.118	0.240
Min.	0.031	0.154	–0.390	–0.630
Max.	0.309	0.236	0.288	1.057
HSCEI	Mean	S.D.	Skew.	Ex. Kurt.
Avg.	0.059	0.172	0.024	–0.002
S.D.	0.020	0.008	0.108	0.255
Min.	0.020	0.149	–0.281	–0.591
Max.	0.158	0.196	0.332	0.853

appear to be normally distributed.

Figure 3 shows the unconditional realized correlation between Cheung Kong Holdings (0001.HK) and Hutchison Whampoa (0013.HK), the two largest stocks in HSI in terms of market capitalization. It appears that normal distribution can serve as a good approximation to the daily realized correlations.

4. DYNAMIC DEPENDENCE OF VOLATILITIES AND CORRELATIONS

After exploring the unconditional distributions of the return generating process, we extend our analysis to the conditional distributions of the volatility processes. Previous studies have shown the existence of long-term dependence in volatility. Here we focus on logarithmic volatilities and correlations since both can be approximated well by normal distributions, with results in Table 4.

Table 4. Dynamic volatility dependence.

Panel A: $lv_{i,t}$			
HIS	Q_{22}	ADF	d_{GPH}
Avg.	1820.540	−1.947	0.584
S.D.	1614.710	0.675	0.143
Min.	301.310	−3.025	0.340
Max.	7007.940	−0.402	0.912
HSCEI	Q_{22}	ADF	d_{GPH}
Avg.	953.975	−2.351	0.559
S.D.	762.417	0.631	0.116
Min.	80.408	−3.207	0.302
Max.	2840.620	−0.620	0.769
Panel B: $\text{Corr}_{i,j,t}$			
HIS	Q_{22}	ADF	d_{GPH}
Avg.	67.600	−3.627	0.233
S.D.	176.480	0.764	0.176
Min.	12.690	−6.047	−0.357
Max.	1999.320	−1.304	0.812
HSCEI	Q_{22}	ADF	d_{GPH}
Avg.	64.815	−3.490	0.275
S.D.	74.803	0.777	0.168
Min.	11.350	−5.666	−0.183
Max.	568.424	−1.406	0.743

4.1. Logarithmic Standard Deviations

In Figure 4, we present the time-series plot of the logarithmic standard deviations of Cheung Kong Holdings (0001.HK), which shows possible positive autocorrelation in Figure 5.

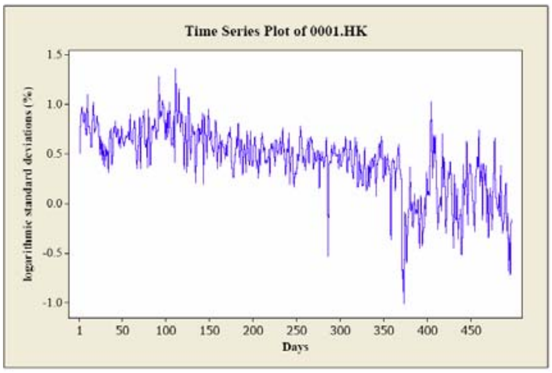

Figure 4. Time Series Plot for the logarithmic standard deviation of Cheung Kong Holdings (0001.HK). The figure shows the time series plot of the daily realized logarithmic standard deviations for Cheung Kong Holdings (0001.HK). The realized volatilities are calculated from five-minute intraday returns.

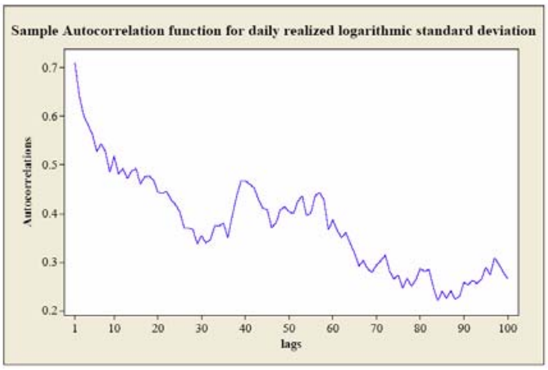

Figure 5. Sample autocorrelations for daily logarithmic standard deviations. The figure shows the sample autocorrelations for the daily realized logarithmic standard deviations for Cheung Kong Holdings (0001.HK), $lv_{0001.HK,t} \equiv \log(v_{0001.HK,t})$, out to a displacement of 100 days. The realized volatilities are calculated from five-minute intraday returns.

We report the values of standard Ljung-Box portmanteau test for the joint significance of the first 22 autocorrelations of $lv_{i,t}$ in column 1 of Panel A. The null hypothesis is significantly rejected for all stocks. In fact, the autocorrelations are systematically above the conventional 95% confidence band, even at a long horizon of 120 trading days (approximately half year). This slow delay has led us to perform augmented Dickey-Fuller tests to detect the existence of unit roots in each of the volatility series. The results are reported in column 2, which shows that not all stocks can reject the null hypothesis at 5% level.

With such findings, we extend the analysis to examine the potential long-memory effect modeled by a FIGARCH model as in Baillie, Bollerslev and Mikkelsen (1996). Using the approach developed in Geweke and Porter–Hudak (1983), the degree of fractional integration for the realized logarithmic volatilities, denoted as d_{GPH}, are shown in the third column in Panel A. The average values are 0.58/0.56 (HSI/HSCEI) with respective standard deviations of 0.14/0.12 (HSI/HSCEI). Therefore, the results seem to suggest the existence of long-memory effects in logarithmic realized standard deviations

4.2. Correlations

Panel B of Table 4 reports the results on the temporal behavior of the daily realized correlations $Corr_{i,j,t}$. Compared to the analysis in section 5.1, the null hypothesis of no autocorrelation is not rejected for quite a number of stocks using Ljung-Box portmanteau test. Similarly, we test for the existence of unit root by augmented Dickey-Fuller test, and we found that again quite a number of correlation pairs cannot reject the null hypothesis. We then fit a FIGARCH model to investigate the long memory effect in the processes. The averages of d_{GPH} are 0.23/0.27 (HSI/HSCEI) with standard deviations of 0.18/0.17 (HSI/HSCEI). Thus, the stationarity assumption for correlations does not seem to hold for some elements in the correlation matrix, although the evidence is much weaker than logarithmic standard deviations.

Figure 6 shows the time-series plot of the realized correlations between Cheung Kong Holdings (0001.HK) and Hutchison Whampoa (0013.HK). Despite the variability, there does not seem to have strong evidence of possible autocorrelation.

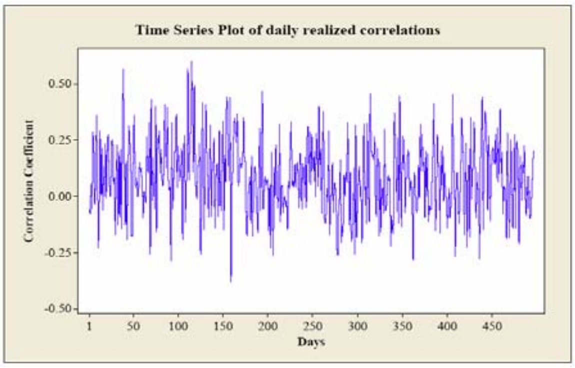

Figure 6. Time series of daily correlations. The figure shows the time series of daily realized logarithmic standard deviations between Cheung Kong Holdings (0001.HK) and Hutchison Whampoa (0013.HK). The sample period is between January 2, 2004 and December 30, 2005, for a total of 496 daily observations. The realized correlations are calculated from five-minute intraday returns.

5. CONCLUSION

In this paper, we analyzed the high-frequency transaction data for selected stocks in HSI and HSCEI over 2004–2005 based on ABDE (2001) approach. The univariate unconditional distributions of daily returns, variances, logarithmic standard deviations, covariances and correlations are examined. It is found that the mean daily realized volatilities of H-shares to be significantly higher than their HSI counterparts, while the correlations between H-shares stay relatively lower than that of HSI stocks. We also investigated the possible long-memory effect in realized logarithmic standard deviations and correlations through standard time series tests and FIGARCH model. It is found that realized daily logarithmic standard deviations demonstrated strong evidence of long-memory, while the long-memory property seems to be weaker in correlations for large portion of stocks in our study.

As mentioned in McAleer and Medeiros (2007), however, it should be noted that microstructure noise could cause severe problems in terms of consistent estimation of the daily realized volatility estimator along the lines of ABDE (2001). Therefore, further investigation on the volatility structures in Hong Kong stock market will be pursued in near future.

6. REFERENCES

Aït-Sahalia, Y., P.A. Mykland and L. Zhang (2005), How often to sample a continuous time process in the presence of market microstructure noise, *Review of Financial Studies* 18, 351-416.

Aït-Sahalia, Y., P.A. Mykland and L. Zhang (2006), Ultra high frequency volatility estimation with dependent microstructure noise, working paper 11380, NBER.

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001), The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003), Modeling and forecasting realized volatility, *Econometrica* 71, 579-625.

Baillie, R.T, T. Bollerslev and H.O. Mikkelsen (1996), Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 74, 3-30.

Barndorff-Nielsen, O.E., P.H. Hansen, A. Lund and N. Shephard (2006a), Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise, unpublished manuscript, Stanford University.

Barndorff-Nielsen, O.E., P.H. Hansen, A. Lund and N. Shephard (2006b), Subsampling realized kernels, unpublished manuscript, Stanford University.

Barndorff-Nielsen, O.E. and N. Shephard (2002), Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society B* 64, 253-280.

Bollerslev, T., R.F. Engle and D.B. Nelson (1994), ARCH models, in R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics* 4, Amsterdam: North-Holland.

Geweke, J. and S. Porter-Hudak (1983), The estimation and application of long memory time-series models, *Journal of Time Series Analysis* 4, 221-238.

Ghysels, E., A. Harvey and E. Renault (1996), Stochastic volatility, in G.S. Maddala and C.R. Rao (eds.), *Handbook of Statistics* 14, Amsterdam, North-Holland.

McAleer, M. and M.C. Medeiros (2007), Realized volatility: A review, *Econometric Reviews*, to appear.

Zhang, L., P.A. Mykland and Y. Aït-Sahalia (2005), A tale of two time scales: Determining integrated volatility with noisy high frequency data, *Journal of the American Statistical Association* 100, 1394-1411.