The realized volatility of FTSE-100 futures prices

Nelson M.P.C. Areal[*]
areal@exchange.lancs.ac.uk

Stephen J. Taylor[†]
S.Taylor@lancaster.ac.uk

Department of Accounting and Finance
Lancaster University
England LA1 4YX

September 2000

Abstract

Five-minute returns from FTSE-100 index futures contracts are used to obtain accurate estimates of daily index volatility from January 1986 to December 1998. These realized volatility measures are used to obtain inferences about the distributional and autocorrelation properties of FTSE-100 volatility. The distribution of volatility measured daily is similar to lognormal whilst the volatility time series has persistent positive autocorrelation that displays long-memory effects. The distribution of daily returns standardized using the measures of realized volatility is shown to be close to normal unlike the unconditional distribution.

[*]This work was supported by the Portuguese Foundation for Science and Technology. Thanks are also due to our colleagues and to participants at meetings of the European Finance Association and the Portuguese Finance Network for helpful comments, particularly to Charlotte Christiansen, João Duque, Mark Shackleton and Gary Xu.

[†]Corresponding author. Phone: +44 1524 593624, Fax: +44 1524 847321.

The realized volatility of FTSE-100 futures prices

1 Introduction

The distribution of asset returns has often been modeled by mixtures of normal distributions that have different volatility parameters. The distributional and autocorrelation properties of volatility are then important to traders when pricing options because different specifications will give different theoretical prices. This paper uses high-frequency returns from FTSE-100 futures contracts to deduce the distribution and the autocorrelations of FTSE-100 volatility.

The influential empirical study of cotton futures prices and trading volumes by Clark (1973) is the foundation for two conjectures: first that an appropriate distribution for daily volatility is lognormal and second that daily returns are conditionally normal given the level of daily volatility. The unconditional distribution of returns is then a leptokurtic mixture distribution that has fat tails relative to the normal distribution. Tauchen and Pitts (1983) provided further theoretical analysis and results for T-Bill future prices. Taylor (1986) extended Clark's model by proposing that the logarithm of volatility follows a first-order autoregressive process, based upon analysis of the autocorrelations of squared daily returns from futures contracts and other assets. However, empirical investigation of Clark's conjectures using daily returns has limited potential to provide decisive conclusions because daily volatility is then an unobservable latent variable.

Andersen, Bollerslev, Diebold and their co-authors have recently used high-frequency returns to obtain accurate estimates of daily volatility. They show that Clark's two distributional conjectures are close to the empirical distributions, both for spot foreign exchange returns (Andersen, Bollerslev, Diebold, and Labys, 2000a,b) and for returns from the Dow Jones Industrial Average (DJIA) index (Ebens, 1999) and its constituent stocks (Andersen, Bollerslev, Diebold, and Ebens, 2000). Furthermore, there is strong empirical evidence for long memory structure in the volatility process that is generally consistent with the stochastic volatility specifications of Breidt, Crato, and De Lima (1998).

This paper refines the methodology of Andersen, Bollerslev, Diebold, and Labys (2000a,b) and applies it to a futures market for the first time. It uses thirteen years of five-minute returns from FTSE-100 index futures contracts to estimate daily volatility and hence describe the distribution and the time series properties of FTSE-100 volatility. The estimates of daily volatility are more accurate than the simple estimates used in previous papers and, furthermore, they take account of the relatively substantial changes in the index during the hours that the futures market is closed. Section 2 describes the FTSE-100 transaction data from which five-minute returns are defined. These are used in Section 3 to estimate the intraday volatility pattern that can be used to improve the accuracy of estimates of realized volatility. These estimates are defined in Section 4, followed by discussion of their distribution and time series properties. The distribution of daily returns standardized by daily realized volatility is described in Section 5 followed by conclusions in Section 6.

2 Data

2.1 Transaction prices

FTSE-100 volatility has been calculated from the record of floor transaction prices for FTSE-100 futures contracts sold on compact disk (CD) by the London International Financial Futures Exchange (LIFFE). Their CDs also include electronic transaction prices for approximately an hour after floor trading stops that are not used in this study; neither are bid and ask floor prices used. Each line in the LIFFE data files provides a time recorded to an accuracy of one second and also a trading volume when the price refers to a transaction.

The longest time period considered is from 2 January 1986 until 29 December 1998. Some days that were not holidays are missing, particularly in the earlier years. A total of 31 days are missing of which 22 are in 1987 and 1988. The CD that we used contains 2.97 million transactions of which 2.85 million were floor transactions. Trading volume was relatively low in the initial years and has increased from 0.12 million contracts in 1986 to 4.59 million contracts in 1998.

At any time most trading activity is concentrated in a single contract, usually nearest to delivery. Returns are calculated from the nearest contract, except on the days before expiry and that include expiry when the next contract is used. Hence returns are nearly always calculated from the prices of the contract that has the highest trading volume.

The record of transactions has been edited in a number of ways. Days that only have transactions for part of the normal trading day are excluded entirely; 28 days are excluded, most being on December 24 and December 31. All floor transactions that are timed outside floor trading hours and all transactions that are recorded as having zero volume are excluded; 5,494 transactions are deleted. On 17 days in 1990 the transaction times are recorded without minutes and seconds. For these days, transaction times from previous days were used to produce approximations to the missing information.

More problematic are a small number of misrecorded transaction prices. On some occasions it is clear from a sequence of three prices that the middle price is almost certainly wrong, for example the sequence 2600, 3600, 2600 is indicative of a mistake in the first digit. We have identified 56 suspicious prices using simple outlier filters and inspection of time series plots of returns and prices. All these transactions have been deleted, the majority being before 1989. Although we are confident that the deletions are mistakes there may be mistakes that have escaped detection.

2.2 Intraday returns

All results have been obtained from five-minute returns, as this frequency is generally acknowledged to be the highest that avoids distortions from microstructure effects such as the bid-ask spread. Two methods have been used in previous work to define returns and both are evaluated here. The first uses the latest price before each five-minute mark whilst the second interpolates between the latest price before and the consecutive price after the five-minute mark. The first method is used here unless stated otherwise.

The number of five-minute returns computed for a day depends on the trading hours for futures contracts and these have changed occasionally. Trading was from 09:35 to 15:35 for the short period from 2 January until 24 April 1986, from 09:05 to 16:05 for the four years from 28 April 1986 until 23 March 1990, from 08:35 to 16:10 for the eight years from 26 March

1990 to 17 July 1998 and from 08:35 to 16:30 for the remaining few months until December 1998. Hence there are 72, 84, 91 or 95 five-minute returns for a day depending on the trading hours.

Table 1 summarizes the statistical properties of the $285,960$ five-minute returns in the complete time series (1986-1998) and the $188,825$ five-minute returns during the eight-year period (1990-1998) when trading was from 08:35 to 16:10. It can be seen first that the skewness and the kurtosis are much higher for the longer period, because of the crash in October 1987, and second that the frequency of zero returns is much higher for the longer period because trading volumes were relatively low before 1990. For these reasons we now discuss the results obtained for the shorter period. The average returns are small and are not significantly different from zero at the 20% level. The distribution of the returns is almost symmetric and has fat tails and a substantial peak at zero. Although 22.3% of the five-minute returns calculated without interpolating prices are zero, only 2.65% of the five-minute intervals contain no transactions.[1] A few returns are substantial with extreme values of -1.7% on 5 December 1997 and 3.0% on 5 October 1990.

The autocorrelations of the returns are very close to zero. Bid-ask bounce might be expected to create negative dependence between consecutive returns but there is no evidence for such an effect in Table 1. The first autocorrelation of five-minute returns calculated from interpolated prices is 0.0356 compared with 0.0013 when the latest prices are used. Hence we prefer to use the latest prices when defining returns and we apply this definition throughout the remainder of the paper. The autocorrelations of absolute returns are positive and significant at many lags, commencing with 0.2345 for consecutive absolute returns.

[1]For interpolated prices the proportion of zero returns with no trades is 1.19%.

3 Intraday volatility

3.1 Intraday volatility multipliers

The FTSE-100 index is more volatile when the market opens, when macroeconomic news is released in the UK and the US and when the American equity markets are open. Intraday volatility is modeled by assuming there is a fixed multiplicative effect, that may vary across the days of the week. Let $r_{t,j}$, $0 \leq j \leq n$, represent a set of $n+1$ intraday returns for day t, such that $j = 0$ represents the closed-market period from the close on day $t-1$ until the open on day t, $j = 1$ represents the five minutes commencing at the open, ..., concluding with $j = n$ representing the final five minutes that end when the floor market closes. Let the daily return r_t and the latent level of volatility σ_t for a day t be defined by

$$r_t = \sum_{j=0}^{n} r_{t,j} \quad \text{and} \quad var(r_t|\sigma_t) = \sigma_t^2.$$

Multiplicative volatility terms are defined by supposing that

$$var(r_{t,j}|\sigma_t) = \lambda_j \sigma_t^2 \quad \text{with} \quad \sum_{j=0}^{n} \lambda_j = 1.$$

Thus λ_j is the proportion of a trading's day total return variance that is attributed to period j, here assuming that intraday returns are uncorrelated and, to simplify notation, that the multipliers are the same for all days t. The proportion of the open-market variance is defined by

$$\kappa_j = \frac{\lambda_j}{1-\lambda_0} \quad \text{with} \quad \sum_{j=1}^{n} \kappa_j = 1.$$

3.2 Estimates

Simple estimates of the variance proportions, following Taylor and Xu (1997), are given by,

$$\hat{\lambda}_j = \frac{\sum_t r_{t,j}^2}{\sum_t \sum_{k=0}^{n} r_{t,k}^2} \quad \textit{and} \quad \hat{\kappa}_j = \frac{\sum_t r_{t,j}^2}{\sum_t \sum_{k=1}^{n} r_{t,k}^2}$$

where the summations over days t are for some appropriate set of days S. The set S might be all days, or all Mondays, etc. [2] We give most attention to the eight-year period from 26 March 1990 to 17 July 1998 during which time n was constant and equal to 91. Although the number of five-minute returns is constant we divided this period in two, from 26/03/1990 to 17/11/1993 and from 18/11/1993 to 17/07/1998, when we estimated the intradaily volatility multipliers, due to changes in the release times of UK macroeconomic announcements from 11:30 to 09:30.[3]

In the remainder of Section 3 we discuss the period from November 1993 to July 1998. The closed-market proportion λ_0 is substantial, particularly when S is restricted to Mondays. For all days $\hat{\lambda}_0$ is 30.7%, increasing to 38.3% when the closed-market period is from Friday's close to Monday's open.

[2] Sets S are restricted to those days for which UK local time is five hours ahead of Eastern US local time, i.e. we exclude a few days every year when local times differ by four hours.

[3] The Central Statistical Office changed their announcement times on 23/08/1993, the Bank of England changed on 03/09/1993, and the labour market statistics started to be released at 09:30 from 18/11/1993.

Figure 1 shows the estimates $\hat{\kappa}_j$, calculated by day of the week. It is immediately apparent that volatility is high at the open with $\hat{\kappa}_1$ almost 8% on Mondays and between 4% and 6% on other days. Volatility then declines until the interval 09:30 to 09:35 when announcements are often made about important UK macroeconomic indicators followed by a generally steady decline until a sharp increase in the interval from 13:30 to 13:35 that often includes major US announcements. Volatility falls after these announcements and subsequently increases for the remainder of the London trading hours, with a local peak from 14:15 to 14:20 on some days (again reflecting US news) and an end-of-week spike from 16:05 to 16:10 on Fridays. Tse (1999) has provided similar figures for a much shorter period, without distinguishing between days of the week. The impact of US macroeconomic news released at 08:30 local time, documented in detail by Ederington and Lee (1993) for US futures contracts, is clearly also important for the UK equity market and usually occurs at 13:30 in London.

Andersen and Bollerslev (1997) use regression methodology to produce smooth variance multipliers. Their Flexible Fourier Functions (FFF) have been used to produce smoothed estimates $\tilde{\lambda}_j$ and $\tilde{\kappa}_j$ as the fitted values when the multipliers $\hat{\lambda}_j$ and $\hat{\kappa}_j$ are regressed on a constant, linear, quadratic, sinusoidal and dummy variables; we use trigonometric functions at 12 frequencies and dummy variables for the five-minute intervals ending at 09:35 or 11:35, 13:35 and 14:20 and we rescale the fitted values to ensure their sums are unity. [4] Figure 2 illustrates the FFF multipliers $\tilde{\kappa}_j$ when they are estimated ignoring day-of-the-week effects. The smooth multipliers $\tilde{\kappa}_j$ are very similar to the simple estimates $\hat{\kappa}_j$.

[4] Slightly different multipliers are used on the few days that the UK is four hours ahead of the Eastern US, when the two US news dummy variables are each moved by one hour.

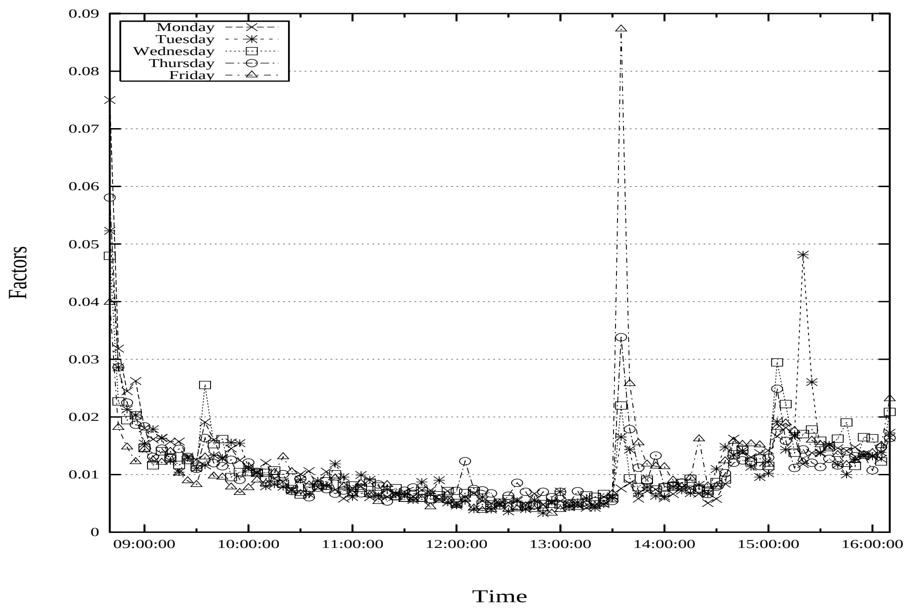

Figure 1: Five-minute open-market variance proportions for the FTSE-100 futures index, by day of the week, for the period from 18/11/1993 to 17/07/1998.

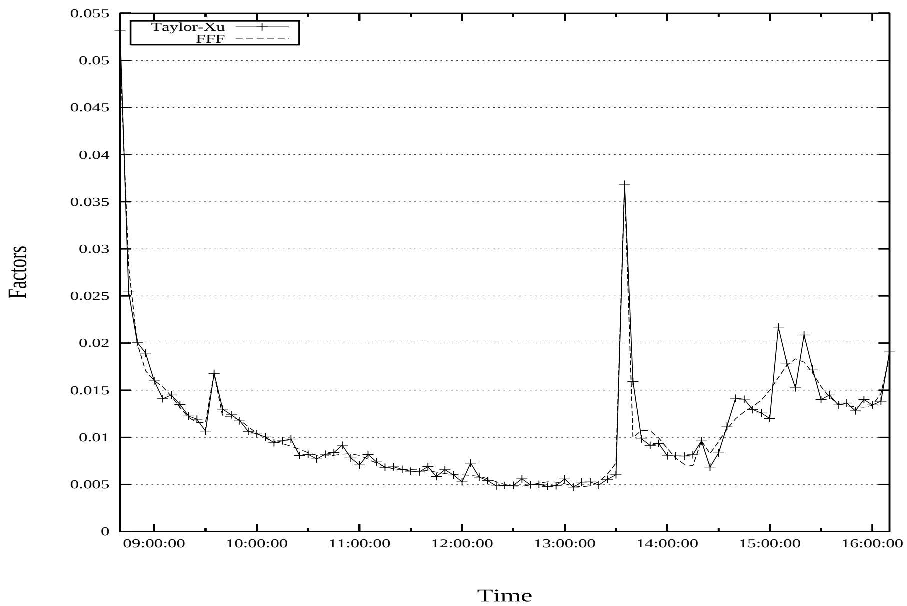

Figure 2: Five-minute fitted open-market variance proportions for the FTSE-100 futures index, using all days of the week, for the period from 18/11/1993 to 17/07/1998.

4 Realized volatility

4.1 Computational methods

The realized variance for trading day t, from the close on day $t-1$ to the close on day t, is estimated by weighting the intraday squared returns, as follows [5]

$$\hat{\sigma}_t^2 = \sum_{j=0}^{n} w_j r_{t,j}^2.$$

To ensure conditionally unbiased estimates when intraday returns are uncorrelated, so that $E[\hat{\sigma}_t^2|\sigma_t^2] = \sigma_t^2$, it is necessary to apply the constraint $\sum_{j=0}^{n} \lambda_j w_j = 1$. Taylor and Xu (1997), Andersen, Bollerslev, Diebold, and Labys (2000a) and related papers simply use $w_j = 1$ for all intraday periods j. The estimate $\hat{\sigma}_t^2$ is then a consistent and unbiased estimate of "integrated" volatility [6] but it will not have the least variance when n is finite. Consequently, we also consider other weights. It is shown in the Appendix that if $\sum_{j=0}^{n} \lambda_j w_j = 1$ then the variance of the estimate $\hat{\sigma}_t^2$ is minimized when

$$w_j = \frac{1}{(n+1)\lambda_j}.$$

In particular, the optimal weight w_0 for the closed market return is much less than for the other returns because λ_0 is substantial. When w_0 is con-

[5]Returns are not mean adjusted because the difference between $E[r_{t,j}^2]$ and $var(r_{t,j})$ is negligible.

[6]When σ follows a diffusion process, the realized volatility $\sum_{j=0}^{n} r_{t,j}^2$ converges to the integrated volatility, $\int_{t-1}^{t} \sigma^2(s)ds$, as $n \to \infty$ when innocuous regularity conditions are assumed (Andersen, Bollerslev, Diebold, and Labys, 2000a).

strained to be zero, the optimal weights become

$$
\begin{aligned}
w_j &= \frac{1}{(1-\lambda_0)n\kappa_j}, \quad 1 \le j \le n \\
&= 0, \qquad\qquad\qquad j = 0
\end{aligned}
$$

Realized variances have been obtained using several sets of weights, in particular results are reported when either

$$w_j = 1, \ \ 0 \le j \le n$$

or

$$
\begin{aligned}
w_j &= \frac{1}{(1-\hat{\lambda}_0)n\tilde{\kappa}_j}, \quad 1 \le j \le n \\
&= 0, \qquad\qquad\qquad j = 0
\end{aligned}
$$

with separate multipliers estimated for each day of the week. These sets of weights are referred to respectively as equal weights and optimal weights.

4.2 Summary statistics

Table 2 summarizes the distribution of daily realized volatility obtained using five-minute returns from March 1990 to July 1998, for both equal weights and optimal weights. The daily numbers have been annualized by multiplying $\hat{\sigma}_t$ by $\sqrt{251}$. The average annual standard deviation equals 14.2% using equal weights and 15.1% using optimal weights. The average value of $251\hat{\sigma}_t^2$ equals $(16.2\%)^2$ using either set of weights. This is close to the annualized variance of daily returns, $(15.5\%)^2$, hence any bias caused by autocorrelation among intraday returns is small. Unless stated otherwise the following remarks apply to the annualized numbers calculated using the optimal weights.

There are several outliers, the most extreme being on 28 October 1997 when annualized volatility equals 152% using equal weights and 81% using optimal weights; it is noteworthy that US volatility was exceptionally high on both 27 and 28 October 1997 (Ebens, 1999). FTSE-100 volatility was high around these dates and exceeded 30% on 23, 28, 29, 30 and 31 October. Three of the highest estimates, all above 50%, are for the day that Sterling left the Exchange Rate Mechanism (ERM) of the European Monetary System and the two following days (16, 17 and 18 September 1992). Also, of the 20 estimates that exceed 40%, 11 are in the period from August to December 1998.

The distribution of annualized volatility is skewed to the right and highly leptokurtic with a sample kurtosis of 19 for the optimal weights. Consequently, Table 2 also presents summary statistics for the logarithms of daily realized volatility, $\ln(\hat{\sigma}_t)$. It can be deduced from the summary statistics for $\ln(\hat{\sigma}_t)$ that $\hat{\sigma}_t$ calculated from the optimal weights is a more accurate estimate of σ_t than $\hat{\sigma}_t$ calculated from equal weights.[7] Figures 3 and 4 are plots of the annualized time series $\sqrt{251}\hat{\sigma}_t$ for the longer time period from January 1986 to December 1998, respectively with a linear scale and a logarithmic scale for the volatility numbers. [8] The exceptionally high volatility during the 1987 crash is clearly visible. [9]

[7]Let $\hat{\sigma}_t^2 = \sigma_t^2(1+u_t)$ with u_t the zero-mean measurement error. Then $\ln(\hat{\sigma}_t^2) = \ln(\sigma_t^2) + u_t - 1/2u_t^2 + \ldots$ and hence a more accurate estimate has a higher value of $E[\ln(\hat{\sigma}_t^2)]$ and a lower value of $var[\ln(\hat{\sigma}_t^2)]$.

[8]Four sets of weights are used, depending on the hours that futures are traded as defined in Section 2.2.

[9]Realized volatility peaked at 365% on Tuesday 20 October 1987 and was also above 100% on 19, 21, 23, and 27 October.

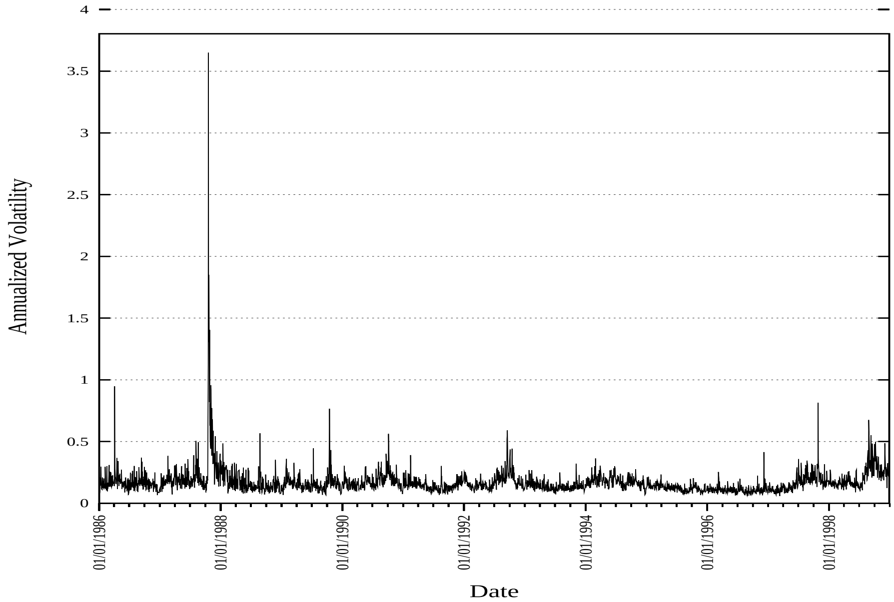

Figure 3: Annualized FTSE-100 volatility calculated each day from optimally weighted squared five minute returns.

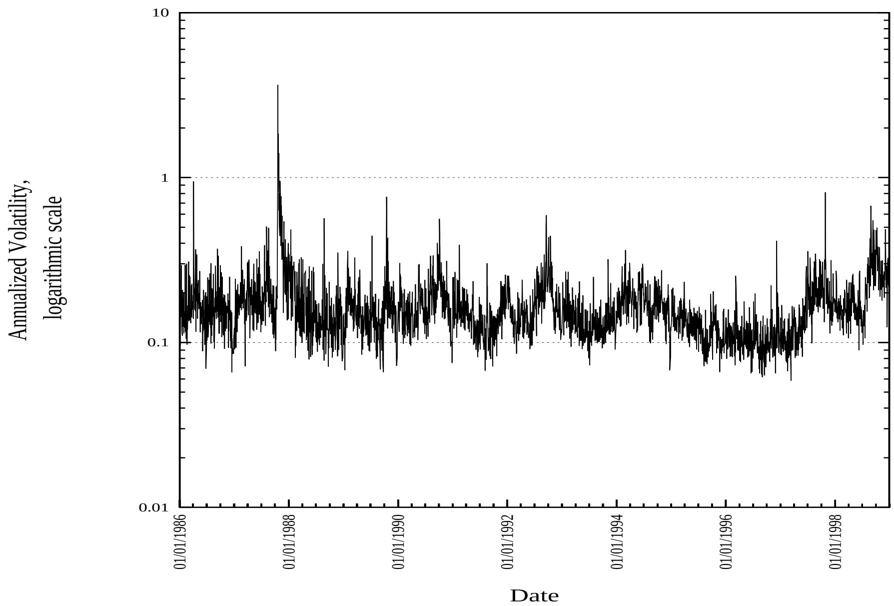

Figure 4: Annualized FTSE-100 volatility calculated each day from optimally weighted squared five minute returns.

4.3 The distribution of daily volatility

The distribution of $\ln(\hat{\sigma}_t)$ from March 1990 to July 1998 is almost symmetric and approximately Gaussian when the optimal weights are used, the skewness and kurtosis then being 0.44 and 3.71 respectively; these moments are however significantly different from the Gaussian values of 0 and 3, the standard errors of the estimates being 0.054 and 0.108 respectively. The kurtosis is much higher at 5.96 when equal weights are used, although this falls to 4.22 if the overnight period is excluded to give estimates of open-market volatility. Figure 5 shows the empirical distribution of $\ln(\hat{\sigma}_t)$ when the optimal weights are used. The continuous curves overlaying the histogram are firstly a normal density that matches the mean and standard deviation (dotted curve) and secondly the density estimate based upon Gaussian kernels and bandwidth equal to 0.1 (solid curve); the kernel estimate is more smooth than that provided by the standard bandwidth of 0.06 for this data (Silverman, 1986, page 48).

There are several possible explanations for the excess kurtosis in the distribution of $\ln(\hat{\sigma}_t)$. First, there is the extreme outlier, second $\ln(\hat{\sigma}_t)$ is an accurate but nevertheless imperfect estimate of $\ln(\sigma_t)$ and third it is, of course, possible that the distribution of $\ln(\sigma_t)$ has excess probability in the right tail relative to the normal distribution, possibly reflecting occasional financial crises.

The skewness and kurtosis estimates are similar to those reported for the open-market period in the US. Our open-market figures of 0.55 and 4.22 for equal weights and 0.46 and 3.81 for optimal weights can be compared with 0.75 and 3.78 for the DJIA index from 1993 to 1998 (Ebens, 1999), and medians of 0.19 and 3.89 for the 30 DJIA stocks during the same period (Andersen, Bollerslev, Diebold, and Ebens, 2000). For the 24-hour spot FX

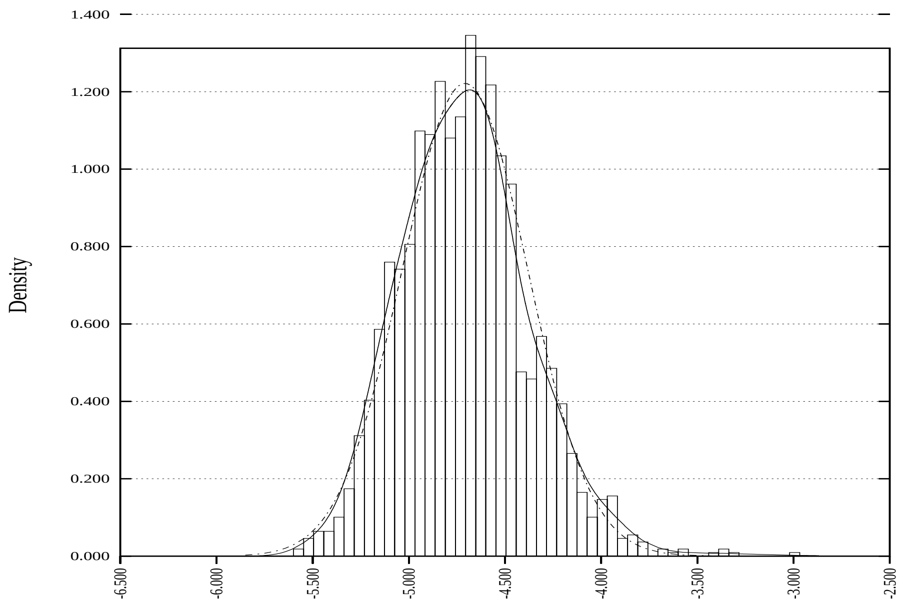

Figure 5: The distribution of the logarithm of realized volatility for the FTSE-100 index from 1990 to 1998.

market the estimates of Andersen, Bollerslev, Diebold, and Labys (2000a) are 0.35 and 3.27 for the DM/$ rate and 0.28 and 3.47 for the Yen/$ rate.

4.4 Temporal dependence

Figure 6 shows the autocorrelations of the time series of daily realized volatility $\ln(\hat{\sigma}_t)$ from March 1990 to July 1998, calculated using the optimal weights. Positive dependence is observed for 180 lags, or about nine months. There is a clear although minor seasonal effect based upon the five days of the week that shows itself on the figure as the local peaks at lags 5, 10, 15, 20 and 25.

The maximum autocorrelation on Figure 6 is at the first lag and equals 0.65. This value is equal to the first-lag autocorrelation of approximately 0.65 reported by both Ebens (1999) for the DJIA and by Andersen, Bollerslev, Diebold, and Labys (2000a) for FX.

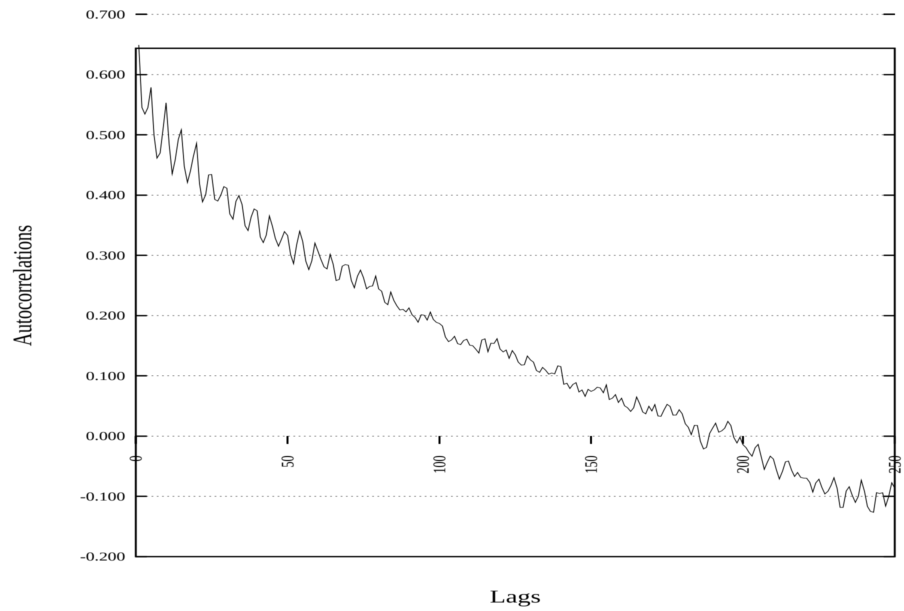

Figure 6: Autocorrelations of the logarithm of realized volatility for the FTSE-100 index from March 1990 to July 1998.

The augmented Dickey-Fuller test provides highly significant evidence that the realized volatility process does not contain a unit root. Hence we consider $I(d)$ processes with $d < 1$. The slow decline in the autocorrelations of the realized volatility series suggests a long memory process, as reviewed by Baillie (1996). Two standard methods provide conclusive evidence that a long memory component exists in the volatility series. These methods have been applied to the series $ln(\hat{\sigma}_t)$ obtained from the optimal weights, adjusted by removing the appropriate day-of-the-week mean $m_{j(t)}$, $j(t) \epsilon \{1, \ldots, 5\}$, from each observation.

First, consider the variance S_T of the sum of T consecutive observations. These variances follow a scaling law for long memory processes such that

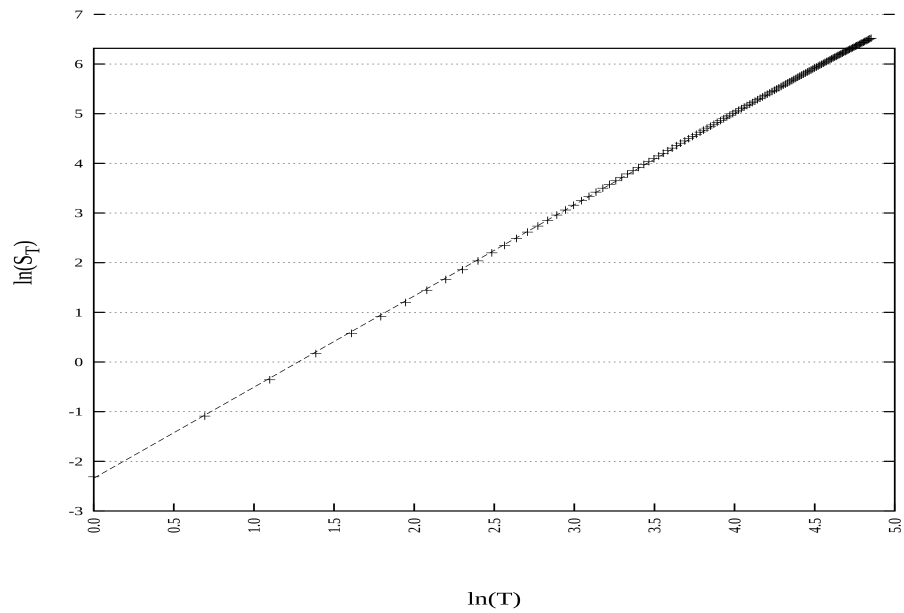

Figure 7: Scaling plot for daily logarithmic realized volatility, S_T is the variance of T consecutive observations.

$$T^{-(2d+1)}S_T \rightarrow \text{ constant, as } T \rightarrow \infty, \text{ with } d > 0,$$

whilst short memory processes have $d = 0$. Figure 7 shows that $ln(S_T)$ is essentially a linear function of $ln(T)$ over the range $1 \leq T \leq 128$. From the slope of this function, d is estimated to be 0.42.

Second, consider the spectral density $f(\omega)$ of the process that generates the observations. For a long memory process,

$$\omega^{2d}f(\omega) \rightarrow \text{ constant, as } \omega \rightarrow 0.$$

The Geweke-Porter-Hudak (GPH) estimate of d is provided by regressing

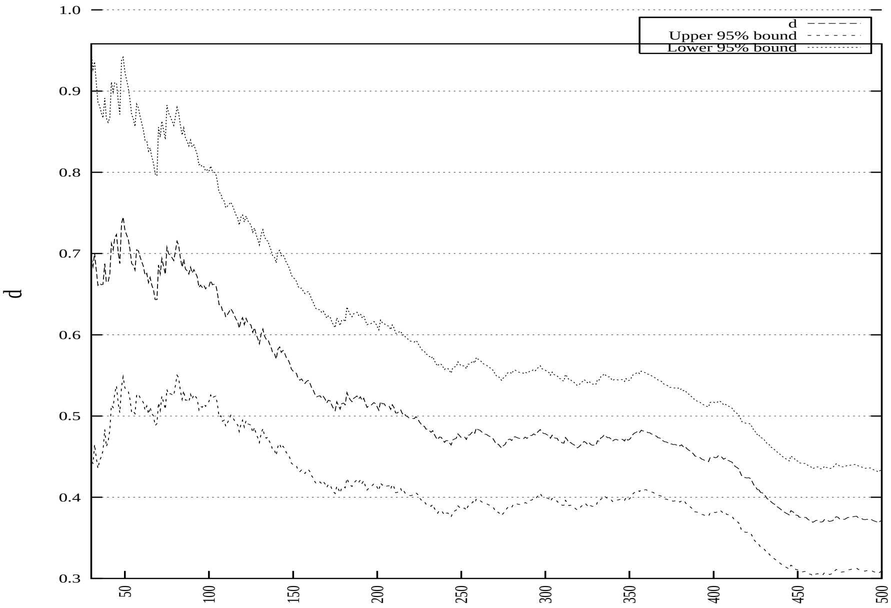

Figure 8: GPH estimates of the degree of fractional integration, d, as a function of the number of periodogram ordinates, n^θ, used in their calculations.

the logarithm of the periodogram estimate of the spectral density against $\ln(\omega)$ over a range of frequencies ω (Geweke and Porter-Hudak, 1983). We use frequencies $\frac{2\pi j}{n}$, with $j = 1, 2, \ldots, n^\theta$, $\theta = 0.8$ and $n = 2075$ observations, to obtain an estimated d equal to 0.43 with a standard error of 0.031. Figure 8 shows the estimated degree of fractional integration d as a function of the number of periodogram ordinates n^θ, and also displays 95% confidence intervals for d.

The estimates of d for the FTSE data are similar to those reported in Ebens (1999) and by Andersen, Bollerslev, Diebold, and Labys (2000a). Both these studies estimate d to be between 0.35 and 0.45.

A simple long memory filter $(1-L)^{0.43}$ explains almost all of the dependence in the realized volatility. The autocorrelations of

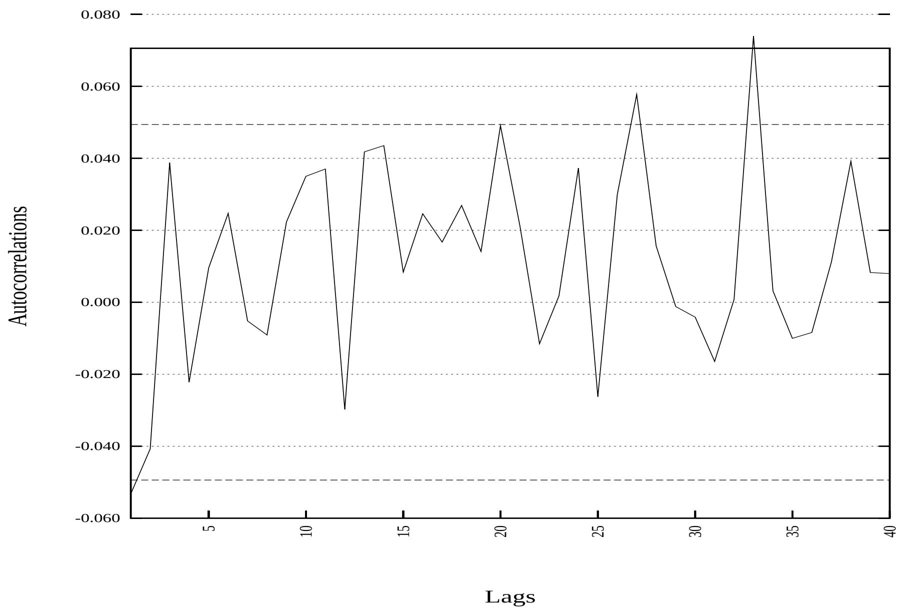

Figure 9: Autocorrelations of $(1-L)^{0.43}[ln(\hat{\sigma}_t)-m_{j(t)}]$.

$$(1-L)^{0.43}[ln(\hat{\sigma}_t)-m_{j(t)}]$$

are near zero, as shown on Figure 9. Table 3 contains some descriptive statistics for the filtered series.

5 The distribution of standardized daily returns

Table 4 summarizes the distribution and autocorrelations of standardized daily returns $r_t^* = \frac{(r_t-\bar{r})}{\hat{\sigma}_t}$ from March 1990 to July 1998, with $\hat{\sigma}_t$ calculated using either equal weights or the optimal weights. The autocorrelations of the r_t^* are very small and are not significantly different from zero. Table 4 also provides the same summary statistics for the daily returns r_t. It is noted that the average daily return is 0.000483 or about 12% per annum. As the average five-minute return is negative (approximately -3% per annum, see Table 1) all of the gains from long futures positions during the sample period occurred when the market was closed.

The kurtosis of the daily returns is 4.81 so that they are not normally distributed. The kurtosis of the standardized daily returns is however 2.77 for the optimal weights, so that standardizing the returns brings the distribution much nearer to the normal. Although near to normal, both the skewness and kurtosis are significantly different from the normal values of 0 and 3 at the 2% significance level. The kurtosis figure of 2.77 is similar to the 2.75 of Ebens (1999) for the DJIA index, less than the median of 3.13 for the DJIA stocks (Andersen, Bollerslev, Diebold, and Ebens, 2000) and more than the 2.41 for FX given by Andersen, Bollerslev, Diebold, and Labys (2000b).

Figure 10 shows the empirical distribution of r_t^* when the optimal weights are used. The continuous curves overlaying the histogram are once more the matched normal density (dotted curve) and the kernel density estimate with bandwidth equal to 0.25 (solid curve). It can be seen that some of the non-normality is simply a consequence of zero returns that reflect the minimum tick equal to 0.5 index points; 37 of the 2075 daily returns are zero.

6 Conclusions

This study of five-minute returns from FTSE-100 futures provides several conclusions about FTSE-100 volatility measured at the daily frequency. Neither the distribution of the logarithm of volatility nor that of returns standardized by realized volatility is exactly normal. A lognormal distribution for

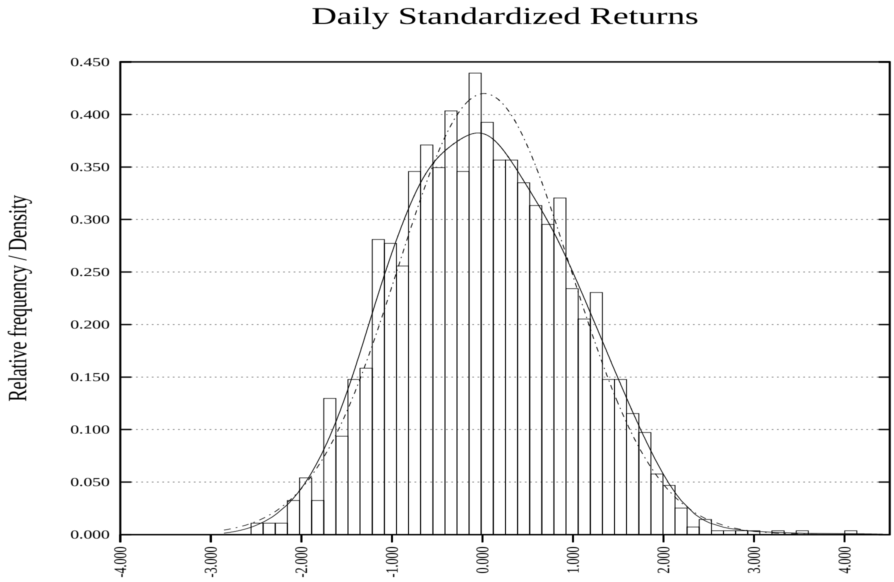

Figure 10: The distribution of the daily standardized returns of the FTSE-100 index futures from March 1990 to July 1998.

volatility is not predicted by any theory. It is nevertheless interesting that the distribution is near to lognormal with the main discrepancy being an excess probability of extremely high levels of volatility. A normal distribution for standardized returns might be expected from the central limit theorem. The divergence of the empirical distribution from the normal is minor and may simply reflect the fact that futures prices are a discrete process. These distributional conclusions are the first reported for the volatility of futures prices using high frequency-data and are also the first that take account of substantial price changes when the market is closed.

The time series behavior of FTSE-100 realized volatility is best described by a long memory process. The logarithms of daily volatility display substantial autocorrelations that do not decrease rapidly towards zero as the time lag increases, as occurs for short memory processes such as ARMA models.

Instead, fractional differencing of daily volatility provides a filtered series that is almost uncorrelated so that a simple long memory process provides a parsimonious model for volatility.

The conclusions about the distribution and the autocorrelation structure of volatility have implications for researchers, traders and regulators. The shape of the unconditional distribution of volatility restricts the diffusion models for volatility that are credible when options are priced with volatility assumed to be stochastic. The temporal behavior of volatility has immediate implications for the term structure of implied volatility that need to be researched. Volatility forecasts can be derived by filtering realized volatility using the fractional differencing filter and their information content can be compared with that of implied volatilities that are known to forecast well in comparisons with short memory ARCH forecasts. The relatively high frequency of very high levels of volatility, and the consequent possibilities of extreme returns, are quantified by the methods in this paper and thus should be useful to risk managers.

Appendix

Note to referees : this appendix can be deleted if the proof is considered to be either too easy or too technical for the journal, in which case the appendix can be made available as a web document.

The optimal weights presented in Section 4.1 are obtained by assuming that squared intraday returns $r_{t,j}^2$ are the product of the multiplier λ_j, the unobservable daily variance σ_t^2 and a residual term $e_{t,j}$ such that

$$r_{t,j}^2 = \lambda_j \sigma_t^2 e_{t,j}$$

with the $e_{t,j}$ independent and identically distributed, with mean 1 and variance v. Then the weights w_j are to be chosen to define

$$\hat{\sigma}_t^2 = \sum_{j=0}^{n} w_j r_{t,j}^2$$

so that both

$$E[\hat{\sigma}_t^2 | \sigma_t^2] = \sigma_t^2, \ i.e. \ \sum_{j=0}^{n} \lambda_j w_j = 1,$$

and the conditional variance of the estimator is minimized, namely

$$\begin{aligned} var(\hat{\sigma}_t^2 | \sigma_t^2) &= \sigma_t^4 var(\sum_{j=0}^{n} \lambda_j w_j e_{t,j}) \\ &= \sigma_t^4 v \sum_{j=0}^{n} \lambda_j^2 w_j^2 \end{aligned}$$

Introducing a Lagrange multiplier θ and defining

$$F = \sum_{j=0}^{n} \lambda_j^2 w_j^2 + \theta(\sum_{j=0}^{n} \lambda_j w_j - 1)$$

then

$$\frac{\partial F}{\partial w_j} = 2w_j\lambda_j^2 + \theta\lambda_j$$

and hence all these partial derivatives can only be zero when $\lambda_j w_j$ is the same for all intervals j, i.e. when

$$w_j = \frac{1}{(n+1)\lambda_j}.$$

References

Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance 4*, 115–158.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2000). The distribution of stock return volatility. *Journal of Financial Economics*, forthcoming.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2000a). The distribution of exchange rate volatility. *Journal of the American Statistical Association*, forthcoming.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2000b). Exchange rate returns standardized by realized volatility are (nearly) gaussian. Working Paper, Department of Finance, Northwestern University.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics 73*, 5–59.

Breidt, F. J., N. Crato, and P. De Lima (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics 83*, 325–348.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica 41*, 135–155.

Ebens, H. (1999). Realized stock volatility. Working paper, Department of Economics, John Hopkins University.

Ederington, L. H. and J. H. Lee (1993). How markets process information: news releases and volatility. *Journal of Finance 49*, 1161–1191.

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.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis.* Chapman and Hall.

Tauchen, G. E. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica 51*, 485–505.

Taylor, S. J. (1986). *Modelling financial time series.* John Wiley. Chichester.

Taylor, S. J. and X. Xu (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance 4*, 317–340.

Tse, Y. (1999). Market microstructure of the FT-SE 100 index futures: an intraday empirical analysis. *Journal of Futures Markets 19*, 31–58.

	86-98		90-98	
	Latest	**Interpolated**	**Latest**	**Interpolated**
Average	-0.00000132	-0.00000137	-0.00000111	-0.00000111
Standard Deviation	0.00103	0.00098	0.00086	0.00082
Skewness	-2.9510	-3.2526	0.2302	0.4817
Kurtosis	462.0078	541.0832	25.2162	40.5652
Minimum	-0.0896 20/10/1987	-0.0914 20/10/1987	-0.0172 05/12/1997	-0.0168 05/12/1997
Maximum	0.0577 20/10/1987	0.0595 20/10/1987	0.0303 05/10/1990	0.0374 05/10/1990
Number of zeros	81 565	43 626	42 135	16 263
Number of observations	285 960	285 960	188 825	188 825
Autocorrelations				
- Returns				
Lag 1	-0.0290	0.0080	0.0013	0.0356
Lag 2	0.0097	0.0168	-0.0175	-0.0142
Lag 3	-0.0282	-0.0302	-0.0159	-0.0165
Lag 4	-0.0128	-0.0119	-0.0085	-0.0090
Lag 5	-0.0214	-0.0231	-0.0037	-0.0039
- Absolute returns				
Lag 1	0.3264	0.3572	0.2345	0.2582
Lag 2	0.2687	0.2973	0.2046	0.2248
Lag 3	0.2721	0.2943	0.1863	0.2022
Lag 4	0.2517	0.2685	0.1695	0.1823
Lag 5	0.2277	0.2344	0.1620	0.1748

Returns are calculated either from the latest prices before five-minute marks or after interpolating between prices before and after the marks. The complete time-series is from January 1986 until December 1988. The shorter time-series is from March 1990 to July 1998.

Table 1: Summary statistics for five-minute FTSE-100 returns.

	Equal weights		Optimal weights	
	Annual Std. Dev.	$ln(\sigma_t)$	**Annual Std. Dev.**	$ln(\sigma_t)$
Average	0.1423	-4.7914	0.1512	-4.7081
Standard Deviation	0.0749	0.3665	0.0562	0.3266
Skewness	6.1626	1.0352	2.4747	0.4362
Kurtosis	78.4180	5.9581	18.5100	3.7122
Minimum	0.0529 02/09/1996	-5.7019 02/09/1996	0.0586 12/03/1997	-5.9999 12/03/1997
Maximum	1.5151 28/10/1997	-2.3473 28/10/1997	0.8145 28/10/1997	-2.9679 28/10/1997
Number of observations	2075	2075	2075	2075
Augmented Dickey-Fuller Test	-7.3283	-5.5959	-5.5364	-4.9600
Autocorrelations				
Lag 1	0.4468	0.5656	0.6109	0.6493
Lag 2	0.3446	0.4876	0.4977	0.5455
Lag 3	0.3404	0.4797	0.4764	0.5345
Lag 4	0.2908	0.4388	0.4598	0.5451
Lag 5	0.2488	0.4511	0.4774	0.5787

Squared five minute returns and the squared closed-market return are weighted and aggregated to define realized daily variances. The weights are either all equal or are optimal choices that minimize the variances of the volatility estimates. Summary statistics are tabulated for both annualized realized volatility (assuming 251 trading days in a year) and the logarithms of daily realized volatility. Asterisks indicate test statistics significant at the 1% level.

Table 2: The distribution of realized volatility for FTSE-100 futures from March 1990 to July 1998.

	Filtered series
Average	0.000072
Standard Deviation	0.205131
Skewness	0.8194
Kurtosis	5.5550
Minimum	-0.7999 23/12/1994
Maximum	1.3226 28/10/1997
Number of observations	2075
Autocorrelations	
Lag 1	-0.0532 (0.0173)
Lag 2	-0.0407 (0.0530)
Lag 3	0.0388 (0.0620)
Lag 4	-0.0222 (0.1892)
Lag 5	0.0096 (0.3522)
Ljung-Box(20)	30.4895 (0.0623)

The volatility series was filtered using the fractional integration filter $(1-L)^{0.43}$. Asterisks indicate test statistics significant at the 1% level. The numbers in parentheses are the levels of significance of the autocorrelation statistics.

Table 3: Descriptive statistics for the series defined by filtering the logarithms of daily realized volatility.

		Standardized daily returns	
	Daily Returns	Equal weights	Optimal weights
Average	0.0005	-0.0035	0.0199
	(0.0122)	(0.4375)	(0.1701)
Standard Deviation	0.0098	1.0003	0.9500
Skewness	0.1614*	0.0064	0.1573*
	(0.0013)	(0.4526)	(0.0017)
Kurtosis	4.8084*	2.4692*	2.7729
	(0.0000)	(0.0000)	(0.0174)
Minimum	-0.0449	-2.9536	-2.5551
	05/10/1992	17/07/1992	02/01/1991
Maximum	0.0496	3.1770	4.1357
	10/04/1992	01/07/1997	30/12/1991
Tests for normality			
χ^2, 50 bins	56.8900	85.3429*	64.5063
	(0.1529)	(0.0005)	(0.0458)
Jarque-Bera	291.7512*	24.3731*	13.0145*
	(0.0000)	(0.0000)	(0.0015)
Kolmogorov-Smirnov	1.1460*	1.3032*	1.1706*
Anderson-Darling	3.5376 *	2.9670*	1.9494*
Number of observations	2075	2075	2075
Autocorrelations			
- Returns			
Lag 1	0.0112	-0.0076	-0.0079
	(0.3041)	(0.3640)	(0.3589)
Lag 2	-0.0114	-0.0171	-0.0293
	(0.3018)	(0.2174)	(0.0913)
Lag 3	-0.0214	-0.0297	-0.0322
	(0.1648)	(0.0878)	(0.0710)
Lag 4	-0.0169	-0.0172	-0.0172
	(0.2202)	(0.2169)	(0.2163)
Lag 5	-0.0333	-0.0162	-0.0306
	(0.0648)	(0.2307)	(0.0814)
Ljung-Box(20)	25.8996	21.6526	22.8192
	(0.1692)	(0.3597)	(0.2978)

Daily futures returns from March 1990 to July 1998 are standardized by dividing mean-adjusted returns by daily realized volatilities. Realized volatilities are the square roots of realized variances, defined by aggregating weighted, squared five-minute returns and the squared closed-market return. The weights are either all equal or are the optimal choices that minimize the measurement error of the realized variances. Asterisks indicate test statistics that are significant at the 1% level. The numbers in parentheses are levels of significance of the test statistics.

Table 4: The distributions of FTSE-100 futures daily returns and daily standardized returns.