Forecasting Volatility Using A Continuous Time Model

Maria Helena Lopes Moreira da Veiga*
Universitat Autònoma de Barcelona,
Department of Economics, Edifici B, 08193 Bellaterra, Spain.
Universidade do Porto,
Faculdade de Economia, R. Dr‘ Roberto Frias, 4200 Porto, Portugal.

September 4, 2003

Abstract

This paper evaluates the forecasting performance of a continuous stochastic volatility model with two factors of volatility (SV2F) and compares it to those of GARCH and ARFIMA models. The empirical results show that the volatility forecasting ability of the SV2F model is better than that of the GARCH and ARFIMA models, especially when volatility seems to change pattern. We use ex-post volatility as a proxy of the realized volatility obtained from intraday data and the forecasts from the SV2F are calculated using the reprojection technique proposed by Gallant and Tauchen (1998).

Key*words*: Efficient Method of Moments (EMM), Reprojection, Factors of Volatility, Fractional Integration, Volatility Forecasting.
JEL classification: C10; G13

1 Introduction

Volatility plays an important role for asset pricing theory as it is directly linked to the risk-return relation. Good measures and forecasts of future volatility are of vital importance for finance theory[1]. One measure of volatility has been "the implied volatility" obtained from Black-Scholes model, but as empirical evidence has been showing, the performance of this model is not the same in all periods. In fact, after the October 1987 stock market crash, the model is not adjusting reality so well and the implied volatility seems to have little predictive

*Email contact: mlopes@idea.uab.es. Grant Sponsors: FCT, Programa Praxis XXI and FEP (Faculdade de Economia do Porto).

[1]Notice that volatility is used as a measure of risk. The higger is the volatility, the higger is the risk and consequently the higger is going to be the expected return.

power relative to the historical volatility, Canina and Figlewski (1993). On the other hand, the standard well known ARCH/GARCH models seem to perform poorly in forecasting volatility[2].

Nowadays with the availability of high-frequency data and the advances of computational tools it is possible to improve volatility forecasting. There are two ways of forecasting volatility. The first way treats volatility as observed and uses intra-period data to get the so called "realized volatility" obtained basically by summing intra-period squared returns and fitting to them models that incorporate the main features of this data, for instance long-memory property. Since much theoretical work assumes that the logarithm of asset prices follow a continuous time model, like a diffusion, one advantage of this procedure is that the realized volatility can be made arbitrarily close to the underlying integrated volatility[3] by reducing the intra-period. However, if we choose tick-by-tick prices we might have to use some forms of interpolation since these prices are not generally available at all equally spaced time. Empirical evidence shows that it can cause negative correlation in the returns series and consequently grow poor the forecasts made by models using this data. Other problems are related with volatility change of pattern due to market microstructures, for instance the existence of lunch periods, the close of the market, etc.. In order to try avoiding these drawbacks, in this paper, we use 15-minutes intraday data. The second way treats volatility as latent in the sense that it can be filtered after estimation. Getting a correct specification is pivotal since volatility estimates are model dependent. In this paper, we follow this second alternative and we fit the continuous time model with two factors of volatility of Gallant and Tauchen (2001) for the return of a Microsoft share.

To sum up, the aim of this paper is to evaluate the volatility forecasting performance of the continuous time stochastic volatility model comparatively to the ones obtained with the traditional GARCH and ARFIMA models. In order to inquire into this, we estimate using the Efficient Method of Moments (EMM) of Gallant and Tauchen (1996) a continuous time stochastic volatility model for the logarithm of asset price and we filter the underlying volatility using the reprojection technique of Gallant and Tauchen (1998). Under the assumption that the model is correctly specified, we obtain a consistent estimator of the integrated volatility by fitting a continuous time stochastic volatility model to the data. The forecasting evaluation for the three estimated models is going to be done with the help of the R^{*2} of the individual regressions of realized volatility[4] on the volatility forecasts obtained from the estimated models. The empirical results indicate the better performance of the continuous time model in the out-of-sample periods compared to the ones of the traditional GARCH and ARFIMA models. Further, these two last models show difficulties in tracking the growth pattern of the realized volatility. This probably is due to the change

[2]See Day and Lewis (1992), Jorian (1996), Pagan and Schwert (1990), West and Cho (1995), etc.

[3]The integral of instantaneous volatility over the period.

[4]Since realized volatility is considered a good measure of volatility. R^{*2} is the corrected R^2 of Anderson and Bollerslev (2002).

of pattern in volatility in this last part of the sample.

The plan of the paper is as follows. Section 2 introduces the concept of realized volatility and the way to calculate it. Section 3 presents the continuous time model and the estimation results. Section 4 evaluates the forecasting performance of the three estimated specifications and Section 5 concludes the paper.

2 Realized Volatility

2.1 Theoretical relation between realized volatility and integrated volatility

Let $r_{t,j}$, $0 \leqslant j \leqslant n$, represent a set of $n+1$ intraday returns for day t. $j = 0$ refers to the closed market period that ranges from day $t-1$ until the open on day t. $j = 1$ is the fifteen minutes commencing at the open and $j = n$ is the last fifteen minutes return before market closes.

Much theoretical work models the logarithm of asset prices (p_k) as a univariate diffusion, Bollerslev, Diebold and Labys (2000),

$$dp_k = \mu_k dt + \sigma_k dW$$

where W is a Wiener process. So the daily return of asset k is given by

$$p_k(t) - p_k(t-1) \equiv r_k(t) = \int_{t-1}^{t} \mu_k(s)ds + \int_{t-1}^{t} \sigma_k(s)dW(s). \tag{1}$$

These authors proved that under innocuous regularity conditions, the realized volatility $\sum_{j=0}^{n} r_{t,j}^2$ converges to the integrated volatility $\int_{t-1}^{t} \sigma_k^2(s)dW(s)$ as n converges to ∞. So, the performance of the realized volatility estimator depends only on the number of observations. For a given sample period the higher the frequency of the data and the larger the number of observations, the better the approximation of the realized volatility estimator to the integrated volatility.

2.2 Data

Realized volatility has been calculated from the intraday 15-minutes price of a share of Microsoft[5], from 10^{th} of April, 1997, till 23^{rd} of February, 2001, according to Nelson and Taylor (2000).

[5]The data was obtain freely from Price-data.com

The models, GARCH, ARFIMA and the continuous time model with two factors of volatility described below use daily data on Microsoft[6] from 13^{th} of March, 1986, till 23^{rd} of February, 2001[7], for 3,778 observations.

2.2.1 Calculating the realized volatility

In this paper the realized variance for the trading day t[8] is calculated as a weighted average of the intraday squared returns. Accordingly to Nelson and Taylor (2000) it is given by

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

and we must impose the constraint $\sum_{j=0}^{n} \lambda_j w_j = 1$ in order to ensure conditionally unbiased estimates when intraday returns are uncorrelated. λ_j represents the proportion of a trading's day total return variance that is attributed to period j. They assume that the $\lambda's$ are equal for all days t. In order that $\hat{\sigma}_t^2$ is a consistent, unbiased and efficient (with the least variance) estimate of the integrated volatility, Nelson and Taylor (2000) deduced that

$$w_j = \frac{1}{(n+1)\lambda_j}. \tag{3}$$

In particular, because the weight w_0 for the closed market return is much less than for the other returns (because λ_0 is very big) they specify w_j as

$$\begin{aligned} w_j &= \frac{1}{(1-\lambda_0)nk_j}, & 1 \leqslant j \leqslant n \\ w_j &= 0 & j = 0, \end{aligned}$$

where k_j is the proportion of the open-market variance given by

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

Natural estimates of these variance proportions are:

[6] Adjusted for stock splits.
[7] See figure 1.
[8] The period ranges from the close on day t-1 to the the close on day t.

$$\hat{\lambda_j} = \frac{\sum_t r_{t,j}^2}{\sum_t \sum_{i=0}^{n} r_{t,i}^2} \quad \text{and} \quad \hat{k_j} = \frac{\sum_t r_{t,j}^2}{\sum_t \sum_{i=1}^{n} r_{t,i}^2},$$

where the sums over days t can be for all days or for particular days[9].

3 The model

Recently researchers tend to model volatility as stochastic. The literature is vast referring to the estimation of models with or without stochastic volatility or with or without jumps, see Bates (2000), Gallant et al. (2001), Gallant and Tauchen (2001), Ghysels et al. (1995), etc..

This paper estimates the stochastic volatility model with two factors of volatility[10] given by:

$$\frac{dP_t}{P_t} = \alpha_{10}dt + \exp(\beta_{10} + \beta_{12}U_{2t} + \beta_{13}U_{3t})dW_{1t} \tag{4}$$

$$dU_{2t} = (\alpha_{20} + \alpha_{22}U_{2t})dt + dW_{2t} \tag{5}$$

$$dU_{3t} = (\alpha_{30} + \alpha_{33}U_{3t})dt + dW_{3t} \tag{6}$$

where P_t is the daily value of a share of Microsoft and W_i with $i = 1, 2$, are Wiener processes.

In this system the instantaneous standard deviation of the rate of return is an exponential function of the factors U_{2t} and U_{3t}. The drifts in equations 5 and 6 allow for mean reversion when $\alpha_{22} \neq 0$ and $\alpha_{33} \neq 0$. Small values of α_{22} and α_{33} mean that a shock to the volatility of the return takes time to dissipate. This is referred in the financial econometrics literature as persistence, and a big percentage of the financial series seem to show this feature, Zaffaroni (2000). β_{10} is also an important parameter since it takes care of the long-run mean of the volatility of the price equation 4.

[9] See figure 3. For more about these estimates, see Taylor and Xu (1997).

[10] Gallant and Tauchen (2001) already estimated this model for a subsample of the data used in this paper.

3.0.2 Identification restrictions

To achieve identification it is necessary to impose some restrictions. In this concrete case for the logarithmic specification we set

$$\alpha_{20} = 0,\ \alpha_{30} = 0.$$

Hence the previous specification becomes:

$$\frac{dP_t}{P_t} = \alpha_{10}dt + \exp(\beta_{10} + \beta_{12}U_{2t} + \beta_{13}U_{3t})dW_{1t} \tag{7}$$

$$dU_{2t} = \alpha_{22}U_{2t}dt + dW_{2t} \tag{8}$$

$$dU_{3t} = \alpha_{33}U_{3t}dt + dW_{3t} \tag{9}$$

We use these restrictions, like in Gallant and Tauchen (2001), first because they are common in previous similar SDE and second because they provide flexibility and numerical stability in the estimation phase.

3.1 The EMM

The model above is estimated using the Efficient Method of Moments (EMM).

Let $\{y_t\}_{t=-\infty}^{\infty}$, $y_t \in \Re^M$, be a multiple, discrete stationary time series and $x_t = (y_{t-L}, ..., y_t)$ a stretch from the previous process with density $p(y_{-L}, ..., y_0|\rho)$ defined over $\Re^l, l = M(L+1)$. ρ is a vector of unknown parameters and $\{\tilde{y}_t\}_{t=-L}^{n}$ the real data from which it is to be estimated. The main problem that makes traditional methods of estimation inviable is that this density is in general not available. However, expectations of the forms

$$E_\rho(g) = \int ... \int g(y_{-L}, ..., y_0)p(y_{-L}, ..., y_0)dy_{-L}....dy_0$$

can be approximated quite well by averaging over a long simulation

$$E_\rho(g) = \frac{1}{N}\sum_{t=1}^{N} g(\hat{y}_{t-L}, ...\hat{y}_{t-1},\ \hat{y}_t).$$

Let $\{\hat{y}_t\}_{t=-L}^{N}$ denote the simulation from $p(y/\ x,\rho)$, where $x = x_{-1} = (y_{-L},...,y_{-1})$, $y = y_0$ and $p(y/\ x,\rho) = p(y_{-L},...,y_0|\rho)/p(y_{-L},...,y_{-1}/\rho)$. Notice that the length of simulation should be large enough for the Monte Carlo error to be negligible.

Gallant and Tauchen (1996) proposed an estimator for the vector of parameters ρ in the situation above. This method relies on a minimum chi-square estimator for the vector of parameters, which permits the optimized chi-square criterion to be used to test the specification adopted. The moment conditions entering the minimum chi-square criterion come from the score vector $\frac{\partial}{\partial\theta}\log f(y_t/x_{t-1},\theta)$ of an auxiliary model $f(y_t/x_{t-1},\theta)$ that closely approximates the true density. If this is true, the EMM estimator will be nearly as efficient as the ML estimator. One commonly used auxiliary model in applications is the SNP density $f_K(y/x,\theta)$ that was proposed by Gallant and Nychka (1987). It has been showed that the efficiency of the EMM estimator can be as close as the efficiency of the ML estimator by making K large enough, Gallant and Long (1995).

3.1.1 Projection step

The first step is to obtain the auxiliary model. Therefore, we use the SNP density that is obtained by expanding in a Hermite expansion the square root of $h(z)$, an innovation density,

$$\sqrt{h(z)} = \sum_{i=0}^{\infty}\theta_i z^i\sqrt{\phi(z)}.$$

Here $\phi(z)$ is the standard normal density function[11]. The reshaped density is given by

$$h_K(z) = \frac{P_K^2(z)\phi(z)}{\int P_K^2(u)\phi(u)du},$$

where

$$P_K(z) = \sum_{i=0}^{K}\theta_i z^i,$$

and $h_K(z)$ integrates to one since it is normalized. The SNP density is, according to the following location-scale transformation $y = \sigma z + \mu$,

$$f_K(y|\theta) = \frac{1}{\sigma}h_K(\frac{y-\mu}{\sigma}).$$

[11]This expansion exists because Hermite functions are dense in L_2 and $\sqrt{h(z)}$ is an L_2 function.

Following our notation, $h(z) = p(x, y|\rho^0)$ is the transition density and ρ^0 is the true vector of parameters. Therefore, the location-scale transformation becomes

$$y = R_x z + \mu_x,$$

where z is an innovation and R_x is an upper triangular matrix. R_x, for a GARCH specification which is the one that model the data used in this paper, is given by

$$vech(R_{x_{t-1}}) = \rho_0 + \sum_{i=1}^{L_r} P_i |y_{t-1-L_r} - \mu_{x_{t-2-L_r+i}}| + \sum_{i=1}^{L_g} diag(G_i) R_{x_{t-2-L_g+i}},$$

where $vech(R)$ is a vector of dimension $M(M+1)/2$ which contains the unique elements of the matrix R, ρ_0 denotes a vector of dimension $M(M+1)/2$, P_1 through P_{L_r} are $M(M+1)/2$ by M matrices and G_1 through G_{L_g} are vectors of length $M(M+1)/2$.

The density function of this innovation is

$$h_K(z|x) = \frac{P_K^2(z,x)\phi(z)}{\int P_K^2(u,x)\phi(u)du},$$

where $P(z, x)$ is a polynomial in (z, x) of degree K and $\phi(z)$ is the multivariate density of M independent standard normal random variables. As before, the polynomial $P_K(z, x)$ equals

$$P_K(z,x) = \sum_{\alpha=0}^{K_z} (\sum_{\beta=0}^{K_x} a_{\beta\alpha} x^\beta) z^\alpha,$$

where α and β are multi-indexes with degrees K_z and K_x, respectively. Since $h_K(z|x)$ is a homogeneous function of the coefficients of $P_K(z, x)$, it is necessary to impose a restriction ($a_{00} = 1$) to have a unique representation.

The location function is linear

$$\mu_x = b_0 + Bx_{t-1},$$

with b_o a vector and B a matrix, both formed of parameters to be estimated.

Taking in account the location-scale transformation the SNP density becomes at last

$$f_{\scriptscriptstyle K}(y|x,\theta) = \frac{h_K[R_x^{-1}(y-\mu_x)|x]}{\det(R_x)}.$$

The maximal number of lags is $L = max(L_u, L_u + L_r, L_p)$. L_u denotes the number of lags in μ_x, $L_u + L_r$ is the number of lags in R_x and finally L_p denotes the number of lags that go into the x part of the polynomial $P_K(z,x)$.

SNP Estimation Results In this subsection of the paper we present the results of the projection step.

The auxiliary model that best fits the raw data is found using the SNP model described in the previous section. The first 47 observations were reserved for forming lags. The values taken by L_u, L_g, L_r, L_p, K_z and K_x were determined by going along a expansion path and the selection criterion used was the BIC (Bayesian Information Criterion), Schwarz (1978).

As always, models that present a small value for the BIC criterion are preferred to the ones with higher values. The expansion path has a tree structure. As Gallant and Tauchen (1996) suggested, better than expanding the entire tree structure is to start expanding L_u keeping $L_r = L_p = K_z = K_x = 0$ till the BIC increases value. The following step is to expand in L_r with $L_p = K_z = K_x = 0$. Next, one expands K_z with $K_x = 0$ and finally L_p and K_x. Sometimes it can happen that the smallest value of the BIC is somewhere inside the tree. So, it is convenient for this reason to expand K_z, L_p and K_x at a few intermediate values of L_r.

The best model according to this procedure[12] has

$$L_u = 1, L_r = 1, L_g = 1, L_p = 1, K_z = 6 \text{ and } K_x = 0$$

and can be characterized as a Semiparametric GARCH.

3.2 The estimation step

In this section the main aims are: first of all estimate the vector of parameters ρ, test if the specification proposed for modeling the data is adequate by using the minimum chi-square criterion, and finally analyze the reasons of the system failure and shed light on the possible modifications that can better fit the data.

[12]This strategy reveals itself reasonable in much applied work, Fenton and Gallant (1996b). Gallant and Tauchen (2001) also arrived to the same specification.

The EMM estimator $\hat{\rho}_n$ is determined as follows. First, we use the score generator determined in the projection step

$$f(y_t|x_{t-1},\theta) \qquad \theta \in \Re^{p_\theta}$$

and the data $\{\tilde{y}_t\}_{t=-L}^{n}$ in order to obtain the quasi-maximum likelihood estimate

$$\tilde{\theta}_n = \arg\max_{\theta\in\Theta} \frac{1}{n}\sum_{t=0}^{n} \log[f(\tilde{y}_t|\tilde{x}_{t-1},\theta)].$$

The information matrix is

$$\tilde{I}_n = \frac{1}{n}\sum_{t=0}^{n}[\frac{\partial}{\partial\theta}\log f(\tilde{y}_t|\tilde{x}_{t-1},\tilde{\theta}_n)][\frac{\partial}{\partial\theta}\log f(\tilde{y}_t|\tilde{x}_{t-1},\tilde{\theta}_n)]'.$$

In the literature it is assumed that $f(y|x,\tilde{\theta}_n)$ is a good approximation to the true density of the data. Otherwise, more complicated expressions for the weighting matrix should be used[13].

Defining the moment conditions by

$$m(\rho,\tilde{\theta}) = E_\rho\{\frac{\partial}{\partial\theta}\log f(y_t|x_{t-1},\tilde{\theta})\},$$

which are obtained by averaging over a long simulation

$$m(\rho,\tilde{\theta}_n) = \frac{1}{N}\sum_{t=0}^{N}[\frac{\partial}{\partial\theta}\log f(\hat{y}_t|\hat{x}_{t-1},\tilde{\theta}_n)],$$

the EMM estimator is obtained by

[13]See Gallant and Tauchen (1996) and Gallant and Tauchen (2001). However, Gallant and Long (1997), Gallant and Tauchen (1999) and Coppejans and Gallant (2002), proved if the auxiliary model corresponds to the SNP density the information matrix above will be the adequate.

$$\hat{\rho}_n = \arg\min m'(\rho, \tilde{\theta}_n)(\tilde{I}_n)^{-1} m(\rho, \tilde{\theta}_n). \tag{10}$$

The asymptotic properties of the estimator are derived in Gallant and Tauchen (1996) and presented below. Define ρ^0 as the true value of the parameter ρ and θ^0 as an isolated solution of the moment conditions $m(\rho^0, \theta) = 0$. Then under regularity conditions it can be shown that

$$\lim_{n\to\infty} \hat{\rho}_n = \rho^0 \ a.s.,$$

$$\sqrt{n}(\hat{\rho}_n - \rho^0) \xrightarrow{D} N\{0, [(M^0)'(I^0)^{-1}(M^0)]^{-1}\},$$

$$\lim_{n\to\infty} \hat{M}_n = M^0 \ a.s. \text{ and}$$

$$\lim_{n\to\infty} \tilde{I}_n = I^0 \ a.s.,$$

where $\hat{M}_n = M(\hat{\rho}_n, \tilde{\theta}_n)$, $M^0 = M(\rho^0, \theta^0)$, $M(\rho, \theta) = (\frac{\partial}{\partial \rho'}) m(\rho, \theta)$ and

$$I^0 = E_{\rho^0}[\frac{\partial}{\partial \theta} \log f(y_0|x_{-1}, \theta^0)][\frac{\partial}{\partial \theta} \log f(y_0|x_{-1}, \theta^0)]'.$$

These asymptotic results permit testing if the model is correctly specified. Under the H_0 that $p(y_{-L}, ..., y_0|\rho)$ is the correct model

$$L_0 = n m'(\hat{\rho}_n, \tilde{\theta}_n)(\tilde{I}_n)^{-1} m(\hat{\rho}_n, \tilde{\theta}_n)$$

follows asymptotically a chi-square with $p_\theta - p_\rho$ degrees of freedom. It is also possible to test restrictions on the parameters, i.e.,

$$H_0 : h(\rho^0) = 0$$

where h is a mapping from R into $\Re^q$ and the test statistic is given by

$$L_h = n[m'(\hat{\rho}_n, \tilde{\theta}_n)(\tilde{I}_n)^{-1} m(\hat{\rho}_n, \tilde{\theta}_n) - m'(\hat{\rho}_n, \tilde{\theta}_n)(\tilde{I}_n)^{-1} m(\hat{\rho}_n, \tilde{\theta}_n)] \qquad \overset{a}{\sim} \chi^2(q)$$

and

$$\hat{\rho}_n = \underset{h(\rho)=0}{\arg\min} m'(\rho, \tilde{\theta}_n)(\tilde{I}_n)^{-1} m(\rho, \tilde{\theta}_n).$$

Finally, it is also possible to obtain confidence intervals for the parameters by computing the standard deviations using numeric methods. These intervals present a drawback because sometimes a parameter approaches a value for which the model is explosive and this fact is not accompanied by an increase in the EMM objective function. Gallant and Tauchen (1996) came up with a solution that consists of inverting the difference test L_h[14]. These "inverted" intervals are not free of problems. In fact, it was shown that they do not present more accurate coverage probabilities, especially when the degrees of freedom are low.

Since

$$\sqrt{n} m(\hat{\rho}_n, \tilde{\theta}_n) \xrightarrow{d} N\{0, I^0 - (M^0)[(M^0)'(I^0)^{-1}(M^0)]^{-1}(M^0)'\},$$

the t-ratios are given by

$$T_n = S_n^{-1} \sqrt{n} m(\hat{\rho}_n, \tilde{\theta}_n),$$

where $S_n = (diag\{\tilde{I}_n - (\hat{M}_n)[(\hat{M}_n)'(\tilde{I}_n)^{-1}(\hat{M}_n)]^{-1}(\hat{M}_n)'\})$. The characteristics of the data are reflected in the different elements of score. If the model fails to fit these characteristics this fact comes out in the large values taken by the t-ratios (of the elements of the score). In this case, the failure can suggest alternative modelizations.

3.2.1 Empirical results

All the estimated results were obtained using the computer package EMM programmed by Gallant and Tauchen (1997) with Fortran 77 available at ftp.econ.duke.edu. The global minima of equations 4 to 6 were found through an exhaustive search grid of the starting values and the help of randomization.

[14] In order to invert the test we select for the interval that values of ρ_i^* for which the H_0: $\rho_i^0 = \rho_i^*$ is not reject under the test L_h.

Table 1 gives a summary of the specification presented in section three and shows the value of the diagnostic test which follows an asymptotic chi-square with $p_\theta - p_\rho$ degrees of freedom. From the table and in particular from the chi-square test, we can verify that the two factors volatility model passes the specfication test without violating any of the moment conditions. Moreover, all coefficients are statistically significant[15] and the first volatility factor is very slow mean reverting while the second is extremely fast mean reverting, as in Gallant and Tauchen (2001).

4 Forecasting

Forecasting using the continuous time stochastic volatility model requires the reprojection step. It allows us to filter the volatility factors U_{2t} and U_{3t} and consequently to obtain a forecast of the underlying integrated volatility for any desired sampling frequency. In fact, as a by-product of the estimation step we obtain a long simulation of the volatility factors $\{\hat{U}_{2t}\}_{t=1}^{N}$ and $\{\hat{U}_{3t}\}_{t=1}^{N}$. Having as the main aim to obtain

$$E(U_{2t}|\{y_\tau\}_{\tau=1}^{t}),$$

$$E(U_{3t}|\{y_\tau\}_{\tau=1}^{t})$$

we start generating simulations of $\{\hat{U}_{2t}\}_{t=1}^{N}$, $\{\hat{U}_{3t}\}_{t=1}^{N}$ and $\{\hat{y}_t\}_{t=1}^{N}$ at the estimated parameter $\hat{\rho}$ and with N equal 100 000. Then we impose the same SNP-GARCH model founded in the projection step, on the simulated values $\hat{y}_t$. According to Gallant and Tauchen (2001), this provides a good representation of the one-step ahead conditional variance $\hat{\sigma}_t^2$ of $\hat{y}_{t+1}$ given $\{\hat{y}_\tau\}_{\tau=1}^{t}$[16]. Then, we run regressions of $\hat{U}_{2t}$ and $\hat{U}_{3t}$ on lags of $\hat{\sigma}_t^2$, $\hat{y}_t$, $|\hat{y}_t|$.

$$\hat{U}_{2t} = \alpha_0 + \alpha_1\hat{\sigma}_{t-1}^2 + ... + \alpha_p\hat{\sigma}_{t-p}^2 + \theta_1\hat{y}_{t-1} + ... + \theta_q\hat{y}_{t-q} + \pi_1|\hat{y}_{t-1}| + ... + \theta_r|\hat{y}_{t-r}|$$

$$\hat{U}_{3t} = \beta_0 + \beta_1\hat{\sigma}_{t-1}^2 + ... + \beta_p\hat{\sigma}_{t-p}^2 + \gamma_1\hat{y}_{t-1} + ... + \gamma_q\hat{y}_{t-q} + \lambda_1|\hat{y}_{t-1}| + ... + \lambda_r|\hat{y}_{t-r}|.$$

With this procedure we obtain calibrated functions inside the simulation that give predicted values of U_{2t} and U_{3t} given $\{y_\tau\}_{\tau=1}^{t-1}$. Finally, we evaluate these

[15] See table 2.

[16] In fact, given the length of simulation, these regressions are as Gallant and Tauchen (2001) say analytic projections.

functions on the observed data series $\{\hat{y}_\tau\}_{\tau=1}^{t-1}$ to obtain forecasts of the volatility factors, $\tilde{U}_{2t}$ and $\tilde{U}_{3t}$. The volatility forecast, for day t +1 will be

$$\exp(\hat{\beta}_{10} + \hat{\beta}_{12}\tilde{U}_{2(t+1)} + \hat{\beta}_{13}\tilde{U}_{3(t+1)}).$$

We are going to split the sample in two subsamples, the first subsample is used to estimate the models and the second part (the out-of- sample period) is used to evaluate the models' forecasts. We use three out-of-sample periods. The first out-of-sample period ranges from the 11^{th} of January 2001 till the 23^{rd} February 2001, the second ranges from the 3^{rd} of January 2000 till the 23^{rd} February 2001 and the third out-of-sample ranges from the 4^{th} of January 1999 till the 31^{st} of December 1999[17].

4.1 Alternative Models

We also tried two different specifications. The first one is the traditional GARCH model. The parameters of this model are estimated with the historical daily data to build out-of-sample volatility forecasts[18].

There is strong empirical evidence that the volatility has long memory, in the sense that the effect of a shock to volatility persists for a long number of periods. The second specification is the ARFIMA model and it tries to fit this feature by modelling volatility as a fractionally integrated process.

Definition 1 *A stationary process* $\{y_t\}$ *is said to be* ***long memory*** [19] *if its ACF (autocorrelation function* γ*) decays toward zero so slowly that*

$$\sum_{u=-n}^{n} |\gamma_y(u)| \uparrow \infty \qquad as \qquad n \to \infty$$

In order to make inferences about the long memory characteristic of the volatility series[20] we will use a formal test.

4.1.1 Testing the existence of long memory

There are many tests that we can apply to check for long memory of volatility. In this paper, we first use the traditional R/S method.

Consider $Y_1, Y_2, ..., Y_n$ the observations in n successive periods and $\bar{Y}$ the

[17] We use this last out-of-sample period because we would also like to see the forecasting performance of the models in a horizon that does not seem to show structural changes in volatility.

[18] See table 1 for the estimation results. We use the GARCH package for Ox to estimate the model available at Jurgen A. Doornik web page.

[19] Zaffaroni (2001) lecture notes.

[20] I use as a proxy of the volatility the squared returns.

empirical average. The adjusted range R is defined as

$$R(n) = \max_{0\leq l\leq n}\{\sum_{i=1}^{l} Y_i - l\bar{Y}\} - \min_{0\leq l\leq n}\{\sum_{i=1}^{l} Y_i - l\bar{Y}\}$$

and an estimate of the variance of the process underlying the data is

$$S^2(n,q) = \sum_{j=-q}^{q} \omega_q(j)\hat{\gamma}(j),$$

where $\hat{\gamma}(j)$ is an estimate of the autocovariance function at lag j and $\omega_q(j)$ are weights. Finally the R/S statistic is then defined as

$$Q(n,q) = \frac{R(n)}{S(n,q)}.$$

Helms et al.(1984) set $q = 0$ and $\omega_0(0) = 1$. The R/S statistic with these restrictions suffers from two disavantages: first its distribution is not known and secondly it can be affected by short-memory components. Lo (1991) modified this statistic by putting $q \neq 0$ in order to deal with these problems. His weights were given by

$$\omega_q(j) = 1 - \frac{j}{q+1} \quad with \quad q < n$$

and q was chosen as the greatest integer less than or equal to

$$(\frac{3n}{2})^{\frac{1}{3}}(\frac{2\hat{\rho}(1)}{1-\hat{\rho}^{2}(1)})^{\frac{2}{3}},$$

with $\hat{\rho}(1)$ as an estimate of the first order autocorrelation of the process.

For short memory processes the values of $Q(n,q)$ converge to n^J. d is the long memory parameter and J is related to it by $J = d + 1/2$. Mandelbrot and Taqqu (1979) also proved that the process has long memory when $J > 1/2$ and their estimator for J was

$$\hat{J} = \frac{\log(R(n)/S(n))}{\log n}.$$

According to the correlogram and the long memory test it seems that the series of squared returns[21] is fractionally integrated, that is

[21] See figure 4 and table 4.

Definition 2 *A stationary process* $\{y_t\}$ *is said to be fractionally integrated with long memory if it can be written as*

$$(1-L)^d \phi(L) y_t = \theta(L)\epsilon_t,$$

where L *is the lag operator,* $\phi(L)$ *and* $\theta(L)$ *are polynomials in the lag operator with roots inside the unit circle,* ϵ_t *are independently and identically distributed as* $N(0, \sigma^2)$ *and*

$$0 < d < \frac{1}{2}.$$

Therefore, we can think of an ARFIMA model as a quite good description of the volatility dynamics.

4.1.2 Estimating the ARFIMA model

We use the ARFIMA package for Ox[22] in order to estimate the parameters of the ARFIMA model. The best model according to the BIC criterion does not have a moving average part and the autorregressive part is of order one[23].

4.2 Evaluating and comparing alternative volatility forecasts

In this subsection we assess the performance of the volatility forecasts generated from the continuous time stochastic volatility model and compare it with the performance of the GARCH and ARFIMA forecasts for the three out-of-sample periods[24].

For the first out-of-sample period we are going to use one-day-ahead volatility forecasts[25] and then we compare them to the estimate of realized volatility determined before. For this, we proceed following the analysis in Andersen and Bollerslev (1997) by regressing the realized volatilities on a constant and on the various model forecasts. Tables 6 to 8 report the estimated regressions for the one-day-ahead out-of-sample forecasts that assumes the following form:

$$\widehat{rvolatility}_{t+1} = \beta_0 + \beta_1 \widehat{\sigma^2_{t+1/t,ARFIMA}} + u_{t+1} \tag{11}$$

$$\widehat{rvolatility}_{t+1} = \beta_0 + \beta_1 \widehat{\sigma^2_{t+1/t,GARCH}} + u_{t+1} \tag{12}$$

[22] Available at Jurgen A. Doornik web page: www.nuff.ox.ac.uk/Users/Doornik.

[23] See table 5.

[24] The first out-of-sample period ranges from the 11^{th} of January 2001 till the 23^{rd} February 2001, the second ranges from the 3^{rd} of January 2000 till the 23^{rd} February 2001 and the third out-of-sample ranges from the 4^{th} of January 1999 till the 31^{st} of December 1999.

[25] In this case the models and the filters have been estimated and computed 27 times.

$$\widehat{rvolatility}_{t+1} = \beta_0 + \beta_1 \widehat{\sigma^2_{t+1/t,SV2F}} + u_{t+1}$$

The analysis of the results is based on the R^{*2} of the regressions above[26] and on the t-statistics for the hypothesis of $\beta_0 = 0$ and/or $\beta_1 = 1$. We use both OLS and instrumental variables (IV) methods of estimation. The use of IV can be justified by the existence of a possible error in the forecast of future volatility that would lead the OLS estimates to be inconsistent. The instruments that were used are the past volatility forecasts for the two first equations and the squared return for the last equation because it seemed more correlated to the volatility forecast than to its past value.

For the considered out-of-sample period we verify that the hypothesis of $\beta_0 = 0$ and $\beta_1 = 1$ are both rejected at a 5% significance level for the GARCH and ARFIMA models. Moreover the coefficients of volatility forecasts of both regression models 12 and 13 show negative signs, which could lead us think that both models are inappropriate to forecast volatility (however both variables are statistically insignificant). These strange results may be explained by a structural change in volatility observed in the out-of-sample period and not taken in to account in both specifications.

Contrarily, the SV2F model seems to forecast much better in the out-of-sample period of 28 days. The empirical results report that the variable volatility forecast is probably an unbiased estimator of future volatility since the hypothesis of $\beta_0 = 0$ and $\beta_1 = 1$ are not reject at a 5% significance level[27]. The R^{*2} is equal to 0.235883, which is larger than the ones observed for the GARCH and the ARFIMA.

The better performance of the continuous stochastic volatility model is due probably to its flexibility and ability to capture volatility persistence. As has been reported in several papers, for instance in Diebold and Inoue (1999), Granger and Hyung (1999), Kim and Kon (1999) and Beine and Laurent (2000), structural change in volatility and persistence in volatility are imperfect substitutes. By this we mean that the persistence captured in a model is strongly reduced when we include structural shifts in the variance. Since we do not allow

[26] Anderson and Bollerslev (2002) show that there is a bias in empirical realized volatility measures built directly from high- frequency data due to the existence of market microstructure frictions. This leads to a downward bias in the R^2 obtained from the above regressions. In fact, they show that these R^2 will under-estimate the true R^{*2} by the multiplicative factor: $Var[RV_t(h)]/Var[IV_t] \approx Var[RV_t(h)]/\{Var[RV_t(h)] - hE[RQ_t(h)]\}$ where RV_t is the realized volatility, IV_t is the integrated volatility and

$$RQ_t(h) \equiv \frac{1}{h}\frac{2}{3}\sum_{i=1}^{1/h} r^{(h)4}_{t-1+ih},$$

with $1/h = 96$ corresponding to the use of "15-minute" returns. For more details please check Anderson and Bollerslev (2002).

[27] See tables 6 and 7.

for shifts in the variance, since they occur over a short period at the very end of sample, and thus cannot be explicitly modeled, the SV2F tries to accommodate the "missing" shifts by allowing that one factor of volatility be extremely slow mean reversing (a sign of strong persistence in volatility). The GARCH and ARFIMA models are not able to account for the apparent switch in volatility.

Next, we use the second out-of-sample period to investigate the forecasting performance of the SV2F model at longer horizons. We denote the whole out-of-sample period as $[t, T]$, where t corresponds to the 3^{rd} of January 2000 and T to the 23^{rd} February 2001. We split the out-of sample period $[t, T]$ into the subsets $[t, t_1]$, $[t_{1+1}, t_2]$ and $[t_{2+1}, T]$ which are used for volatility forecasting. Notice that t_1 corresponds to the 18^{th} of May 2000 and t_2 to the 4^{th} of October 2000. In other words, we estimate the SV2F model and calculate the volatility factors U_{2t} and U_{3t} three times, at $t-1$ (12^{th} of December 1999), t_1 and t_2. Therefore, the volatility forecast for day $t+7$ is for example

$$\exp(\hat{\beta}_{10} + \hat{\beta}_{12}\tilde{U}_{2(t+7)} + \hat{\beta}_{13}\tilde{U}_{3(t+7)}).$$

Remember that the $\beta's$ and the calibrated coefficients of $\tilde{U}_2$ and $\tilde{U}_3$ remain the same till the next estimation date.

Table 8 reports the forecasting results for the SV2F, GARCH and ARFIMA models. Once more the GARCH and the ARFIMA models perform poorly. The hypothesis of $\beta_0 = 0$ and/or $\beta_1 = 1$ are both rejected at a 5% significance level for these two models and the coefficients of volatility forecasts of both regression models 12 and 13 show negative signs as before.

The volatility forecasting performance based on the stochastic volatility model improves over the other two although we can no longer claim that the volatility forecast is an unbiased estimator of future volatility. Observing once more figure 3, we see that the out-of-sample period of 289 periods corresponds exactly to the part of sample where volatility pattern seems to change. This probably explains the poor performance of both GARCH and ARFIMA. The SV2F model performs better due, as it was explained before, to its ability of capturing volatility persistence.

Finally, we also evaluate the forecasting performance of the continuous time model in the out-of-sample period that ranges from the 4^{th} of January 1999 till the 31^{st} of December 1999. We choose this period because it precedes the period of volatility's pattern change. As before, we denote the whole out-of-sample period as $[t, T]$, where t corresponds to the 4^{th} of January, 1999 and T to the 31^{st} December, 1999. We split the out-of sample period $[t, T]$ into the subsets $[t, t_1]$, $[t_{1+1}, t_2]$ and $[t_{2+1}, T]$ which are used for volatility forecasting[28]. Notice that t_1 corresponds to the 4^{th} of May 1999 and t_2 to the 2^{nd} of September 1999. In other words, we estimate the SV2F model and calculate the volatility factors U_{2t} and U_{3t} three times, at $t-1$ (31^{st} of December 1998), t_1 and t_2.

[28] We use, this time, 10-days-ahead forecasts.

Considering this out-of-sample period we observe[29], for the continuous time model, that the hypothesis of $\beta_0 = 0$ is not rejected at a 1% significance level and the hypothesis of $\beta_1 = 1$ is not rejected at any conventional significance levels. The GARCH and ARFIMA models perform worse accordingly to R^{*2} and the previous hypothesis of $\beta_1 = 1$ is rejected at all conventional significance levels.

Moreover, following the analysis in Anderson, Bollerslev, Diebold and Labys (2001), we also focus our forecasting evaluation on regressions of the realized volatility on a constant, on the SV2F model forecasts and on the other benchmark model's forecasts:

$$\widehat{rvolatility}_{t+1} = \beta_0 + \beta_1 \widehat{\sigma^2_{t+1/t,SV2F}} + \beta_2 \widehat{\sigma^2_{t+1/t,GARCH}} + u_{t+1} \qquad (13)$$

$$\widehat{rvolatility}_{t+1} = \beta_0 + \beta_1 \widehat{\sigma^2_{t+1/t,SV2F}} + \beta_2 \widehat{\sigma^2_{t+1/t,ARFIMA}} + u_{t+1} \qquad (14)$$

Table 10 reports the empirical results. On including both the SV2F and the GARCH or the ARFIMA forecasts in the same regression, the estimates of the coefficients β_2 in equations 13 and 14 are not different from zero statistically and the hypothesis of $\beta_0 = 0$ and/or $\beta_1 = 1$ in both regressions are not rejected at any conventional significance levels. Futhermore, the inclusion of the GARCH or ARFIMA forecasts does not improve significantly the R^{*2} relatively to the one based only on the SV2F forecasts. So, according to these results, when there is not an aparent change of pattern in volatility, it seems that the SV2F volatility forecast is an unbiased estimator of future volatility.

5 Conclusion

In this paper we evaluate the predictive ability of the continuous time stochastic volatility model with two factors of volatility (SV2F) and compare its volatility forecasts to the forecasts obtained from the traditional GARCH model and ARFIMA models. We choose as a proxy of ex-post volatility the realized volatility obtained from the intraday returns. We argue that this is a good measure of ex-post volatility because much theoretical work models the logarithm of asset prices as a univariate diffusion and it has been shown that under innocuous regularity conditions, the realized volatility converges to the integrated volatility. We have been careful in order to avoid microstructures problems by considering only 15-minutes observations.

The main contributions of this paper include: First, the computation of the realized volatility accordingly to Nelson and Taylor (2000) in order to ensure

[29] See table 9.

conditionally unbiased estimates when intraday returns are uncorrelated. Secondly, we apply the reprojection technique proposed by Gallant and Tauchen (1998) to obtain volatility forecasts from the SV2F model and finally, we compare the forecasting performance of this last model with two others. The empirical results show that the volatility forecasting performance of the stochastic volatility model is significantly better than that of the other two models, which both perform poorly in short and mid-ranges forecast horizons.

Acknowledgement 3 *I thank my advisor Michael Creel for introducing me to the idea of Efficient Method of Moments and for his advice and encouragement during the research. I also thank George Tauchen and Tim Bollerslev for helpful remarks during my stay at Duke University and seminar participants at the Duke and UAB Universities.*

References

[1] Anderson, T. G.and Bollerslev, T. (1997), "Answering the Critics:Yes, ARCH Models Do Provide Good Volatility Forecasts", *NBER Working paper*, 6023, NBER.

[2] Anderson, T. G., Bollerslev, T.,Diebold, F.X. and Labys, P. (2000), "Exchange Rate Returns Standardized by Realized Volatility are (nearly) Gaussian", *Multinational Finance Journal*, 4, 159 -179.

[3] Anderson, T. G., Bollerslev, T.,Diebold, F.X. and Labys, P. (2001), "Modeling and Forecasting Realized Volatility", *Working Paper 8160*, NBER.

[4] Anderson, T. G.and Bollerslev, T. (2002), "Correcting the Errors: A Note on Volatility Forecast Evaluation Based on High-Frequency Data and Realized Volatilities", *Discussion paper*, Duke University.

[5] Areal, N.M.P.C. and Taylor, S.J. (2000), "The Realized Volatility of FTSE-100 Future Prices", *Discussion Paper*, Lancaster University.

[6] Bates, D. (2000), "Post-'87 Crash Fears in the S&P 500 Futures Option market", *Journal of Econometrics*, 94, 181-238.

[7] Beine, M. and Laurent, S. (2000), "Structural Change and Long Memory in Volatility: New Evidence from Daily Exchange Rates", *Discussion Paper*, Université of Liège.

[8] Campbell, J., A. Lo and C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press.

[9] Canina and Figlewski (1993), "The Informational Content of Implied Volatility", *Review of Financial Studies*.

[10] Chernov, M., Gallant, A. R., Ghysels, E. and G. Tauchen (2001), "Alternative Models for Stock Price Dynamics", *Discussion Paper*, University of North Caroline at Chapel Hill.

[11] Coppegans, M. and Gallant, A.R. (2002), "Cross-Validated SNP Density Estimates", *Journal of Econometrics*, 110, 27 - 65.

[12] Day, T. E. & Lewis, C. M. (1992), "Stock Market Volatility and the Information Content of Stock Index Options", *Journal of Empirical Finance* 5, 131-154.

[13] Diebold, F.X., Inoue,A. (1999), "Long Memory and Structural Change", *Manuscript*, Departmant of Finance, Stern School, NYU, May 1999.

[14] Doornik, J.A. and Ooms, M. (2001), "A Package for Estimating, Forecasting and Simulating ARFIMA Models: ARFIMA Package 1.01 for Ox", *Guide*, Erasmus University.

[15] Fenton, V.M. and Gallant, A. R. (1996), "Qualitative and Asymptotic Performance of SNP Density Estimators", *Journal of Econometrics*, 74, 74 -118.

[16] Gallant, A. R. and Long, J. R.(1995), "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square", *Biometrika*, 84, 125-141.

[17] Gallant, A. R. and G. Tauchen (1996), "Which Moments to Match?", *Econometric Theory*, 12, 657- 681.

[18] Gallant, A. R., D. Hsieh and G. Tauchen (1997), "Estimation of Stochastic Volatility Models with Diagnostics", *Journal of Econometrics,* 81, 159-192.

[19] Gallant, A.R., C.-T. Hsu and G. Tauchen (1999), "Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance", *Review of Economics and Statistics,* (forthcoming).

[20] Gallant, A.R. and J.R. Long (1997), "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square", *Biometrika*, 84, 125-141.

[21] Gallant, A. R. and G. Tauchen (1989), "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications", *Econometrica*, 57, 1091-1120.

[22] Gallant, A.R. and G. Tauchen (1993), "SNP: A Program for Nonparametric Time Series Analysis, Version 8.3", User's Guide, *Discussion Paper*, University of North Carolina at Chapel Hill.

[23] Gallant, A. R. and G. Tauchen (1997), "EMM: A Program for Efficient Method of Moments Estimation", Version 1.4, User's guide, *Discussion Paper*, University of North Carolina at Chapel Hill.

[24] Gallant, A. R. and G. Tauchen (1998), "Reprojecting Partially Observed Systems with Application to Interest rate Diffusions", *Journal of American Statistical Association,* 93, 10-24.

[25] Gallant, A. R. and G. Tauchen (1999), "The Relative Efficiency of Methods of Moments Estimators", *Journal of Econometrics* 92, 149-172.

[26] Gallant, A. R. and G. Tauchen (2001), "Efficient Method of Moments", *Discussion Paper*,University of North carolina at Chapel Hill.

[27] Granger, C.W.J., Hyung, N. (1999), "Occasional Structural Breaks and Long Memory", UCSD, *Discussion Paper 99-14, June 1999.*

[28] Gysels, E., Harvey, A. and Renault, E. (1995), "Stochastic Volatility", in *Handbook of Statistics*,14, Statistical Methods in Finance, G.S. Maddala and C. Rao (eds), North Holland, Amsterdam.

[29] Helms, B., Kaen, F., and Rosenman, R. (1984), "Memory in Commodity Futures Contracts", *Journal of Futures Markets, 4, 559-567.*

[30] Hol, E. and Koopman, S. J. (2000), "Forecasting the Variability of Stock Index Returns with Stochastic Volatility Models and Implied Volatility", *Discussion Paper*, Tinbergen Institute.

[31] Jiang, G. J. and Sluis, P. J. (1998), "Forecasting Volatily under Multivariate Stochastic Volatility Model via Reprojection", *Discussion paper*, Tilburg University.

[32] Jorion, P. (1996), "Risk and Turnover in the Foreign Exchange Market", *in* J.A. Frankel, G. Galli & A. Giovannini, eds, "The Microstruture of Foreign Exchange Markets", The University of Chicago Press, Chicago.

[33] Kim, D., Kon, S.J (1999), "Structural Change and Time Dependence in Models of Stock Returns", *Journal of Empirical Finance*, 6, 283-308.

[34] Laurent, S. and Peters, J.P (2002), "A Tutorial for G@rch 2.3, a Complete Ox Package for Estimating and Forecasting ARCH Models",
www.crest.fr/pageperso/lfa/sebastien.laurent/sebatien.laurent.htm.

[35] Lo, A.W. (1991), "Long Term Memory in Stock Market Prices", *Econometrica 59, 1279-1313.*

[36] Mandelbrot, B.B. and Taqqu, M (1979), "Robust R/S Analysis of Long-Run Serial Correlation", *Proceedings of the 42nd Session of the International Statistical Institute*, International Statistical Institute.

[37] Nelson, D. B. (1990), "ARCH Models as Diffusion Approximations", *Journal of Econometrics* 45, 7-39.

[38] Neslson, A. and Taylor, S. J.(2000), "The Realized Volatility of FTSE -100 Future Prices", *Discussing paper*, Department of Accounting and Finance, Lancaster University.

[39] Pagan, A. R. and Schwert, G.W. (1990), "Alternative Models for Conditional Stock Volatility", *Journal of Economics and Business Statistitcs* 9, 63-71.

[40] Schwarz, G. (1978), "Estimating the Dimension of a Model", *The Annals of Statistics*, vol 6, no. 2, pp 461-464.

[41] Taylor, S. J. and Xu, X. (1997), "The Incremental Volatility Information in One Million Foreign Exchange Quotations", *Journal of Empirical Finance 4, 317-340.*

[42] West, K. D. and Cho, D. (1995), "The Preditive Ability of Several Models of Exchange Rate Volatility", *Journal of Econometrics,* 69, 367-391.

[43] Zaffaroni, P. (2000), "Time Series Models of Changing Volatility", *Lecture Notes.*

Figures and Tables

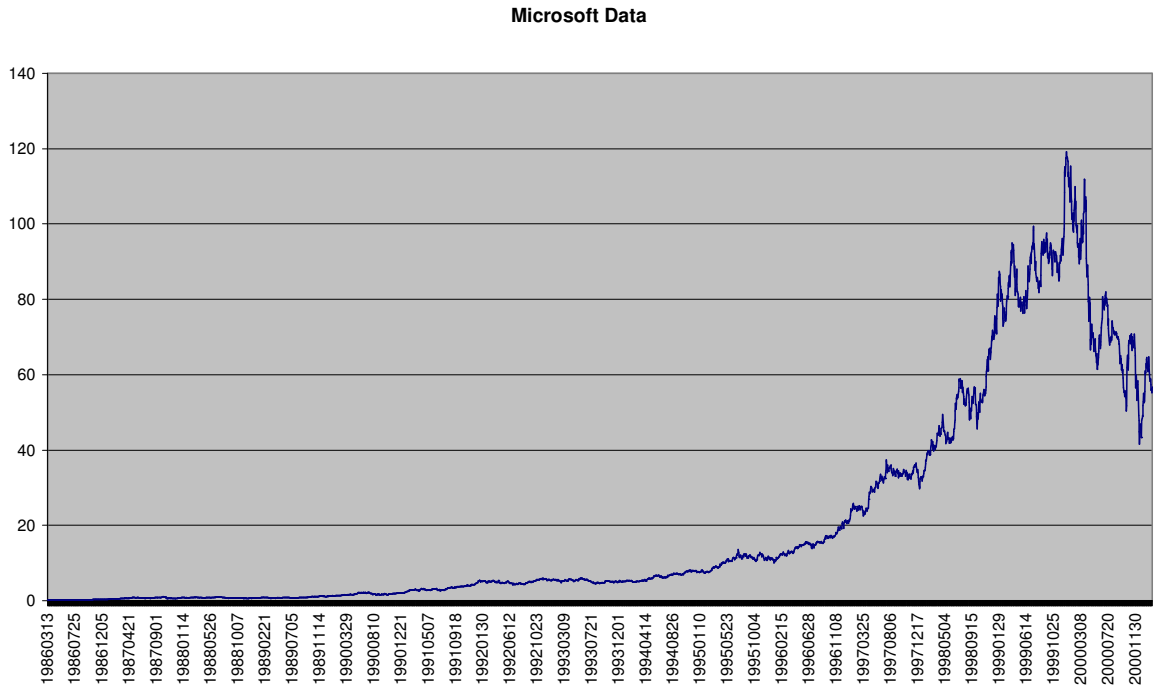

Figure 1

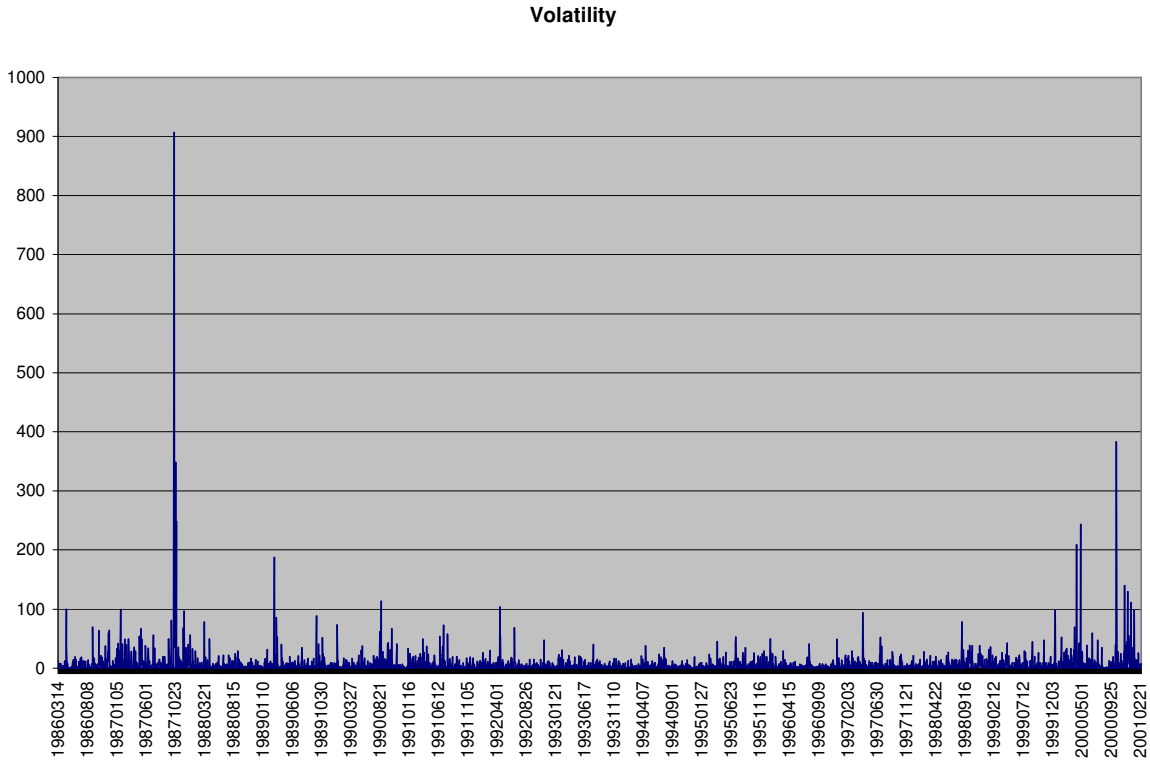
Volatility
1000
900
800
700
600
500
400
300
200
100
0
19860314
19860808
19870105
19870601
19871023
19880321
19880815
19890110
19890606
19891030
19900327
19900821
19910116
19910612
19911105
19920401
19920826
19930121
19930617
19931110
19940407
19940901
19950127
19950623
19951116
19960415
19960909
19970203
19970630
19971121
19980422
19980916
19990212
19990712
19991203
20000501
20000925
20010221

Figure 2

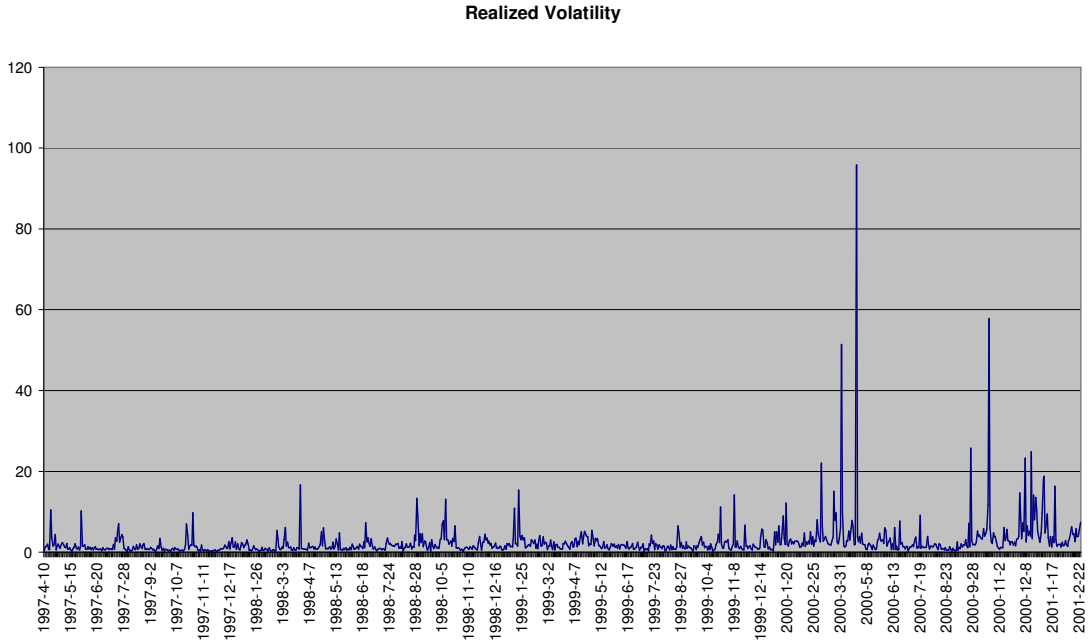
Realized Volatility
120
100
80
60
40
20
0
1997-4-10
1997-5-15
1997-6-20
1997-7-28
1997-9-2
1997-10-7
1997-11-11
1997-12-17
1998-1-26
1998-3-3
1998-4-7
1998-5-13
1998-6-18
1998-7-24
1998-8-28
1998-10-5
1998-11-10
1998-12-16
1999-1-25
1999-3-2
1999-4-7
1999-5-12
1999-6-17
1999-7-23
1999-8-27
1999-10-4
1999-11-8
1999-12-14
2000-1-20
2000-2-25
2000-3-31
2000-5-8
2000-6-13
2000-7-19
2000-8-23
2000-9-28
2000-11-2
2000-12-8
2001-1-17
2001-2-22

Figure 3

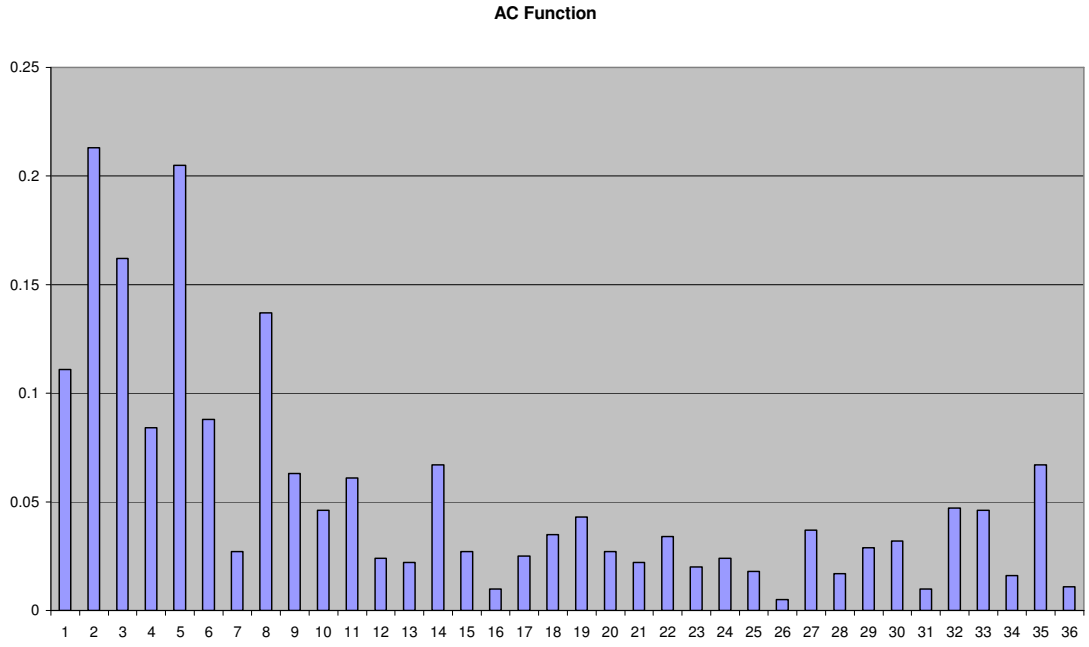

Figure 4

Model	$\boldsymbol{\alpha}_{10}$	$\boldsymbol{\alpha}_{22}$	$\boldsymbol{\alpha}_{33}$	$\boldsymbol{\beta}_{10}$	$\boldsymbol{\beta}_{12}$	$\boldsymbol{\beta}_{13}$	**N**	$\boldsymbol{\chi}^2$	**df**	**p-val**
SV2F	*	*	*	*	*	*	100k	6.704	5	0.2436

Table 1:*is used for free parameters. 100k refers to a simultion of length 100 000 at step size $\Delta = 1/6048$,corresponding to 24 steps per day and 252 trading days per year.

Speficication	$\boldsymbol{\alpha}_{10}$	$\boldsymbol{\alpha}_{22}$	$\boldsymbol{\alpha}_{33}$	$\boldsymbol{\beta}_{10}$	$\boldsymbol{\beta}_{12}$	$\boldsymbol{\beta}_{13}$
SV2F						
Estimate	0.42419	-0.00027568	-89.2101	-0.1099	0.006289	-4.6277
Std. Dev.	0.0736671	0.00014924	3.932837	0.008741	0.001034	0.0755
95%Lower	0.2693015	-0.00049470	-97.15399	-0.123322	0.004320	-4.77841
95% Upper	0.5793685	-0.00007876	-81.432454	-0.097090	0.008299	-4.480246

Table 2: Estimates, Standard Deviations and Confidence Intervals

	Estimates	Std. Error	t-value	Prob
Cst(M)	0.186230	0.035534	5.241	0.000
Cst(V)	0.267139	0.062602	4.267	0.000
GARCH(beta1)	0.864093	0.020700	41.74	0.000
ARCH(alpha1)	0.096255	0.013620	7.067	0.000

Table 3

Long Memory Test	q=0	$q=q^*$
Q	1116.861	680.330
J	0.852075	0.79189
d	0.352075	0.291894

Table 4

	Estimates	Std. Error	t-value	Prob	BIC
ARFIMA(0,d,0)					
d parameter	0.126917	0.01142	11.1	0.000	3.178017
ARFIMA(1,d,0)					
d parameter	0.233008	0.01821	12.8	0.000	3.163838
AR-1	-0.186391	0.02271	-8.21	0.000	
ARFIMA(2,d,0)					
d parameter	0.240679	0.02533	9.50	0.000	3.165289
AR-1	-0.195044	0.03005	-6.49	0.000	
AR-2	-0.009916	0.02259	-0.439	0.661	
ARFIMA(0,d,1)					
d parameter	0.265111	0.02659	9.97	0.000	3.164190
MA-1	-0.216276	0.03214	-6.73	0.000	
ARFIMA(0,d,2)					
d parameter	0.243477	0.03137	7.76	0.000	3.165364
MA-1	-0.197413	0.03562	-5.54	0.000	
MA-2	0.0227693	0.01933	1.18	0.239	
ARFIMA(1,d,1)					
d parameter	0.242072	0.02843	8.51	0.000	3.165280
AR-1	-0.138536	0.1095	-1.27	0.206	
MA-1	-0.057881	0.1298	-0.446	0.656	

Table 5

Dependent variable RV n = 28	**Estimates**	**Std. Error**	**T-value**	**Prob**	**R^{*2}**
ARFIMA					**0.054**
β_0	6.018	2.135	2.819	0.0091	
β_1	-0.219	0.180	-1.212	0.2362	
GARCH					**0.028**
β_0	5.129	1.944	2.6378	0.0139	
β_1	-0.148	0.171	-0.866	0.3946	
SV2F					**0.24**
β_0	0.634	1.127	0.562	0.5786	
β_1	2.668	0.942	2.833	0.0088	

Table 6: OLS estimation

Dependent variable RV n = 28	**Estimates**	**Std. Error**	**T-value**	**Prob**
ARFIMA				
β_0	7.710	2.944	2.619	0.0148
$\beta 1$	-0.370	0.254	-1.458	0.1574
GARCH				
β_0	6.928	2.933	2.944	0.0069
β_1	-0.318	0.212	-1.503	0.1455
SV2F				
β_0	-5.225	2.951	-1.770	0.0884
$\beta 1$	8.097	2.648	3.057	0.0051

Table 7: IV Estimation

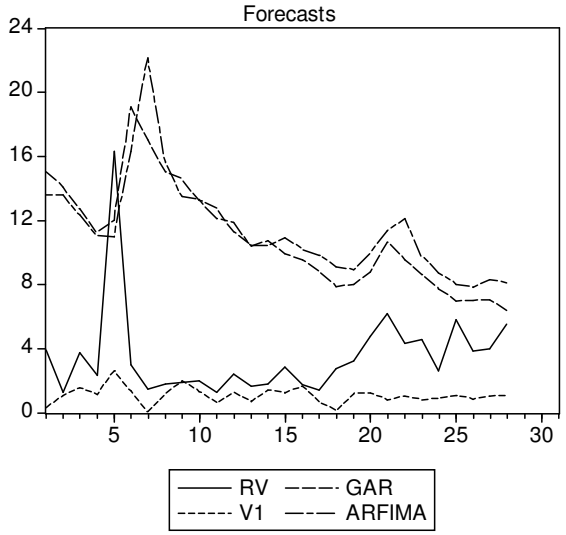

Figure 5

Dependent variable RV n = 289	**Estimates**	**Std. Error**	**T-value**	**Prob**	**R*²**
GARCH					**0.004**
β_0	10.601	4.726	2.243	0.0257	
β_1	-1.027	0.733	-1.401	0.1622	
ARFIMA					**0.03**
β_0	10.66	2.396	4.450	0.0000	
β_1	-0.933	0.292	-3.194	0.0016	
SV2F					**0.12**
β_0	-14.79	0.731	-20.23	0.0000	
β_1	18.86	0.234	80.63	0.0000	

Table 8: OLS estimation with Newey-West HAC Standard Errors

Dependent variable RV n = 252	**Estimates**	**Std. Error**	**T-value**	**Prob**	**R*^{2}**
GARCH (10 days ahead)					**0.01**
β_0	1.016	0.833	1.220	0.2237	
β_1	0.176	0.154	1.144	0.2539	
ARFIMA (10 days ahead)					**0.041**
β_0	-0.179	1.267	-0.142	0.8876	
β_1	0.385	0.231	1.667	0.0968	
SV2F (10 days ahead)					**0.10**
β_0	0.947	0.460	2.059	0.0405	
β_1	0.940	0.484	1.942	0.0533	

Table 9: OLS estimation with Newey-West HAC Standard Errors

Dependent variable RV n = 252	**Estimates**	**Std. Error**	**T-value**	**Prob**	**Adj.R*^{2}**
regression 13					**0.103**
β_0	0.406	0.926	0.438	0.6618	
β_1	0.948	0.485	1.957	0.0516	
β_2	0.098	0.134	0.727	0.4679	
regression 14					**0.114**
β_0	-0.685	1.281	-0.535	0.5929	
β_1	0.870	0.453	1.920	0.0560	
β_2	0.306	0.190	1.608	0.1091	

Table 10: OLS estimation with Newey-West HAC Standard Errors