ARE FEEDBACK FACTORS IMPORTANT IN MODELLING FINANCIAL DATA? [1]
Helena Veiga[2]

Abstract

This paper provides empirical evidence that continuous time models with one factor of volatility are, in some circumstances, able to fit the main characteristics of financial data and reports insights about the importance of introducing feedback factors for capturing the strong persistence caused by the presence of changes in the variance. We use the Efficient Method of Moments (EMM) by Gallant and Tauchen (1996) to estimate and to select among logarithmic models with one and two stochastic volatility factors (with and without feedback).

Keywords: Volatility Factors, Feedback, Persistence, Changes in Variance, Efficient Method of Moments and Reprojection.

Are feedback factors important in modelling financial data?

Helena Veiga*

Statistics Department,
Universidad Carlos III de Madrid,
C/ Madrid 126, 28903 Getafe (Madrid), Spain

Abstract

This paper provides empirical evidence that continuous time models with one factor of volatility are, in some circumstances, able to fit the main characteristics of financial data and reports insights about the importance of introducing feedback factors for capturing the strong persistence caused by the presence of changes in the variance. We use the Efficient Method of Moments (EMM) by Gallant and Tauchen (1996) to estimate logarithmic models with one and two stochastic volatility factors (with and without feedback) and to select among them.

JEL Classification: G13, C14, C52, C53
Keywords: Volatility Factors, Feedback, Persistence, Changes in Variance, Efficient Method of Moments and Reprojection.

1   Introduction

Gallant and Tauchen (2001) and Chernov et al. (2003) propose several models in continuous time and evaluate the importance of several volatility factors to the modelization of equity returns. Both papers provide empirical evidence that continuous time stochastic volatility models with one volatility factor are not able to capture simultaneously extra kurtosis and volatility persistence. They argue that the introduction of a second factor leads to specialization: one factor is going to be slow mean reverting while the other accommodates the fat tails of returns distribution.

This paper provides empirical evidence that continuous time models with one factor of volatility are, in some circumstances, able to fit the main characteristics of financial data. The success of these models in fitting the features of

data is due to the introduction of a feedback feature that is able to capture the strong persistence caused by the presence of changes in variance. The estimated models are direct extensions of Gallant and Tauchen's (2001) model, by including the feedback feature. Chernov et al. (2003) also present logarithmic models in continuous time but our specifications differ from the previous because we do not allow for stochastic instantaneous expected returns. One advantage of these logarithmic specifications is that they allow the volatility to be state dependent and the pricing formulas may be computed by simulation. Chernov et al. (2003) consider this an advantage when compared to the risk-neutral measure transformations used by the affine models.

The empirical results report the one factor logarithmic volatility model without feedback does not fit the Microsoft data which confirms the previous findings. Nevertheless, we get a new result with the introduction of the feedback factor. The model, now, does pass the specification test. This feature reveals to be of extreme importance because it allows not only to capture the low variability of the volatility factor when the factor is itself low, volatility clustering, but also to fit the increase in volatility persistence that occurs when there are apparent changes in variance. The introduction of a second factor of volatility with feedback does not seem relevant for the Microsoft data.

This paper uses EMM (Efficient Method of Moments) by Gallant and Tauchen (1996) that is based on two compulsory phases: Projection that consists of projecting the observed data onto a transition density that is a good approximation of the distribution implicit in the true data generating process, and Estimation that consists of estimating the parameters of the model with the help of the score generator. This score enters the moment conditions in which we replace the parameters of the auxiliary model by their quasi-MLEs obtained in the projection step and the estimates of the proposed model are obtained by minimizing the GMM criterion function. Since the minimized criterion function scaled by the number of observations follows asymptotically a chi-square distribution, we can apply diagnostic tests that help explaining the reasons for the failure of the model. Finally, the optional step called reprojection, is a post-estimation simulation analysis that allows to filter volatility, to obtain the density implicit in the model and to forecast volatility, see Gallant and Tauchen (1998).

The reason for the choice of EMM is the presence of unobserved variables in the proposed model, which makes the likelihood for the entire state vector frequently not feasible. Nevertheless, the simulation of the evolution of the state vector is quite possible and the EMM is based on this. Aït-Sahalia (1996a, 1996b) also develop an alternative estimation strategy for estimation stochastic differential equations. This method differs from the EMM because the moment functions are computed directly from the data rather than simulated and the full observation of the state is necessary in order to estimate all the parameters. Brandt and Santa-Clara (2002) also apply the simulated likelihood estimation procedures to multivariate diffusion processes. Nevertheless, these procedures have difficulties to deal with latent variables and moreover, the simulations have to be performed for every conditioning variable and for every parameter value.

The paper is organized as follows. Section two presents and characterizes the models we study. Section three covers the projection, estimation and reprojection steps and reports the empirical results for the Microsoft data. Section four concludes the paper.

2 Continuous Time Stochastic Volatility Logarithmic Models

Consider the following specification:

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

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

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

where P_t is an asset price that evolves in continuous time and W_i with $i = 1, 2, 3$ are three independent wiener processes. This specification extends the model in Gallant and Tauchen (2001) since it includes the feedback features $\beta_{22}U_2$ and $\beta_{33}U_3$ in the equations 2 and 3, respectively and it can be seen as a general specification that nests several others. The first is the one volatility factor logarithmic model, denoted $L1$, where $\beta_{13} = 0$ and $\beta_{22} = 0$, the second is the two volatility factor logarithmic model denoted $L2$, where $\beta_{13} \neq 0$, $\beta_{22} = 0$ and $\beta_{33} = 0$, the third is the one factor logarithmic volatility model with feedback, $L1F$, with $\beta_{13} = 0$ and $\beta_{22} \neq 0$, and finally, the logarithmic model with two factors of volatility and feedback denominated $L2F$, where $\beta_{13} \neq 0$, $\beta_{22} \neq 0$ and $\beta_{33} \neq 0$. Looking at the system, we observe that the volatility factors of equation 1 present drifts and volatilities that are linear functions of themselves, respectively. Moreover, the drifts in equations 2 and 3 allow for mean reversion when $\alpha_{ii} \neq 0$ for $i = 2, 3$. A small value for α_{ii} means that a shock to volatility takes time to dissipate. Finally, β_{10} is also an important parameter because it takes care of the long-run mean of the volatility of the price equation.

The empirical results reveal that the feedback feature is crucial in capturing the strong persistence in volatility observed for the very last part of the sample (see Figure 1), where at least a change in variance seems to occur. In order to investigate the truth of our suspicions we apply the Inclán and Tiao (1994) test for the detection of changes in variance. There are several alternatives in the literature, such as: Kokoszha (2000) that is, as the previous, a CUSUM type test for single breaks and differs from the first because it assumes strongly dependent returns with finite fourth order moments and Lavielle and Moulines (2000) that propose a Least Squares type test for multiple breaks. Andreou and

Ghysels (2002) analyze the finite sample properties of these tests and show they can have important distortions in terms of size and/or power. We choose Inclán and Tiao (1994) test because it performs better, in finite samples, comparatively to the previous tests even in non-independent settings. Inclán and Tiao (1994) test statistic is defined as: $\sqrt{\frac{T}{2}}\max_{k}|D_k|$, where $D_k = \frac{C_k}{C_T} - \frac{k}{T}$, $C_k = \sum_{j=1}^{k} x_j$ and x is, in our case, the sample volatility in the daily price of a share of Microsoft, adjusted for stock splits, from 13th of March, 1986 till 23rd of February, 2001, including 3778 observations, see Figures 1 and 2. If we are interested in finding multiple changes in the variance, the solution is an iterative scheme based on successive applications of D_k to pieces of the series, splitting consecutively after finding a change point.[1] Figure 3 gives an idea how the algorithm works and Table 1 resumes the changes in variance found in the series. We observe that the Inclán and Tiao (1994) algorithm detects eight change points in the last part of the sample and according to Beine and Laurent (2003), Diebold and Inoue (2001) and Granger and Hyung (1999) these change points are going to generate extra persistence if we discard them in the specification of models.

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

$$\alpha_{20} = 0, \alpha_{30} = 0, \beta_{20} = 1, \beta_{30} = 1.$$

These restrictions are the minimum necessary to achieve identification, they are common in previous similar SDE and they provide flexibility and numerical stability in the estimation phase. After imposing the restrictions, the system becomes:

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

$$dU_{2t} = \alpha_{22}U_{2t}dt + (1 + \beta_{22}U_{2t})dW_{2t} \tag{5}$$

$$dU_{3t} = \alpha_{33}U_{3t}dt + (1 + \beta_{33}U_{3t})dW_{3t}. \tag{6}$$

3 Efficient method of moments (EMM)

In this Section we estimate our model and the benchmark models using EMM by Gallant and Tauchen (1996). EMM is based on two compulsory phases: Projection that consists of projecting the observed data onto a transition density,

---

[1]See Appendix B for details on the ICSS algorithm.

that is a good approximation of the distribution implicit in the true data generating process. The simulated density is denominated the auxiliary model and its score is called "the score generator for EMM". The advantage is that the score has an analytical expression. The second phase consists of estimating the parameters of the model with the help of the score generator. This score enters the moment conditions in which we replace the parameters of the auxiliary model by their quasi-MLEs obtained in the projection step. Then, the estimates of the proposed model are obtained by minimizing the GMM criterion function. Finally, we can apply diagnostic tests that help explaining the reasons for the failure of a model.[2]

In the projection step, we proceed carefully along an expansion path with tree structure and the selected model comes out to be a semiparametric GARCH (auxiliary model), as in Gallant and Tauchen (2001).[3]

3.1   EMM Estimation Results

All the estimated results are obtained using the computer package EMM programmed by Gallant and Tauchen (1996) with Fortran 77 available at ftp.econ. duke.edu. The global minima of equations 4 and 6 are found through an exhaustive search grid of the starting values and the help of randomization.

Table 2 provides a summary of the specifications 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 observe that the results for the model with one volatility factor and without feedback confirm prior findings in the literature. The model is sharply rejected at a 5% significance level; but a new result appears when we introduce the feedback factor. The model does pass the specification test and the feedback feature is not only significant but also vital for fitting the moment conditions, see Tables 2 and 4. The conditional variance is now much better accommodated. Moreover, the coefficients are statistically significant and the feedback factor has a negative value, implying the variability of the volatility factor to be low when itself is low (volatility clustering). Another interesting estimate is α_{22}, the parameter of mean reversion. Looking at Table 3, we observe that its value is inferior to unity which implies that shocks to volatility take time to dissipate.

---

[2] See appendix A for asymptotic distributions of EMM estimators.

[3] The auxiliary model has the following parametrization $L_u = 1, L_r = 1, L_g = 1, L_p = 1, K_z = 6$ and $K_x = 0$. 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, we expand 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. For more details in the selection of a auxiliary model check the Gallant, Rossi and Tauchen (1992).

Through Figure 1 we also observe possible changes in volatility, specially for the last part of the sample, that are confirmed by the Inclán and Tiao (1994) test. Recent studies, for instance: Beine and Laurent (2003), Diebold and Inoue (2001) and Granger and Hyung (1999) report that there is an increase in volatility persistence if we do not account for the possibility of change points. In order to investigate this, we consider the sample used in Gallant and Tauchen (2001) that ranges from March 13, 1986 till November 16, 2000 and our sample. We compute the autocorrelation functions of the squared observations and of the absolute values and we observe specially for the latter sample that the ACF decays slowly towards zero, see Figures 5 and 6. We also compare their $L1$ model estimates with ours, and we verify that the parameter of mean reversion, α_{22}, in their case is greater than one in absolute value, which means fast mean reversion and consequently low persistence in volatility. In contrast, the same specification estimated considering the sample used in this paper reports an empirical result for that parameter of -0.902 much smaller in absolute value than the previous one (see Table 3). Both evidences are signals of an increase in persistence in the presence of changes in volatility.

Although the frequency of data is daily, it is scaled so that the coefficients are on an annually basis. That is, a value of 0.4102 for α_{10} represents an annual average rate of return equal to 41.02%. The step size is $\Delta = 1/6048$, which corresponds to 24 steps per day and 252 trading days per year.

Since the feedback factor reveals itself of extreme importance we estimate a two factor logarithmic volatility model incorporating this feature. Analyzing the results we can say that the parameter β_{13} is not significant, and consequently, the second factor of volatility. Finally, we estimate the $L2$ specification as in Gallant and Tauchen (2001) and the empirical results show that this model is another possible candidate to modelize the data. Here, the extra persistence is capture by the first volatility factor which is extremely slow mean reverting.

From the estimation step, we get two candidates, $L1F$ and $L2$ that are going to be tested in the reprojection step.

3.2 The reprojection step

The reprojection step allows us to filter the volatility factors U_{2t} and U_{3t} for any desired sampling frequency. In fact, as a by-product of the estimation step we obtain a long simulation of the volatility factors $\{\widehat{U}_{2t}\}_{t=1}^{N}$ and $\{\widehat{U}_{3t}\}_{t=1}^{N}$ and $\{\widehat{y}_t\}_{t=1}^{N}$ at the estimated vector of parameters $\widehat{\rho}$. Then, we impose the same SNP-GARCH model founded in the projection step, on the simulated values $\widehat{y}_t$. According to Gallant and Tauchen (2001), this provides a good representation of the one-step ahead conditional variance $\widehat{\sigma}_t^2$, given the past information. We proceed by running regressions of $\widehat{U}_{2t}$ and $\widehat{U}_{3t}$ on $\widehat{\sigma}_t^2$, $\widehat{y}_t$ and $|\widehat{y}_t|$ and lags of these series:

$$
\begin{aligned}
\widehat{U}_{2t} = & \ \alpha_0 + \alpha_1\widehat{\sigma}_t^2 + \alpha_2\widehat{\sigma}_{t-1}^2 + ... + \alpha_p\widehat{\sigma}_{t-p}^2 + \theta_1\widehat{y}_t + \theta_2\widehat{y}_{t-1} + ... \\
& +\theta_q\widehat{y}_{t-q} + \pi_1|\widehat{y}_t| + \pi_2|\widehat{y}_{t-1}| + ... + \pi_r|\widehat{y}_{t-r}| + u_t,
\end{aligned}
$$

$$\widehat{U}_{3t} = \beta_0 + \beta_1 \widehat{\sigma}_t^2 + \beta_2 \widehat{\sigma}_{t-1}^2 + ... + \beta_p \widehat{\sigma}_{t-p}^2 + \gamma_1 \widehat{y}_t + \gamma_2 \widehat{y}_{t-1} +$$
$$... + \gamma_q \widehat{y}_{t-q} + \lambda_1 |\widehat{y}_t| + \lambda_2 |\widehat{y}_{t-1}| + ... + \lambda_r |\widehat{y}_{t-r}| + \mu_t.$$

With this procedure we obtain calibrated functions inside the simulation that gives predicted values of U_{2t} and U_{3t} given $\{y_\tau\}_{\tau=1}^t$. In fact, given the length of the simulation, these regressions are as Gallant and Tauchen (2001) say analytic projections. Finally, we evaluate these functions on the observed data series $\{\widehat{y}_\tau\}_{\tau=1}^t$ to obtain reprojected values of the volatility factors, $\widetilde{U}_{2t}$ and $\widetilde{U}_{3t}$.

Figures 7, 8 and 9 show the reprojected volatility factors of models $L2$ and $L1F$, respectively. As it was expected, $\widetilde{U}_{3t}$ for the $L2$ is quite choppy while $\widetilde{U}_{2t}$ is sightly slower moving, see Figures 7 and 7.1. Moreover, the increase in volatility, in the last part of the sample and the crash of 1987 are attributed in its majority to the fast mean reverting factor, U_{3t}, which suggests that both events are temporary. Finally, the reprojected volatility factor of $L1F$ model is the most "alive" of the three, it tracks quite well the volatility pattern and it is also able to capture some extra noise. Therefore, we can conclude that for the data and sample used, the $L1F$ specification works quite well in modelling volatility.

4   Conclusion

This paper studies four systems of SDE : $L1$, $L1F$, $L2$ and $L2F$. From the diagnostics at the estimation step two models seem to fit the data, $L1F$ and $L2$. One possible reason for the failure of the model with one volatility factor and without feedback could be its inadequacy to model the strong persistent caused by changes in variance. This drawback, however, is overcome by introducing feedback. It allows for volatility clustering and consequently it is able to capture the strong persistence. The model, now, seems to fit all the score moment conditions associated with the GARCH parameters, as well as, the score moment conditions corresponding to the Hermite parameters responsible for the tail behavior. The second selected candidate in the estimation step is the logarithm model with two factors of volatility.

Reprojection assumes an important role in the model selection since it gives more tools for comparing among models. By computing the reprojected volatility factors implied by the previous specifications, we observe that there is no advantage in estimating the two factor stochastic volatility model for this sample. The $L1F$ model is able to reproject volatility quite well and without missing the stock market crash of 1987.

Relatively to the more complicated specification, $L2F$, the empirical results show that the second factor is not significant when we introduce feedbacks into the specifications of volatility factors.

Appendix A: the estimation step

The EMM estimator $\widehat{\rho}_n$ 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 $\{\widetilde{y}_t\}_{t=-L}^{n}$ in order to obtain the quasi-maximum likelihood estimate

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

with information matrix

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

In the literature it is assumed that $f(y|x, \widetilde{\theta}_n)$ is a good approximation to the true density of the data. Otherwise, more complicated expressions for the weighting matrix should be used.[4]

Defining the moment conditions by

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

which are obtained by averaging over a long simulation

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

the EMM estimator is given by

$$\widehat{\rho}_n = \arg\min m'\left(\rho, \widetilde{\theta}_n \right) (\widetilde{I}_n)^{-1} m(\rho, \widetilde{\theta}_n). \qquad (7)$$

The asymptotic properties of the estimator are derived in Gallant and Tauchen (1996) and are presented below. Defining 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

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

---

[4] 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.

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

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

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

where $\widehat{M}_n = M(\widehat{\rho}_n, \widetilde{\theta}_n)$, $M^0 = M(\rho^0, \theta^0)$, $M(\rho, \theta) = \left(\frac{\partial}{\partial \rho'}\right) m(\rho, \theta)$ and $I^0 = E_{\rho^0}\left[\frac{\partial}{\partial \theta} \log f(y_0|x_{-1}, \theta^0)\right]\left[\frac{\partial}{\partial \theta} \log f(y_0|x_{-1}, \theta^0)\right]'$. These asymptotic results allow testing the specification. Therefore under H_0, $p(y_{-L}, ..., y_0|\rho)$ is the correct model and $L_0 = nm'(\widehat{\rho}_n, \widetilde{\theta}_n)(\widetilde{I}_n)^{-1} m(\widehat{\rho}_n, \widetilde{\theta}_n)$ follows asymptotically a chi-square with $p_\theta - p_\rho$ degrees of freedom. Finally, 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

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

and

$$\widehat{\widehat{\rho}}_n = \arg\min_{h(\rho)=0} \ m'(\rho, \widetilde{\theta}_n)(\widetilde{I}_n)^{-1} m(\rho, \widetilde{\theta}_n).$$

Appendix B: Iterated Cumulative Sums of Squares (ICSS) Algorithm (Inclán and Tiao(1994))

Step0. Let $t_1 = 1$.

Step1. Calculate $D_k(a[t_1 : T])$. Let $k^*(a[t_1 : T])$ be the point at which $\max_k D_k(a[t_1 : T])$ is obtained, and let

$$M(t_1 : T) = \max_{\substack{k \\ t_1 \leqslant k \leqslant T}} \sqrt{\frac{(T - t_1 + 1)}{2}} |D_k(a[t_1 : T])|.$$

If $M(t_1 : T) > D^*$, consider that there is a change at $k^*(a[t_1 : T])$ and go to step 2a. Otherwise, there is no evidence of changes in variance and the algorithm stops.

Step 2a. Let $t_2 = k^*(a[t_1 : T])$. Evaluate $D_k(a[t_1 : t_2])$. If $M(t_1 : t_2) > D^*$, there is a new change point and we should repeat step 2a till $M(t_1 : t_2) < D^*$.

Step 2b. Now we should do the same search starting from the first change found in step 1 till the end of the series. Define a new value for t_1 and let $k^*(a[t_1 : T]) + 1$. Evaluate $D_k(a[t_1 : T])$ and repeat step 2b until $M(t_1 : T) < D^*$. See Inclán and Tiao (1994) for more details.

References

[1] Aït-Sahalia, Yacine (1996a): Testing Continuous-Time Models of the Spot Interest Rate, *Review of Financial Studies 9*, 385-342.

[2] Aït-Sahalia, Yacine (1996b): Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica* 64, 527-560.

[3] Andreou, E. and Ghysels, E. (2002): Detecting Multiple Breaks in Financial Market Volatility Dynamics, *Journal of Applied Econometrics* 17, 579-600.

[4] Beine, M. and Laurent, S. (2003): Central bank Interventions and Jumps in Double Long Memory Models of Daily Exchange Rates, *Journal of Empirical Finance* 10, 641-660.

[5] Brandt, M. and Santa-Clara, P. (2002): Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rates Dynamics in Incomplete Markets, *Journal of Financial Economics* 63, 161-210.

[6] Chernov, M., Gallant, A. R., Ghysels, E. and G. Tauchen (2003): Alternative Models for Stock Price Dynamics, *Journal of Econometrics* 116, 225-257.

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

[8] Diebold, F.X. and Inoue, A. (2001): Long Memory and Regime Switching, *Journal of Econometrics* 105, 131-159.

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

[10] Gallant, A. R., Rossi, P.E. and Tauchen, G. (1992): Stock Prices and Volume, *The Review of Financial Studies* 5, 199-242.

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

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

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

[14] Gallant, A. R. and G. Tauchen (2001): Efficient Method of Moments, *Discussion paper*, University of North Carolina at Chapel Hill.

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

[16] Inclán, C. and Tiao, G. C. (1994): Use of Cumulative Sums of squres for retrospective Detection of Changes of Variance, *Journal of the American Statistical Association* 89, 913-923.

[17] Kokosza, P. and Leipus, R. (2000): Change-Point Estimation in ARCH Models, *Bernoulli* 6, 1-28.

[18] Lavielle, M. and Moulines, E. (2000): Least-Squares Estimation of An Unknown Number of Shifts in Time Series, *Journal of Time Series Analysis* 20, 33-60.

[19] Schwarz, G. (1978), "Estimating the Dimension of a Model", *The Annals of Statistics* 6, 461-464.

5   Figures and Tables
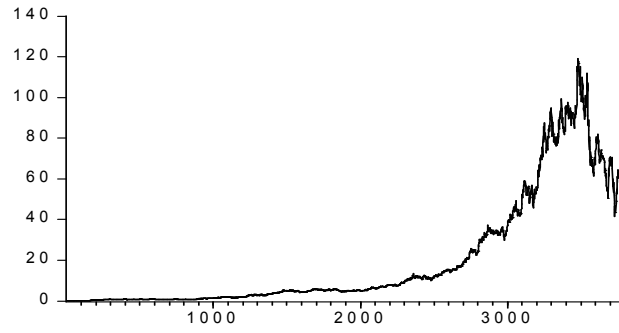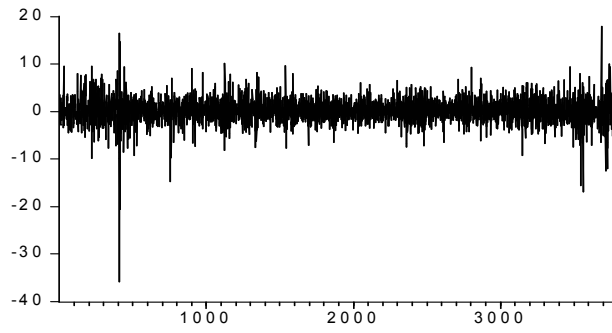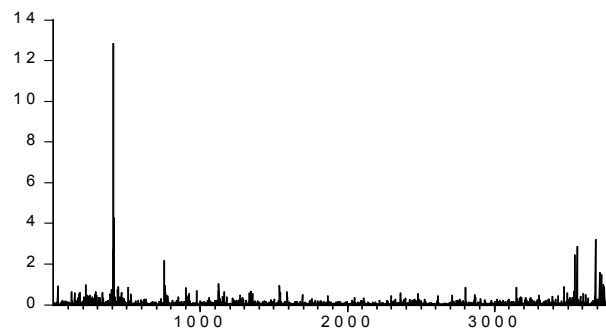
Figure 1. a) Price of a microsoft share, b) daily returns in % and c) daily volatility in %.

Series: RET	
Sample 2 3778	
Observations 3777	
Mean	0.150291
Median	0.000000
Maximum	17.86918
Minimum	-35.82807
Std. Dev.	2.541242
Skewness	-0.864767
Kurtosis	17.99882
Jarque-Bera	35874.54
Probability	0.000000
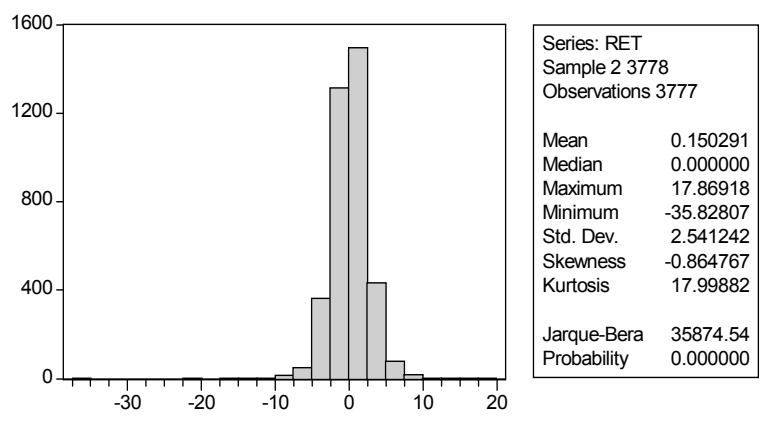
Figure 2. Histogram of returns.
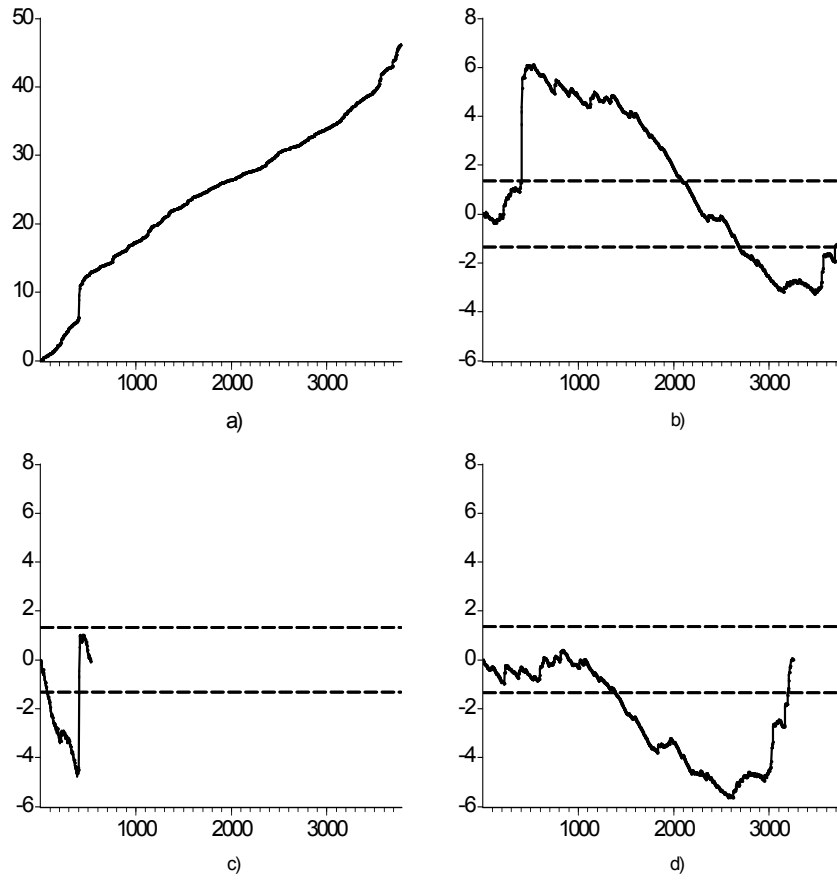
Figure 3. a) Cumulative sum of squares, b) Dk plot with k=1,...,3777: the estimate of the change point is k*= 528 and corresponds to 1988/04/14, c) Dk plot with k=1,...,528: the estimate of the change point corresponds to 1987/09/18 and d) Dk plot with k= 529,...,3777: the estimate of the change point corresponds to 1998/08/26.

	Detected changes in variance
At the beginning of the sample	1986/10/03; 1987/05/12;1987/09/18; 1988/04/14
In the middle of the sample	1992/07/02; 1994/08/01; 1995/04/13; 1996/01/19
	1996/03/26; 1996/07/10; 1996/08/02; 1996/12/06
At the end of the sample	1998/08/26; 1999/04/27; 1999/12/10; 2000/03/03
	2000/04/25; 2000/07/24; 2000/10/13; 2001/01/03

Table1

14

Model	α_{10}	α_{22}	α_{33}	β_{10}	β_{12}	β_{13}	β_{22}	β_{33}	N	χ^2	df	p-val
L1	*	*		*	*				100k	16.65	7	0.02
L1F	*	*		*	*		*		100k	10.64	6	0.10
L2	*	*	*	*	*	*			100k	6.70	5	0.24
L2F	*	*	*	*	*	*	*	*	100k	2.02	4	0.73

Table 2: *is used for free parameters. 100k refers to a simulation of length 100,000 at step size $\Delta = 1/6048$, corresponding to 24 steps per day and 252 trading days per year.

Model	α_{10}	α_{22}	α_{33}	β_{10}	β_{12}	β_{13}	β_{22}	β_{33}
L1								
Estimate	0.42	-0.90		-0.12	0.43			
Std. Dev.	0.07	0.10		0.003	0.04			
95% Lower	0.27	-1.05		-0.13	0.37			
95% Upper	0.57	-0.76		-0.12	0.49			
L1F								
Estimate	0.41	-0.16		-0.11	-0.16		-0.22	
Std. Dev.	0.04	0.002		0.003	0.003		0.004	
95% Lower	0.33	-0.16		-0.12	-0.17		-0.23	
95% Upper	0.50	-0.16		-0.10	-0.15		-0.22	
L2								
Estimate	0.42	-0.0003	-89.2	-0.12	0.006	-4.63		
Std. Dev.	0.07	0.0002	3.93	0.0087	0.001	0.08		
95% Lower	0.27	-0.0005	-97.2	-0.12	0.004	-4.78		
95% Lower	0.58	-0.0001	-81.4	-0.097	0.008	-4.48		
L2F								
Estimate	0.42	-2.16	0.21	-0.13	0.72	-0.50	0.73	7.19
Std. Dev.	0.08	0.08	2.69	0.09	0.04	6.67	0.62	0.64
95% Lower	0.41	-2.16	0.17	-0.13	0.72	-0.61	0.73	7.18
95% Lower	0.42	-2.16	0.25	-0.12	0.727	-0.41	0.74	7.20

Table 3: Estimates, Standard Deviations and Confidence Intervals.

Coefficient	L1	L1F	L2
Location Function:			
b_0	0.07	0.36	0.10
b_1	1.69	0.84	1.24
Scale Function:			
τ_0	1.82	0.38	0.76
τ_{gz}	2.54	0.13	1.03
τ_{gx}	2.32	0.28	0.92
Hermite Polynomial:			
$a_{0,1}$	0.08	0.45	0.36
$a_{0,2}$	1.87	1.21	-0.02
$a_{0,3}$	-0.07	0.54	0.50
$a_{0,4}$	1.81	1.98	0.15
$a_{0,5}$	-0.42	0.50	0.42
$a_{0,6}$	1.23	1.86	-0.03

Table 4: Scores Diagnostic.
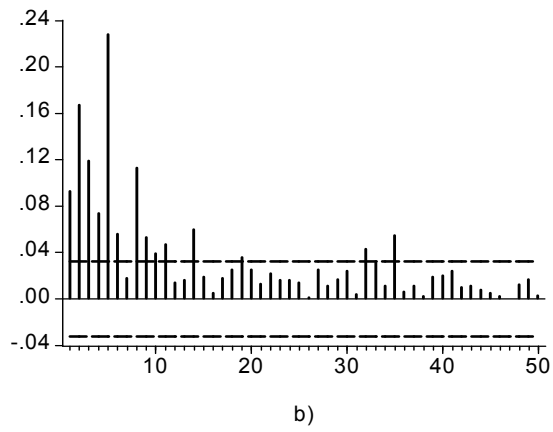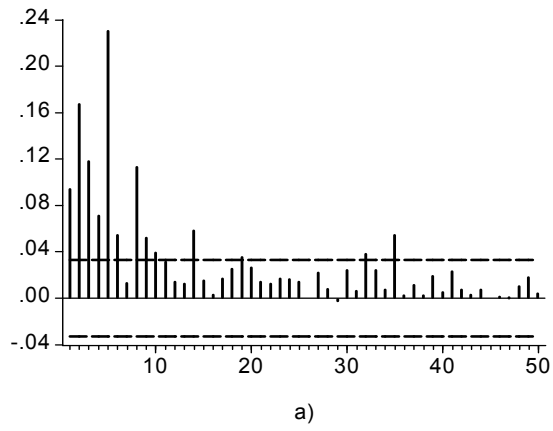
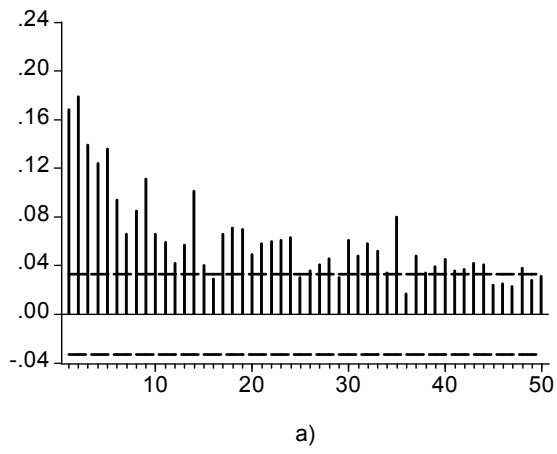Figure 5: Autocorrelation functions of squared observations. a) shorter sample and b) larger sample.
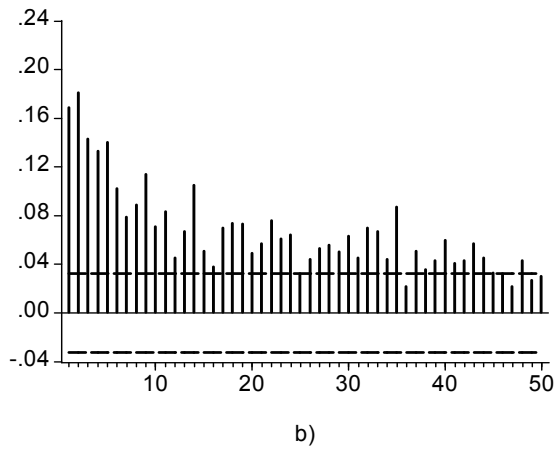
Figure 6: Autocorrelation functions of squared observations. a) shorter sample and b) larger sample.
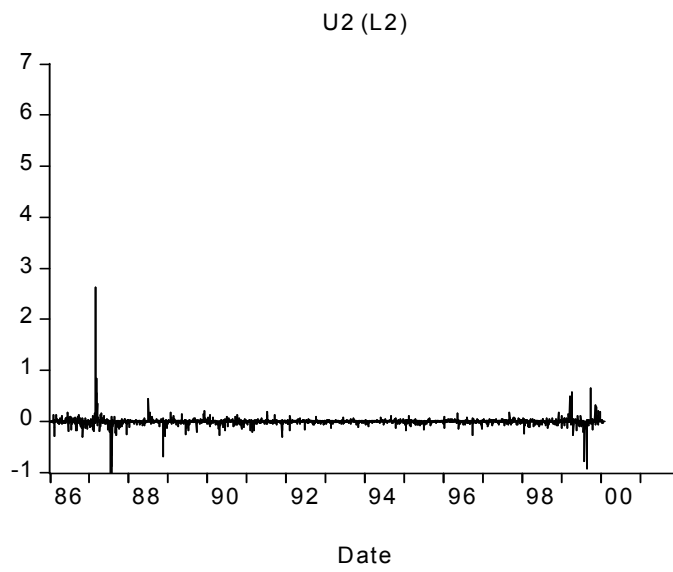
U2 (L2)

Figure 7

U3 (L2)

Figure 8

U2(L1F)
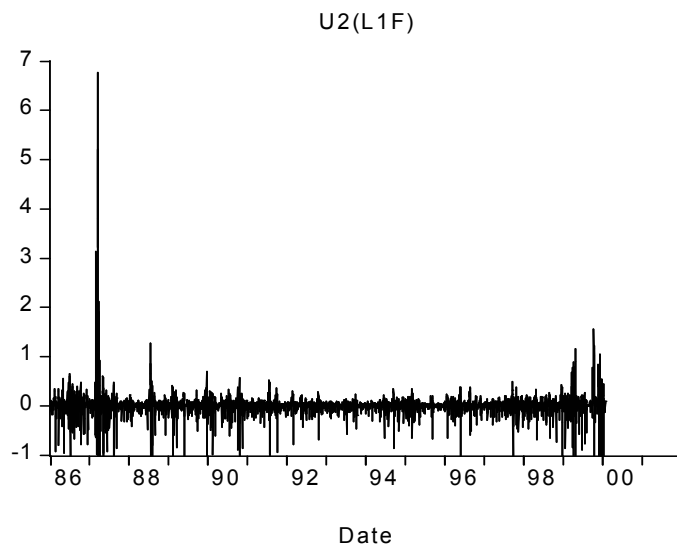

Figure 9