



Working Paper 06-13(03)
Statistics and Econometrics Series
Febrero 2006

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

A TWO FACTOR LONG MEMORY STOCHASTIC VOLATILITY MODEL¹

Helena Veiga²

Abstract

In this paper we fit the main features of financial returns by means of a two factor long memory stochastic volatility model (2FLMSV). Volatility, which is not observable, is explained by both a short-run and a long-run factor. The first factor follows a stationary AR(1) process whereas the second one, whose purpose is to fit the persistence of volatility observable in data, is a fractional integrated process as proposed by Breidt et al. (1998) and Harvey (1998). We show formally that this model (1) creates more kurtosis than the long memory stochastic volatility (LMSV) of Breidt et al. (1998) and Harvey (1998), (2) deals with volatility persistence and (3) produces small first order autocorrelations of squared observations. In the empirical analysis, we use the estimation procedure of Gallant and Tauchen (1996), the Efficient Method of Moments (EMM), and we provide evidence that our specification performs better than the LMSV model in capturing the empirical facts of data.

Keywords: Two Volatility Factors, Long Memory, Fractional Integration, EMM, Reprojection.

¹ I thank Michael Creel for introducing me to the idea of Efficient Method of Moments and for his constant advice. I am grateful to George Tauchen for his helpful remarks during my stay at Duke University. I also thank Esther Ruiz and seminar participants at the Universidad Complutense for their valuable comments. This research has been made possible thanks to the financial support of Fundação para a Ciência e Tecnologia. The usual disclaimer applies.

² Departamento de Estadística, Universidad Carlos III de Madrid.

A Two Factor Long Memory Stochastic Volatility Model

Helena Veiga*

*Statistics Department,
Universidad Carlos III de Madrid,
C/ Madrid 126, 28903 Getafe (Madrid), Spain.
E-mail: mhveiga@est-econ.uc3m.es*

Abstract

In this paper we fit the main features of financial returns by means of a two factor long memory stochastic volatility model (2FLMSV). Volatility, which is not observable, is explained by both a short-run and a long-run factor. The first factor follows a stationary AR(1) process whereas the second one, whose purpose is to fit the persistence of volatility observable in data, is a fractional integrated process as proposed by Breidt et al. (1998) and Harvey (1998). We show formally that this model (1) creates more kurtosis than the long memory stochastic volatility (LMSV) of Breidt et al. (1998) and Harvey (1998), (2) deals with volatility persistence and (3) produces small first order autocorrelations of squared observations. In the empirical analysis, we use the estimation procedure of Gallant and Tauchen (1996), the Efficient Method of Moments (EMM), and we provide evidence that our specification performs better than the LMSV model in capturing the empirical facts of data.

Keywords: Two Volatility Factors, Long Memory, Fractional Integration, EMM, Reprojection.

JEL Classification: C14.

1 Introduction

The most important feature of the conditional return distribution is its second moment dynamics. This fact has led to an enormous literature on the modelling

*I thank Michael Creel for introducing me to the idea of Efficient Method of Moments and for his constant advice. I am grateful to George Tauchen for his helpful remarks during my stay at Duke University. I also thank Esther Ruiz and seminar participants at the Universidad Complutense for their valuable comments. This research has been made possible thanks to the financial support of Fundação para a Ciência e Tecnologia. The usual disclaimer applies.

of return volatility that had as its starting point the ARCH model of Engle (1982). The main aim of this model was to fit volatility clustering and the fat tails of the return distributions. Posteriori models were able to deal with more complex features of financial time series data, e.g. the asymmetric responses of volatility to shocks and the persistence of volatility processes. With respect to the persistence of volatility Ding et al. (1993), de Lima and Crato (1994), and Bollerslev and Mikkelsen (1996) suggested that volatility could be model as a fractional integrated process. The existence of a fractional root in the volatility process would generate more persistence and consequently lead to autocorrelation functions of squared returns that decay slowly towards zero. In the light of this finding, Breidt et al. (1998) and Harvey (1998) proposed a new time series representation of persistence in conditional volatility denoted long memory stochastic volatility model (LMSV). The LMSV incorporates an ARFIMA process in a standard stochastic volatility scheme.

In this paper, we model volatility persistence by assuming that the volatility of asset returns shows a long memory feature captured by a fractionally integrated process as in Breidt et al. (1998) and Harvey (1998). We also incorporate a short-run factor into the volatility specification. This factor helps to generate extra kurtosis, to accommodate the volatility persistence and to obtain small values for the first-order autocorrelation of squared returns. The introduction of this extra volatility factor is justified by the results of Andersen et al. (2002), Chernov and Ghysels (2000), Chernov et al. (2003), Eraker et al. (2003), Gallant and Tauchen (2001), Jones (2003), Merville and Piepeta (1989) and Pan (2002). These authors found that stochastic volatility models with one volatility factor are not able to characterize all moments of asset return distributions. In particular, the fat tails of the return distribution are captured rather poorly. Additionally, Merville and Piepeta (1989) found that some shocks that drive volatility away from its long-term value are temporal because there is a force pulling volatility back to its long-term value. These two findings may suggest that volatility can be decomposed into a transitory and a persistent component.

More recently, the use of fractional integrated processes to generate persistence has been questioned. Liu (2000), Breidt et al. (1998), and Morana and Beltrati (2004) argued that persistence might be generated by a regime switching model. We justify our choice based on the results of Bollerslev and Mikkelsen (1996), who proved that fractional integration was an adequate proceeding for generating persistence. Moreover, our model (denoted 2FLMSV) has further advantages. First, it is well defined in the mean square sense and consequently many of its stochastic characteristics are easy to establish. Second, it generalizes the LMSV model and inherits most of its statistical properties. Finally, it incorporates the possibility of a leverage effect by allowing for a correlation between the returns and the changes in volatility, [see, Taylor (1994) and Harvey and Shephard (1996)].

In our empirical analysis, we test the performance of the 2FLMSV model by fitting it to the returns of the S&P 500 composite index. We use the efficient method of moments (EMM) of Gallant and Tauchen (1996) because of its testing advantages and the impossibility of applying maximum likelihood estimation in

models with latent variables.¹ Our results evidence that the 2FLMSV model fits volatility persistence and the fat tails of the distribution of returns better than the LMSV of Breidt et al. (1998), and Harvey (1998).

The paper is organized as follows: In the next Section, we present the models and derive the main statistical properties of the 2FLMSV model. We run a Monte Carlo experiment in section 3. Afterwards, we describe our estimation technique and report the corresponding results. Finally, we conclude. Proofs and Figures are relegated to the Appendix.

2 Long Memory Stochastic Volatility Models

2.1 The Asymmetric LMSV Model

In this Subsection we review the asymmetric LMSV model of Ruiz and Veiga (2006) and present some of its most important statistical properties. This model is an extension of the LMSV specification of Breidt et al. (1998) and Harvey (1998). Formally, let the return of a financial asset at time t , y_t , satisfy

$$y_t = \sigma_t \varepsilon_t, \quad \text{where} \quad \sigma_t = \sigma \exp(h_{1t}/2). \quad (1)$$

In equation 1, σ denotes a scale parameter, σ_t^2 is the conditional variance of y_t , ε_t is $NID(0, 1)$ and h_{1t} is a fractional integrated Gaussian noise process given by

$$(1 - L)^d h_{1t} = \varepsilon_t \quad \text{and} \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2). \quad (2)$$

In equation 2, L stands for the lag operator, d is the parameter of fractional integration and h_1 is an unobservable latent variable that is weakly stationary in the range $d \in (0, 0.5)$. Hosking (1981) showed that h_{1t} has the following $MA(\infty)$ representation for $d < 0.5$:

$$h_{1t} = (1 - L)^{-d} \varepsilon_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k},$$

where $\psi_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)}$ and $\Gamma(\cdot)$ is the Gamma function. Observe that the coefficients ψ_k converge hyperbolically to zero.²

¹Further estimation techniques are the procedure based on a spectral regression proposed by Geweke and Porter-Hudak (1983) and a frequency-domain estimator for the fractionally integrated stochastic volatility model (LMSV) suggested by Breidt et al. (1998). The latter estimator is consistent but the asymptotic distribution is not known. Wright (1999) proposed a new estimator for the LMSV model based on the minimum distance estimator (MDE) proposed by Tieslau et al. (1996). It consists of minimizing a quadratic distance function and estimates autocorrelations at various lags. The estimator is \sqrt{T} -consistent and asymptotically normal, provided that the parameter of fractional integration is smaller than 0.25.

²For $d > -1$ the binomial expansion of $(1-L)^d$ is given by: $(1-L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-L)^k = 1 - dL - \frac{d(1-d)}{2!}L^2 - \frac{d(1-d)(2-d)}{3!}L^3 - \dots$. Remember that $\left[\frac{\Gamma(a+x)}{\Gamma(b+x)}\right] \approx x^{a-b}$. So, when $k \rightarrow \infty$, $\psi_k \approx \frac{k^{d-1}}{\Gamma(d)}$.

Ruiz and Veiga (2006) assumed additionally that $(\varepsilon_t, \varepsilon_{t+1})'$ follow the bivariate normal distribution

$$\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t+1} \end{pmatrix} \sim NID \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \delta\sigma_\varepsilon \\ \delta\sigma_\varepsilon & \sigma_\varepsilon^2 \end{pmatrix} \right), \quad (3)$$

where δ , the correlation between ε_t and ε_{t+1} , induces correlation between returns and changes in volatility, [see, Taylor (1994) and Harvey and Shephard (1996)]. Therefore, if $\delta = 0$, the model simplifies to the LMSV specification of Breidt et al. (1998) and Harvey (1998). Moreover, note that the series of returns is a martingale difference and, consequently, an uncorrelated sequence.

Another interesting point is the behavior of the h_{1t} autocorrelation function (ACF). Baillie et al. (1996) computed the moments of h_{1t} from the properties of the lognormal distribution and obtained for $k \geq 1$ the following results:

$$\sigma_{h1}^2 = \sigma_\varepsilon^2 \frac{\Gamma(1-2d)}{[\Gamma(1-d)]^2},$$

$$\gamma(k) \equiv cov(h_{1t}, h_{1t+k}) = \sigma_\varepsilon^2 \frac{\Gamma(1-2d)\Gamma(k+d)}{\Gamma(d)\Gamma(1-d)\Gamma(k+1-d)}$$

and

$$\rho(k) \equiv corr(h_{1t}, h_{1t+k}) = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)} = \prod_{1 \leq i \leq k} \frac{i+d-1}{i-d}.$$

Note that $\rho(k)$ can be written as $\rho(k) \approx \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1}$ as $k \rightarrow \infty$. Since the autocorrelation function decays hyperbolically towards zero, the effect of a shock to volatility takes time to dissipate. This property of volatility is called long memory.³

Ruiz and Veiga (2006) derived the variance and the autocorrelation functions of y_t^2 for the case when y_t follows an asymmetric LMSV model, e.g. equations 1 to 3. They obtained for $k \geq 1$ that

$$var(y_t^2) = \sigma^4 \exp(\sigma_{h1}^2) [K_\varepsilon \exp(\sigma_{h1}^2) - 1]$$

and

$$corr(y_t^2, y_{t+k}^2) = \frac{\exp(\sigma_{h1}^2 \rho(k)) (1 + \delta^2 \sigma_\varepsilon^2 \psi_{k-1}^2) - 1}{K_\varepsilon \exp(\sigma_{h1}^2) - 1},$$

where K_ε is the kurtosis of ε_t . Furthermore, the excess of kurtosis of y_t was shown to be equal to $EK = 3[\exp(\sigma_{h1}^2) - 1]$. Figure 1 reports the autocorrelation

³We can say that a process y_t displays long memory if its spectrum $f(\lambda)$ has the following asymptotic decay: $f(\lambda) \approx C(\lambda) |\lambda|^{-2d}$ as $\lambda \rightarrow 0^+$ whenever $d \neq 0$ and $C(\lambda) \neq 0$. If $d > 0$, then the autocorrelations are not summable, i.e. $\lim_{n \rightarrow \infty} \sum_{k=-\infty}^n |\rho(k)| = \infty$, and the spectrum diverges at 0, [see, Breidt et al. (1998)]. Brockwell and Davies (1991) and Beran (1994) provided a more restrictive definition of long-memory that is more suitable for *ARFIMA*(p, d, q) processes.

functions of y_t^2 for two LMSV processes until lag 50. We observe that processes with larger d 's are simultaneously able to generate more persistence and capture more kurtosis.

2.2 The Asymmetric 2FLMSV Model

We propose a long memory stochastic volatility model with two volatility factors. It is denoted 2FLMSV. The objective of the first factor is to capture persistence in volatility and is similar in spirit to Breidt et al. (1998) and Harvey (1998)'s volatility process. The second factor accommodates the short run dynamics and helps generating extra kurtosis. Formally, we have now that

$$y_t = \varepsilon_t \sigma \exp\left(\frac{h_{1t} + h_{2t}}{2}\right), \quad (4)$$

where ε_t is again $NID(0, 1)$. As before, σ is a scale parameter. The fractional integrated process h_{1t} is the same as in the base model and thus given by equation 2. Moreover, we assume that h_{2t} follows the following AR(1) process:

$$h_{2t} = \phi h_{2t-1} + \eta_t, \quad (5)$$

where $\eta_t \sim NID(0, \sigma_\eta^2)$ and $|\phi| < 1$. The last condition guarantees the stationarity of the process. The errors η_t , ε_t and ε_t are contemporaneously independent for all t and it is assumed that h_1 and h_2 are unobservable latent variables. Finally, $(\varepsilon_t, \varepsilon_{t+1})'$ follow the bivariate normal distribution presented in equation 3. Hence, the 2FLMSV model is fully specified by the set of equations 4, 2, 5 and 3.

Next, we derive closed form expressions for the main moments and the autocorrelation functions of y_t^2 and $|y_t|$ in order to gain insights about the model's ability to capture the empirical dynamics of the returns distribution. In particular, we are interested in the second moment, because it turned out to be the dominant time varying feature of the return distribution.

Proposition 1 *In the 2FLMSV model, the first, the second and the fourth unconditional moment of y_t^2 exist and have the following form:*

$$E(y_t^2) = \sigma^2 \exp\left(\frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{2}\right),$$

$$E(y_t^4) = \sigma^4 K_\varepsilon \exp[2(\sigma_{h_1}^2 + \sigma_{h_2}^2)]$$

and

$$\text{var}(y_t^2) = \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) [K_\varepsilon \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) - 1].$$

Proof. See the Appendix. ■

It is easy to obtain from Proposition 1 a closed-form expression for the excess of kurtosis of y_t . It is equal to

$$EK = 3[\exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) - 1].$$

A direct consequence of this result is that the 2FLMSV model is able to generate higher kurtosis than the LMSV of Breidt et al. (1998) and Harvey (1998). Next, we calculate the autocorrelation function of y_t^2 .

Proposition 2 For $k \geq 1$, the autocorrelation function of y_t^2 in the 2FLMSV model is given by

$$\text{corr}(y_t^2, y_{t+k}^2) = \frac{\exp\left(\sigma_{h_1}^2 \rho(k) + \sigma_{h_2}^2 \phi^k\right) (1 + \delta^2 \sigma_\epsilon^2 \psi_{k-1}^2) - 1}{K_\epsilon \exp\left(\sigma_{h_1}^2 + \sigma_{h_2}^2\right) - 1}.$$

Proof. See the Appendix. ■

Figure 2 reports the autocorrelation functions of y_t^2 for several 2FLMSV processes. We observe that the 2FLMSV model generates extra kurtosis in comparison to the LMSV (e.g. if we compare the first panel of Figure 1 with the first four panels in Figure 2, then we see that the excess of kurtosis increased from 0.348 to at least 0.522). Additionally, we see that both models perform equally well in capturing persistence and that the 2FLMSV model creates smaller first order autocorrelations for some parameter values ($\{d = 0.45, \sigma_\epsilon^2 = 0.1, \sigma_\eta^2 = 0.1, \phi = 0.1\}$ and $\{d = 0.45, \sigma_\epsilon^2 = 0.05, \sigma_\eta^2, \phi = 0.1\}$).

Proposition 3 For $k \geq 1$, the k -th order autocorrelation of $|y_t|$ in the 2FLMSV model is given by

$$\text{corr}(|y_t|, |y_{t+k}|) = \frac{\exp\left(\frac{\sigma_{h_1}^2 \rho(k) + \sigma_{h_2}^2 \phi^k}{4}\right) \left(1 + \frac{\sqrt{\pi}}{\sqrt{2}} \frac{\delta \sigma_\epsilon}{2} \psi_{k-1}\right) - 1}{\frac{\pi}{2} \exp\left(\frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{4}\right) - 1}.$$

Proof. See the Appendix. ■

One empirical evidence of the ACF of the S&P 500 squared returns presented in Figure 7 is that the first order autocorrelation is smaller than the second order autocorrelation. Looking at Figure 3, we observe that the introduction of negative asymmetry leads the model to replicate this feature [see, e.g. the last four panels when $\delta = -0.2$]. Finally, Figure 4 shows the relationship between the first order autocorrelations of absolute and squared returns. The difference between the two is known as Taylor effect. When the correlation between the level and volatility noises is positive, i.e. $\delta > 0$, the Taylor effect is stronger for higher values of δ . This corresponds to area B in Figure 3. On the other hand, if $\delta < 0$, the Taylor effect disappears and the $\text{corr}(|y_t|, |y_{t+1}|)$ is always smaller than $\text{corr}(y_t^2, y_{t+1}^2)$. Note that if $\delta = 0$, then there is no asymmetry in the model.

3 A Monte Carlo Experiment

In this Section, we simulate several 2FLMSV processes with parameter values that reproduce the properties of real series of daily financial returns. For this

purpose, we generate 1000 replicates of size $T = 500, 1000$ and 5000 . Tables 1 and 2 report the empirical relative biases and standard deviations of the sample autocorrelations of squared returns for lags $k = 1, 10$ and 50 . For the majority of models the relative biases are severe. They are extremely severe for the autocorrelation of order 50 in models whose parameter d is 0.2. We also observe that these biases decrease dramatically with the sample size. The same happens with the Monte Carlo standard deviations. The best fits occur for models with the parameters values $\{\phi, d, \sigma_\epsilon^2, \sigma_\eta^2, \delta\} = \{0.5, 0.2, 0.1, 0.1, 0.5\}$ and $\{0.5, 0.2, 0.1, 0.1, 0.5\}$ and with a sample size of $T = 5000$. The existence of these biases may increase the difficulty of identifying long memory as it was argued by Pérez and Ruiz (2003). Finally, we also observe that the autocorrelation functions (in particular, the first order autocorrelations of squared observations) are in majority smaller than the respective theoretical values. Pérez and Ruiz (2003) found similar evidence for LMSV models.

4 Empirical Example

In this Section, we evaluate the performance of our model in capturing the empirical features of financial data. We use daily close price data on the S&P 500 composite index from January 3, 1928 to February 19, 2002. This makes up to a total of 18 609 observations.

Figure 5 plots the price level and the returns on the index (adjusted for dividends and splits) over the sample period. Figure 6 provides some summary statistics of the data. The average return is 0.022 per day and the daily variance is 1.3631. Moreover, the distribution of returns is negatively skewed and the kurtosis is also quite high.

Finally, we compute the correlograms of squared and absolute returns series (Figure 7). We verify that the autocorrelation function of squared observations converges faster towards zero than the autocorrelation function of absolute returns. Finally, we also observe that $\rho(1)$ is smaller than $\rho(2)$ in both cases.

4.1 Detecting Long Memory

In order to justify our suspicion that volatility follows a fractional integrated process, we use two main tests. The first one is the traditional R/S statistic corrected for short-memory components, [see, Lo (1991)]. The second one is a Wald type test in the time domain similar to the Dickey-Fuller approach, [see, Dolado et al. (2002)].

Table 3 reports the results of the R/S test. Since the relation between J and d is given by $J = d + 1/2$, [see, Mandelbrot and Taqqu (1979)] and the estimated value of J , \hat{J} , is larger than 0.5, we conclude that the parameter of fractional integration d is strictly positive.

Dolado et al. (2002) tested the null hypothesis of a fractional integrated process of order d_0 , $FI(d_0)$, versus a fractional integrated process of order d_1 , $FI(d_1)$, with $d_1 < d_0$. The test is the t-statistic associated to the coefficient of

$\Delta^{d_1}y_{t-1}$ in a regression of $\Delta^{d_0}y_t$ on $\Delta^{d_1}y_{t-1}$ and some lags of $\Delta^{d_0}y_t$. In our case, we consider two different null hypotheses: $d_0 = 0.3$ and $d_0 = 0.4$. The t-statistics are normally distributed because the process is stationary under the null hypothesis. The parameter d_1 is estimated by fitting an ARFIMA(1,d,0) to the squared returns series. Fractional integration is not rejected at the 5% significance level for the squared returns [see, Table 4].

4.2 Efficient Method of Moments (EMM)

Now, we estimate the LMSV and the 2FLMSV models using EMM of Gallant and Tauchen (1996). EMM is based on two compulsory phases: The first phase (Projection) 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. The simulated density is denominated the auxiliary model and its score is called "*the score generator for EMM*". The advantage of this method is that the score has an analytical expression. In the projection step, we proceed carefully along an expansion path with tree structure and the selected model comes out to be a semiparametric ARCH (auxiliar model), as in Gallant et al. (1997).⁴In the second phase the parameters of the models are estimated 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 models are obtained by minimizing the GMM criterion function. Finally, EMM provides us diagnostic tests that explain the reasons for the failure or success of a model.

4.3 Empirical Results

We start by estimating the benchmark model, the LMSV model of Breidt et al. (1998) and Harvey (1998). The estimated specification is a slightly modified version of the model presented in Subsection 2.1. Following Gallant et al. (1997) we have that

$$y_t - \mu_y = c_1(y_{t-1} - \mu_y) + c_2(y_{t-2} - \mu_y) + r_y \varepsilon_t \exp(h_{1t}/2),$$

and

$$(1 - L)^d h_{1t} = r_{h1} \tilde{\varepsilon}_t.$$

At this instance, we suppose that errors are Gaussian; that is, ε_t is $NID(0, 1)$, $\tilde{\varepsilon}_t$ is $NID(0, 1)$, h_{1t} is stationary and ε_t and $\tilde{\varepsilon}_t$ are mutually independent for all t . The change in errors notation helps detecting which parameters are separately

⁴The score generator is the semiparametric density (SNP) proposed by Gallant and Tauchen (1989) with the following tuning parameters: $L_u = 2$, which means two lags in the linear part of the SNP model, $L_r = 28$ that corresponds to twenty eight lags in the ARCH part, $L_p = 1$, one lag in the polynomial part and finally $K_z = 8$, which corresponds to a polynomial part of degree 4 in z . Therefore, the selected auxiliary model is a semiparametric ARCH that accounts for the full complexity of the data.

identified and the parallelism is easy to establish since $r_{h1}\tilde{\varepsilon}_t$ corresponds to ε_t , in our previous notation. We also introduce two lags of the dependent variable because time-series are usually correlated. Finally, μ_y denotes the mean of y_t .

We use the same estimation procedure of Gallant et al. (1997). Since the fractional integrated process of equation 2 can be written as a moving average of infinite order for $|d| < 0.5$, that is

$$h_{1t} = (1 - L)^{-d}\varepsilon_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} \quad \text{with} \quad \psi_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)},$$

and the Cholesky factorization of the covariance matrix of h_{1t} is impossible to compute, we truncate the infinite moving average and trim off the first 10 000 realizations. Consequently, due to this truncation procedure, the generated process is going to be stationary for $|d| < 1$ but, it is still able to generate high volatility persistence, [as it has been argued by Bollerslev and Mikkelsen (1996)].

Table 5 reports the results of the specification test: the null hypothesis of correct specification is sharply rejected. Looking at the EMM quasi-t-ratios \widehat{T}_n plotted in Figure 8 we observe that the model has difficulties in matching the features of the polynomial part of the SNP score (a_{20} until a_{90}). This means that either the specification $\exp(h_{1t}/2)$ is incorrect and/or ε_t is not Gaussian. We also observe that the scores of the ARCH specification (r_{24} until r_{28}) are bigger than 2. This may indicate that the conditional variance is poorly fitted.

Since the exponential transformation does not seem to be a problem in Gallant et al. (1997), we apply a spline error transformation to the Gaussian innovation in order to improve the fit of the polynomial part of the SNP score. The model becomes,

$$y_t - \mu_y = c_1(y_{t-1} - \mu_y) + c_2(y_{t-2} - \mu_y) + r_y \varepsilon_t \exp(h_{1t}/2),$$

$$\varepsilon_t = T_z(z_t),$$

$$T_z(z_t) = b_{z0}(b_c, b_d) + b_{z1}(b_c, b_d)z_t + b_{z2}(b_c, b_d)z_t^2 + b_{z3}(b_c, b_d)z_t \max(0, z_t)$$

and

$$(1 - L)^d h_{1t} = r_{h1}\tilde{\varepsilon}_t.$$

Liu (2000) defined b_{z0} , b_{z1} , b_{z2} and b_{z3} as functions of b_c and b_d such that if b_c and b_d are equal to 0, $b_{z0} = b_{z2} = b_{z3} = 0$ and $b_{z1} = 1$. Moreover, he used $b_{z0} = (b_c + 0.5b_d)/s$, $b_{z1} = 1/s$, $b_{z2} = b_c/s$ and $b_{z3} = b_d/s$, where $s = (b_c + 0.5b_d)^2 + 1 + 3b_c^2 + 1.5b_d^2 + 2[(b_c + 0.5b_d)b_c + 0.5(b_c + 0.5b_d)b_d + 0.7979b_d + 1.5b_d b_c]$. These restrictions on the b 's allow to identify all the parameters of the model and make the expected value of $T_z(z)$ and its variance to be 0 and 1, respectively.

Table 6 and Figure 8 reveal that this change reduces the value of the EMM objective function. The moments of the polynomial part of the SNP score are now better fitted. Although there is an improvement, the model is still not able to fit all the kurtosis of data. With respect to the scores of the ARCH

specification, we encounter the same problem as before and therefore the idea of a possible misspecification in the conditional variance process is reinforced.

We have proved in Subsection 2.2 that the introduction of an extra volatility factor allows the model to generate extra kurtosis. Having in mind this purpose, we estimate the following specifications

$$\begin{aligned}
y_t - \mu_y &= c_1(y_{t-1} - \mu_y) + c_2(y_{t-2} - \mu_y) + r_y \varepsilon_t \exp\left(\frac{h_{1t} + h_{2t}}{2}\right), \\
\varepsilon_t &= T_z(z_t), \\
T_z(z_t) &= b_{z0}(b_c, b_d) + b_{z1}(b_c, b_d)z_t + b_{z2}(b_c, b_d)z_t^2 + b_{z3}(b_c, b_d)z_t \max(0, z_t), \\
(1 - L)^d h_{1t} &= r_{h1}(\tilde{\varepsilon}_t + \delta \varepsilon_{t-1}), \text{ with } \delta = 0 \text{ and } \delta \neq 0 \\
&\text{and} \\
h_{2t} &= \phi_1 h_{2t-1} + \phi_2 h_{2t-2} + r_{h2} \tilde{\varepsilon}_t.
\end{aligned}$$

We impose the same restrictions as before to the spline error, $T_z(z_t)$ in order to identify the parameters of the model.

The empirical results for the non-asymmetric 2FLMSV model are reported in Tables 6 and 7 and Figure 9. They reveal that, (1) the fit is better, (2) the fatness of the tails are much better accommodated and (3) the model is able to deal much better with volatility persistence than the LMSV model with spline errors. Looking carefully at the quasi t-ratios we observe that the majority of them are smaller than 2 in absolute value, and therefore, the model does better than the corresponding benchmark model. Furthermore, all coefficients have the expected signs and are statistical significant at the 5% significance level. Nevertheless, we still reject the specification. Finally, we re-estimate the spline 2FLMSV specification introducing a leverage effect. In Tables 6 and 7 we observe that the fit does not improve much and that the estimate of leverage effect, δ , is not significant. This is not really a surprise, because the spline transformation introduces already an asymmetry into the model.

The fact that the 2FLMSV specification is still rejected by the specification test can be due to an over rejection problem that characterizes the χ^2 test. Chumacero (1997) studies the small sample properties of EMM estimators for the ARSV model with the help of a Monte Carlo experiment and confirms that inference based on the over identifying restrictions test as well as other χ^2 statistics show important over rejections. In fact, if we compute the reprojected volatility using the reprojection technique of Gallant and Tauchen (1998) [see, Figure10] and compare it to the one-step-ahead conditional volatility computed on the observed data, then we observe that the reprojected volatility closely encompasses the empirical volatility without missing the volatility cycles of S&P 500 (the initial high volatility, the period of low volatility at the middle of the sample, the stock market crash of October 1987 and the high volatility at the end of sample).⁵ This is important because the main role of these models is to pro-

⁵The reprojected volatility is calculated with the reprojection technique of Gallant and Tauchen (1998). Thus, as a by-product of the estimation step, we obtain a long simulation of

duce accurate future values of volatility that can be applied in areas such as risk management and asset pricing.

5 Conclusion

In this paper, we propose a two factor long memory stochastic volatility model (2FLMSV) as an alternative to the LMSV model of Breidt et al. (1998) and Harvey (1998). We still model volatility persistence by assuming that the volatility of returns shows a long memory feature captured by a fractionally integrated process. The innovation is that we introduce a short run volatility factor that allows the model to generate extra kurtosis and to accommodate volatility persistence.

In the first step of our analysis, we derive the most important moments and the autocorrelation functions of squares and absolute values of y_t (y_t follows a 2FLMSV process). Afterwards, we apply the efficient method of moments of Gallant and Tauchen (1996) in order to compare the 2FLMSV empirically with the LMSV model.

Our results evidence that the short run volatility factor seems to improve the EMM criterion (a similar result was found by Liu (2000)) and that the long memory stochastic volatility model with two factors of volatility creates more kurtosis than the benchmark model.

References

- [1] Andersen, T. G., Benzoni, L. and Lund, J. (2002): An Empirical Investigation of Continuous-Time Equity Return Models, *Journal of Finance* **57**, 1239-1284.
- [2] Baillie, R., Bollerslev, T. and Mikkelsen, H.O. (1996): Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* **52**, 91-113.
- [3] Beran, J. (1994): Statistics for Long-Memory Processes, Chapman and Hal, New York.

y_t at the estimated parameter vector with a simulation length equal 100 000. If we impose a SNP-GARCH model on the simulated values of y_t (\hat{y}_t), we obtain a good representation of the one-step-ahead conditional variance that we denoted $\hat{\sigma}_t^2$. Therefore, regressing $\hat{\sigma}_t^2$ on lags of $\hat{\sigma}_t^2$, \hat{y}_t , $|\hat{y}_t|$, such as $\hat{\sigma}_t^2 = \alpha_0 + \alpha_1 \hat{\sigma}_{t-1}^2 + \dots + \alpha_p \hat{\sigma}_{t-p}^2 + \theta_1 \hat{y}_{t-1} + \dots + \theta_q \hat{y}_{t-q} + \pi_1 |\hat{y}_{t-1}| + \dots + \pi_r |\hat{y}_{t-r}| + u_t$, gives us a calibrated function inside the simulation that provides predicted values of the conditional variance. Then, if we evaluate this function on the observed data series we obtain estimates of the conditional variance, $\hat{\sigma}_t^{*2}$.

The empirical volatility (the volatility obtained directly from the data) is computed by taking the square root of a moving average of squared residuals, $[(m+1)^{-1} \sum_{j=0}^m \hat{e}_{t-j}^2]^{1/2}$, $m = 4$ or $m = 26$, from the estimation of the AR(1) model $y_t = \alpha_0 + \alpha_1 y_{t-1} + e_t$, [see, Gallant and Tauchen (1998)].

- [4] Bollerslev, T. and Mikkelsen, H.O. (1996): Modeling and Pricing Long Memory in Stock Market Volatility, *Journal of Econometrics* **73**, 151-184.
- [5] Breidt, F., Crato, N. and de Lima, P.J.F. (1998): On The Detection and Estimation of Long Memory in Stochastic Volatility, *Journal of Econometrics* **83**, 325-348.
- [6] Brockwell, P.J. and Davis, R.A. (1991): Time Series: Theory and Methods, 2nd ed., Springer Verlag, Berlin.
- [7] Chernov, M., Gallant, A. R., Ghysels, E. and Tauchen, G. (2003): Alternative Models for Stock Price Dynamics, *Journal of Econometrics* **116**, 225-257.
- [8] Chernov, M. and Ghysels, E. (2000): A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation, *Journal of Financial Economics* **56**, 407-458.
- [9] Chumacero, R. A. (1997): Finite Sample Properties of the Efficient Method of Moments, *Studies in Nonlinear Dynamics and Econometrics* **2**, n° 2.
- [10] de Lima, P.J.F. and Crato, N. (1994): Long-Range Dependence in the Conditional Variance of Stock Returns, *Economic Letters* **45**, 281-285.
- [11] Ding, Z., Engle, R.F. and Granger, C.W.J. (1993): A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance* **1**, 83-106.
- [12] Dolado, J., Gonzalo J. and Mayoral, L. (2002): A Fractional Dickey-Fuller Test for Unit Roots, *Econometrica* **70**, 1963-2006.
- [13] Engle, R.F. (1982): Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation, *Econometrica* **50**, 987-1008.
- [14] Eraker, B., Johannes, M. and Polson, N. (2003): The Impact of Jumps in Returns and Volatility, *Journal of Finance* **53**, 1269-1300.
- [15] Gallant, A.R., Hsieh, D. and Tauchen, G. (1997): Estimation of Stochastic Volatility Models with Diagnostics, *Journal of Econometrics* **81**, 159-192.
- [16] Gallant, A. R. and Tauchen, G. (1996): Which Moments to Match, *Econometric Theory* **12**, 657-681.
- [17] 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.
- [18] Gallant, A. R. and Tauchen, G. (2001): Efficient Method of Moments, *Discussion Paper*, University of North Carolina at Chapel Hill.

- [19] Geweke, J. and Porter-Hudak, S. (1983): The Estimation and Application of Long-Memory Time Series Models, *Journal of Time Series* **4**, 221-238.
- [20] Harvey, A. (1998): Long Memory in Stochastic Volatility, In J. Knight and S. Satchell (eds.), *Forecasting Volatility in Financial Markets*. Oxford: Butterworth-Heinemann.
- [21] Harvey A.C. and Shephard, N.G. (1996): Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns, *Journal of Business and Economic Statistics* **14**, 429-434.
- [22] Hosking, J.R.M. (1981): Fractional Differencing, *Biometrika* **68**, 165-176.
- [23] Jones, C. (2003): The Dynamics of Stochastic Volatility, *Journal of Econometrics* **116**, 181-224.
- [24] Liu, M. (2000): Modeling Long Memory in Stock Market Volatility, *Journal of Econometrics* **99**,139-171.
- [25] Lo, A.W. (1991): Long Term Memory in Stock Market Prices, *Econometrica* **59**, 1279-1313.
- [26] 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.
- [27] Merville, L.J. and Pieptea, D.R. (1989): Stock-Price Volatility, Mean-Reverting Diffusion and Noise, *Journal of Financial Economics* **24**, 193-214.
- [28] Morana, C. and Beltratti, A. (2004): Structural Change and Long-Range Dependence in Volatility If Exchange Rates: Either, Neither of Both?, *Journal of Empirical Finance* **11**, 629-658.
- [29] Pan, J. (2002): The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, *Journal of Financial Economics* **63**, 3-50.
- [30] Pérez, A. and Ruiz, E. (2003): Properties of the Sample Autocorrelations of Nonlinear Transformations in Long-Memory Stochastic Volatility Models, *Journal of Financial Econometrics* **1**, 420-444.
- [31] Ruiz, E. and Veiga, H. (2006): Modelling Long-Memory Volatilities with Leverage Effect: A-LMSV Versus FIEGARCH, *Manuscript*, Universidad Carlos III de Madrid.
- [32] Taylor, S.J. (1994): Modelling Stochastic Volatility: A Review and Comparative Study, *Mathematical Finance* **4**, 183-204.
- [33] Tieslau, M.A., Schmidt, P. and Baillie, R.T. (1996): A Minimum Distance Estimator for Long-Memory Processes, *Journal of Econometrics* **71**, 249-264.

[34] Wright, J. H. (1999): A New Estimator of the Fractionally Integrated Stochastic Volatility Model, *Economics Letters* **63**, 295-303.

Appendix: Proofs

Proof of Proposition 1

(1) Taking squares in equation 4 yields $y_t^2 = \varepsilon_t^2 \sigma^2 \exp(h_{1t} + h_{2t})$. Now, we apply expectations to the latter equation to obtain that

$$E(y_t^2) = E[\varepsilon_t^2 \sigma^2 \exp(h_{1t} + h_{2t})] = \sigma^2 \exp\left(\frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{2}\right).$$

The last step of the former calculations follows because $E[\exp(h_{1t})] = \exp\left(\frac{\sigma_{h_1}^2}{2}\right)$, $E[\exp(h_{2t})] = \exp\left(\frac{\sigma_{h_2}^2}{2}\right)$ and the processes h_{1t} and h_{2t} are not correlated. \square

(2) Taking both sides of equation 4 to the power 4 and applying expectations afterwards yields

$$E(y_t^4) = E[\varepsilon_t^4 \sigma^4 \exp(0.5(h_{1t} + h_{2t}))^4] = E(\varepsilon_t^4) \sigma^4 E[\exp(2(h_{1t} + h_{2t}))].$$

Finally, since $E(\varepsilon_t^4) = K_\varepsilon$ and $E[\exp(2(h_{1t} + h_{2t}))] = \exp\left(2^2 \frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{2}\right)$ we obtain that $E(y_t^4) = K_\varepsilon \sigma^4 \exp[2(\sigma_{h_1}^2 + \sigma_{h_2}^2)]$. \square

(3) We have that $\text{var}(y_t^2) = E(y_t^4) - E(y_t^2)^2$. If we replace $E(y_t^4)$ and $E(y_t^2)$ by the expressions obtained in part (1) and (2) of this proof, then we see that

$$\begin{aligned} \text{var}(y_t^2) &= \sigma^4 K_\varepsilon \exp[2(\sigma_{h_1}^2 + \sigma_{h_2}^2)] - \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) \\ &= \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) [K_\varepsilon \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) - 1]. \end{aligned}$$

This completes the proof of Proposition 1. \square

Proof of Proposition 2

Taking squares in equation 4 yields $y_t^2 = \varepsilon_t^2 \sigma^2 \exp(h_{1t} + h_{2t})$. If we apply that $\text{cov}(y_t^2, y_{t+k}^2) = E(y_t^2 y_{t+k}^2) - E(y_t^2)E(y_{t+k}^2)$, then we obtain

$$\begin{aligned} \text{cov}(y_t^2, y_{t+k}^2) &= E[\varepsilon_t^2 \sigma^2 \exp(h_{1t} + h_{2t}) \varepsilon_{t+k}^2 \sigma^2 \exp(h_{1t+k} + h_{2t+k})] - \\ &\quad - E[\varepsilon_t^2 \sigma^2 \exp(h_{1t} + h_{2t})] E[\varepsilon_{t+k}^2 \sigma^2 \exp(h_{1t+k} + h_{2t+k})]. \end{aligned}$$

We use part (1) of Proposition 1 to yield that

$$\begin{aligned} \text{cov}(y_t^2, y_{t+k}^2) &= E[\varepsilon_t^2 \sigma^2 \exp(h_{1t} + h_{2t}) \varepsilon_{t+k}^2 \sigma^2 \exp(h_{1t+k} + h_{2t+k})] - \\ &\quad - \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2). \end{aligned} \quad (6)$$

In the next step we develop the expression $E(y_t^2 y_{t+k}^2)$. We start by adding and subtracting $\psi_{k-1} \epsilon_{t+1}$ to the process h_{1t} . We obtain that $E(y_t^2 y_{t+k}^2)$ is equal to

$$\sigma^4 E[\varepsilon_t^2 \exp(h_{1t} + h_{1t+k} - \psi_{k-1} \epsilon_{t+1} + \psi_{k-1} \epsilon_{t+1}) \exp(h_{2t} + h_{2t+k}) \varepsilon_{t+k}^2].$$

If we apply then that $E(\varepsilon_{t+k}^2) = 1$, $E[\exp(h_{2t} + h_{2t+k})] = \exp(\sigma_{h_2}^2 + \sigma_{h_2}^2 \phi^k)$ and $E[\exp(h_{1t} + h_{1t+k} - \psi_{k-1} \epsilon_{t+1})] = \exp(\sigma_{h_1}^2 + \sigma_{h_1}^2 \rho_{h_1}(k) - 0.5 \psi_{k-1}^2 \sigma_\epsilon^2)$, then we can reduce the former equation to

$$\begin{aligned} E(y_t^2 y_{t+k}^2) &= \sigma^4 E[\varepsilon_t^2 \exp(\psi_{k-1} \epsilon_{t+1})] \exp(\sigma_{h_1}^2 + \sigma_{h_1}^2 \rho_{h_1}(k) - 0.5 \psi_{k-1}^2 \sigma_\epsilon^2) \cdot \\ &\quad \cdot \exp(\sigma_{h_2}^2 + \sigma_{h_2}^2 \phi^k). \end{aligned}$$

Now, observe that $E[\varepsilon_t^2 \exp(\psi_{k-1} \epsilon_{t+1})] = E[E(\varepsilon_t^2 | \epsilon_{t+1}) \exp(\psi_{k-1} \epsilon_{t+1})]$. Since the process $\varepsilon_t | \epsilon_{t+1}$ is distributed $N(\epsilon_{t+1} \cdot \delta / \sigma_\epsilon, 1 - \delta^2)$, we can calculate immediately that $E(\varepsilon_t^2 | \epsilon_{t+1}) = (1 - \delta^2) + \frac{\delta^2}{\sigma_\epsilon^2} \epsilon_{t+1}^2$. Consequently, the expectation $E[E(\varepsilon_t^2 | \epsilon_{t+1}) \exp(\psi_{k-1} \epsilon_{t+1})]$ has to be equal to $(1 - \delta^2) \exp(\psi_{k-1}^2 \frac{\sigma_\epsilon^2}{2}) + \frac{\delta^2}{\sigma_\epsilon^2} E[\epsilon_{t+1}^2 \exp(\psi_{k-1} \epsilon_{t+1})]$. Once we have calculated the last expectation we can see that $E[E(\varepsilon_t^2 | \epsilon_{t+1}) \exp(\psi_{k-1} \epsilon_{t+1})] = \exp(\psi_{k-1}^2 \frac{\sigma_\epsilon^2}{2}) (1 + \delta^2 \sigma_\epsilon^2 \psi_{k-1}^2)$. We can now replace the expectation in the original expression. As a result, we observe that

$$\begin{aligned} E(y_t^2 y_{t+k}^2) &= \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_1}^2 \rho_{h_1}(k) - \psi_{k-1}^2 \frac{\sigma_\epsilon^2}{2}) \exp(\psi_{k-1}^2 \frac{\sigma_\epsilon^2}{2}) \cdot \\ &\quad \cdot (1 + \delta^2 \sigma_\epsilon^2 \psi_{k-1}^2) \exp(\sigma_{h_2}^2 + \sigma_{h_2}^2 \phi^k). \end{aligned}$$

After rearranging terms we see that

$$E(y_t^2 y_{t+k}^2) = \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) \exp(\sigma_{h_1}^2 \rho_{h_1}(k) + \sigma_{h_2}^2 \rho_{h_2} \phi^k) (1 + \delta^2 \sigma_\epsilon^2 \psi_{k-1}^2).$$

If use the last expression in equation 6, we obtain for $k \geq 1$ that

$$\begin{aligned} \text{cov}(y_t^2, y_{t+k}^2) &= \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) \exp(\sigma_{h_1}^2 \rho_{h_1}(k) + \sigma_{h_2}^2 \rho_{h_2} \phi^k) \cdot \\ &\quad \cdot (1 + \delta^2 \sigma_\epsilon^2 \psi_{k-1}^2) - \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) \\ &= \sigma^4 \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) \cdot \\ &\quad \cdot [\exp(\sigma_{h_1}^2 \rho_{h_1}(k) + \sigma_{h_2}^2 \rho_{h_2} \phi^k) (1 + \delta^2 \sigma_\epsilon^2 \psi_{k-1}^2) - 1] \end{aligned}$$

For $k \geq 1$, the autocorrelation function is then given by

$$\text{corr}(y_t^2, y_{t+k}^2) = \frac{\exp(\sigma_{h_1}^2 \rho_{h_1}(k) + \sigma_{h_2}^2 \rho_{h_2} \phi^k)(1 + \delta^2 \sigma_\epsilon^2 \psi_{k-1}^2) - 1}{K_\epsilon \exp(\sigma_{h_1}^2 + \sigma_{h_2}^2) - 1}.$$

This completes the proof of Proposition 2. \square

Proof of Proposition 3

This proof follows the very same steps as the proof of Proposition 2. Additionally, it is required to know that the expected value of a chi-squared variable X with v degrees of freedom to the power a , $E(X_v^a)$, and the expected value of the absolute value of random variable $Y \sim N(0, 1)$ to the power b , $E(|Y|^b)$, are equal to $2^a \frac{\Gamma(\frac{a+\frac{v}{2}}{2})}{\Gamma(\frac{v}{2})}$ and $E(X_1^{b/2})$, respectively. Using these results we obtain after some computations that

$$E(|y_t|) = \sigma E\left(|\varepsilon_t| \exp\left(\frac{h_{1t} + h_{2t}}{2}\right)\right) = \sigma \frac{\sqrt{2}}{\sqrt{\pi}} \exp\left(\frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{8}\right)$$

and

$$\text{var}(|y_t|) = \sigma^2 \frac{2}{\pi} \exp\left(\frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{4}\right) \left[\frac{\pi}{2} \exp\left(\frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{4}\right) - 1 \right].$$

For $k \geq 1$, the covariance function for $|y_t|$ is computed in the same way as the one in the proof of Proposition 2. The difference is that now $E[|\varepsilon_t| \exp(\psi_{k-1} \varepsilon_{t+1})] = E[E(|\varepsilon_t| | \varepsilon_{t+1}) \exp(\psi_{k-1} \varepsilon_{t+1})] = \exp(\psi_{k-1}^2 \frac{\sigma_\epsilon^2}{8}) (\frac{\sqrt{2}}{\sqrt{\pi}} + \frac{\delta \sigma_\epsilon}{2} \psi_{k-1})$. Consequently, we yield

$$\text{cov}(|y_t|, |y_{t+k}|) = \sigma^2 \left(\frac{2}{\pi} + \frac{\sqrt{2}}{\sqrt{\pi}} \frac{\delta \sigma_\epsilon}{2} \psi_{k-1} \right) \exp\left(\frac{\sigma_{h_1}^2 (1 + \rho(k)) + \sigma_{h_2}^2 (1 + \phi^k)}{4}\right) - \frac{2}{\pi}$$

and

$$\text{corr}(|y_t|, |y_{t+k}|) = \frac{\exp\left(\frac{\sigma_{h_1}^2 \rho(k) + \sigma_{h_2}^2 \phi^k}{4}\right) \left(1 + \frac{\sqrt{\pi}}{\sqrt{2}} \frac{\delta \sigma_\epsilon}{2} \psi_{k-1}\right) - 1}{\frac{\pi}{2} \exp\left(\frac{\sigma_{h_1}^2 + \sigma_{h_2}^2}{4}\right) - 1}.$$

This completes the proof of Proposition 3. \square

Figures and Tables

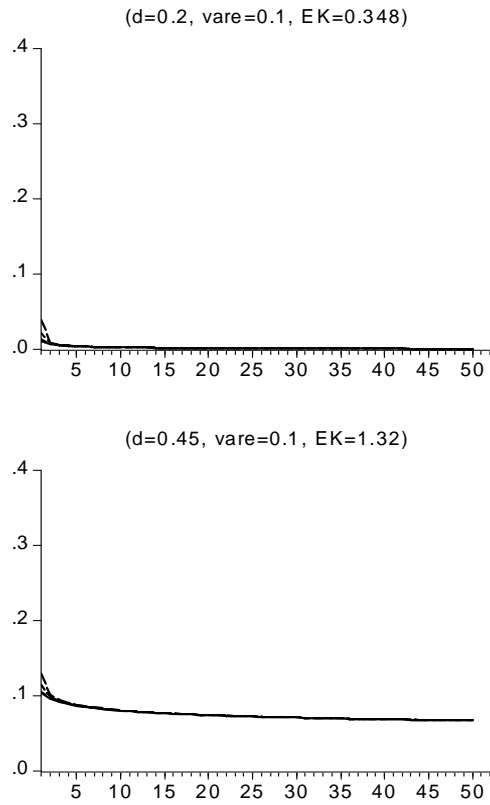


Figure 1: Autocorrelations of y_t^2 in LMSV processes. (continuous line ($\delta = 0$), dotted ($\delta = 0.2$), dotted discontinuous ($\delta = 0.5$) and discontinuous ($\delta = 0.8$)).

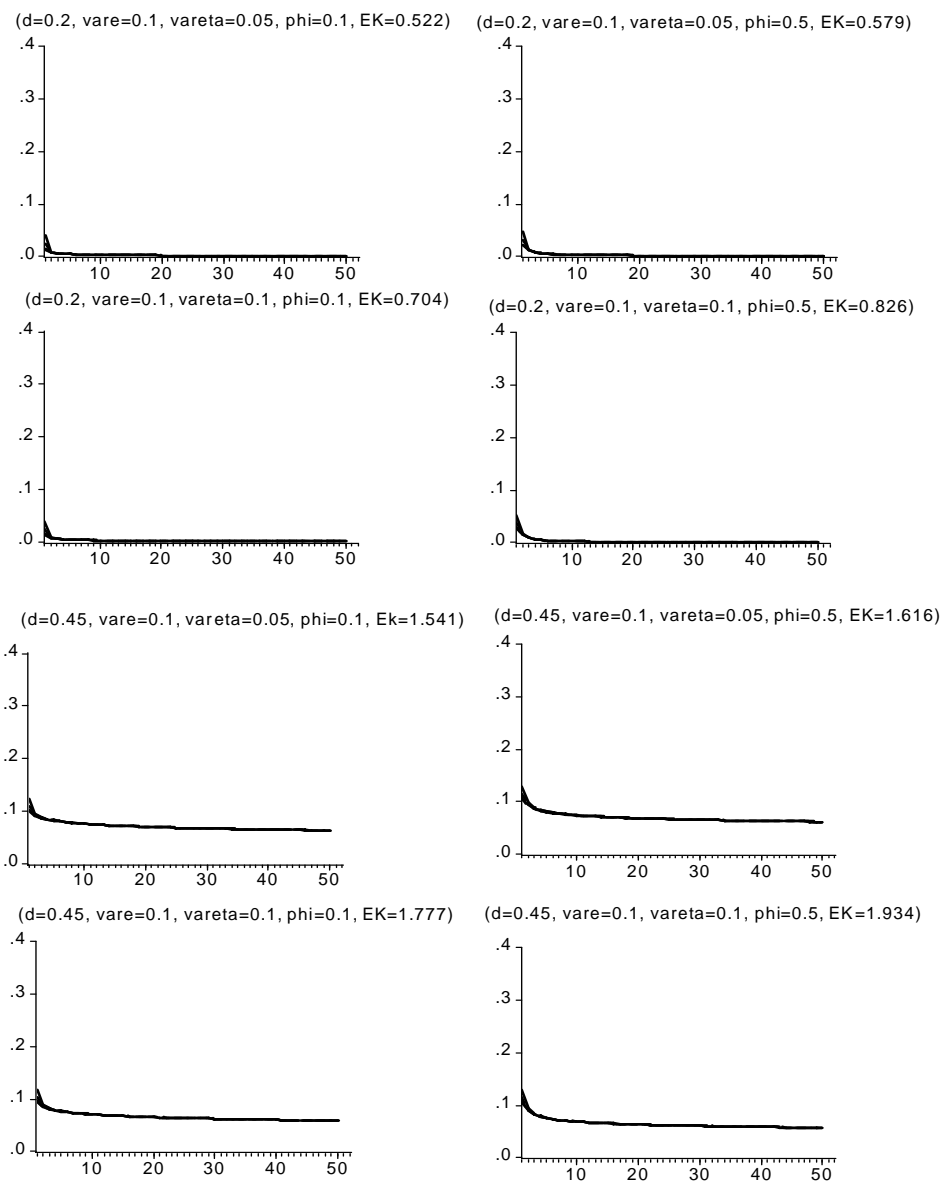


Figure 2: Autocorrelations of y_t^2 in 2FLMSV processes. (continuous line ($\delta = 0$), dotted ($\delta = 0.2$), dotted discontinuous ($\delta = 0.5$) and discontinuous ($\delta = 0.8$)).

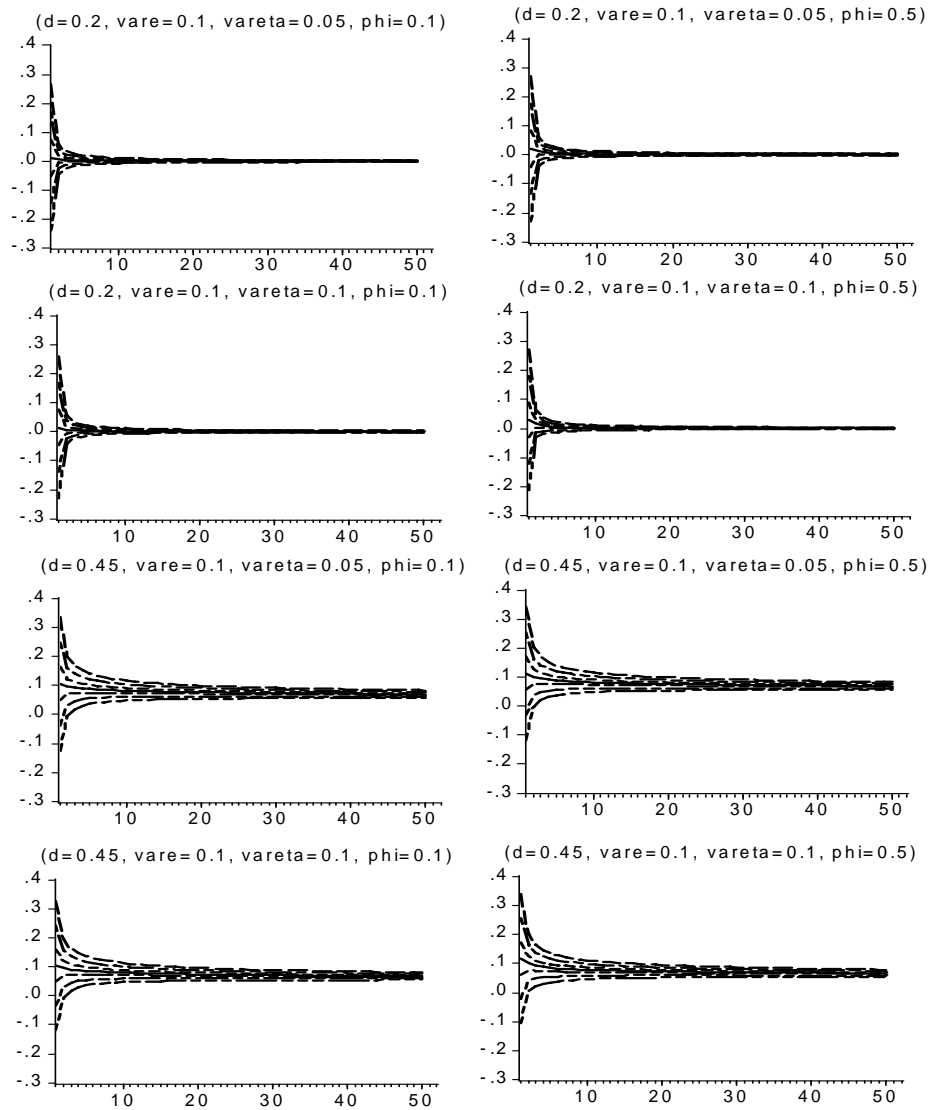


Figure 3: Autocorrelations of $|y_t|$ in 2FLMSV processes. (continuous line ($\delta = 0$), dotted ($\delta = 0.2$), dotted discontinuous ($\delta = 0.5$), discontinuous ($\delta = 0.8$), two dotted discontinuous ($\delta = -0.2$), large discontinuous ($\delta = -0.5$) and three dotted discontinuous ($\delta = -0.8$)).

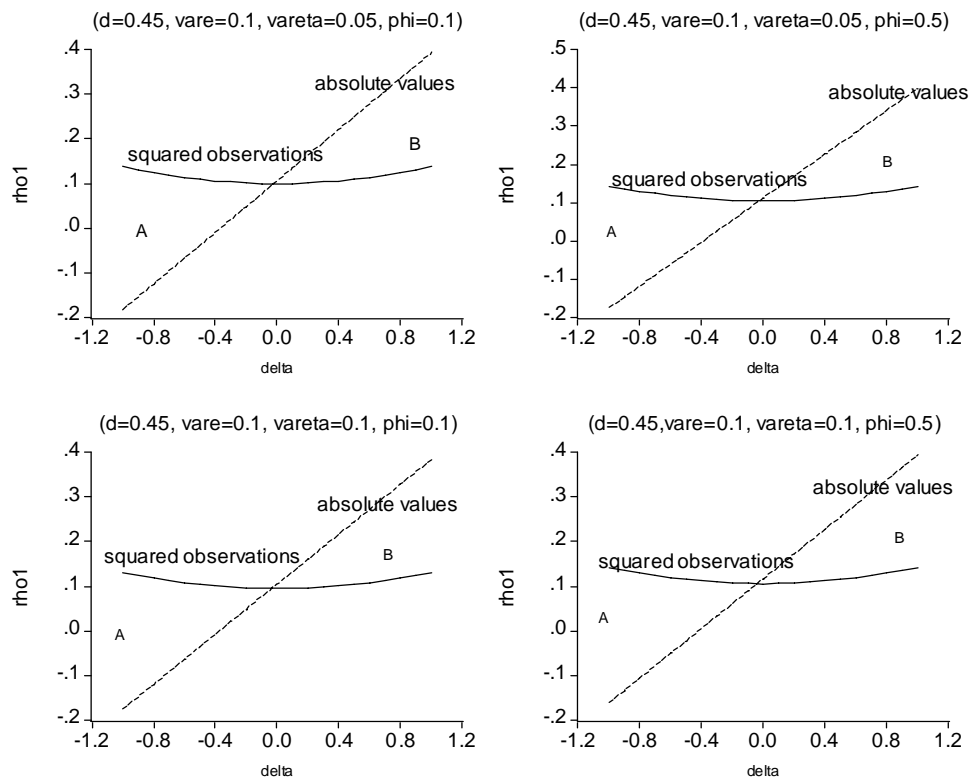


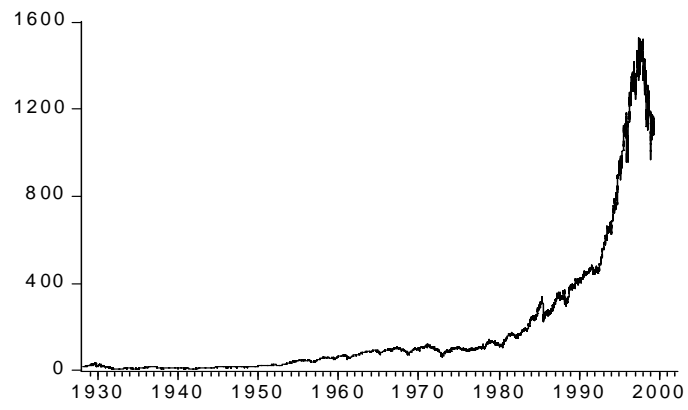
Figure 4: The relationship between the autocorrelation of order 1 and δ for several 2FLMSV models.

		T=500			T=1000			T=5000		
$\{\phi, d, \sigma_\epsilon^2, \sigma_\eta^2, \delta\}$	Series Lag	ACF	R. Bias	Std. dev.	R. Bias	Std. dev.	R. Bias	Std. dev.		
$\{0.1, 0.2, 0.1, 0.05, 0\}$	y_t^2	1	0.0131	-0.8535	0.0456	-0.6686	0.0331	-0.6444	0.0141	
		10	0.0028	-0.0012	0.0436	-0.5681	0.0302	-0.5592	0.0143	
		50	0.0011	-3.4123	0.0434	-2.6413	0.0303	-0.8087	0.0143	
$\{0.5, 0.2, 0.1, 0.05, 0\}$	y_t^2	1	0.0209	-0.3223	0.0482	-0.1949	0.0347	-0.1697	0.0148	
		10	0.0027	-1.4899	0.0435	-0.5579	0.0304	-0.5281	0.0142	
		50	0.0010	-3.2604	0.0433	-2.7189	0.0304	-0.8274	0.0142	
$\{0.1, 0.2, 0.1, 0.1, 0\}$	y_t^2	1	0.0141	-0.7531	0.0455	-0.5795	0.0332	-0.5377	0.0142	
		10	0.0026	-1.5542	0.0432	-0.6058	0.0304	-0.5344	0.0143	
		50	0.0010	-3.5666	0.0434	-2.9295	0.0300	-0.8261	0.0142	
$\{0.5, 0.2, 0.1, 0.1, 0\}$	y_t^2	1	0.0285	-0.0858	0.0505	0.0203	0.0365	0.0015	0.0157	
		10	0.0025	-1.6517	0.0432	-0.6008	0.0316	-0.0012	0.0141	
		50	0.0009	-3.4477	0.0433	-3.0784	0.0302	-0.0008	0.0142	
$\{0.1, 0.45, 0.1, 0.05, 0\}$	y_t^2	1	0.1000	-0.8376	0.0487	-0.7947	0.0349	-0.7842	0.0156	
		10	0.0757	-0.9163	0.0453	-0.8549	0.0323	-0.8450	0.0149	
		50	0.0633	-1.0009	0.0444	-0.9682	0.0301	-0.9194	0.0146	
$\{0.5, 0.45, 0.1, 0.05, 0\}$	y_t^2	1	0.1054	-0.6408	0.0534	-0.5959	0.0380	-0.5835	0.0171	
		10	0.0742	-0.9220	0.0450	-0.8566	0.0327	-0.8459	0.0149	
		50	0.0620	-1.0001	0.0441	-0.9702	0.0303	-0.9220	0.0146	
$\{0.1, 0.45, 0.1, 0.1, 0\}$	y_t^2	1	0.0955	-0.6875	0.0522	-0.6238	0.0373	-0.6041	0.0173	
		10	0.0710	-0.8021	0.0475	-0.7091	0.0346	-0.6879	0.0160	
		50	0.0593	-0.9637	0.0453	-0.9044	0.0306	-0.8352	0.0150	
$\{0.5, 0.45, 0.1, 0.1, 0\}$	y_t^2	1	0.1058	-0.5238	0.0567	-0.4627	0.0404	-0.4430	0.0189	
		10	0.0682	-0.8069	0.0470	-0.7079	0.0350	-0.6855	0.0162	
		50	0.0570	-0.9635	0.0449	-0.9051	0.0308	-0.8365	0.0151	

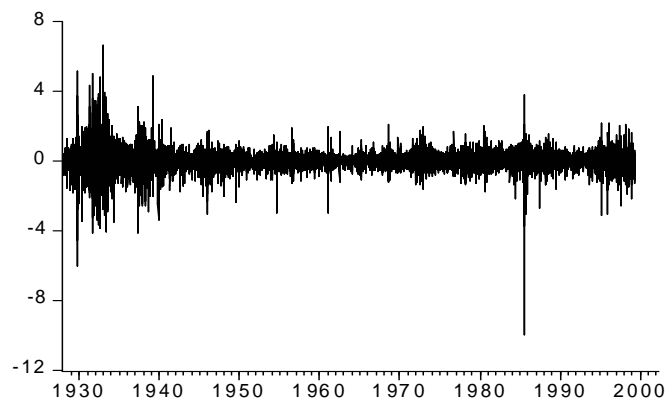
Table 1: Monte Carlo finite sample relative biases and standard deviations of sample autocorrelations of y_t^2 in 2FLMSV models together with their ACF values, for $k=1, 10$ and 50 .

$\{\phi, d, \sigma_\epsilon^2, \sigma_\eta^2, \delta\}$	Series Lag	ACF	T=500		T=1000		T=5000		
			R. Bias	Std. dev.	R. Bias	Std. dev.	R. Bias	Std. dev.	
$\{0.1, 0.2, 0.1, 0.05, 0.5\}$	y_t^2	1	0.0233	-0.8974	0.0458	-0.7769	0.0333	-0.7709	0.0143
		10	0.0028	-1.3359	0.0439	-0.4856	0.0303	-0.4227	0.0144
		50	0.0011	-3.3780	0.0432	-2.6818	0.0302	-0.6558	0.0143
$\{0.5, 0.2, 0.1, 0.05, 0.5\}$	y_t^2	1	0.0311	-0.1962	0.0504	-0.0953	0.0365	-0.0686	0.0156
		10	0.0028	-1.7127	0.0432	-0.8207	0.0307	-0.6880	0.0142
		50	0.0010	-3.2711	0.0432	-3.1150	0.0302	-1.0144	0.0142
$\{0.1, 0.2, 0.1, 0.1, 0.5\}$	y_t^2	1	0.0237	-0.8340	0.0457	-0.7147	0.0333	-0.6983	0.0143
		10	0.0026	-1.4420	0.0435	-0.5197	0.0305	-0.1627	0.0143
		50	0.0010	-3.5308	0.0434	-2.9699	0.0299	-0.6716	0.0142
$\{0.5, 0.2, 0.1, 0.1, 0.5\}$	y_t^2	1	0.0380	-0.3047	0.0508	-0.2160	0.0366	-0.0075	0.0157
		10	0.0025	-1.5461	0.0435	-0.5220	0.0309	-0.0008	0.0142
		50	0.0009	-3.4230	0.0433	-3.1517	0.0301	-0.0007	0.0142
$\{0.1, 0.45, 0.1, 0.05, 0.5\}$	y_t^2	1	0.1095	-0.8252	0.0496	-0.7750	0.0357	-0.7663	0.0162
		10	0.0759	-0.8878	0.0465	-0.8268	0.0329	-0.8076	0.0159
		50	0.0633	-0.9918	0.0443	-0.9584	0.0299	-0.8984	0.0148
$\{0.5, 0.45, 0.1, 0.05, 0.5\}$	y_t^2	1	0.1150	-0.6474	0.0545	-0.5960	0.0388	-0.5863	0.0176
		10	0.0744	-0.8950	0.0461	-0.8300	0.0334	-0.8094	0.0155
		50	0.0620	-0.9917	0.0442	-0.9621	0.0302	-0.9013	0.0148
$\{0.1, 0.45, 0.1, 0.1, 0.5\}$	y_t^2	1	0.1045	-0.6629	0.0538	-0.5884	0.0389	-0.5683	0.0185
		10	0.0712	-0.7495	0.0494	-0.6574	0.0357	-0.6173	0.0170
		50	0.0594	-0.9472	0.0452	-0.8857	0.0304	-0.7972	0.0154
$\{0.5, 0.45, 0.1, 0.1, 0.5\}$	y_t^2	1	0.1148	-0.5183	0.0583	-0.4471	0.0420	-0.4286	0.0199
		10	0.0684	-0.7556	0.0487	-0.6574	0.0361	-0.6151	0.0172
		50	0.0570	-0.9481	0.0450	-0.8885	0.0307	-0.7984	0.0155

Table 2: Monte Carlo finite sample relative biases and standard deviations of sample autocorrelations of y_t^2 in 2FLMSV models together with their ACF values, for $k=1, 10$ and 50 .



a)



b)

Figure 5: a) S&P 500 Index and b) Daily returns in percentage.

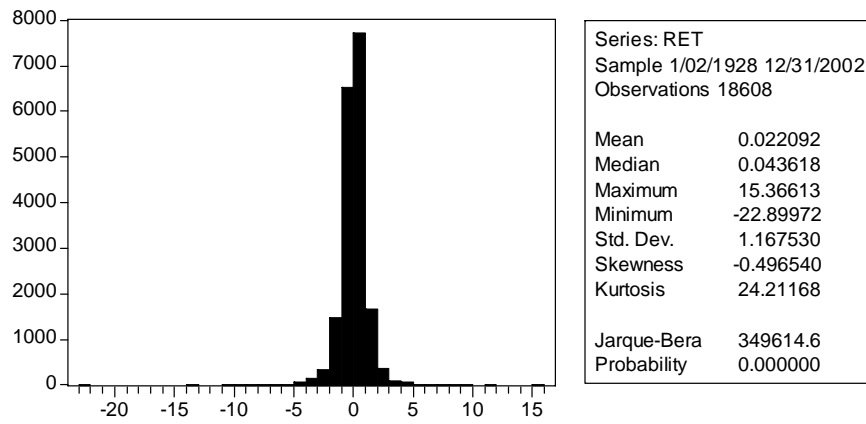
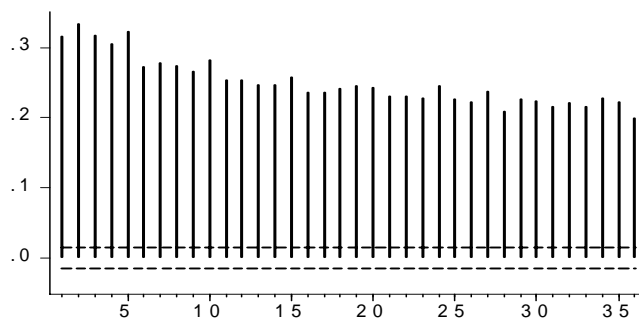
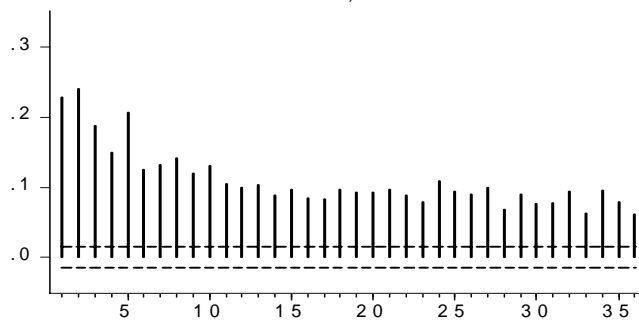


Figure 6: Histogram of returns.



a)



b)

Figure 7: Autocorrelation functions: a) absolute returns of S&P 500 and b) squared returns of S&P 500.

R/S	$\mathbf{q} = \mathbf{0}$	$\mathbf{q} = \mathbf{q}^*$
Q	1432.63	592.93
\tilde{J}	0.739	0.649
\hat{d}	0.239	0.149

Table 3

S&P 500	$t_{H_0:d_0=0.3}$	$t_{H_0:d_0=0.4}$	\hat{d}_1
Squared Returns	-1.649476	-4.049*	0.2632

Table 4: * means that the null hypotheses is rejected.

Model	χ^2	df	$p - value$
LMSV, Gaussian error	214.40	27	<0.0001
2FLMSV, Gaussian error	102.64	24	<0.0001

Table 5: χ^2 is the value of the EMM criterion, which follows a χ^2 statistic with degree of freedom of df . $L\epsilon$ is the autocorrelation order of the error of the fractional integrated process for the volatility factor.

Model	χ^2	df	$p - value$
LMSV, spline error	82.48	25	<0.0001
2FLMSV, spline error	46.32	22	≈ 0.002
Asymmetric 2FLMS, spline error	46.25	21	0.001

Table 6: χ^2 is the value of the EMM criterion, which follows a χ^2 statistic with degree of freedom of df .

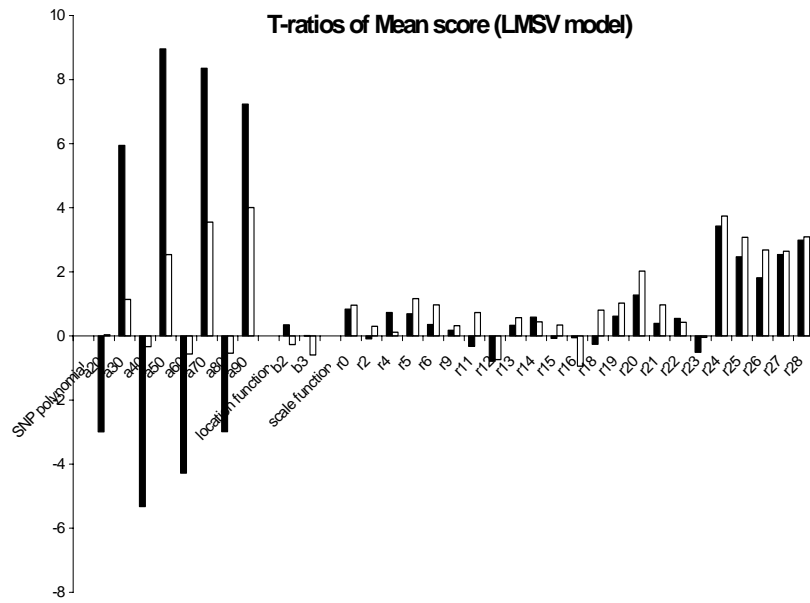


Figure 8: The black bars correspond to the Gaussian LMSV and the white bars to the spline LMSV.

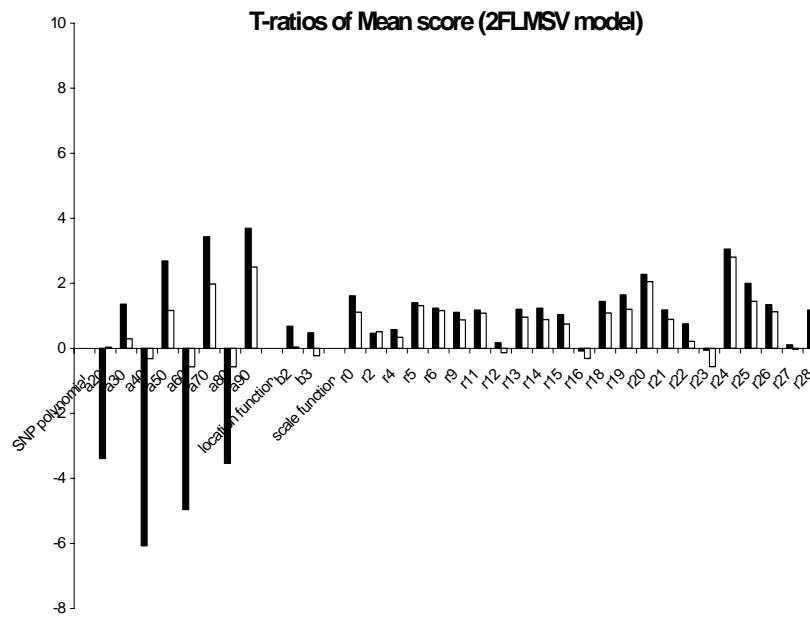


Figure 9: The black bars correspond to the Gaussian 2FLMSV and the white bars to the spline 2FLMSV.

	μ_y	c_1	c_2	r_y	r_{h2}	ϕ_1	ϕ_2	r_{h1}
Spline								
LMSV	0.073	0.113	-0.042	0.844				0.00001
95%Lower	0.058	0.098	-0.058	0.780				0.00001
95% Upper	0.089	0.129	-0.027	0.912				0.00001
2FLMSV	0.071	0.111	-0.046	0.403	0.343	0.403	0.376	0.00001
95%Lower	0.071	0.111	-0.046	0.403	0.319	0.293	0.280	0.00001
95% Upper	0.071	0.111	-0.046	0.403	0.386	0.403	0.452	0.00001
asymmetric 2FLMSV	0.071	0.111	-0.046	0.403	0.344	0.403	0.376	0.00001
95%Lower	0.071	0.111	-0.046	0.403	0.318	0.293	0.282	0.00001
95% Upper	0.071	0.111	-0.046	0.403	0.386	0.293	0.452	0.00001

Table 7: Fitted parameter values (Spline errors) and confidence intervals for these estimates.

	d	bc	bd	δ
Spline				
LMSV	0.351	-0.131	0.189	
95%Lower	0.351	-0.158	0.146	
95% Upper	0.351	-0.108	0.234	
2FLMSV	0.834	-0.065	0.064	
95%Lower	0.834	-0.083	0.043	
95% Upper	0.834	-0.053	0.093	
asymmetric 2FLMSV	0.834	-0.065	0.064	-0.37E-12
95%Lower	0.834	-0.086	0.043	0.000
95% Upper	0.834	-0.053	0.093	0.000

Table 7 (cont.)

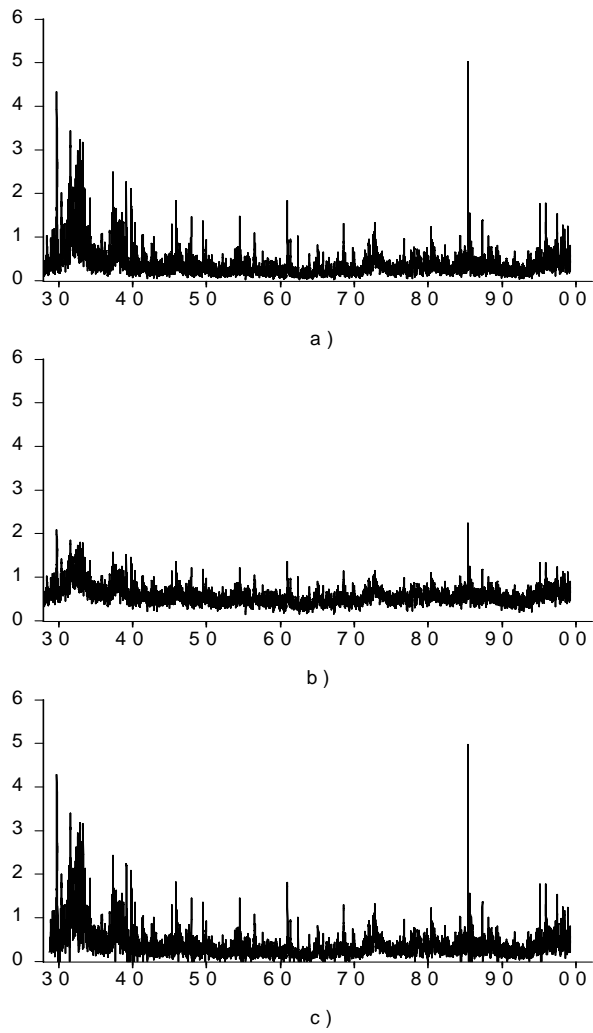


Figure 10: Plots of one-step-ahead volatilities: a) from equally weighted MA(4) of squared AR(1) residuals; b) from equally weighted MA(26) of squared AR(1) residuals; and the reprojected volatility from the 2FLMSV model.

