

Estimating smooth functions of sample mean in diffusion processes: a MBB approach

Giuseppina Albano, Francesco Giordano, Cira Perna

Dipartimento di Scienze Economiche e Statistiche, Università di Salerno

E-mail: pialbano@unisa.it, giordano@unisa.it, perna@unisa.it

The present work focuses on the inference in stochastic volatility models. More precisely, estimation of suitable functions of the mean vector of the increment stock price is performed without estimating in advance the parameters of the model. A moving block bootstrap (MBB) approach is then suggested in order to estimate the variance of those functions and properties of the involved estimators are discussed. Simulations on the model are also performed.

Keywords: Stochastic volatility, blockwise bootstrap, diffusion processes.

1. Introduction

The importance of conditional volatility in finance has led researchers (e.g., Hull and White, 1987; Longstaff and Schwartz, 1992; Heston, 1993) to extend early asset pricing theories (e.g., Black and Scholes, 1973; Vasicek, 1977) to the case in which volatility evolves in a non deterministic way. Empirically, stochastic volatility is well captured by the ARCH-type models introduced by Engle (1982). From a continuous time perspective, the initial contribution of Nelson (1990) established that ARCH models can be seen as approximations of diffusion processes. Furthermore the continuous time approach provided by diffusion processes can be useful when data are observed at non-regular intervals. These reasons justify the extensive use of stochastic volatility models in finance to describe some empirical facts of the stock and the derivative prices. The general stochas-

tic evolution for the stock and its volatility is described by two stochastic differential equations. In this way this approach permits to handle with Ito's calculus which provides methodologies already well-established in literature.

In this paper we focus our attention on the following stochastic volatility model:

$$\begin{aligned} dY_t &= \sigma_t dW_{1,t} \\ d\sigma_t^2 &= (\omega - \vartheta \sigma_t^2) dt + \alpha \sigma_t^2 dW_{2,t}. \end{aligned} \tag{1}$$

where the diffusion Y_t is the centered log-price of the stock and W_1 and W_2 are two independent Brownian motions. The model (1) is known in literature as GARCH diffusion since, under suitable assumptions on the parameters, it comes out in the well known paper by Nelson (1990) as the continuous limit in law of a suitable GARCH model.

In estimating the parameters α , ϑ , ω in the model (1), methods based on classical maximum likelihood or conditional moments do not work (see, for example, Figà-Talamanca, 2009). In a previous paper (Albano *et al.*, 2010(a)), using relations between the moments of volatility and the increments of the log-stock price Y_t , a method based on unconditional moments was proposed. Naturally, if the aim is estimating any function of the process, for example the variance or the autocovariance functions, a plug-in method can be used, after estimating the parameters α , ω and ϑ in (1). Unfortunately, this procedure is not always efficient, since long time series are generally needed.

This paper focuses on the inference on smooth functions of the volatility, such as autocovariance functions. We propose a method able to estimate them avoiding the first step of the estimation of the parameters.

The paper is organized as follows: in Section 2 theoretical properties of the process (1) are presented and in Section 3 inference for the parameters ω , ϑ and α in (1) and for smooth functions of the mean of the increment process of Y_t is discussed. In Section 4 a bootstrap procedure for estimating the variance of the involved estimators is proposed. Fi-

nally, in Section 5 simulation results are presented and some concluding remarks are discussed in Section 6.

2. Preliminary results

In this section we investigate some theoretical properties of the process (1) which are useful for the results on the inference discussed in the following sections.

Firstly, it can be shown (e.g. Capasso and Bakstein, 2005) that, if ω and α in (1) are positive constants, then there exists a strong solution to (1). Moreover, if the volatility at initial time t_0 , i.e. σ_0^2 , is a random variable independent from $W_{2,t}$, by Ito's formula it is possible to obtain the explicit expression of the volatility:

$$\sigma_t^2 = \omega F^{-1}(t, W_{2,t}) \int_0^t F(s, W_{2,s}) ds + F^{-1}(t, W_{2,t}) \sigma_0^2 \quad \forall t \geq 0, \quad (2)$$

where $F(t, W_{2,t}) = \exp\{(\vartheta + \frac{\alpha^2}{2})t - \alpha W_{2,t}\}$. For simplicity, in (2) we have assumed $t_0 = 0$. From (2) it is easy to see that the volatility process $\{\sigma_t^2\}$ is non negative for all $t \geq 0$. Moreover, after some cumbersome calculations, it is also possible to obtain the following approximation for the stochastic integral in (2):

$$\begin{aligned} F^{-1}(t, W_{2,t}) \int_0^t F(s, W_{2,s}) ds &= \int_0^t \exp\left\{-\left(\vartheta + \frac{\alpha^2}{2}\right)s + \alpha W_{2,s}\right\} ds \\ &\approx \frac{1 - e^{-\vartheta t}}{\vartheta}. \end{aligned} \quad (3)$$

More precisely, setting $I_t = \int_0^t \exp\left\{-\left(\vartheta + \frac{\alpha^2}{2}\right)s + \alpha W_{2,s}\right\} ds$, it can be proved that:

$$\begin{aligned} \mathbb{E} I_t &= \frac{1 - e^{-\vartheta t}}{\vartheta}, \\ \mathbb{E}(I_t I_{t-k}) &= \frac{1 - e^{-\vartheta t}}{\vartheta} \frac{1 - e^{-\vartheta(t-k)}}{\vartheta}. \end{aligned}$$

From (3) the volatility process in (2) can be written as:

$$\sigma_t^2 \approx \sigma_0^2 \Lambda(t) + \frac{\omega}{\vartheta} (1 - e^{-\vartheta t}) \quad (4)$$

where $\Lambda(t) \sim LN\left(-(\vartheta + \frac{\alpha^2}{2})t, \alpha^2 t\right)$ and LN is the lognormal distribution.

Stationarity and ergodicity of the process (Y_t, σ_t^2) is guaranteed if the initial value (Y_0, σ_0^2) is assumed to be independent from the two-dimensional brownian motion $(W_{1,t}, W_{2,t})$.

For the volatility process σ_t^2 , if $\frac{2\vartheta}{\alpha^2} > -1$ and $\omega > 0$,

$$\sigma_t^2 \xrightarrow{d} Inv\Gamma\left(1 + \frac{2\vartheta}{\alpha^2}, \frac{2\omega}{\alpha^2}\right)$$

where $Inv\Gamma$ is the inverse Gamma distribution.

3. Inference on the model

Let $Y_0, Y_\delta, \dots, Y_{h\delta}, \dots, Y_{n\delta}$ be observations on the GARCH diffusion model (1) with frequency δ , and let $\sigma_0^2, \sigma_\delta^2, \dots, \sigma_{h\delta}^2, \dots, \sigma_{n\delta}^2$ be the corresponding volatilities. Let X_t be the increment process of Y_t , i.e.

$$X_t = Y_t - Y_{t-1} = \sqrt{\sigma_t^2 \delta} Z_t \quad (5)$$

where $Z_t \stackrel{i.i.d.}{\sim} N(0, 1)$ ($t = 0, \delta, \dots, n\delta$). From (4) it is easy to obtain the following recursive relation for the volatility:

$$\sigma_{h\delta}^2 = e^{-(\vartheta + \frac{\alpha^2}{2})\delta + \alpha W_\delta} \sigma_{(h-1)\delta}^2 + \frac{\omega}{\vartheta} (1 - e^{-\vartheta \delta}) \quad h = 1, 2, 3, \dots, n. \quad (6)$$

In the following, we illustrate a method to estimate the parameters ϑ , α and ω in the GARCH diffusion model (1) and we propose a procedure to estimate directly any smooth function of the moments of X_t which avoids the first step of the estimation of the parameters.

3.1. Inference on the parameters in the GARCH diffusion model

From (6) it can be proved that (Albano *et al.*, 2010(a)) the asymptotic moments of the volatility are:

$$\lim_{h \rightarrow \infty} \mathbb{E}[\sigma_{h\delta}^2] = \frac{\omega}{\theta}, \quad (7)$$

$$\lim_{h \rightarrow \infty} \mathbb{E}[\sigma_{h\delta}^4] = \frac{\omega^2}{\theta^2} \frac{1 - e^{-2\theta\delta}}{1 - e^{(-2\theta + \alpha^2)\delta}}, \quad (8)$$

$$\lim_{h \rightarrow \infty} \mathbb{E}[\sigma_{h\delta}^2 \sigma_{(h-1)\delta}^2] = e^{-\theta\delta} E\sigma_{t\delta}^4 + \frac{\omega^2}{\theta^2} (1 - e^{-\theta\delta}). \quad (9)$$

Moreover, from (5) we find that they are linked with the corresponding ones of the increment process X_t in the following way:

$$\begin{aligned} \lim_{h \rightarrow \infty} \mathbb{E}X_{h\delta}^2 &= \delta \lim_{h \rightarrow \infty} \mathbb{E}\sigma_{h\delta}^2 \\ \lim_{h \rightarrow \infty} \mathbb{E}X_{h\delta}^4 &= 3\delta^2 \lim_{h \rightarrow \infty} \mathbb{E}\sigma_{h\delta}^4 \\ \lim_{h \rightarrow \infty} \mathbb{E}[X_{h\delta}^2 X_{(h-k)\delta}^2] &= \delta^2 \lim_{h \rightarrow \infty} \mathbb{E}[\sigma_{h\delta}^2 \sigma_{(h-k)\delta}^2]. \end{aligned} \quad (10)$$

Then, if there exists the second moment of the volatility, the method based on the moments of the volatility process suggests the following estimators for θ , ω and α^2 :

$$\begin{aligned} \hat{\theta} &:= f_1(M_2, M_4, E_1) = \frac{1}{\delta} \log \frac{\hat{\gamma}(0)}{\hat{\gamma}(1)}, \\ \hat{\omega} &:= f_2(M_2, M_4, E_1) = M_2 \hat{\theta} \\ \hat{\alpha}^2 &:= f_3(M_2, M_4, E_1) = \frac{1}{\delta} \log \left\{ e^{2\hat{\theta}\delta} \left(1 - \frac{M_2^2}{M_4} \right) + \frac{M_2^2}{M_4} \right\}. \end{aligned} \quad (11)$$

where the statistics M_2 , M_4 and E_1 are defined as follows:

$$M_2 = \frac{1}{n\delta} \sum_{t=1}^n X_t^2, \quad M_4 = \frac{1}{3n\delta^2} \sum_{t=1}^n X_t^4, \quad E_1 = \frac{1}{n\delta^2} \sum_{t=1}^n X_t^2 X_{t-1}^2 \quad (12)$$

and $\hat{\gamma}(0)$ and $\hat{\gamma}(1)$ are the sample variance and covariance of $\{\sigma_{h\delta}^2\}$:

$$\hat{\gamma}(0) = M_4 - M_2^2 \quad \hat{\gamma}(1) = E_1 - M_2^2. \quad (13)$$

Moreover, assuming the existence of the eighth moment of the increment process, the strong consistency and the asymptotic normality for the proposed estimators can be proved (Albano *et al.*, 2010(a)).

3.2. Inference on smooth functions of the mean

Let $\mathbf{v} = (v_1, v_2, \dots, v_m)$ ($m \geq 1$) be a mean vector of the increment process X_t defined in (5), i.e.

$$v_i = \mathbb{E}[X_t^{j_1} X_{t-1}^{j_2} \cdots X_1^{j_d}] \quad (j_1, j_2, \dots, j_d \in \mathbb{N}_0; \quad i = 1, \dots, m).$$

Let $\mathbf{V}_n = (V_1^{(n)}, V_2^{(n)}, \dots, V_m^{(n)})$ ($m \geq 1$) be the corresponding sample mean vector, i.e. $V_i^{(n)}$ is the sample estimator of v_i . Let us consider a function $H : \mathbb{R}^m \rightarrow \mathbb{R}$ and suppose that the parameter of interest θ is a function of \mathbf{v} , i.e.

$$\theta = H(\mathbf{v}).$$

A natural estimator for $H(\mathbf{v})$ is $H(\mathbf{V}_n)$.

In the following we will suppose that the following assumptions hold:

- A1. the function H has continuous partial derivatives with respect all the components in a neighborhood \mathcal{C} of \mathbf{v} ;
- A2. the gradient \mathbf{A}_v of H is non-null in \mathcal{C} .

Under these assumptions, it can be shown (see, for example, Serfling, 1980) that

$$\sqrt{n}[H(\mathbf{V}_n) - H(\mathbf{v})] \xrightarrow{d} N(0, \mathbf{A}_v^T \Sigma_{\mathbf{v}} \mathbf{A}_v) \quad (14)$$

and

$$n[\text{var } H(\mathbf{V}_n)] - \mathbf{A}_v^T \Sigma_{\mathbf{v}} \mathbf{A}_v \rightarrow 0, \quad n \rightarrow \infty \quad (15)$$

i.e. $H(\mathbf{V}_n)$ is asymptotically normal and it is a consistent estimator of $H(\mathbf{v})$. In (14) and (15) $\Sigma_{\mathbf{v}}$ represents the asymptotic covariance matrix of the vector \mathbf{V}_n .

An example. Let us focus on the autocovariance functions $\gamma_{\sigma^2}(k)$ with lag k ($k \geq 0$) of the volatility process σ_t^2 . Moreover, from (5) it is

easy to see (Albano et al., 2010(a)) that it is related to the autocovariance function of the square of the increment process X_t^2 . So a natural estimator for $\gamma_{\sigma^2}(k)$ is

$$\hat{\gamma}_{\sigma^2}(k) = \begin{cases} M_4 - M_2^2 & k = 0 \\ E_k - M_2^2 & k > 0 \end{cases} \quad (16)$$

where M_2, M_4 are defined in (12) and

$$E_k = \frac{1}{n\delta^2} \sum_{t=k+1}^n X_t^2 X_{t-k}^2.$$

Further, the variance of $\hat{\gamma}_{\sigma^2}(k)$ is such that

$$n \text{var} \hat{\gamma}_{\sigma^2}(k) \rightarrow \mathbf{A}_{\mathbf{v}}^T \Sigma_{\mathbf{v}} \mathbf{A}_{\mathbf{v}},$$

with

$$\mathbf{v} = (v_1, v_2) = \begin{cases} (\mathbb{E}X_t^2, \mathbb{E}X_t^4) & k = 0 \\ (\mathbb{E}X_t^2, \mathbb{E}X_t^2 X_{t-k}^2) & k > 0 \end{cases}$$

and $\mathbf{A}_{\mathbf{v}} = (1, -2v_1)^T$. □

From (14) and (15) it is evident that to obtain an estimation of the variance of $H(\mathbf{V}_n)$ it is necessary a preliminary estimation of $\Sigma_{\mathbf{v}}$. For particular choices of the function H , explicit expressions for $\Sigma_{\mathbf{v}}$ can be obtained and a plug-in procedure can be improved to estimate the asymptotic variance of $H(\mathbf{V}_n)$ (see, for example, Genon and Catalot, 2000 and Figà-Talamanca, 2009). Further, in such kind of procedures, since the process $\{X_t\}$ has two independent noises, for small samples, the estimation could not capture the dependence on $\{X_t\}$. Finally, finding explicit expressions for $\Sigma_{\mathbf{v}}$ involves some cumbersome calculations which, moreover, depend on the particular choice of the smooth function H . To avoid such kind of problems, we propose a moving block bootstrap procedure, which is independent from the particular choice of H and applicable in a more general context.

4. MBB for estimating the variance

In order to estimate Σ_v we use a bootstrap procedure for dependent data. This class of resampling procedures comes out as an extension of the classical bootstrap for independent data first proposed by Efron (1979). The main advantage is that MBB is able to preserve the dependence structure of the original data in the bootstrap samples. More precisely, in nonparametric schemes, blocks of consecutive observations are randomly resampled with replacement from the original time series and assembled by joining the blocks in random order to obtain a simulated version of the original series (Kunsch, 1989; Politis, 1992). These approaches, known as blockwise bootstrap or moving block bootstrap, generally work satisfactory and enjoy the properties of being robust against misspecified models. In order to illustrate the procedure in our context, let us consider the centered and scaled estimator V_n given by

$$T_n = \sqrt{n}(V_n - v).$$

Suppose that $b = \lfloor n/l \rfloor$ blocks are resampled so the resample size is $n_1 = bl$. If V_n^* is the sample mean of the n_1 bootstrap observations based on the MBB, the block bootstrap version of T_n is:

$$T_n^* = \sqrt{n_1}(V_n^* - \mathbb{E}_* V_n^*)$$

where \mathbb{E}_* denotes the conditional expectation given the observations $\chi_n = \{X_1, X_2, \dots, X_n\}$.

In the following we will assume for simplicity that $n_1 \approx n$, a reasonable choice in the case of long time series.

Using the geometrically α -mixing condition for the process $\{X_t\}$ as proved in (Genon and Catalot, 2000) and choosing the length l of the blocks such that

$$A3. \quad l \rightarrow \infty; \quad l^{-1} + \frac{l}{n} \rightarrow 0 \quad n \rightarrow \infty,$$

we have (Lahiri, 2003) that

$$var_* T_n^* \xrightarrow{p} \Sigma_v. \quad (17)$$

Now, let $T_{1n} = \sqrt{n}[H(\mathbf{V}_n) - H(\mathbf{v})]$ be the centered and scaled estimator of the smooth function $H(\mathbf{v})$; then the block bootstrap version of T_{1n} is

$$T_{1n}^* = \sqrt{n_1}[H(\mathbf{V}_n^*) - H(\mathbb{E}_* \mathbf{V}_n^*)].$$

Under the assumptions A1, A2 and A3 it can be proved (Albano *et al.*, 2010(b)) that the bootstrap distribution of T_{1n}^* converges to the distribution of T_{1n} , i.e. to the normal distribution. So, for the inference on T_{1n} through T_{1n}^* we just need to estimate the variance of T_{1n} . Moreover, it can be proved (Albano *et al.*, 2010(b)) that

$$\text{var}_*(T_{1n}^*) \xrightarrow{p} \mathbf{A}_v^T \Sigma_v \mathbf{A}_v \quad \text{as } n \rightarrow \infty.$$

So, setting $T'^* = \hat{\mathbf{A}}_v^T \mathbf{V}_n^* \hat{\mathbf{A}}_v$, where $\hat{\mathbf{A}}_v$ is the sample estimator of the gradient \mathbf{A}_v , from (17) we obtain:

$$n \text{var}_*(T'^*) = \hat{\mathbf{A}}_v^T (n \text{var}_* \mathbf{V}_n^*) \hat{\mathbf{A}}_v \xrightarrow{p} \mathbf{A}_v^T \Sigma_v \mathbf{A}_v \quad (18)$$

from which we can estimate the variance $\text{var}_* T_{1n}^*$ using the bootstrap variance \mathbf{V}_n^* and the sample estimator $\hat{\mathbf{A}}_v$.

5. Simulations

In order to evaluate the performance of the proposed procedure, a small Monte Carlo experiment has been performed. The parameters in (1) are fixed as $\vartheta = 0.6, \omega = 0.5$, while three different values for α are chosen: $\alpha = \{0, 0.2, 0.4\}$. The value $\alpha = 0$ corresponds to a deterministic volatility. The frequency of the observation is fixed as $\delta = 1/4$. Three different time series lengths are considered $n = \{500, 1000, 2000\}$, and for each length $N = 300$ Monte-Carlo runs are generated. In Figure 1 the distribution of the statistics M_2 , M_4 and E_1 are reported. Straight line indicates the corresponding true value of the involved moment. It is evident that the widths of the corresponding box plots become smaller and smaller as the length of the time series increases. These empirical results confirm the consistency of the sample mean vector (M_2, M_4, E_1) . Moreover also the bias seems to be slight for the three statistics. When α

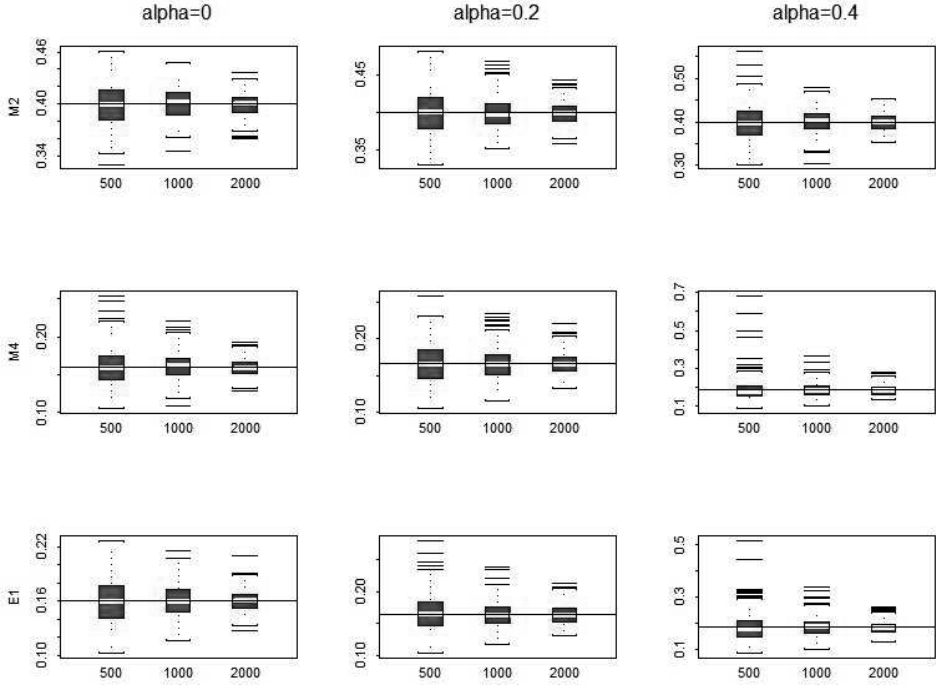


Figure 1. Box plots for M_2 , M_4 and E_1 for $\alpha = 0$ (left), $\alpha = 0.2$ (center) and $\alpha = 0.4$ (right). The straight line represents the true value of the corresponding moment.

increases the box plots show a smaller variability, so the estimates seem to work satisfactory when the stochastic component of the volatility is not negligible. Let $A_n^{(j)}$ ($j = 1, 2, 3$) be the rescaled statistics of M_2 , M_4 and E_1 , respectively.

In Table 1 the root mean square error

$RMSE := \sqrt{E_N[\text{var}_*(A_n^j) - v^{(j)}]}$ ($j = 1, 2, 3$) is reported for the three different time series lengths.

It is evident that it tends to decrease as the length of the time series increases. Now, let us focus on the variance of the volatility process, $\gamma_{\sigma^2}(0)$. Also in this case, the results reported in Figure 2 confirm the consistency of the proposed estimator $\hat{\gamma}_{\sigma^2}(0)$, for the three values of α . Moreover, the root mean square error of $\sqrt{n}\hat{\gamma}_{\sigma^2}(0)$ seems to decrease as n increases.

Table 1. Root mean square error of A_n^j ($j = 1, 2, 3$) for $\alpha = 0, 0.2$ and 0.4 .

		α		
n		0	0.2	0.4
M_2	500	0.0678	0.1180	0.4512
	1000	0.0490	0.0850	0.3359
	2000	0.0333	0.0731	0.3052
M_4	500	0.2789	0.2724	5.1979
	1000	0.1302	0.1922	1.5314
	2000	0.0863	0.1412	1.1772
E_1	500	0.1453	0.3310	2.0566
	1000	0.1207	0.1872	1.3008
	2000	0.0746	0.1659	1.2919

6. Concluding remarks

The present paper proposes a method for estimating smooth functions of the mean in stochastic volatility models, avoiding the preliminary estimation of the parameters of the model. A MBB approach is then performed to evaluate the variance of the proposed estimators. The procedure yields satisfactory results in a simulation study for finite sample sizes.

Several different aspects should be further explored to get a better insight of the suggested procedure.

Table 2. Root mean square error of $\hat{\gamma}_{\sigma^2}(0)$ for $\alpha = 0, 0.2$ and 0.4 .

	α		
n	0	0.2	0.4
500	0.1404	0.0998	3.7975
1000	0.0506	0.0851	0.8231
2000	0.0386	0.0635	0.6784

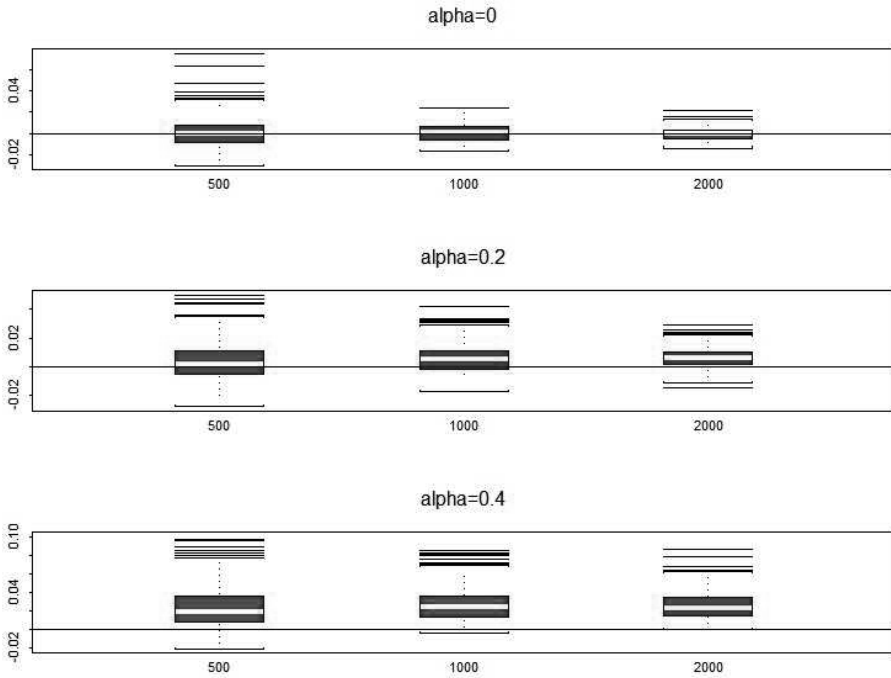


Figure 2. Box-plots for $\hat{\gamma}_{\sigma^2}(0)$ for $\alpha = 0$ (top), $\alpha = 0.2$ (center) and $\alpha = 0.4$ (bottom). The straight line represents the zero.

References

Albano G., Giordano F., Perna C. (2010a), On the estimation in continuous limit of GARCH processes, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, C. Perna, M. Sibillo (Eds), Springer, to appear.

Albano G., Giordano F., Perna C. (2010b), Continuous time processes for the analysis of volatility, *to appear*.

Black F., Scholes M. (1973), The price of options and liabilities, *Journal of Political Economy*, 81, 637-659.

Capasso V., Bakstein D. (2005), *An introduction to continuous-time stochastic processes: theory, models, and applications to biology, finance and engineering*, Birkhauser, Boston.

Efron B. (1979), Bootstrap methods: another look at the jackknife, *Annals of Statistics*, 7, 1-26.

Engle R.F. (1982), Autoregressive conditional heteroskedasticity with esti-

mates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1008.

Figá-Talamanca G. (2009), Testing volatility autocorrelation in the constant elasticity of variance stochastic volatility model, *Computational Statistics and Data Analysis*, 53, 2201-2218.

Genon-Catalot V., Jeantheau T., Larédo C. (2000), Stochastic volatility models as hidden Markov models and statistical applications, *Bernoulli*, 6, 1051-1079.

Heston S.L. (1993), A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies*, 6, 327-343.

Hull J.C., White A. (1987), The pricing of options on assets with stochastic volatilities, *Journal of Finance*, 42, 281-300.

Kunsch H.R., (1989), The jackknife and bootstrap for general stationary observations, *Annals of Statistics*, 17, 1217-1241.

Lahiri S.N., (2003), Resampling methods for dependent data, *Springer-Verlag*, New York.

Longstaff F.A. Schwartz E.S. (1992), Interest rate volatility and the term structure: a two-factor general equilibrium model, *Journal of Finance*, 47, 1259-1282.

Nelson D.B. (1990), ARCH models as diffusion approximations. *Journal of Econometrics*, 45, 7-38.

Politis D.N. (1992), Romano, J.P., A general resampling scheme for triangular arrays of alpha-mixing random variables with application to the problem of spectral density estimation, *Annals of Statistics*, 20, 1985-2007.

Serfling R.J. (1980), *Approximation Theorems of Mathematical Statistics*, J. Wiley & Sons, New York.

Vasicek O. (1977), An equilibrium characterization of the term structure, *Journal of Financial Economics*, 5, 177-188.