Nonparametric estimation of stochastic volatility models ⋆

Roberto Renò [a,1]

[a] *Dipartimento di Economia Politica, Università di Siena*
Piazza S.Francesco 7, 53100, Siena, Italy

Abstract

This letter introduces nonparametric estimators of the drift and diffusion coefficient of stochastic volatility models which exploit techniques for estimating integrated volatility with high-frequency data. The performance of the proposed estimators is assessed on simulations of two popular stochastic volatility models.

JEL Classification code: C14
Keywords: Nonparametric estimation, stochastic volatility, realized volatility

1 Introduction

In this letter we propose nonparametric estimators for stochastic volatility models of the type,

$$\begin{aligned} dX_t &= \mu(X_t)dt + \sigma_t dW_{1,t} \\ d\sigma_t^2 &= m(\sigma_t^2)dt + \Lambda(\sigma_t^2)dW_{2,t}, \end{aligned} \tag{1}$$

that is, estimators of the functions $m(\cdot), \Lambda^2(\cdot)$, based on a partially observed variable X_t, which consists of n equally spaced observations in an interval $[0, T]$. [2] dW_1 and dW_2 are possibly correlated Brownian motions. Such models

⋆ I wish to acknowledge Roberto Casarin for his suggestions, and Emma Thorley for reviewing the manuscript.

[1] E-mail: reno@unisi.it. Telephone: +39 0577 232649. Fax: +39 0577 232661.

[2] For the sake of brevity, throughout this letter we skip technical details regarding the assumptions on the drift and the diffusion of the SDEs involved.

are routinely adopted in finance to model stock prices, stock indexes, currency prices and the short term interest rate. Typically, stochastic volatility models are empirically found to outperform constant volatility models. Standard references are Hull and White (1987) and Andersen and Lund (1997) for stock prices and interest rates respectively, but these are only an extract from a very large quantity of literature. Ghysels et al. (1996) review stochastic volatility models.

The typical approach to stochastic volatility models is to parametrically specify the drift $m(\cdot)$ and the diffusion $\Lambda(\cdot)$. Two popular examples are the affine Heston (1993) model, and the GARCH continuous time model studied, among others, by Drost and Werker (1996).

The purpose of this letter is to propose a nonparametric method for estimating the real functions $m(\cdot)$ and $\Lambda^2(\cdot)$, and to provide preliminary evidence, via simulation, of the reliability of the method. We introduce the estimators in Section 2 and describe our simulation experiments in Section 3. Conclusions and possible directions for future work are outlined in Section 4

2 The proposed estimators

Nonparametric estimation has been worked out for single-factor models of the type:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \tag{2}$$

Florens-Zmirou (1993) and Jiang and Knight (1997) provide estimators of the diffusion coefficient $\sigma(\cdot)$; see also Stanton (1997). An extension has recently been provided by Bandi and Phillips (2003), who estimate the drift coefficient as well.

While nonparametric estimators are interesting in themselves, despite some practical issues like the choice of the bandwidth parameter and the small-sample biases (Pritsker, 1998; Chapman and Pearson, 2000), one important limitation is that they have been studied for single-factor models only. In practice, it is well known that single-factor models are too naive, both for stock prices and spot rate modeling. Motivated by this consideration, in this work we adopt nonparametric techniques, building on the work of Florens-Zmirou (1993) and Bandi and Phillips (2003), to estimate stochastic volatility models of the type (1).

The idea is the following. If we could observe the process σ_t^2, then the results on nonparametric estimation of univariate diffusions would yield suitable estimators. We then replace σ_t^2, which is unobservable, with a consistent estimator of it. The proposed estimators are:

$$\hat{m}_{HF}(x) = \frac{n \sum_{i=0}^{n-1} K\left(\frac{\tilde{\sigma}^2_{iT/n} - x}{h_{n,T}}\right)(\tilde{\sigma}^2_{(i+1)T/n} - \tilde{\sigma}^2_{iT/n})}{T \sum_{i=0}^{n-1} K\left(\frac{\tilde{\sigma}^2_{iT/n} - x}{h_{n,T}}\right)}, \quad (3)$$

$$\hat{\Lambda}^2_{HF}(x) = \frac{n \sum_{i=0}^{n-1} K\left(\frac{\tilde{\sigma}^2_{iT/n} - x}{h_{n,T}}\right)(\tilde{\sigma}^2_{(i+1)T/n} - \tilde{\sigma}^2_{iT/n})^2}{T \sum_{i=0}^{n-1} K\left(\frac{\tilde{\sigma}^2_{iT/n} - x}{h_{n,T}}\right)}, \quad (4)$$

where $K(\cdot)$ is a kernel[3], $h_{n,T}$ is a bandwidth parameter, and we define $\tilde{\sigma}_t^2$ as:

$$\tilde{\sigma}_t^2 = \frac{1}{\Delta^-_{n,T} + \Delta^+_{n,T}} \int_{t-\Delta^-_{n,T}}^{t+\Delta^+_{n,T}} \sigma^2(s) ds, \quad (5)$$

where $\Delta^{\pm}_{n,T}$ vanish when $n, T \to \infty$. By the almost sure continuity of $\sigma(t)$, we know indeed that, as $n, T \to \infty$, almost surely $\tilde{\sigma}_t^2 \to \sigma_t^2$. Here, the subscript HF stands for *high frequency*, since the advantage of this procedure is that a consistent estimate of the integrated volatility $\tilde{\sigma}_t^2$ can be achieved with high frequency data. The most popular estimator for the integrated volatility is realized volatility, see Andersen et al. (2003). There are also many alternatives for estimating the integrated volatility: see Andersen et al. (2003) for a review.

Thus, we want to use a combination of daily and high frequency data. We then implement our estimators on simulated time series.

3 Performance on simulated data

To check the reliability of the proposed estimators, we test them on simulated time series of the GARCH(1,1) continuous-time model (Drost and Werker, 1996), using the parameters estimated in Andersen and Bollerslev (1998) on currency rates, and the Heston (1993) diffusion model, using the parameters obtained by Jones (2003) on stock index prices. Both models have a linear

[3] A kernel is a symmetric real function which integrates to one and is such that $\int x^2 K(x) < \infty$.

drift,

$$m(\sigma^2) = k(\omega - \sigma^2),$$

while the variance is:

$$\begin{aligned} \Lambda(\sigma^2) &= \eta\sigma^2 \quad (GARCH) \\ \Lambda(\sigma^2) &= \eta\sqrt{\sigma^2} \ (HESTON), \end{aligned} \tag{6}$$

where k, ω, η are positive constants. We simulate the GARCH and Heston diffusion models by a second-order Euler discretization scheme, then we compute the estimators (3),(4) on simulated paths. To estimate the integrated volatility on each day, we use realized volatility on a grid of m intraday observations. We use the following bandwidth:

$$h_{n,T} = h_s \cdot \hat{\sigma} \cdot n^{-\frac{1}{5}}, \tag{7}$$

where h_s is a real constant set equal to 1.06, and $\hat{\sigma}$ is the sample standard deviation of realized volatilities. We use $\Delta^+_{n,T} = 0, \Delta^-_{n,T} = T/n$ and the Gaussian kernel.

Figure 1 shows the results for the estimation of the variance for the Heston and GARCH models, for $m = 432$, while Figure 2 shows the estimation of the drift for the GARCH model only. The figures show that the estimators are capable of reconstructing the generated drift and diffusion functions.

3.1 Small sample bias correction

When the number of intraday observations m is smaller, a bias is introduced in the estimation of the diffusion coefficient, due to the fact that the integrated volatility is measured with an error. We can improve the performance of our estimator, in small samples, in the following way. We are using $(\tilde{\sigma}^2_{(i+1)T/n} - \tilde{\sigma}^2_{iT/n})^2$ as an estimator of the local variance. This is asymptotically consistent, but in small samples it is correct only if the drift is zero. For interest rate models, the drift is very small when compared to the diffusion part; for stochastic volatility, the drift is very large as widely documented by the presence of heteroskedasticity; see e.g. Bollerslev et al. (1992, 1994). To correct the estimator, instead of (4), we can use:

$$\hat{\Lambda}^2_{HF,correct}(x) = \frac{n \sum_{i=0}^{n-1} K\left(\frac{\tilde{\sigma}^2_{iT/n} - x}{h_{n,T}}\right) \Delta\sigma^2}{T \sum_{i=0}^{n-1} K\left(\frac{\tilde{\sigma}^2_{iT/n} - x}{h_{n,T}}\right)}, \tag{8}$$

where

$$\Delta\sigma^2 = \left(\tilde{\sigma}^2_{(i+1)T/n} - \tilde{\sigma}^2_{iT/n} - \frac{T}{n}\hat{m}_{HF}\left(\tilde{\sigma}^2_{iT/n}\right)\right)^2. \tag{9}$$

The benefit of this correction is illustrated in Figure 3, with $m = 288$.

4 Conclusions and directions for future work

In this letter we propose, to our knowledge for the first time, recipes to nonparametrically estimate the drift and the diffusion coefficient of a stochastic volatility model. The proposed estimators are devised by borrowing from the theory of nonparametric estimation of univariate processes, and from the econometric literature on estimating quadratic variation with high-frequency data. Evidence on simulated data shows that the proposed estimators are able to reproduce the drift and the diffusion coefficients of the generated models.

The results briefly presented in this letter are encouraging, but they are also preliminary in many respects. First, we can extend the estimators to allow for jumps in the equation driving the observable variable, as well as to estimate leverage when the observable variable and the latent volatility factor are correlated. Second, the choice of the bandwidth parameter can be refined, for example using automated techniques instead of the simple rule of thumb (7) adopted here. Finally, technical conditions for convergence should be worked out, paying particular attention to the conditions on the convergence rate of $\Delta^{\pm}_{n,T}$ in (5) and the possibility of holding T fixed when estimating $\Lambda^2(\cdot)$. Once this program has been completed, the methodology can be used in many potential applications, such as option pricing and hedging and model validation. Research on this topic is in progress.

References

Andersen, T. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Andersen, T., T. Bollerslev, and F. X. Diebold (2003). Parametric and nonparametric volatility measurement. In L. P. Hansen and Y. Ait-Sahalia (Eds.), *Handbook of Financial Econometrics.* Amsterdam: North-Holland.

Andersen, T. and J. Lund (1997). Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics 77*, 343–377.

Bandi, F. and P. Phillips (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica 71*(1), 241–283.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). ARCH modelling in finance: a selective review of the theory and empirical evidence. *Journal of Econometrics 52*, 5–59.

Bollerslev, T., R. Engle, and D. Nelson (1994). Arch models. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume IV. Amsterdam: North-Holland.

Chapman, D. and N. Pearson (2000). Is the short rate drift actually nonlinear? *Journal of Finance 55*(1), 355–388.

Drost, F. and B. Werker (1996). Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics 74*, 31–57.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability 30*, 790–804.

Ghysels, E., A. Harvey, and E. Renault (1996). Stochastic volatility. In G. Maddala (Ed.), *Handbook of Statistics*, Volume 14, pp. 119–191. Amsterdam: North-Holland.

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

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatility. *Journal of Finance 42*, 281–300.

Jiang, G. and J. Knight (1997). A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. *Econometric Theory 13*, 615–645.

Jones, C. (2003). The dynamics of stochastic volatility: evidence from underlying and options markets. *Journal of Econometrics 116*, 181–224.

Pritsker, M. (1998). Nonparametric density estimation and tests of continuous time interest rate models. *Review of Financial Studies 11*(3), 449–487.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance 52*, 1973–2002.

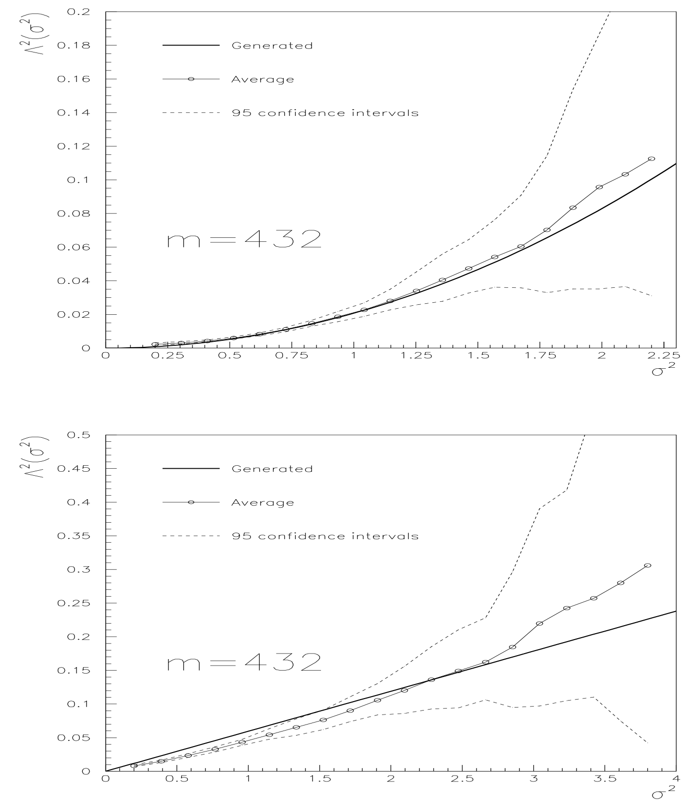

Fig. 1. Shows the estimate (dots) on simulated data of the diffusion coefficient $\Lambda^2(\sigma^2)$ compared to the generated models (solid line). The dashed lines are 95 % confidence bands. Top: generated model is GARCH. Bottom: generated model is Heston. Results are obtained with $1,000$ replications.

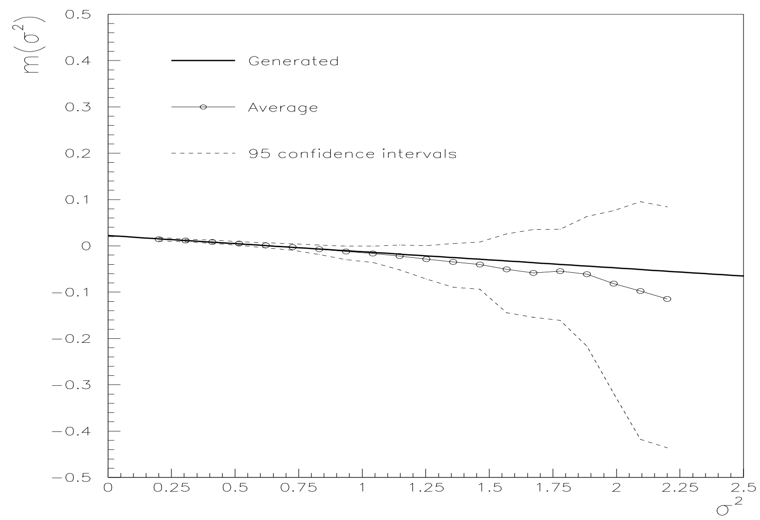

Fig. 2. Shows the estimate (dots) on simulated data of the drift function $m(\sigma^2)$ compared to the generated GARCH model (solid line).

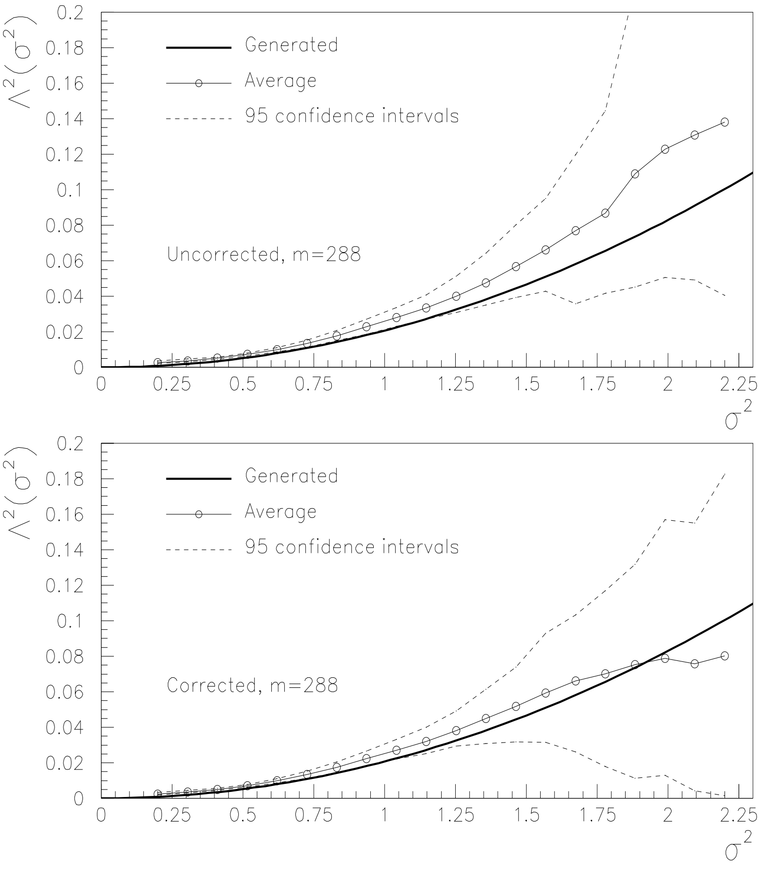

Fig. 3. Top: shows the estimated diffusion coefficient on simulated data according to the GARCH model with no correction. Bottom: shows the estimated diffusion coefficient computed according to the correction (9).