SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2011-4

Testing the local volatility assumption: a statistical approach

Mark Podolskij and Mathieu Rosenbaum

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Testing the local volatility assumption: a statistical approach *

Mark Podolskij[1] Mathieu Rosenbaum[2]

[1] University of Heidelberg and CREATES
m.podolskij@uni-heidelberg.de

2 École Polytechnique Paris
mathieu.rosenbaum@polytechnique.edu

13 January 2011

Abstract

In practice, the choice of using a local volatility model or a stochastic volatility model is made according to their respective ability to fit implied volatility surfaces. In this paper, we adopt an opposite point of view. Indeed, based on historical data, we design a statistical procedure aiming at testing the assumption of a local volatility model for the price dynamics, against the alternative of a stochastic volatility model.

Key words: Local Volatility Models, Stochastic Volatility Models, Test Statistics, Semi-Martingales, Limit Theorems.

JEL Classification: C10, C13, C14.

1 Introduction

It is well known that the Black-Scholes model does not allow for important stylized facts of asset returns such as heavy tails, gain/loss asymmetry or leverage effect. A classical way to obtain these features of financial data is to use models in which the volatility, that is the diffusion coefficient of the log price, is itself a random process. However, the success of these models with random volatility is clearly not due to their statistical properties. Indeed, practitioners essentially focus on the fact that they enable to fit the implied volatility surface much better than the Black-Scholes model does. Two main types of models in which the

*We would like to thank Holger Dette for turning our attention to the statistical problem discussed in this paper. Mark Podolskij acknowledges financial support from CREATES funded by the Danish National Research Foundation.

volatility is a random process are particularly used: models with local volatility and models with stochastic volatility.

In local volatility models, the volatility is assumed to depend on time and of the present value of the asset only. They can be written on the form

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t, \quad \sigma_t = \sigma(X_t, t),$$

where X_t represents the price of the asset at time t, W_t is a Brownian motion and $\sigma(x,t)$ is a deterministic function. Therefore, in these models, the price X_t and the volatility σ_t/X_t can both be stochastic. However, there is only one factor of randomness, the Brownian motion W_t. In particular, the Black-Scholes model belongs to the class of local volatility models. In that case, $\sigma(X_t, t) = \sigma X_t$, where σ is a positive constant and so the volatility is constant. Among the other local volatility models, the most famous one is probably the constant elasticity of variance model (CEV model) introduced in [6]. In the CEV model, the function $\sigma(x,t)$ is of the form $\alpha x^{1+\beta}$, where $\alpha > 0$ and $0 \leq \beta \leq 1$ are constants.

A first interest of local volatility models is that in spite of the random nature of the volatility, they remain arbitrage free and complete. Mostly, the popularity of these models is due to the work of Dupire. He showed in [10] that provided the market is arbitrage free, one can find a local volatility function which enables to fit exactly the implied volatility surface of the European call options, see also [7]. However, local volatility models have several drawbacks. In particular, since only a finite number of strikes and maturities are available on the market, the derivation of the local volatility function requires interpolation-type methods, leading sometimes to highly unstable results. Moreover, such models do not allow for relevant smile dynamics. In particular, it is shown in [13] that under this specification for the volatility, the spot volatility smile moves in the opposite direction as the underlying. Also, in these models, the forward smile has a flattening dynamic, see [12]. These facts are not in agreement with the behaviors observed on the market. Thus, one source of randomness is not always enough to manage the smile risk.

To remedy this, in so-called stochastic volatility models, one therefore increases the dimension of the underlying Brownian motion. Indeed, they can be written as

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t,$$

where the process σ_t satisfies

$$\mathrm{d}\sigma_t^2 = \overline{a}_t \mathrm{d}t + \overline{\sigma}_t \mathrm{d}W_t + \overline{v}_t \mathrm{d}V_t,$$

with W and V two independent Brownian motions and $\overline{v}_t$ a non degenerate process. Various specifications for stochastic volatility models have been proposed in the literature and are largely used in practice. Let us cite among others the works by Hull and White [15], Heston [14] and the SABR model introduced in [13]. Despite incomplete, these models are very popular among practitioners since they enable to obtain more suitable smiles and smile dynamics. Also, pricing formulas can be semi explicit, as in the case of the Heston model, see [14].

Thus, it is quite clear that the practical relevance of a model is essentially assessed through the lenses of derivatives pricing and hedging. Historical data from the underlying are hardly taken into account. In this paper, we adopt an opposite point of view. Indeed, we are interested in what historical data have to say about volatility model selection. More precisely, we want to know if the historical data are "compatible" with a local volatility model.

This question might appear surprising. Indeed, smile dynamics already seem to give an answer. However we want to stress the following fact: option prices, and more generally implied quantities, are clearly the most important data to deal with when one wants to hedge derivatives. Nevertheless, from a statistical point of view, options data are in practice not "fully reliable", in the sense that their link with historical data through the no free lunch assumption remains arguable. Therefore, although our question is probably not of very first interest from a classical mathematical finance perspective, it is however very natural and important if one really wants to understand the dynamics of asset returns.

Our way to partially answer the preceding question is the following. We use historical data of an asset X regularly observed over a fixed time period $[0,T]$:

$$X_0, X_{\Delta_n}, X_{2\Delta_n}, \dots$$

and work in a high frequency context, which means our asymptotic is that the time span between two observations Δ_n goes to zero. From these data, we build a test statistics S_n whose behavior is approximately standard Gaussian when X belongs to a given class Θ_0 of processes following local volatility models. Therefore, thanks to this statistics, we can construct a statistical test with level α through the rejection area $\{S_n^2 > u_{1-\alpha}\}$, where $u_{1-\alpha}$ is the $1-\alpha$ quantile of a $\chi^2(1)$ distribution. Since our aim is to compare local volatility models and stochastic volatility models, in order to get a meaningful test, we will also impose that our test statistics diverges to infinity when the data generating process belongs to a class Θ_1 of stochastic volatility models. The "tricky" part of the paper will be the construction of this test statistics S_n. Then, deriving its asymptotic behavior when $X \in \Theta_0$ or $X \in \Theta_1$ will be essentially a direct use of recent general results about semi-martingales.

Remark that this kind of tests is in the spirit of several procedures recently developed in the literature. The papers [8], [9], for instance, deal with goodness-of-fit tests for the local volatility models in the high-frequency framework. On the other hand, Aït-Sahalia and Jacod have addressed the following testing problems for Itō semi-martingales: Is the jump part of the semi-martingale equal to zero? Do the jumps have finite or infinite activity? Is the Brownian part equal to zero? Such kind of test procedures is also designed in [11] in order to assess the multifractal nature of data from a semi-martingale. Note that as in all these works, since we want to distinguish between two very large classes, we will only be able to give pointwise results and so the level of the test will not hold uniformly over Θ_0 (which means that the supremum over all $X \in \Theta_0$ of the probabilities of being in the rejection area will not be controlled).

In this paper, we treat the situation where the null hypothesis is that the asset price follows a local volatility model against the alternative that it follows a stochastic volatility model. One can also ask about the case where the null and alternative hypothesis are switched. We highlight the answer to this problem, which is in fact an easier one, in Remark 3.2.

The paper is organized as follows. In Section 2, we propose a test procedure in the (unrealistic) case where both X and σ^2 are observed at times $0, \Delta_n, 2\Delta_n \ldots$ In Section 3, we explain how to deduce from it a feasible test statistics when the volatility process is not observed. A simulation study is performed in Section 4 and the proofs are relegated to Section 5.

2 Case with observed volatility

In this section, we focus on the case where both the price X and the process σ^2 are observed at regular times. Remark that it implies that the volatility process σ_t/X_t is also observed at regular times. The results in the case where the volatility is not observed will naturally follow from those obtained in this section.

2.1 Statistical problem

On a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ we consider a one dimensional continuous Itō process of the form

$$\mathrm{d}X_t = a_t \mathrm{d}t + \sigma_t \mathrm{d}W_t, \quad t \in [0, T],$$

where a is an adapted càdlàg drift process, σ is a positive adapted càdlàg process and W denotes a standard Brownian motion. The process X is assumed to be observed at equidistant time points $t_i = i\Delta_n$, $i = 0, \ldots, [T/\Delta_n]$ and $\Delta_n \to 0$.

Our aim is to decide on the basis of observations $(X_{i\Delta_n}, \sigma^2_{i\Delta_n})$, $i = 0, \ldots, [T/\Delta_n]$, whether the process σ is a deterministic function of X (Markov diffusion case) or if it follows a stochastic volatility model. Formally, the null hypothesis is given as

H_0 : The process σ_t^2 has the form $\sigma_t^2 = z(X_t)$ for a positive function $z \in C^2(\mathbb{R})$ with $z' > 0$.

while the alternative is given by

H_1 : The process σ_t^2 follows a stochastic volatility model $\mathrm{d}\sigma_t^2 = \overline{a}_t \mathrm{d}t + \overline{\sigma}_t \mathrm{d}W_t + \overline{v}_t \mathrm{d}V_t$ for two independent Brownian motions W and V and some adapted càdlàg processes $\overline{a}$, $\overline{\sigma}$ and $\overline{v}$ with $\overline{v}$ being non-vanishing on some measurable set $A \subset [0,T]$ with $\lambda(A) > 0$ (a.s.). Moreover, the processes $\overline{\sigma}$ and $\overline{v}$ are Itō semi-martingales.

Some remarks are in order.

Remark 2.1. By Itō's formula we obtain under H_0 that

$$\mathrm{d}\sigma_t^2 = \big(a_t z'(X_t) + \frac{\sigma_t^2}{2} z''(X_t)\big)\mathrm{d}t + \sigma_t z'(X_t)\mathrm{d}W_t.$$

That is, under H_0, σ^2 follows a stochastic volatility model as in H_1, but with $\overline{v} = 0$. Note however that the models in H_0 and H_1 do not contain all Itō semi-martingale processes for σ^2: the model $\mathrm{d}\sigma_t^2 = \overline{a}_t \mathrm{d}t + \overline{\sigma}_t \mathrm{d}W_t$ with $\overline{\sigma}_t \neq \sigma_t z'(X_t)$ belongs neither to H_0 nor to H_1.

Remark 2.2. The ad-hoc assumption $z' > 0$ in the null hypothesis enables us to provide a test procedure. It looks rather strange, but it is satisfied for many price models. This is in particular true in the Black-Scholes and CEV models mentioned in the introduction. Indeed, in these cases $z(x) = x^p$ for some $p > 0$ and $X_t > 0$.

2.2 Building the test statistics

The starting idea for the construction of our test is quite simple. For any $s, t > 0$, since z is increasing, we have under H_0

$$(X_t - X_s)(\sigma_t^2 - \sigma_s^2) = (X_t - X_s)\big(z(X_t) - z(X_s)\big) \geq 0.$$

Under H_1, the above positivity cannot hold a.s. as long as $\overline{v}$ is non-vanishing (in the sense of H_1).

The first step for our procedure is to choose a function $g : \mathbb{R} \to \mathbb{R}$ with the following properties:

(i) $g \in C_p^1(\mathbb{R})$ (C^1-functions of polynomial growth) with $g(x) > 0$ for $x > 0$ and $g(x) < 0$ for $x < 0$.

(ii) $g^+ = \max(g, 0) \in C^1(\mathbb{R})$.

A simple example of such a function is given by $g(x) = x^3$ (or more generally by $g(x) = x^{2k+1}$, $k \geq 1$). Our test statistic will be based on the quantity $M_{n,T}$, with for $t \leq T$,

$$M_{n,t} = \Delta_n \Big(\sum_{i=2}^{[t/\Delta_n]} g^+\big(\frac{X_{i\Delta_n} - X_{(i-2)\Delta_n}}{\sqrt{2\Delta_n}} \cdot \frac{\sigma^2_{i\Delta_n} - \sigma^2_{(i-2)\Delta_n}}{\sqrt{2\Delta_n}}\big) - \sum_{i=1}^{[t/\Delta_n]} g\big(\frac{X_{i\Delta_n} - X_{(i-1)\Delta_n}}{\sqrt{\Delta_n}} \cdot \frac{\sigma^2_{i\Delta_n} - \sigma^2_{(i-1)\Delta_n}}{\sqrt{\Delta_n}}\big)\Big).$$

Thanks to the properties of g, the products of increments involved in $M_{n,t}$ are always positive under H_0. Since $g^+ = g$ on $\{g(x) > 0\} = \{x > 0\}$, this will imply that, under H_0, $M_{n,T}$ goes to 0 in probability (see Corollary 2.5). On the other hand, under the alternative, we will have $\lim_{n\to\infty} M_{n,t} > 0$.

Remark 2.3. Notice that we compare two estimators at frequencies Δ_n and $2\Delta_n$ in the definition of $M_{n,t}$. When the process σ^2 is fully observed it is more natural to use the same frequency Δ_n. Indeed, in this case we would accept the null hypothesis only if $M_{n,t} = 0$ identically. However, in practice, the process σ^2 is not observed (see Section 3). Thus, we need to use an estimator $\hat{\sigma}^2$ instead of the true process σ^2 in the definition of $M_{n,t}$. Now, if we would use the same frequency in our test statistic, say Δ_n, the asymptotic results would solely come from the approximation error when replacing σ^2 by its empirical analogue $\hat{\sigma}^2$. In such situation, the obtained central limit theorems would be typically *infeasible* (i.e. the asymptotic results cannot be used for statistical inference). The reason is the following: the resulting centered and properly normalized statistics would essentially be an odd functional of the price process X and in this case central limit theorems are known to be infeasible. We refer to [18] for asymptotic results for general functionals of continuous semi-martingales.

2.3 Law of large numbers

Before we proceed with the weak law of large numbers for $M_{n,t}$, we need to introduce some notation. For k a positive integer, Ψ a 2×2-matrix and f a function from $(\mathbb{R}^2)^k$ to $\mathbb{R}$, we set

$$\rho_\Psi^{\otimes k}(f) = \mathbb{E}[f(\Psi U_1, \ldots, \Psi U_k)],$$

where $U_1, \ldots, U_k$ are iid bidimensional Gaussian vector with covariance matrix equal to identity. If $k = 1$, we simply write $\rho_\Psi(f)$. We also use the following representation of (X, σ^2):

$$\mathrm{d}\begin{pmatrix} X_t \\ \sigma_t^2 \end{pmatrix} = \begin{pmatrix} a_t \\ \overline{a}_t \end{pmatrix}\mathrm{d}t + \Sigma_t \mathrm{d}\begin{pmatrix} W_t \\ V_t \end{pmatrix}, \qquad \Sigma_t = \begin{pmatrix} \sigma_t & 0 \\ \overline{\sigma}_t & \overline{v}_t \end{pmatrix}.$$

Note that this representation holds under H_1, and under H_0 with $\overline{v} = 0$ and $\overline{\sigma}_t = \sigma_t z'(X_t)$. We have the following law of large numbers.

Theorem 2.4. *Let $h(x,y) = g(xy)$. Under H_0 and under H_1, it holds that*

$$M_{n,t} \to M_t = \int_0^t \rho_{\Sigma_u}(h^+ - h)\mathrm{d}u,$$

in probability, uniformly over compact sets in $[0,T]$.

Let $A = \{s,\ \overline{v}_s \neq 0\} \subset [0,t]$ (then $A = \emptyset$ under H_0 and $\lambda(A) > 0$ (a.s.) under H_1). Under H_1, since $\mathbb{P}(U_1U_2 > 0) > 0$ and $\mathbb{P}(U_1U_2 < 0) > 0$ for any normal variable (U_1, U_2) with correlation $|\rho| < 1$, we have

$$\int_A \rho_{\Sigma_u}(h^+ - h)\mathrm{d}u > 0.$$

On the other hand, it holds that $M_t = 0$ under H_0. Therefore, $M_{n,T}$ has different behaviors under H_0 and under H_1. Thus we have the following corollary which will enable us to construct a test statistics in the remaining part of the section.

Corollary 2.5. *We have the following convergence in probability: under H_0, $M_{n,T} \to 0$ and under H_1, $M_{n,T} \to M_T > 0$.*

2.4 Central limit theorem

Before defining the test statistics, we first give a general central limit theorem associated to the preceding law of large numbers. Let $f_i : (\mathbb{R}^2)^2 \to \mathbb{R}$, $i = 1, 2$, be defined by

$$f_1\big((x_1,x_2),(y_1,y_2)\big) = g^+\big(\frac{(x_1+y_1)(x_2+y_2)}{2}\big),\ f_2\big((x_1,x_2),(y_1,y_2)\big) = g(x_1x_2).$$

Define also $\tilde{f}_i : (\mathbb{R}^2)^3 \to \mathbb{R}$, $i = 1, 2$, by

$$\tilde{f}_i\big((x_1,x_2),(y_1,y_2),(z_1,z_2)\big) = f_1\big((x_1,x_2),(y_1,y_2)\big)f_i\big((y_1,y_2),(z_1,z_2)\big).$$

Finally, we set $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$ and $\Delta_i^{n,2} X = X_{i\Delta_n} - X_{(i-2)\Delta_n}$. We have the following result which will be our basis for the test statistics.

Theorem 2.6. *Under H_0 and under H_1, the process*

$$\Delta_n^{-1/2}\begin{pmatrix} \Delta_n \sum_{i=2}^{[t/\Delta_n]} g^+\Big(\frac{\Delta_i^{n,2} X \Delta_i^{n,2} \sigma^2}{2\Delta_n}\Big) - \int_0^t \rho_{\Sigma_u}(h^+)\mathrm{d}u \\ \Delta_n \sum_{i=1}^{[t/\Delta_n]} g\Big(\frac{\Delta_i^n X \Delta_i^n \sigma^2}{\Delta_n}\Big) - \int_0^t \rho_{\Sigma_u}(h)\mathrm{d}u \end{pmatrix}$$

converges stably in law towards a process V_t that is, conditionally on $\mathcal{F}$, a centered Gaussian process with independent increments, such that $(i,j = 1,2)$

$$\Theta_{ij,t} = \mathbb{E}[V_{i,t}V_{j,t}|\mathcal{F}] = \int_0^t R^{ij}_{\Sigma_u}\mathrm{d}u,$$

where

$$R_\Sigma^{11} = \rho_\Sigma^{\otimes 2}(f_1^2) + 2\rho_\Sigma^{\otimes 3}(\tilde{f}_1) - 3\big(\rho_\Sigma^{\otimes 2}(f_1)\big)^2$$
$$R_\Sigma^{12} = \rho_\Sigma^{\otimes 2}(f_1 f_2) + \rho_\Sigma^{\otimes 3}(\tilde{f}_2) - 2\rho_\Sigma^{\otimes 2}(f_1)\rho_\Sigma^{\otimes 2}(f_2)$$
$$R_\Sigma^{22} = \rho_\Sigma^{\otimes 2}(f_2^2) - \big(\rho_\Sigma^{\otimes 2}(f_2)\big)^2.$$

Consequently, we deduce that

$$\Delta_n^{-1/2}\big(M_{n,t} - \int_0^t \rho_{\Sigma_u}(h^+ - h)\mathrm{d}u\big)$$

converges stably in law towards a mixed normal random variable with conditional variance equal to $\Theta_{11,t} - 2\Theta_{12,t} + \Theta_{22,t}$.

2.5 The test statistics

In order to obtain a formal test statistics, we need to estimate the conditional covariance matrix Θ. We give such estimates in the following theorem. For a function k, we define

$$k[i,n] = k\Big(\frac{\Delta_i^n X \Delta_i^n \sigma^2}{\Delta_n}\Big), \quad k[i,n,2] = k\Big(\frac{\Delta_i^{n,2} X \Delta_i^{n,2} \sigma^2}{2\Delta_n}\Big).$$

We have the following result.

Theorem 2.7. *Consistent estimates for the terms of asymptotic covariance matrix in Theorem 2.6 are given by*

$$\hat{\Theta}_{11,t} = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \big\{g^+[i+1,n,2](g^+[i+1,n,2] + 2g^+[i+2,n,2] - 3g^+[i+3,n,2])\big\}$$
$$\hat{\Theta}_{12,t} = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \big\{g^+[i+1,n,2](g[i,n] + g[i+1,n] - 2g[i+2,n]\big\}$$
$$\hat{\Theta}_{22,t} = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \big\{g[i,n](g[i,n] - g[i+1,n]\big\}.$$

We can now define our test statistics S_n by

$$S_n = \frac{\Delta_n^{-1/2} M_{n,T}}{\sqrt{\hat{\Theta}_{11,T} - 2\hat{\Theta}_{12,T} + \hat{\Theta}_{22,T}}}.$$

We have the following corollary.

Corollary 2.8. *Under* H_0*,* S_n^2 *converges in law to a* $\chi^2(1)$ *distribution. Under* H_1*,* S_n^2 *converges in probability to infinity.*

Therefore, we reject the null hypothesis at level $\alpha \in (0,1)$ when $S_n^2 > u_{1-\alpha}$, where $u_{1-\alpha}$ denotes the $1-\alpha$ quantile of a $\chi^2(1)$ distribution.

3 Non observed volatility

Of course, in practice, we only observe over $[0, T]$ the sample

$$X_0, X_{\Delta_n}, X_{2\Delta_n}, \dots$$

The process σ^2 is not observed and needs to be locally estimated. To adapt the test statistics built in Section 2, the idea is to use a subsample

$$(X_0, \hat{\sigma}_0^2), (X_{\Delta'_n}, \hat{\sigma}^2_{\Delta'_n}), (X_{2\Delta'_n}, \hat{\sigma}^2_{2\Delta'_n}), \dots$$

with a slower frequency Δ'_n (instead of Δ_n), where $\hat{\sigma}_s^2$ is a consistent estimate of the true value σ_s^2 based on the price observations $(X_{i\Delta_n})_{i\geq 1}$. If the estimation accuracy of the volatility σ^2 is sufficiently good compared to Δ'_n, the asymptotic results of Section 2 remain valid when Δ_n is replaced by Δ'_n and σ^2 is replaced by $\hat{\sigma}^2$. More precisely, using classical localization procedures, we immediately deduce the following result.

Theorem 3.1. *Assume that there exists an estimator $\hat{\sigma}^2$ of σ^2 such that the sequence*

$$u_n^{-1} \sup_{s\in[0,T]} |\hat{\sigma}_s^2 - \sigma_s^2|$$

is tight, for some u_n tending to zero such that $(\Delta'_n)^{-1} u_n \to 0$. If we replace in the definition of $M_{n,t}$ and in all the results of Section 2 the increments $\Delta_i^n \sigma^2$ by their empirical counterpart $\Delta_i^n \hat{\sigma}^2$, then all the results of Section 2 still hold provided we also replace Δ_n by Δ'_n. In particular, the test statistic

$$\hat{S}_n = \frac{\Delta_n'^{1/2} \left(\sum_{i=2}^{[T/\Delta'_n]} g^+\left(\frac{\Delta_i'^{n,2} X \Delta_i'^{n,2} \hat{\sigma}^2}{2\Delta'_n} \right) - \sum_{i=1}^{[T/\Delta'_n]} g\left(\frac{\Delta_i'^{n} X \Delta_i'^{n} \hat{\sigma}^2}{\Delta'_n} \right) \right)}{\sqrt{\hat{\Theta}_{11,T} - 2\hat{\Theta}_{12,T} + \hat{\Theta}_{22,T}}}, \quad (3.1)$$

where the quantities $\hat{\Theta}_{kl,T}$ are defined as in Theorem 2.7 with (Δ_n, σ^2) replaced by $(\Delta'_n, \hat{\sigma}^2)$, converges in distribution to a standard normal variable under H_0 and diverges to infinity in probability under H_1.

Consistent pointwise estimates $\sigma^2_{n,s}$ of the process σ_s^2 can be obtained using realized variance over some local window around the time s, that is

$$\sigma^2_{n,s} = \frac{1}{2k_n\Delta_n} \sum_{i=[s/\Delta_n]-k_n}^{[s/\Delta_n]+k_n} (\Delta_i^n X)^2 \to \sigma_s^2, \qquad k_n\Delta_n \leq s \leq T - k_n\Delta_n, \quad (3.2)$$

in probability, provided $k_n \to \infty$ with $k_n\Delta_n \to 0$; see for example [4], [19] (the estimates $\sigma^2_{n,s}$ for $s \in [0, k_n\Delta_n)$ (resp. $s \in (T - k_n\Delta_n, T]$) are obtained similarly by using k_n increments of X on the right hand side of s (resp. on the left hand side of s)). The estimator $\sigma^2_{n,s}$ is probably the most intuitive one; however, one can also use any type of kernel estimators to get a proxy for σ_s^2.

The uniform bound u_n for the estimator $\hat{\sigma}^2 = \sigma_n^2$ is given by $u_n = (k_n \Delta_n)^{1/2}$, where k_n satisfies $k_n \le c\Delta_n^{-1/2}$ for some constant $c > 0$ (see e.g. [4]). The restriction on the window size k_n comes from the smoothness of the process σ^2. We see that the best attainable rate is essentially $u_n = \Delta_n^{1/4}$, which implies that Δ_n' has to converge to 0 slower than $\Delta_n^{1/4}$.

Remark 3.2. If we consider the null hypothesis of a stochastic volatility model (as described in Section 2) against the alternative of a local volatility model, the testing procedure is somewhat easier (in particular, we do not require the monotonicity assumption on the function $z \in C^2(\mathbb{R})$). Recall that the dynamics of the bivariate process (X_t, σ_t^2) is given by

$$\mathrm{d}\begin{pmatrix} X_t \\ \sigma_t^2 \end{pmatrix} = \begin{pmatrix} a_t \\ \overline{a}_t \end{pmatrix}\mathrm{d}t + \Sigma_t \mathrm{d}\begin{pmatrix} W_t \\ V_t \end{pmatrix}, \qquad \Sigma_t = \begin{pmatrix} \sigma_t & 0 \\ \overline{\sigma}_t & \overline{v}_t \end{pmatrix},$$

where $\overline{v} = 0$ in the local volatility case while $\overline{v}$ is a non-degenerate process for the stochastic volatility model. Thus, our test problem is equivalent to testing whether the bivariate process (X_t, σ_t^2) is generated by two independent Brownian motions (null hypothesis) or not (the alternative). The latter testing problem has been discussed in [17] for a general d-dimensional continuous Itō semi-martingale. Since the volatility process σ^2 is not observed, it has to be estimated by some $\hat{\sigma}^2$ (say, by σ_n^2) and then we can apply the procedure proposed in [17] for $d = 2$ (again we have to use a subsample $(X_{i\Delta_n'}, \hat{\sigma}^2_{i\Delta_n'})$ as above).

4 A simulation study

We give in this section some numerical results about our test procedure with $g(x) = x^3$. In the following, we work with equidistant observations

$$(X_0, X_{1/n}, \ldots, X_1),$$

in one of the two following model:

- The Black-Scholes model:

$$\mathrm{d}X_t = \sigma X_t \mathrm{d}W_t.$$

The value of the parameters in the simulations are $X_0 = 1$ and $\sigma = 0.2$.
- The non correlated Heston model:

$$\mathrm{d}X_t = \sigma_t \mathrm{d}W_t, \quad \sigma_t = s_t X_t,$$
$$\mathrm{d}s_t^2 = (a - k s_t^2)\mathrm{d}t + \varepsilon s_t \mathrm{d}V_t,$$

where W et V are two non correlated Brownian motions. The value of the parameters in the simulations are $X_0 = 1$, $s_0^2 =, 0.04$, $a = 0.02$, $k = 0.5$ and $\varepsilon = 0.1$, so that the Feller condition is satisfied.

4.1 Observed volatility

We begin with the case where σ_t is observed at the same instants as X_t. The obtained results for the behavior of our test over 3000 simulations with $n = 1025$ and $n = 131\ 073$ are given in Figure 1 and Table 1.

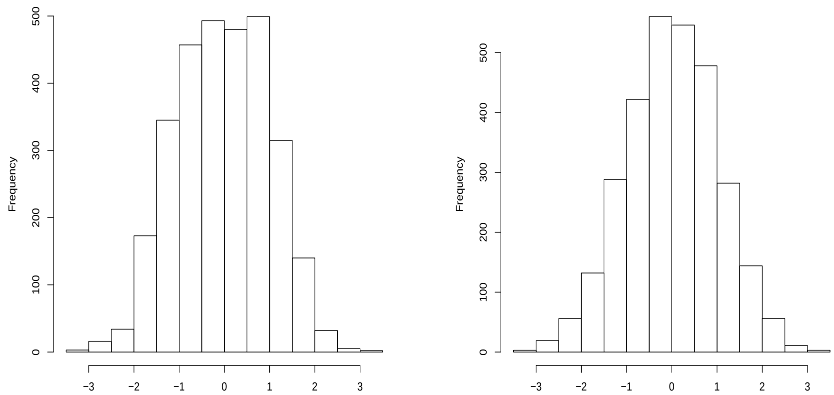

Figure 1: *Histogram of the test statistics in the Black-Scholes case, when the volatility is observed, for* $n = 1025$ *(left) and* $n = 131\ 073$ *(right).*

Simulated Process	Black-Scholes		Heston	
Number n of data	1025	131 073	1025	131 073
Level of the test				
10%	9.7%	10.4%	51.9%	100%
5%	3.7%	5.5%	39.9%	99.9%
1%	0.8%	1.0%	19.9%	99.8%

Table 1: *Percentage of rejection of the local volatility assumption.*

The results are quite satisfactory. Indeed, under H_0, for both values of n, the distributions of the test statistics are quite close to standard Gaussian and the rejection rates are not far from the theoretical ones. Also, the obtained empirical powers are quite reasonable.

4.2 Estimated volatility

We now assume that we observe n data from the price X and estimate the process σ^2 at $n' < n$ equidistant points. The squared volatility at time t is estimated using the estimator (3.2) on log price data and then converting it in order to obtain estimates for σ^2. In the simulations, we use $n = 131\ 073$,

$n' = 1025$, $k_n = 1000$ and compute the test statistic $\hat{S}_n^2$ introduced in (3.1). The value $k_n = 1000$ enables to obtain satisfying estimates σ^2 in the Heston case. Remark that σ^2 is particularly well estimated in the Black-Scholes case since the volatility is constant. The results from 3000 simulations are given in Figure 2 and Table 2.

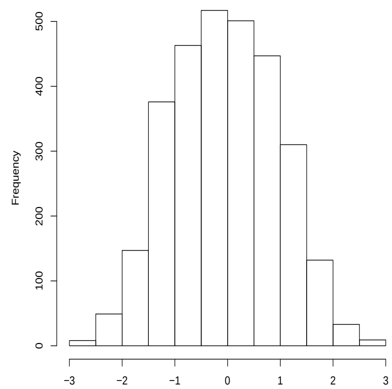

Figure 2: *Histogram of the test statistics in the Black-Scholes case, when the volatility is not observed.*

Simulated Process	Black-Scholes	Heston
Number n of data	131 073	131 073
Number n' of volatility estimates	1025	1025
Level of the test		
10%	9.2%	47.7%
5%	3.9%	36.1%
1%	0.7%	17.2%

Table 2: *Percentage of rejection of the local volatility assumption.*

Here again the results are fairly satisfactory. Indeed, they are of the same order of magnitude as those obtained for $n = 1025$ when the volatility is observed.

5 Proofs

5.1 Proof of Theorem 2.4

Let $Y = (X, \sigma^2)$. For $i = 1, 2$, we write

$$V(f_i, \Delta_n)_t = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} f_i\Big(\frac{\Delta_i^n Y}{\sqrt{\Delta_n}}, \frac{\Delta_{i+1}^n Y}{\sqrt{\Delta_n}}\Big),$$

where the f_i are defined before Theorem 2.6. Note that

$$\begin{aligned} \Delta_n \sum_{i=1}^{[t/\Delta_n]} f_1\Big(\frac{\Delta_i^n Y}{\sqrt{\Delta_n}}, \frac{\Delta_{i+1}^n Y}{\sqrt{\Delta_n}}\Big) &= \Delta_n \sum_{i=2}^{[t/\Delta_n]+1} g^+\Big(\frac{\Delta_i^{n,2} X \Delta_i^{n,2} \sigma^2}{2\Delta_n}\Big) \\ \Delta_n \sum_{i=1}^{[t/\Delta_n]} f_2\Big(\frac{\Delta_i^n Y}{\sqrt{\Delta_n}}, \frac{\Delta_{i+1}^n Y}{\sqrt{\Delta_n}}\Big) &= \Delta_n \sum_{i=1}^{[t/\Delta_n]} g\Big(\frac{\Delta_i^n X \Delta_i^n \sigma^2}{\Delta_n}\Big). \end{aligned}$$

From Theorem 6.1 in [16] (see also [5]), we obtain that for $i = 1, 2$,

$$\Delta_n V(f_i, \Delta_n)_t \to \int_0^t \rho_{\Sigma_u}^{\otimes 2}(f_i) \mathrm{d}u,$$

in probability, uniformly over compact sets in $[0, T]$. The result follows remarking that $\rho_{\Sigma_u}^{\otimes 2}(f_1) = \rho_{\Sigma_u}(h^+)$ and $\rho_{\Sigma_u}^{\otimes 2}(f_2) = \rho_{\Sigma_u}(h)$.

5.2 Proof of Theorem 2.6

Now consider

$$V(f, \Delta_n)_t = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} f_i\Big(\frac{\Delta_i^n Y}{\sqrt{\Delta_n}}, \frac{\Delta_{i+1}^n Y}{\sqrt{\Delta_n}}\Big),$$

where $f\big((x_1, x_2), (y_1, y_2)\big)$ is equal to $\begin{pmatrix} f_1\big(x_1, x_2), (y_1, y_2)\big) \\ f_2\big(x_1, x_2), (y_1, y_2)\big) \end{pmatrix}$. Since for $i = 1, 2$, $f_i\big(-x_1, -x_2), (-y_1, -y_2)\big) = f_i\big(x_1, x_2), (y_1, y_2)\big)$, from Theorem 7.1 in [16], we obtain the following stable convergence in law:

$$\Delta_n^{-1/2}\Big(\Delta_n V(f, \Delta_n)_t - \int_0^t \rho_{\Sigma_u}^{\otimes 2}(f)\Big) \to V(f)_t,$$

where $V(f)$ is, conditionally on $\mathcal{F}$, a centered Gaussian process with independent increments such that for $i = 1, 2$ and $j = 1, 2$,

$$\mathbb{E}[V(f_i)_t V(f_j)_t | \mathcal{F}] = \int_0^t R_{\Sigma_u}^{ij}(f) \mathrm{d}u,$$

where $R_\Sigma^{ij}(f)$ is equal to

$$\sum_{l=-1}^{1} \big(\mathbb{E}[f_i(\Sigma U_2, \Sigma U_3) f_j(\Sigma U_{l+2}, \Sigma U_{l+3})]\big) - 3\mathbb{E}[f_i(\Sigma U_1, \Sigma U_2)]\mathbb{E}[f_j(\Sigma U_1, \Sigma U_2)],$$

with U_1, U_2, U_3 some iid bidimensional Gaussian vector with covariance matrix equal to identity. After straightforward computations and using that f_2 only depends on the first variable, we obtain the expression for the asymptotic conditional covariance matrix in the first part of Theorem 2.6. An application of the Δ-method (for stable convergence) gives the second statement.

5.3 Proof of Theorem 2.7

For the proof of Theorem 2.7, simply remark that for $i = 1, 2$,

$$\big(\rho_\Sigma^{\otimes 2}(f_i)\big)^2 = \rho_\Sigma^{\otimes 4}(f'_i),$$

with $f'_i : (\mathbb{R}^2)^4 \to \mathbb{R}$ such that

$$f'_i\big((x_1, x_2), (y_1, y_2), (x'_1, x'_2), (y'_1, y'_2)\big) = f_i\big((x_1, x_2), (y_1, y_2)\big) f_i\big((x'_1, x'_2), (y'_1, y'_2)\big)$$

and that

$$\rho_\Sigma^{\otimes 2}(f_1)\rho_\Sigma^{\otimes 2}(f_2) = \rho_\Sigma^{\otimes 4}(f'_{12}),$$

with $f'_{12} : (\mathbb{R}^2)^4 \to \mathbb{R}$ such that

$$f'_{12}\big((x_1, x_2), (y_1, y_2), (x'_1, x'_2), (y'_1, y'_2)\big) = f_1\big((x_1, x_2), (y_1, y_2)\big) f_2\big((x'_1, x'_2), (y'_1, y'_2)\big).$$

Then the result follows directly from Theorem 6.1 in [16].

References

[1] Aït-Sahalia, Y. and J. Jacod (2009). Testing for jumps in a discretely observed process. *The Annals of Statistics*, **37** 184–222.

[2] Aït-Sahalia, Y. and J. Jacod (2010). Is Brownian motion necessary to model high frequency data? *The Annals of Statistics*, **38** 3093–312.

[3] Aït-Sahalia, Y. and J. Jacod (2008). Testing whether jumps have finite or infinite activity. *The Annals of Statistics*, to appear.

[4] Alvarez, A., Panloup, F., Pontier, P. and N. Savy (2010). Estimation of the instantaneous volatility. *Statistical Inference for Stochastic Processes*, to appear.

[5] Barndorff-Nielsen, O.E., S.E. Graversen, J. Jacod, M. Podolskij and N. Shephard (2006): A central limit theorem for realised power and bipower variations of continuous semimartingales. *In: From Stochastic Calculus to Mathematical Finance. Festschrift in Honour of A.N. Shiryaev, Y. Kabanov, R. Liptser and J. Stoyanov (Eds.), Springer, Heidelberg* 33–68.

[6] Cox, J.C. (1975). Notes on options pricing I: constant elasticity of variance diffusions. *Working paper, Stanford University.*

[7] Derman, E. and I. Kani (2010). Riding on a smile. *Risk*, **7(2)** 32–39.

[8] Dette, H. and M. Podolskij (2008). Testing the parametric form of the volatility in continuous time diffusion models - an empirical process approach. *Journal of Econometrics*, **143** 56–73.

[9] Dette, H., Podolskij, M. and M. Vetter (2006). Estimation of integrated volatility in continuous time financial models with applications to goodness-of-fit testing. *Scandinavian Journal of Statistics*, **33** 259–278.

[10] Dupire, B. (1994). Pricing with a smile. *Risk*, **7(1)** 18–20.

[11] Duvernet, L., Robert, C.Y. and M. Rosenbaum (2010). Testing the type of a semi-martingale: Itō against multifractal. *Electronic Journal of Statistics*, **4** 1300–1323.

[12] Gatheral, J. (2006). The volatility surface: a practitioner's guide. *Wiley.*

[13] Hagan, P.S., Kumar, D., Lesniewski A.S. and D. E. Woodward. (2002). Managing smile risk. *Wilmott Magazine*, September, 84-108.

[14] Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Finan. Stud.*, **6** 327–343.

[15] Hull, J.C. and A. White (1987). The pricing of options on assets with stochastic volatilities. *J. Finance*, **42** 281–300.

[16] Jacod, J. (2008). Statistics and high frequency data. Lecture notes.

[17] Jacod, J., Lejay, A. and D. Talay (2008). Estimation of the Brownian dimension of a continuous Itō process. *Bernoulli*, **14** 469–498.

[18] Kinnebrock, S. and M. Podolskij (2008). A note on the central limit theorem for bipower variation of general functions. *Stochastic Processes and Their Applications* , **118** 1056–1070 .

[19] Kristensen, D. (2010). Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach. *Econometric Theory*, **26** 60–93.

Research Papers 2011

2010-66: Ole E. Barndorff-Nielsen, David G. Pollard and Neil Shephard: Integer-valued Lévy processes and low latency financial econometrics

2010-67: Shin Kanaya and Dennis Kristensen: Estimation of Stochastic Volatility Models by Nonparametric Filtering

2010-68: Dennis Kristensen and Anders Rahbek: Testing and Inference in Nonlinear Cointegrating Vector Error Correction Models

2010-69: Søren Johansen: The analysis of nonstationary time series using regression, correlation and cointegration –with an application to annual mean temperature and sea level

2010-70: Søren Johansen and Morten Ørregaard Nielsen: A necessary moment condition for the fractional functional central limit theorem

2010-71: Nektarios Aslanidis and Isabel Casas : Modelling asset correlations during the recent financial crisis: A semiparametric approach

2010-72: Søren Johansen and Katarina Juselius: An invariance property of the common trends under linear transformations of the data

2010-73: Peter Sandholt Jensen and Allan H. Würtz: Estimating the effect of a variable in a high-dimensional regression model

2010-74: Peter R. Hansen, Asger Lunde and Valeri Voev: Realized Beta GARCH: A Multivariate GARCH Model with Realized Measures of Volatility and CoVolatility

2010-75: Laurent A.F. Callot: A Bootstrap Cointegration Rank Test for Panels of VAR Models

2010-76: Peter R. Hansen, Asger Lunde and James M. Nason: The Model Confidence Set

2011-01: Cristina Amado and Timo Teräsvirta: Modelling Volatility by Variance Decomposition

2011-02: Timo Teräsvirta: Nonlinear models for autoregressive conditional heteroskedasticity

2011-03: Roxana Halbleib and Valeri Voev: Forecasting Covariance Matrices: A Mixed Frequency Approach

2011-04: Mark Podolskij and Mathieu Rosenbaum: Testing the local volatility assumption: a statistical approach