Rapporto n. 231

Non Parametric Estimation Of Diffusion Coefficient: An Empirical Evidence

using Option Pricing on S&P 500

Romuald N. Kenmoe S

Luglio 2012

NonParametric Estimation Of Diffusion Coefficient:An Empirical Evidence using Option Pricing on S&P 500

Romuald N. Kenmoe S*

Abstract

In this paper, we discuss the estimation of the diffusion coefficient of an Itô process from high-frequency data using a nonparametric approach by Nadayara-Watson estimator. The principal purpose is to estimate the diffusion coefficient using selective estimators of spot volatility proposed by several authors, which are based on the observed trajectories. In general, statistical and econometrical criteria are used for comparing spot volatility estimators used in nonparametric estimators. We want to resort to merely financial metrics to achieve the same task. More precisely, the accuracy of different spot volatility estimates is measured in terms of how accurately they can reproduce market option prices. The model is implemented using S&P 500 data, and successively, we used it to estimate european call option prices written on the S&P 500 index. The estimation results are compared to well-known parametric alternative available in the literature. Empirical results not only provide strong evidence that most traditional pricing model are mispecified, but also confirm that the nonparametric model generates significantly different prices of common derivatives.

1   Introduction

We are concerned with the problem of confronting several spot volatilities and we analyse their capability as nonparametric estimator of diffusion coefficient. Since the early nineties, many authors have been questioned about the best way of estimating diffusion coefficient, namely volatility. The main motivation is the fact that it is the substratum in practically every financial application.

In large part of the continuous-time finance literature, the model specified as the underlying process of the state of log-asset prices, exchange rates, or spot interest rates is a time-homogeneous Itô diffusion process represented by the following stochastic differential equation (SDE):

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \tag{1.1}$$

with initial condition $X_0 = x_0$, where W_t is a standard real Brownian motion and the real function $\mu(x)$ and $\sigma(x)$ are such that a single solution X_t of the stochastic differential equation (1.1) exists. Our specific problem is to estimate the diffusion term $\sigma(x)$ when we observe a discrete realization of the process X_t, viz $n + 1$ observations $\hat{X}_0,...,\hat{X}_n$ at times $t_0=0 < t_1 < ... < t_n = T$ in the interval $[0, T]$.

The theory is constructed on Nadaraya-Watson estimators type which are given by the following formula :

---

*University of Milano-Bicocca, r.kenmoesiyou@campus.unimib.it

$$\hat{\sigma}^2(x) = \frac{\sum_{i=1}^{n-1} K\left(\frac{x-\hat{X}_i}{h}\right)\tilde{\sigma}_i^2}{\sum_{i=1}^{n-1} K\left(\frac{x-\hat{X}_i}{h}\right)} \qquad (1.2)$$

where $\tilde{\sigma}_i$ is a consistent estimate of the volatility at time t_i and h is a smoothing parameter (Silverman, 1986). The procedure studied is two-step. In the first step, we estimate the times series $\tilde{\sigma}_i$, which is implement in the second step using equation (1.2).

In many studies, the authors are concerned with estimating the parameters of equation (1.1), especially the diffusion coefficient. Returns and Volatilities are directly related to asset allocation, risk management and option pricing, proprietary trading. To achieve these objectives, the stochastic dynamics of the underlying state variables have to be specified correctly. For instance, option pricing theory allows us to value stock, index options or any value of a general asset and hedge against the risk of option writers once the model for the dynamic of underlying state variables is available. See the books of mathematical finance by Bjork (2010), Willmott (1998) among others. Althouh many of the stochastic models in use are simple and convenient ones to facilitate mathematical derivations and statitical inference, they are not derived from any economics theory and hence cannot be expected to fit time series financial data. Thus while the pricing theory gives relativately beautiful formulas when the underlying dynamics is correctly specified, it offers no or little guidance in choosing or validating the model. There is always the risk that misspecification due to parametric approach leads to erroneous valuation and hedging strategies. Furthermore there do not always exist the closed form solutions for the state variable and the derivative pricing specified by the function $\sigma(r) = \sigma(r; \overrightarrow{\theta})$ with $\overrightarrow{\theta}$ being a vector of real parameters. Hence, there are genuine needs for flexible stochastic modeling. Nonparametric approaches offer a complete and aesthetic treatement for tackling the above problems.

This paper studies the efficiency of some estimators of the spot volatilities which has been proposed in the recent literature and that use high frequency data. It is well known that financial high-frequency data evidence microstructure effects which render the classical estimator of the volatility inappropriate, namely the "realized volatility". Therefore, it is necessary to use volatility estimates which are robust in the presence of those effects. The article examines some of those estimators comparing their performance with pure financial criteria, namely in term of their ability to working out the price of options written on the S&P 500. Precisely, the use of the studied estimators of the spot volatilty permits, by means of a Nadayara and Watson regression type, to estimate the functional form of the diffusion coefficient in a local volatility model and we successively use it for pricing of the derivatives by Dupire's equation. This approach is based on the estimation of the volatility of the underlying asset, which is different from the classical techniques of derivatives pricing based exclusively on the partial differential equations (PDE) for the contingent claim. Furthermore, this allows to take into account large information contained in the high-frequency times series of the underlying asset which are generally neglected and can be of high interest when pricing "out of the money " options or when less information is available for options similar to those we want to evaluate. The two principal contributions of this article are: firstly, the comparison of different estimators of the spot volatility in term of option pricing. Secondly, we compare the result of this approach with those of classical (parametric) approach based on PDE, and successively, with the prices estimated using only daily data (low frequency).

Assuming that the underlying asset price follows a diffusion process, by imposing suitable conditions on the kernel, we can obtain the nonparametric volatility function of the underlying asset-return process. The constructed volatility will be the continuous-state analog to the im-

plied binomial tree proposed by Rubinstein (1994) and Derman and Kani (1994), and the implied volatility functions in Dupire (1994), and Dumas, Fleming and Whaley (1995). The prices obtained with the nonparametric approach are consistent with options pricing theory. Furthermore, the estimated volatility prevailing in our models evidences the main features of the true implied volatility and is in line generally with those in the literature for the different classes of options in which our sample is divided.

In this paper, we compare the competing nonparametric volatility estimators in term of pricing of the European call option written on the S&P 500. To our knowledge, this is the first investigation in this direction. Although many important contributions have been done on the stochastic volatility models, authors have focused on parametric estimations of volatility and then used it for the options pricing, see Hull and White (1987) and Heston (1993) among others. Stanton (1997), Jiang and Knight (1997), and Renò (2008) proposed nonparametric paradigm in estimating the diffusion coefficient, but they use it to evaluate the dynamic of the spot rate of interest rate. Jiang (1998) used nonparametric estimation of the drift, volatility and market price of risk for pricing the options written on interest rates. Ait-sahalia and Lo used the nonparametric estimate of the state price density to investigate the issue of pricing options written on stock indices. We present a different approach which is based on PDE's and use different nonparametric estimate of the volatility.

We have structured this paper as follows. In Section 2, we show how to estimate volatility from discrete observation in a close interval $[0, T]$ and we also present the whole spot volatility estimators inherent to this issue. In Section 3, we present a synthetic exposition of Dupire's formula and the connection with our models. Furthermore, we discuss the implementation technique for computing option price using nonparametric approach. In section 4, we recall some well-known models used to confront our proposed model, while in Section 5 we present empical results and compare our results with other models available in the literature. Finally, we summarise our findings in section 6.

2   Volatility Estimation

Let's consider the following SDE:

$$\begin{cases} dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \\ X_0 = x_0 \end{cases} \tag{2.1}$$

defined over the interval $[0, T]$, in the filtered probability space $(\Omega, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ satisfying the usual conditions.

The nonparametric estimator for the diffusion function $\sigma^2(.)$ of a general diffusion process is based on observing X_t at $\{t_1, t_2, ..., t_n\}$ in the time interval $[0, T]$. For simplicity, we will discuss only the equispaced data case. Subsequently, we let $\{X_{\triangle_n}, X_{2\triangle_n}, ..., X_{n\triangle_n}\}$ be n equispaced observations at discrete points $\{t_1 = \triangle_n, t_2 = 2\triangle_n, ..., t_n = n\triangle_n\}$, where $\triangle_n = \frac{T}{n}$. The following assumption is in force.

Assumption 1. Given the SDE (2.1), we have that,

1. $X_0 \in \mathcal{L}^2(\Omega)$ *is independent of* W_t, $t \in [0, T]$ *and measurable with respect to* $\mathcal{F}_0$

2. $\mu(x)$ *and* $\sigma(x)$ *are defined on a compact interval* I. $\mu(x)$ *is once continuously differentiable, while* $\sigma(x)$ *is twice continuously differentiable.*

3

3. *A constant K exists such that $0 < \sigma(x) \le K$ and $|\mu(x)| \le K$.*

4. *(Feller condition for non-explosion). Given:*

$$S(\alpha) = \int_0^\alpha e^{\int_0^y -\frac{2\mu(x)}{\sigma(x)}dx} dy \qquad (2.2)$$

$$V(\alpha) = \int_0^\alpha S'(\alpha) \int_0^y \frac{2}{S'(x)\sigma^2(x)} dx \qquad (2.3)$$

then $V(\alpha)$ diverges at the boundaries of I.

Assumption 1 guarantees the existence and the uniqueness of a strong solution. Requiring the Feller condition allows us to deal with the models that do not satisfy global Lipschitz and growth conditions (e.g., square process, see Ait-Sahalia,1996). Moreover, the Feller condition is both necessary and sufficient for recurrence in I; more details can be found in Bandi and Philips (2003): this will be the only condition required of our process in order to construct our model. On the other hand, one can demand global Lipschitz and growth conditions for μ and σ (Karatzas and Shreve, 1998). It is well known that index price stock are nonstationary, this particular feature of stock options renders the nonparametric estimation very challenging.

We assume further $K(.) \in L^2(\mathbb{R})$ to be a bounded kernel, that is, $\int_{-\infty}^\infty K(x)dx = 1$, continuously differentiable and with bounded and positive first derivative, with $lim_{x\to\infty}K(x) = lim_{x\to-\infty}K(x) = 0$. We opt to use the gaussian kernel, and therefore we set

$$K(s) = \frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}} \qquad (2.4)$$

We also define a sequence of bandwiths h_n such that: as $n \to \infty$, we have $h_n \to 0$ and $nh_n \to \infty$.

A very common value for h_n used in applications (Scott, 1992), (Silverman, 1986) is the following

$$h_n = h_s \hat\sigma n^{-\frac{1}{5}} \qquad (2.5)$$

where h_s is a real constant to be tuned, and $\hat\sigma$ is the standard deviation of analysed sample, $\hat\sigma^2 = Var[\hat{X}_i]$. We construct our nonparametric estimator as

$$S_n(x) = \frac{\sum_{i=1}^n K\left(\frac{x - X_{i\triangle_n}}{h}\right)\tilde\sigma_i^2}{\sum_{i=1}^n K\left(\frac{x - X_{i\triangle_n}}{h}\right)} \qquad (2.6)$$

where $\tilde\sigma_i$ is a consistent estimate of the spot volatility, and the first factor at the numerator and the denominator can be used to have a discrete approximation of the kernel function.

Several ways to estimate the spot volatilities can be found in the literature. In this article, we study six known methods, the first five are of realized volatility type while the last one is based on Fourier estimator. The realized volatility is based on the original idea of Merton (1980), who observed that the variance over a fixed interval could be estimated arbitrarily, although accurately, as the sum of squared realizations, provided the data are available at a sufficiently high sampling frequency. The Fourier type estimator was introduced by Malliavin and Mancino (2002) and it is based on Fourier analysis.

4

Florens-Zmirou (F-Z) [1993]
$\tilde{\sigma}_{Fz}^2(t_k') = \frac{(X(t_k')-X(t_{k-1}'))^2}{\triangle'}$
Comte and Renault (C-R) [1998]
$\tilde{\sigma}_R^2(t_k') = \frac{1}{m}\sum_{j=1}^{m}\frac{(X(t_{k-m/2+j}')-X(t_{k-m/2+j-1}'))^2}{\triangle'}$
Ogawa and Sanfelici (O-S) [2010]
$\tilde{\sigma}^2(t_k') = \frac{G(\rho)^{-1}}{L}\sum_{i=1}^{L}\frac{(\triangle_{k+l}'\bar{X})^2}{\triangle'}.$
Foster and Nelson (F-N) [1996]
$\tilde{\sigma}_E^2(t_k') = (1-\lambda)\sum_{j=1}^{i}\omega_i[X(t_{k-j+1}')-X(t_{k-j}')]^2.$
Andreou and Ghysels (A-G) [2002]
$\tilde{\sigma}_W^2(t_k') = \sum_{j=1}^{n_L}\omega_j[X(t_{k-j+1}')-X(t_{k-j}')]^2.$
Malliavin and Mancino (M-M) [2000]
$\tilde{\sigma}_{Fr}(t) = \lim\limits_{M\to\infty}\sum_{k=0}^{M}\left[a_k(\sigma^2)cos(kt)+b_k(\sigma^2)sin(kt)\right].$

Table 1: Specification of the Spot Volatilities

Different realized volatility type models can be obtained using different specifications for $\tilde{\sigma}_i^2$. There have been quite a few number of models discussed in the literature. Ogawa and Sanfelici (2010) listed almost all type of spot volatility estimators presented in the literature to confront them to their real time scheme estimator. Therefore, we borrow sometime from them to construct our kernel based volatility estimate. All the spot volatilities used in this paper are listed in Table 1, including the Fourier estimator.

Andersen and Bollerslev (1998) showed that it is possible to estimate daily volatility using intraday transactions, and that this estimates are by far more accurate than just using the daily squared return. Using high frequency data for this purpose renders the variance estimate more precise. Nevertheless, using the naive realized volatility estimator to implement our nonparametric estimator can be misleading, because intraday data exhibit pronounced seasonalities and microstructure contaminations that could severely distort the estimate. The analyzed spot volatilities used in this paper can mitigate this problem as they have been proposed for dealing with microstructure effects. Ait-Sahalia et al., (2005), Bandi and Russell (2006) proposed techniques for computing realized volatility when data are contaminated, see Ogawa and Sanfelici (2010) for the same topic with the Fourier estimator.

3   From Asset Dynamics to Options Pricing

The nonparametric estimates (2.6) of the diffusion coefficient can be seen as a local volatility since the estimated volatilities are constructed on the local trajectories of the underlying process. These nonparametric coefficient can be used to derive the corresponding price of European call options, based on Dupire's approach which can been resumed as follow.

Suppose that the dynamics of an underlying asset follow the equation,

$$\frac{dS_t}{S_t} = \mu(t, S_t)dt + \sigma(t, S_t)dW$$

where the local volatility function is parametrized as : $\sigma : [0,T]\times]0,\infty[\to \mathbb{R}_+$. Then the no-

arbitrage price of the European call option, satisfies the generalized Dupire partial differential equation

$$\begin{cases} \frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 C}{\partial K^2} + (r-D)(C - K\frac{\partial C}{\partial K}) & \forall (T,K) \in]0,T[\times\mathbb{R}_+ \\ C(0,K) = (S-K)^+ & \forall K \in \mathbb{R}_+ \end{cases} \tag{3.1}$$

where T is the time to expiry, K is the strike and D is the dividend rate. The initial condition is the pay off function. Theoretically, we can use this equation to find out the local volatility $\sigma(.,.)$ from a collection of option prices $(T,K) \to C(T,K)$ observed for a continuum of values of (T,K) by means of the Dupire's formula

$$\sigma(T,K) = \sqrt{\frac{2(C_T + rKC_K)}{K^2 C_{KK}}}$$

and this volatility is unique as Dupire assumes that there is a bijection between the call price $C(T_i, K_i)$ and the local volatility $\sigma(T_i, K_i)$.

When expressing the option price as a funtion of futures price $F_T = S_0 exp\{\int_0^T (r(t) - D(t))dt\}$, we would get the same expression except for the drift. That is

$$\frac{\partial C}{\partial T} = \frac{\sigma^2(K,T)}{2}\frac{K^2 \partial C}{\partial K^2} \tag{3.2}$$

where C now represents $C(F_T, K, T)$. Inverting this gives

$$\sigma^2(K,T,S_0) = \frac{\frac{\partial C}{\partial T}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}} \tag{3.3}$$

The right hand side of equation (3.3) can be computed from known European option prices.

Despite the fact that the theory ensures that there exists a unique local volatility, it is non-trivial problem to recover it from real option data. This drawback derives from the assumption of a well defined European option prices space, which is not the case on real markets. In practice, only a finite number of options $[C(T_i, K_j)]_{i,j}$ is available for different maturities T_i and strike K_j. This renders the problem strongly underdeterminate. For being close to the theory, practitioners resort to smoothness procedures or interpolation techniques of the implied volatility surface in order to obtain a continuous and smooth collection $\hat{C}(T,K)$ to which the Dupire's formula will be applied.

The set of prices $\hat{C}(T,K)$ on which the estimation of the volatility is based may not be a perfect observation of the market price, but a reconstruction (non uniqueness). Therefore, even in the theoretical case where the observed option prices come from a diffusion model, the reconstituted collection $\hat{C}(T,K)$ can differ from the theoretical price $C(T,K)$ obtained from the diffusion model. The difference between the reconstructed collection price $\hat{C}(T,K)$ using smoothness or interpolation can have a complex dependence in (T,K). We will not focus on that aspect here.

To avoid misspecification, related to the use of interpolation techniques, we back out the local volatility function $\sigma(.)$ from high frequency time series of the underlying through the kernel estimator (2.6) and then solve (3.2) to compute the option prices. An alternative way to determine the local volatility can be reached by following Derman and Kani, 1994 (DK) approach. The method presented in this section allows estimation of rather much smoother local volatility functions. From a numerical standpoint, these two approaches (DK and Dupire) are further different from our non-parametric in the sense that we are not imposing the recombination of a computational tree, nor constrained to evaluate numerically the differential of the prices of the option $C_{K,T}(t,S)$.

6

We use the simplified Dupire equation (3.2) where we suppose that the dividend yield $D(T)$ and the risk-free rate $r(T)$ are zero, namely we consider future prices. The same consideration can be used in the original Black & Scholes equation, but for computational convenience, we have chosen the Dupire equation. Given the above restriction, the Dupire equation can be written as

$$\begin{cases} \frac{\partial C}{\partial T} = \frac{\sigma^2(K)}{2} \frac{\partial^2 C}{\partial K^2} \\ C(K,0) = (F-K)^+ \end{cases}. \tag{3.4}$$

where $C(F,t,K,\tau)$ denotes the premium at time t for a given futures price F of an European call of strike K and time to expiry τ. This problem is a parabolic equation with initial condition $C(K,0)$ which can be interpreted as the price today of a call with strike K and immediate maturity. Both derivatives should be positive to avoid arbitrage. We will use equation (3.4) to determine the option prices using our nonparametric estimates for $\hat{\sigma}^2(K)$ obtained by means of kernel regression. The lack of term structure for our local volatility estimates will produce a poor estimation of options with longer maturties. One advantage of this forward equation-type is that all cross-section option series with the same maturity can be valued contemporaneously. A different approach is to solve the Black & Scholes PDE as many time as options of different exercise prices are; for the two approaches we obtain the same results. But the second approach will be much time consuming.

The new equation can be interpreted in another way. If $\sigma^2(K,T)$ is known, it establishes a relationship between the price as of today of call options of varying maturities and strikes.

Equation (3.4) can be seen as the opposed of the classical Black & Scholes partial differential equation which involves, for a fixed option (K and T fixed), derivatives with respect to the current time and value of the spot price. This happens if we set the interest rate equal to zero in the Black & Scholes equation, therefore we retrieve the following:

$$-\frac{\sigma^2(F,t)}{2} \frac{\partial^2 C}{\partial F^2} = \frac{\partial C}{\partial t}. \tag{3.5}$$

Equation (3.4) and (3.5) can be thought as operating in the same space of functionals. However, the relationship is not always true, as (3.5) applies to any contingent claim, though (3.4) holds because the intrinsic value of a call happens to be the second integral of a Dirac function.

In order to discretize the Dupire equation (3.4), we should work on a bounded open set $\mathcal{O} = (K_{min}, K_{max})$ chosen carefully in order to reduce the approximation error of the algoritm. We also need to specify the boundary conditions(Tavella and Randall, 2000). In our case, we shall impose the linearity conditions. These conditions are suitable for instance when a finite difference solution is required (e.g., when there are discrete dividends, when we want to use local volatility surface $\sigma(F,t)$ in our model, or when the strike price resets periodically), and simple, yet consistent boundary conditions are hard to define. It is found to be that, in a large number of option structures far from the strike price, or other such 'interesting' regions, the option value is nearly linear with respect to the spot. This observation is true for many exotic and path-dependent options, and plain vanilla options. This allows us to use rather small computational intervals (K_{min}, K_{max}).

Once the problem has been localized, we restrict K to belong to the interval $[K_{min}; K_{max}]$ and τ varying in the interval $[0; T]$, and obtain the problem

$$\begin{cases} \frac{\partial C}{\partial \tau} - \frac{\sigma^2(K)K^2}{2} \frac{\partial^2 C}{\partial K^2} = 0, & in \quad]K_{min}; K_{max}[\times [0; T], \\ \frac{\partial^2 C}{\partial K^2}(K_{min}, \tau) = 0, & if \quad \tau \in [0; T], \\ \frac{\partial^2 C}{\partial K^2}(K_{max}; \tau) = 0, & if \quad \tau \in [0; T], \\ C(K,0) = f(K) = (S_0 - K)^+ & for \ K \in]K_{min}; K_{max}[. \end{cases} \tag{3.6}$$

We discretize the problem (3.6) by finite difference method. Using the $\theta-$schemes, the time discretization of this equation can be defined. Suppose that $\theta \in [0, 1]$, and let k be the time-step such that $T = Nk$. We approximate the solution C at time nk by the c_h^n, with $c_h^n \in \mathbb{R}^d$ and the sequence $(c_h^n)_{n=0,\dots,N}$ are the solution of the recursive equation

$$
\begin{cases}
c_h^0 = f \\
\frac{c_h^{n+1}-c_h^n}{k} + (1-\theta)\tilde{A}^n c_h^n + \theta \tilde{A}^{n+1} c_h^{n+1} = 0 \ \ if \ \ 0 \le n \le N-1
\end{cases}
\tag{3.7}
$$

where $\tilde{A}^n = \tilde{A}_h(nk)$ is defined by the relation $\tilde{A}_h c_h = \frac{\hat{\sigma}^2(K_i)K_i^2}{2}\frac{c_{h,i-1}^{n-1}-2c_{h,i}^n+c_{h,i+1}^{n-1}}{h^2}$.
The previous system is solved by forward induction.
We can obtain two different schemes type according to the value of θ.

- if $\theta = 0$, the scheme is explicit,

- if $0 < \theta \le 1$, the scheme is implicit.

Therefore, we have to resolve a linear system of the form

$$
M^{n+1}c^{n+1} = q^n
\tag{3.8}
$$

where
$M^n = I - \theta k \tilde{A}^n$
$q^n = (I + (1-\theta)k\tilde{A}_h^{\ n})C_h^n$
with M^n an (d,d) tridiagonal matrix for any n. To solve this system, we can triangularize it at every time-step using the pivoting method.

4 Competitive methods

In order to assess the performance of our model in option pricing we will compare it with popular models: the original Black-Scholes model (B&S) (1973), Merton's jump diffusion model (M) (1976) and the Heston's stochastic volatility model (H) (1993).

In the B&S model, the underlying stock price is assumed to have the dynamic of the geometric Brownian motion diffusion process of the form

$$
dS_t = \mu S_t dt + \sigma S_t dW_t
$$

where σ and μ are the (constant) volatility and drift of the underlying asset.

In the Merton model the stock price underlying is assumed to follow a jump diffusion process

$$
dS_t = (\mu - \lambda_j k_j)S_t dt + \sigma S_t dW_t + (J_t - 1)S_t dq
$$

where q is a Poisson process uncorrelated with W and λ_j the intensity, which is the rate at which jumps occur, J_t is proportional increase in the stock price at time t and $k_j = E(J_t - 1)$ stands for the average jump size. The closed-form solution for the price of European call exists in the special case that the logarithmic of J_t is Normally distributed, with standard deviation δ_j,

$$
C_M = \sum_{n=0}^{\infty} \frac{e^{-\lambda'\tau}(\lambda'\tau)^n}{n!} C_{BS}(S, \tau; \sigma_n, r_n)
\tag{4.1}
$$

8

where $\lambda' = \lambda_j(1 + k_j)$, $\sigma_n^2 = \sigma_j^2 + n\delta_j^2/\tau$ and $r_n = r - \lambda_j k_j + nlog(1 + k_j)/\tau$

Finally, the Heston model assumes that the spot price at time t follows the diffusion

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_t^1$$

where W^1 is a Wiener process. The volatility V_t follows an Ornstein-Uhlenbeck process

$$d\sqrt{V_t} = -\gamma\sigma(t)dt + \delta(t)dW_t^2$$

and W^2 is another Wiener process such that W^1 and W^2 are correlated with correlation ρ. Let $x = \sqrt{V_t}$ and apply the Ito's formula in $f(x) = x^2$. The result is

$$dV_t = [\delta^2 - 2\gamma V_t]dt + 2\delta\sqrt{V_t}dW_t^2$$

Then if we let, $k = 2\gamma$, $\theta = \frac{\delta^2}{2\gamma}$, and $\sigma = 2\delta$ we end up with the Heston model where

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_t^1 \tag{4.2}$$

$$dV_t = k[\theta - V_t]dt + \sigma\sqrt{V_t}dW_t^2 \tag{4.3}$$

$$corr(dW^1, dW^2) = \rho dt. \tag{4.4}$$

V is the implied spot variance of the returns, k is the mean-reversion speed, θ is the long-run variance and σ is the volatility parameter of the diffusion volatility V_t. A closed form solution for this model is available in Heston (1993) and can be implemented by numerical integration of the characteristic function.

5 Empirical findings

In this section, we concentrate on the comparative analysis of all the nonparametric volatility estimators listed in the Table 1 and successively, we make the comparison of our kernel method with the classical calibration of popular models such as the Black & Scholes, the Jump-Diffusion model and the Heston model. All our empirical results will be based on the S&P 500 data set obtained from Chicago Mercantile Exchange (CME) for the sample period spans from January 2, 1990 to December 30, 1994.

5.1 Data description and model implementation

We focus on the options of the S&P 500 index, which are the most actively traded European-styled contracts. Therefore, S&P 500 options and options on S&P 500 futures have been analysed by Bates (1998), Dumas *et al*, Rubinstein (1994), Sanfelici (2007). These considerations provide us a motivation to apply on it our nonparametric estimators of the diffusion coefficient which is successively used to price S&P 500 call options obtained from the CBOE (Chicago Board Option Exchange). The analysed sample spans the period from January 4, 1993 to Decemder 31, 1993 (253 days).

Table 2 describes some statistics of our data set. During the one year considered period the variation exhibited by short-term interest rates is small in size: they range from 2.85 from 3.21%.

Variable	Mean	Std.dev	Min	Max
Call price, C($)	18.29	17.32	0.0249	68.60
Implied volatility (%)	11.01	2.98	5.07	36.83
τ(days)	79.96	66.56	1	350
K (index points)	445.52	27.48	375	550
F (index points)	454.87	10.21	429.18	474.21
r(%)	3.06	0.08	2.85	3.21

Table 2: Summary statistics for the sample of trade CBOE daily call option prices on the S&P 500 index in the period January 4,1993 to December 31,1993 (13 078).τ denotes the times to maturity, r the riskless rate, K the strike price and F the S&P 500 futures value implied from the call and put prices. ' Std.dev.' denotes the sample standard deviation of the variable. During this period, the average daily value of the variable of the S&P 500 index was 451.66

The options in our sample vary considerably in price and terms; for instance, the time-to-maturity varies from 1 to 350 days, with a median of 66 days.

The average total daily volume during the considered period was 65476 contracts. Following the CBOE practice, the expiration months are the three near term months followed by three additional months from the March quaterly cycle (March, June, September, December). The options are European, they expire on the third Friday of the month and the underlying asset is an index, the most likely case for which a lognormal assumption (with continuous dividend stream) can be justified. By the simple effect of diversification, jumps are less likely to occur in the index than in the individual equities. This feature allows us to think that the market is as close as the theoretical assumptions underlying the Black & Scholes model.

The beginning sample contains 16963 call options, we take the average of bid and ask price as our raw data. We consider observations with the time-to-maturity longer or equal to one day. Options having implied volatility greater than 70% and price less than 0.02 or greater than 70.00 were cancelled out. After that, we remain with a final sample of 13078 observations.

When applying our nonparametric approach to the raw data, we have to face two important problems. Firstly, in-the-money options are very infrequently traded with respect to at-the-money and out-the-money options, and hence they are notoriously unreliable. There is an unbalance daily volume for out-of-the-money contracts and the volume of in-the-money contract of the same magnitude.

Secondly, the index typically pays dividend and the future rate of dividend payment is difficult, if not impossible, to determine. The daily dividend provided by Standard and Poor's on the S&P 500 is by nature forward-looking, and there is no reason to assume that the actual dividends recorded ex post correctly reflect the expected future dividend at the time the options is priced.

We can tackle this uncomfortable drawback following Ait-sahalia and Lo (1998) and Sanfelici (2007). Based on the fact that all options are recorded at the same time on each day, we require only the temporally matched index price per day. To get around the unobservability of the dividend rate $\delta_{t,\tau}$, we deduce the futures price $F_{t,\tau}$ and S_t for each maturity τ. By the spot-futures parity, $F_{t,\tau}$ are linked through the relationship

$$F_{t,\tau} = S_t e^{(r_{t,\tau} - \delta_{t,\tau})}. \tag{5.1}$$

To derive the implied futures, we resort to the put-call parity relation, which must be satisfied if arbitrage opportunity are to be eliminated, independently of any parametric option-pricing model.

$$C(S_t, K, \tau, r_{t,\tau}, \delta_{t,\tau}) + Ke^{-r_{t,\tau}} = P(S_t, K, \tau, r_{t,\tau}, \delta_{t,\tau}) + F_{t,\tau}e^{-r_{t,\tau}} \tag{5.2}$$

where C and P denote respectively the call and the put price of actively traded options with the same strike K and the time-to-expiration τ. From this expression, we require reliable call and put prices at the same strike price K and time-to-expiration τ. For this purpose, we must use calls and puts that are closest to at-the-money. It is well known that in-the money options are illiquid relative to the out-of -the-money counterparts, hence any matched pair that is not at-the-money would have one potentially unreliable price.

We divide our data set into several categories as in Bakshi *et al.* (1997) and Sanfelici (2007). Following these authors the division criteria is according to moneyness S_t/K or the time-to-expiration τ. We say that a call option is *at-the-money* (ATM) if $0.98 < S_t/K \leq 1.02$, *out-of-the money* (OTM) if $S_t/K \leq 0.98$ and *in-the-money* (ITM) if $S_t/K > 1.02$. A finer repartition resulted in six moneyness categories. When classifying our data by term of expiration, we say an option contract has a *short term maturity* if $\tau < 60$ days, *medium term maturity* if $60 \leq \tau \leq 180$ days and *long term maturity* if $\tau > 180$ days. From empirical data, the proposed moneyness and maturity classification produce 18 categories which are reported in the Table 3 . ITM and ATM options account for respectively 50% and 24% of the total sample, while the short-term and the medium term take up respectively around 47% and 43%. The average price varies from \$0.19 for short-term deep OTM options to \$47, 29 for long-term deep ITM calls. For empirical study, each class can be used since the options are partitioned quite uniformly in the different moneyness categories and included nearly 1500 observations each and every.

We use the formula (2.6) to derive our nonparametric local volatility function from S&P 500 futures prices. Successively we evaluate the different approach by substituting each estimate in the equation (3.4) for pricing European calls.

We want to check if our nonparametric estimators of volatility verify the empirical facts exhibited by implied volatility when applied in the stock data. The Table 4 summarizes the main features of the implied volatility, which is obtained by inverting the Black & Scholes formula from each option price in our sample and successively producing an average 'implied' volatility for each moneyness-maturity category. The results show that the average B-S implied volatility tend to decrease monotically as the call options move from deep ITM to ATM and then this variation is stronger for the short-term options, which evidences a slight smile; this suggests that the short-term options are proned to severe mispricing. From the smile evidence, we can observe a negatively skewed implicit return distributions with excess kurtosis. Therefore, any acceptable model intended to price options written on S&P 500 have to be consistent with these features.

Many approaches have been proposed to extract local volatility function: Derman and Kani (1994), and Rubinstein (1994) proposed the non arbitrage binomial o trinomial tree model where the volatility function is obtained at the end of each node inverting the corresponding call price. In the continous time approach proposed by Dupire, B (1994), it is retrieved by means of the equation (3.4), where it is assumed to know the prices of options of all strikes and maturities via the implied volatility surface. Loosely speaking, we know the quantities $C_{K,T}(0, S)$ as a function of K and T, and therefore it is quite immediate to evaluate (numerically) the derivatives of the observed option prices with respect to the maturity and the strike price.

Our approach is completely different, in the sense that the volatility is estimated from the underlying index dynamics and does not rely on option prices, we construct our nonparametric estimator directly on the S&P 500 data using the Nadayara -Watson formula which has been explained in the previous section, we derive the nonparametric estimator by substituting in the latter the corresponding spot volatility estimators reported in the Table 1. Therefore we need to choose cleverly the parameter used to construct our local volatility function by means of our nonparametric estimator. We can assume that what we observe in the finite interval $[0, T]$ is

Moneyness S/K		<60	60-180	>180	Subtotal
OTM	<0,96	0.19	0.81	2.35	1.01
		478	979	436	1893
	0.96-0.98	0.78	2.90	10.43	2.42
		738	640	114	1492
ATM	0,98-1,00	2.66	6.52	13.50	4.66
		861	647	58	1566
	1.00-1.02	7.56	11.78	17.75	9.66
		874	656	56	1586
ITM	1.02-1,04	14.64	18.09	25.21	16.99
		788	611	143	1542
	>1.04	33.46	39.01	47.29	36.97
		2416	2192	391	4999
Subtotal		16.56	19.42	21.78	18.29
		6155	5725	1198	13078

Table 3: The reported numbers are, respectively, the average quoted bid-ask mid point price and the total of observations for each moneyness-maturity category. S denotes the spot S&P 500 index level and K is the exercise price.

a part of stationary process. Furthermore, if the kernel function is locally Lipschitz, it can be applied to the nonstationary data. Moreover, the other correction is to choose the bandwidth, we find that the smoothness parameters which provide acceptable results are contained in the interval $17 \leq h \leq 25$, out of this band we obtain unsastifying results. An higher value of h will lead to a smoother estimate of the density in the tails of the distribution where fewer data are available. We finally maintain $h = 20$ because it is the value of the bandwidth which give appealing outcomes when applying on every our six classes. We have used tick-by-tick data of S&P 500 from which we have extracted all price with observation frequency one minute in order to mitigate the effect of microstructure noise. Futhermore, the theory assumes that when the distance between two observations is close to zero, an arbitrary precision in the estimate of the spot volatility can be reached. Broadly speaking, these facts contribute to the improvement of our local volatility function. The estimator $\tilde{\sigma}(K, \tau)$ evidences a strong volatility smile in general. We will turn on this aspect later.

We describe the procedure used to estimate the other structural parameters of the alternative models. A well consolidated practice (Bakshi et al. 1997, Sanfelici 2007) is to compute option-implied parameters by implementing each model in the two steps as explained below.

1. Collect m option prices taken from the same point in time t (or same day) for any m greater than or equal to one plus the number of parameters to be estimated. For each $j = 1, ..., m$, let τ_j and K_j be respectively the time to expiration and the strike price of the jth option; let $C_j(t; \tau_j, K_j)$ be its observed price and $\hat{C}_j(t; \tau_j, K_j)$ its price worked out by the model with S_t and $r_{t,\tau}$ taken from the market. The difference $C_j - \hat{C}_j$ is a function of the values taken by $\Phi = \{\sigma\}$ in the B-S model, by $\Phi = \{\sigma_j, \lambda_j, \delta_j, k_j\}$ in the J-D model and $\Phi = \{\sqrt{V}, \sigma_\nu, k_j, \theta_\nu, \rho\}$ in the Heston model. For each j, we define

$$\epsilon_j(\Phi) = C_j(t; \tau_j, K_j) - \hat{C}_j(t; \tau_j, K_j).$$

Moneyness		Time to maturity			
S/K		< 60	60-180	> 180	Total
OTM	<0.96	8.51	8.00	8.53	8.25
	0.96-0.98	7.98	8.53	9.91	8.36
ATM	0.98-1.00	8.57	9.21	10.03	8.89
	1.00-1.02	9.80	9.98	10.45	9.90
ITM	1.02-1.04	11.43	10.80	11.04	11.15
	>1.04	11.78	13.02	12.49	13.83
Total		11.47	10.64	10.41	11.01

Table 4: Average B-S implied volatilities for different moneyness maturity categories. S denotes the spot S&P 500 index level and K is the exercise price.

2. Find the parameter vector Φ to solve the nonlinear least-squares problem at time t

$$SSE(t) = \min_{\Phi} \sum_{j=1}^{m} [\epsilon_j(\Phi)]^2 \tag{5.3}$$

Go back to Step 1 until the two steps have been repeated for each day in the sample.

The objective function (5.3) is defined as the sum of squared pricing error and may force the estimation to assign more weight to relative expensive options (e.g., ITM and long-term options). In the calibration procedure, when the $\frac{F_{t,\tau_j}}{K_j}$ does not belong to the grid G, the value C_j is approximated by the quadratic interpolation of the nearest grid points.

5.2   Model Comparison Results

Turning back to the behaviour exhibited by our nonparametric approach, the noteworthy fact is that for all the spot volatilities used in our model, the relative local volatility function presents a monotonically decreasing trend from deep in the money to the deep out-of-the-money. This is consistent with the empirical fact underlying the options written on stock index and it is in line with other methods existing in the literature. If we focus our analisys on all options available, the trend of any estimator is almost the same, albeit the Fourier estimator tends to generate the highest skew with respect to the striking price when confronting with other estimators. We can notice in particular that our estimate smile is exaggeratedly asymmetric, pointing to the empirical fact that the European call options written on the stock index lost their U-shape after the 1987 crash. This particular characteristic inherent to our models is consistent with empirical findings and contrasts stochastic volatility models which typically produce symmetric smiles. This is another example among others that nonparametric approach can suggest an information for constructing parametric models, if not to valid them. Much evidence can be found in Figures 6.1-6.6

Table 5 reports the average absolute error (APE), the percentage pricing error (PPE) and the daily averaged mispricing index (MISP)

$$APE = \frac{1}{N} \sum_{i=1}^{N} \left| C_i - \hat{C}_i \right| \tag{5.4}$$

$$PPE = \frac{1}{N} \sum_{i=1}^{N} \frac{\left| C_i - \hat{C}_i \right|}{C_i} \tag{5.5}$$

$$MISP = \frac{\sum\limits_{i=1}^{N}(C_i - \hat{C}_i)/C_i}{\sum\limits_{i=1}^{N}\left|(C_i - \hat{C}_i)/C_i\right|} \qquad (5.6)$$

where C_i is the observed call price available on the market, $\hat{C}_i$ is the call price worked out using nonparametric techniques, and N is the sample dimension. The mispricing index ranges from -1 and 1 and indicates on average, the overpricing (when it is negative) and the underpricing (when it is positive) induced by the model. In general, our model tends to overprice the options irrespectively on the class where it is applied. When the model underprices options, the bias is higher in magnitude. The APE index tends to be high when applied to all data available. The ATM and OTM classes have the lowest APE, but when moves to ITM class this index increases considerably, this is probably due to the fact that ITM options are illiquid, viz difficult to price. The PPE index presents a very alternated behaviour. When we confront our result using MISP index, we notice that, our estimators tend to overprice call price in general in all the analysed classes. But things change when we analyse the ITM class which exhibit very high value. In this case the MISP is positive for all the estimators, suggesting that this loss function underprices in-the-money options.

To gain a better insight into the different nonparametric models performance, we apply our model each time using one of the six alternative sets from the whole sample: ITM, ATM, OTM, short-term, meduim-term, long-term categories. The error tends to be small in the whole sample with respect to the other categories in terms of MISP, while the ITM class presents the lowest PPE index. When ATM options are priced, the resulting estimates do not significantly differ from their counterparts for the whole data set. OTM call options are associated with relative low pricing error, while ITM options correspond to higher mispricing error, indicating that, for the illiquid ITM calls to be priced properly, the volatility of the underlying needs to be higher than for all options of any maturity to be priced.

When analysing with respect to time to expiration, short-term option seem to be more challenging, they are associated with the highest pricing error. However, the structure of mispricing by term to expiration is very similar; this can be due to the fact that our model does not vary according to the time parameter. This can limit our nonparametric approach to be a true pricing engin.

In this last part we compare the results obtained with the nonparametric approach with those computed with some popular methods used by academicians and practitioners. That is the Black & Scholes, the Jump diffusion and the Heston model. In this comparison we have considered only the Fourier estimator. We do not compare them to the other estimators studied in this paper, because numerical results obtained from them have the same features. Therefore, the comments relative to Fourier estimator are also relevant for the remainder estimators

The analysed loss functions (5.4)-(5.6) suggest that for all the three alternative models, there is a positive mispricing index for all categories except for the ATM and the ITM classes in general. However, when using the MISP the price worked out using nonparametric approach tends to overestimate the option price for almost all categories apart from All options and ITM classes, as shown in table 6. The result however shows that the alternative models systematically underprice call price in general, except for the ITM calls for B-S and the ATM calls for J-D and the H model. Nevertheless as one can immagine, both the J-D and the H model are more accurate. To sum up, the nonparametric model represents a real improvement with respect to B-S in term of pricing

	All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
Florens-Zmirou							
APE	0.89	1.34	1.12	0.54	0.44	0.68	0.83
PPE	0.15	0.05	0.19	0.48	0.44	0.43	0.24
MISP	0.13	0.99	0.59	0.19	-0.90	-0.88	-0.39
Comte and Renault	All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
APE	0.90	1.56	1.11	0.59	0.44	0.69	0.83
PPE	0.15	0.05	0.15	0.48	0.44	0.43	0.24
MISP	0.12	0.99	0.59	0.18	-0.90	-0.88	-0.39
Foster and Nelson	All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
APE	1.36	1.54	1.76	0.80	0.38	0.84	1.33
PPE	0.21	0.06	0.25	0.58	0.14	0.16	0.27
MISP	0.96	0.99	0.98	0.41	-0.32	0.01	0.02
Andreou and Ghysels	All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
APE	0.97	1.59	1.22	0.55	0.41	0.66	0.90
PPE	0.14	0.06	0.16	0.92	0.37	0.35	0.24
MISP	0.39	0.91	0.74	-0.45	-0.85	-0.80	-0.24
Ogawa and Sanfelici	All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
APE	0.99	1.77	0.91	0.54	0.27	0.63	0.47
PPE	0.34	0.04	0.18	0.95	0.32	0.89	0.19
MISP	0.63	0.59	0.98	-0.55	-0.95	-0.56	-0.25
Malliavin and Mancino	All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
APE	0.69	0.18	0.89	0.53	0.39	0.64	0.96
PPE	0.15	0.05	0.16	0.80	0.32	0.30	0.21
MISP	-0.64	0.92	-0.21	-0.36	-0.82	-0.71	-0.04

Table 5: Average absolute error (APE) Mispricing index (MISP), and percentage (PPE) pricing errors between the market price and the kernel based-price.

		All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
APE	M-M	0.69	0.18	0.89	0.53	0.39	0.64	0.96
	B-S	0.53	0.32	0.41	0.26	0.33	0.69	1.15
	J-D	0.20	0.14	0.20	0.11	0.12	0.17	0.19
	H	0.14	0.01	0.20	0.11	0.13	0.15	0.18
PPE	M-M	0.29	0.05	0.16	0.80	0.32	0.30	0.21
	B-S	0.40	0.01	0.12	0.53	0.27	0.54	0.90
	J-D	0.10	0.006	0.05	0.20	0.08	0.07	0.10
	H	0.08	0.006	0.05	0.19	0.10	0.07	0.08
MISP	M-M	0.59	0.92	-0.21	-0.36	-0.82	-0.71	-0.04
	B-S	0.90	-0.09	0.59	0.74	0.84	0.88	0.82
	J-D	0.40	0.32	-0.003	0.13	0.19	0.39	0.43
	H	0.17	0.03	-0.03	0.02	0.39	0.45	0.29

Table 6: Average absolute error (APE) Mispricing index (MISP), and percentage (PPE) pricing errors between the market price and the kernel based-price for the Fourier using intraday data and B-S; J-D; H model.

		All options	ITM	ATM	OTM	SHORT	MEDIUM	LONG
APE	Foster and Nelson	28.66	4.79	5.42	6.24	26.24	23.42	19.05
	B-S	0.53	0.32	0.41	0.26	0.33	0.69	1.15
	J-D	0.20	0.14	0.20	0.11	0.12	0.17	0.19
	H	0.14	0.01	0.20	0.11	0.13	0.15	0.18
PPE	Foster and Nelson	31.12	0.21	2.46	19.53	35.27	19.06	9.14
	B-S	0.40	0.01	0.12	0.53	0.27	0.54	0.90
	J-D	0.10	0.006	0.05	0.20	0.08	0.07	0.10
	H	0.08	0.006	0.05	0.19	0.10	0.07	0.08
MISP	Foster and Nelson	-0.99	-0.26	-0.94	-0.70	-0.99	-0.99	-0.98
	B-S	0.90	-0.09	0.59	0.74	0.84	0.88	0.82
	J-D	0.40	0.32	-0.003	0.13	0.19	0.39	0.43
	H	0.17	0.03	-0.03	0.02	0.39	0.45	0.29

Table 7: Average absolute error (APE) Mispricing index (MISP), and percentage (PPE) pricing errors between the market price and the kernel based-price for the Rolling estimator using daily data and B-S; J-D; H model.

	All options
	CPUtime
F-Z	20.24
C-R	28.48
F-N	27.20
A-G	26.20
O-S	30.26
M-M	43.34
B-S	18.29
J-D	1.657e+03
H	3.834e+03

Table 8: CPU time in seconds for the computed European Call price on S&P 500 for different nonparametric methods and the alternative counterparts

properties. However, the performance is worse than the J-D and H, this can be justified by the fact that, these two models use many parameters which can produce an overfitting and in the other hand, our kernel estimator is not time dependent.

The Table 7 reports the results computed using daily data for the estimator proposed by Foster and Nelson. This estimator has been choosen because it has been originally constructed for low frequency data (daily, weekly and monthly). From the results, we notice that the pricing error increase exaggeratedly showing that when daily data are used for computing option prices much information contained in high frequency data are neglected.

Another important issue in the model calibration is the computational cost. The computational cost of the nonparametric is much lower than using the closed-form solution for the J-D model and the latter is cheaper than the H model. Therefore, The J-D model represents a good trade-off between performance and computational cost. The Black and Scholes model is the cheapest and the Fourier estimator evidences the highest computational time among our nonparametric estimates, as shown in Table 8.

In the end, our empirical evidence indicates that taking stochastic volatility into account gives the best improvement over the B-S formula. However, we can conclude that the nonparametric model contribute to explaining from theoretical and quantitative standpoint the strong pricing biases inducted in the B-S model.

6   Conclusion

Since the seminal work of Black & Scholes on options pricing, many researchers proposed sophisticated works in order better to evidence the empirical facts exhibited by the market. Following the local volatility approach pioneered by Derman and Kani (1994), Dupire (1994), and Rubinstein (1994) who supposed that the volatility is a deterministic function of asset price and time, we have studied a new approach and used it for evaluating European call option prices.

In this paper, we have studied a new technique for camparing the nonparametric estimation using a Nadayara-Watson kernel regression, which is based essentially on finance methods. We have used almost all types of spot volatility existing in the literature for computing the option price of European call options. We have shown that our method can be classified in the class of local volatility function which has the particularity of being complete. Relatively to other local volatility function proposed earlier, our method is easy to manage and the computation cost is very low.

The volatility curve obtained with our approach is consitent with the actual market. The Fourier estimator shows to exhibit better the market features compared with other competing volatility estimators analysed.

We compare the nonparametric model to well known popular models, such as Black & Scholes models itself, the Jump-Diffusion model and the Heston's model with stochastic volatility. The price worked out by nonparametric model is obtained using the Dupire's equation which the volatility is construted directly on the S&P 500 index future price spanning from January 4, 1993 to December 31, 1993. To sum up, the nonparametric model represents a real improvement with respect to B-S in term of pricing properties. However, the performance is worse than the J-D and H.

Second, the computational cost of the nonparametric is less than using the closed-form solution for the J-D model and which is cheaper than the H model. Therefore, The J-D model represents a good trade-off between performance and computational cost. The Fourier estimator is the most time consuming among all the studied nonparametric estimators.

In the end, our empirical evidence indicates that taking stochastic volatility into account gives the best improvement over the B-S formula. However, we can conclude that the nonparametric model contribute to explaining from theoretical and quantitative standpoint the strong pricing biases inducted in the B-S model.

A serious limitation to this approach is that it is not varying with respect to the time parameter, including the time variations in the Nadayara-Watson estimator, may be crucial to obtain correct specification. This will permit our model to be a real pricing engin, further investigations in that sense will be the object of future researches.
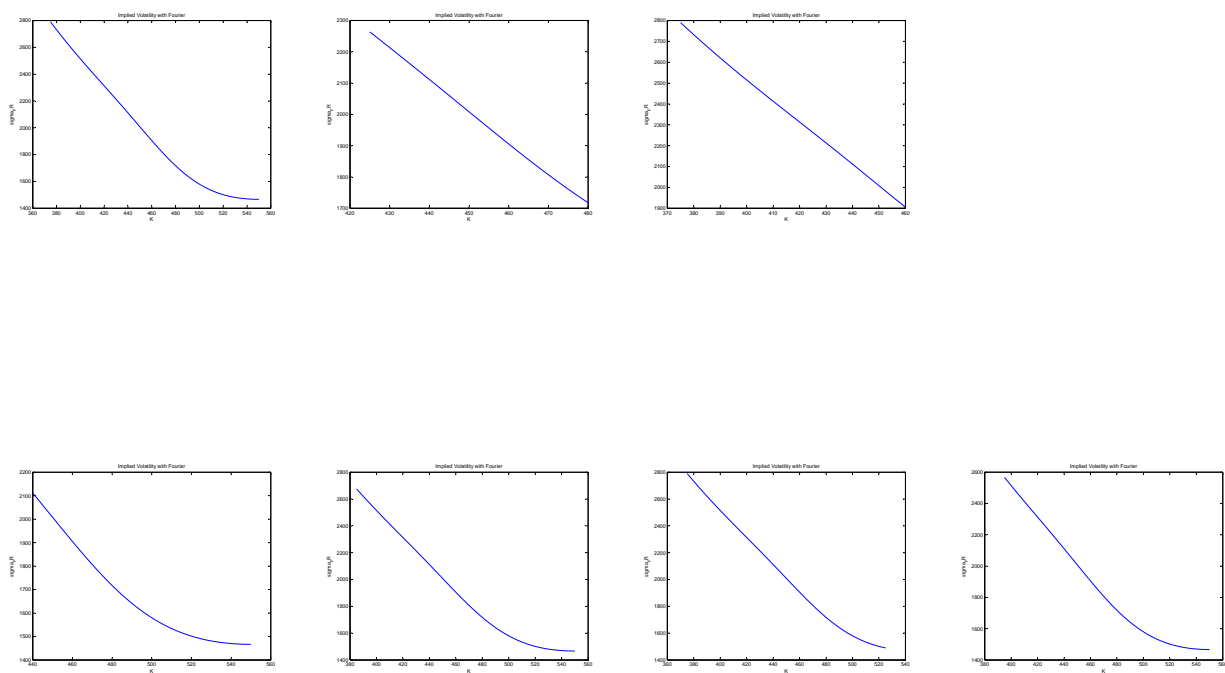
Figure 6.1: Local volatility Curves for Nonparametric Estimator with Mancino and Maliavin spot volatilty.The curve relative to all classes are reported
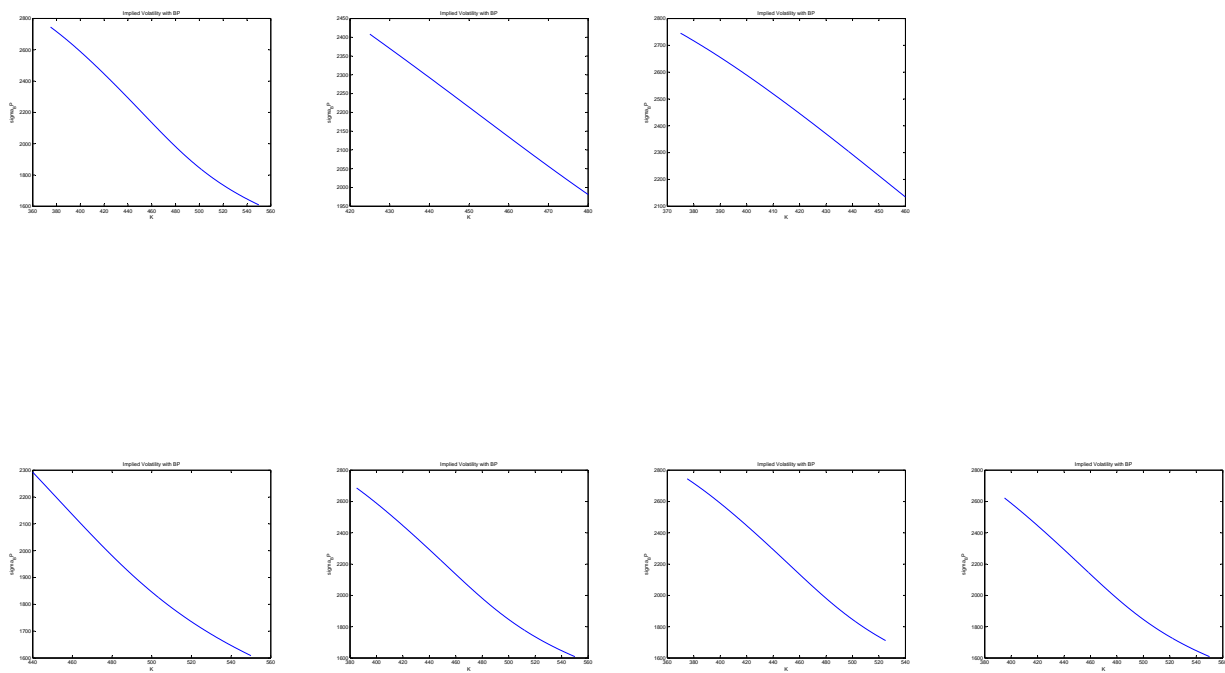
Figure 6.2: Local volatility Curves for Nonparametric Estimator with Comte and Renault spot volatilty.The curve relative to all classes are reported
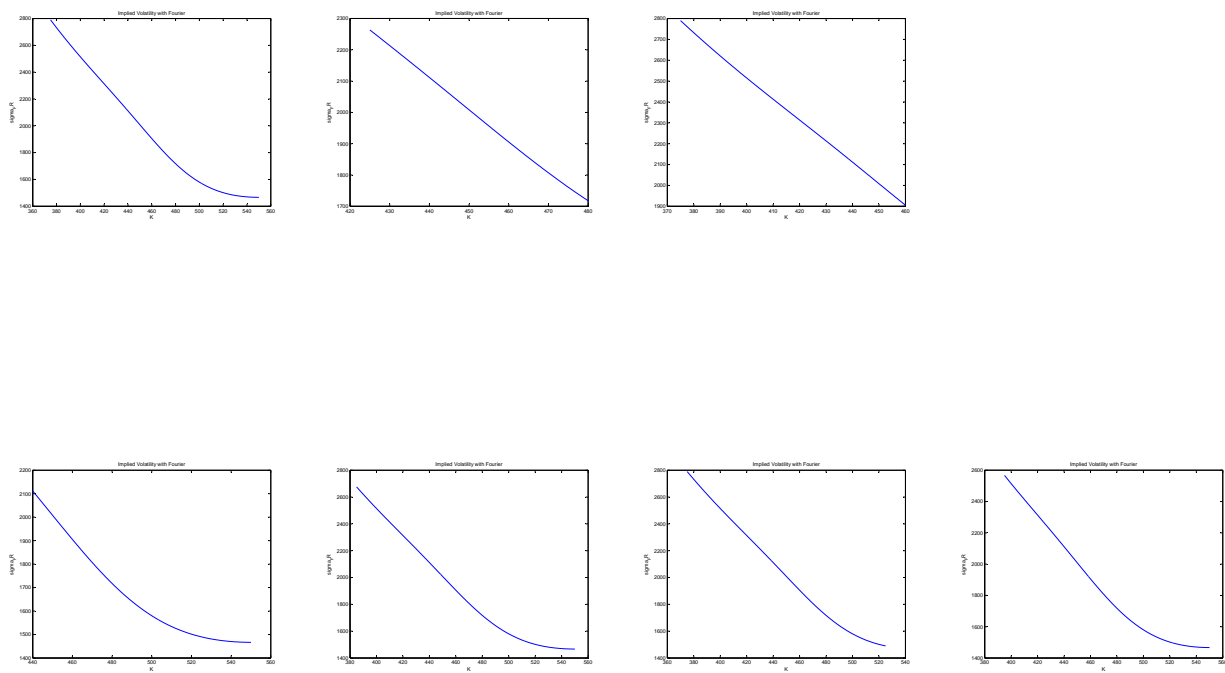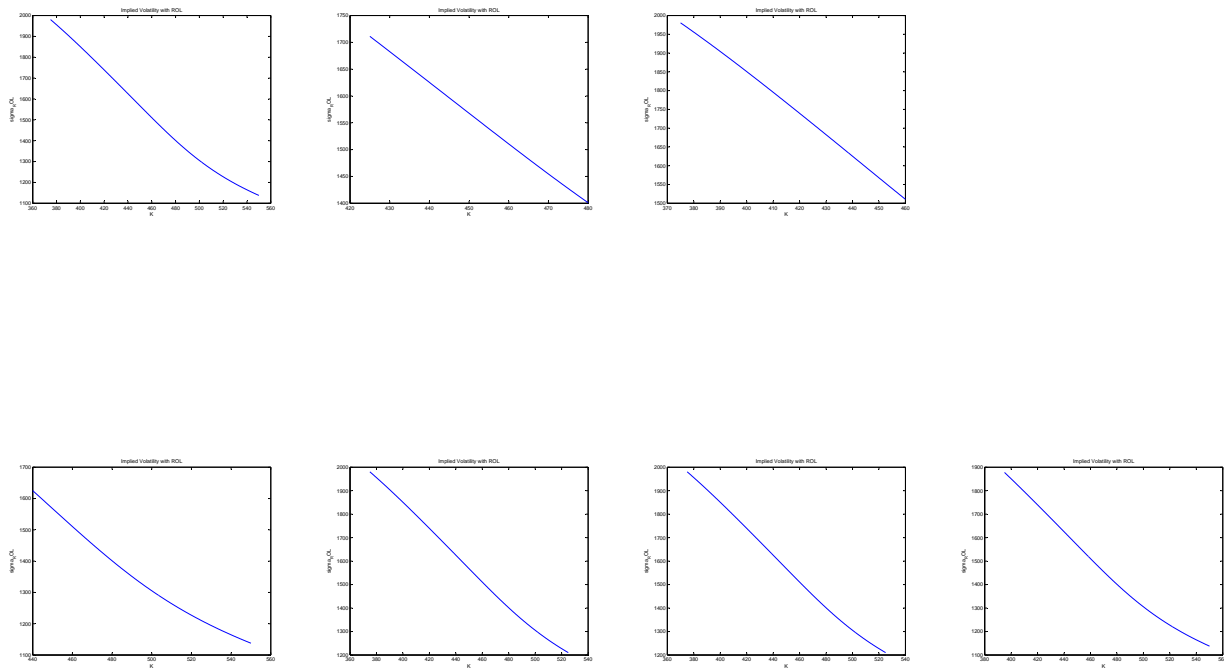
Figure 6.3: Local volatility Curves for Nonparametric Estimator with Mancino and Maliavian spot volatilty.The curve relative to all classes are reported

Figure 6.4: Local volatility Curves for Nonparametric Estimator with Andreou and Ghysels spot volatilty.The curve relative to all classes are reported
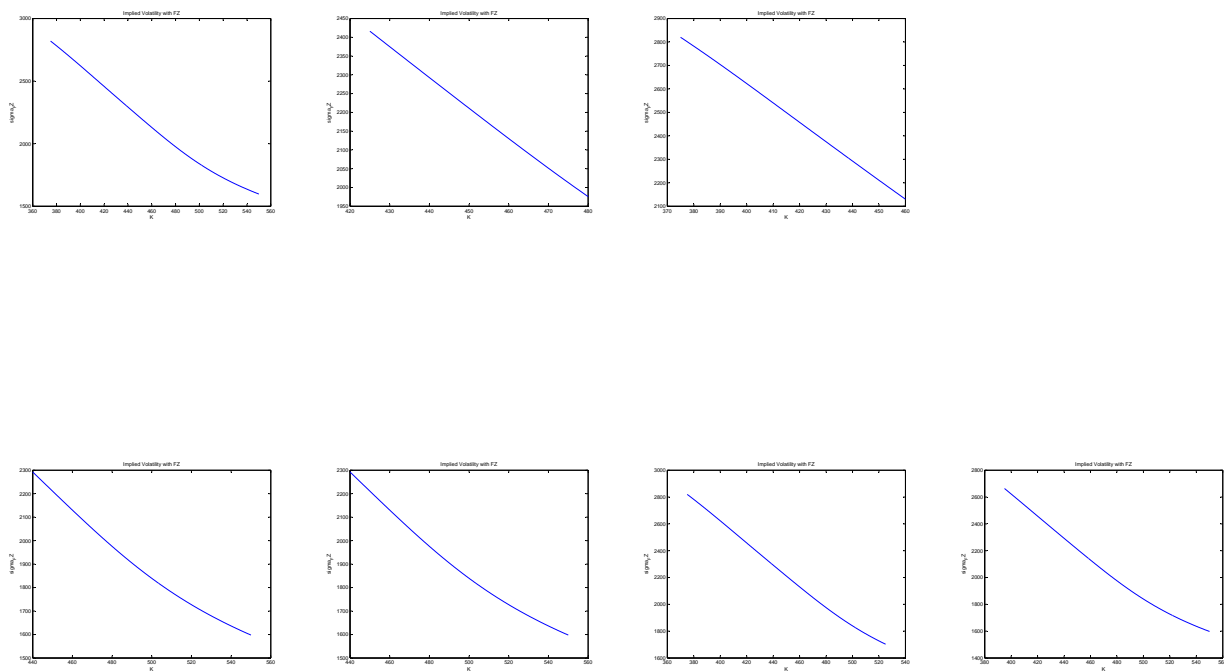
Figure 6.5: Local volatility Curves for Nonparametric Estimator with Florens-Zmirou spot volatilty.The curve relative to all classes are reported
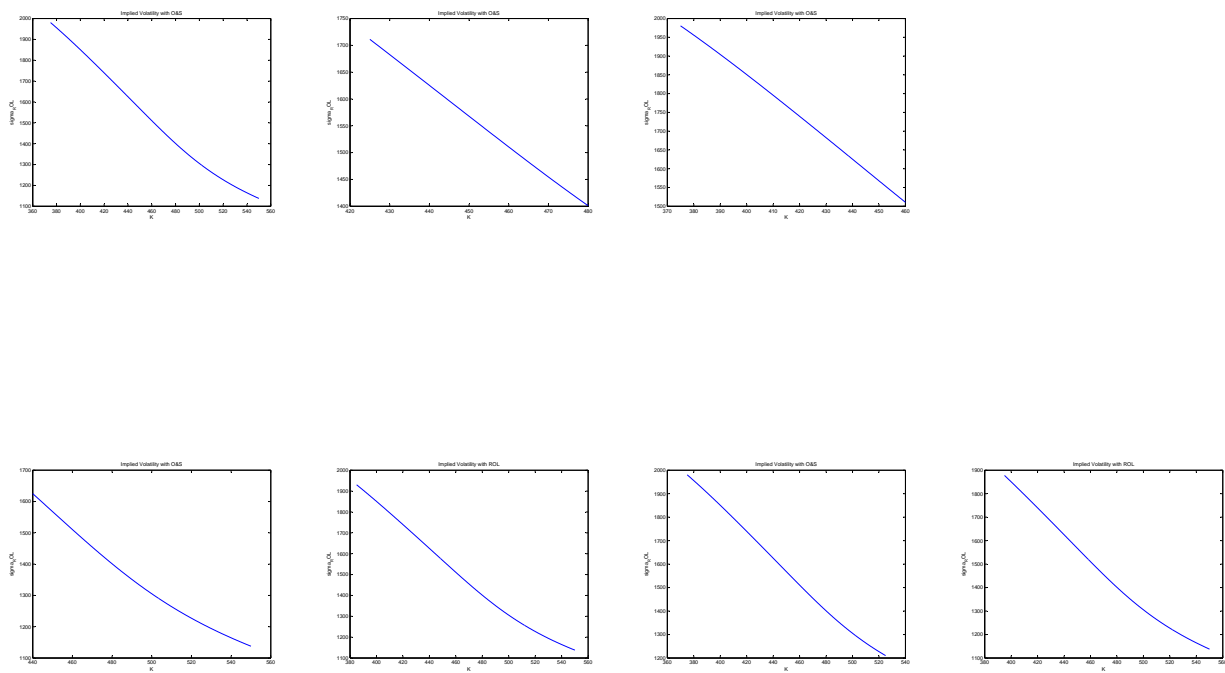
Figure 6.6: Local volatility Curves for Nonparametric Estimator with Ogawa and Sanfelici spot volatilty.The curve relative to all classes are reported

References

[1] Y. Ait-Sahalia and A.W Lo. A nonparametric estimation of state-price densities implicit in financial asset prices. *Econometrica*, 53(2):499–547, 1998.

[2] Yacine Ait-Sahalia. A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. *Econometrica*, 64(3):527–560, 1996.

[3] T.G. Andersen, T. Bollerslev, and F.X. Diebold. Parametric and nonparametric volatility measurement. Technical report, National Bureau of Economic Research, 2002.

[4] E. Andreou and E. Ghysels. Rolling-sample volatility estimators. *Journal of Business and Economic Statistics*, 20(3):363–376, 2002.

[5] O.E. Barndorff-Nielsen. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280, 2002.

[6] E. Barucci and R. Renò. On measuring volatility and the garch forecasting performance. *Journal of International Financial Markets, Institutions and Money*, 12:180–200, 2002a.

[7] E. Barucci and R. Reno. On measuring volatility of diffusion processes with high frequency data. *Economics Letters*, 74(3):371–378, 2002.

[8] T. Björk. *Arbitrage theory in continuous time.* Oxford University Press, USA, 2009.

[9] F Black and M. Scholes. The pricing of options and corporate liabilities. *J.Polit. Econ.*, 81:637–659, 1973.

[10] J.M Volle S. Bonnans, B. Cognet. Estimation de la volatilité locale d'actifs financier par une methode numerique. 2002.

[11] D.T. Breeden and R.H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651, 1978.

[12] K.C. Chan, G.A. Karolyi, F.A. Longstaff, and A.B. Sanders. An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, pages 1209–1227, 1992.

[13] D.A. Chapman, J.B. Long Jr, and N.D. Pearson. Using proxies for the short rate: when are three months like an instant? *Review of Financial Studies*, 12(4):763–806, 1999.

[14] F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4):291–323, 1998.

[15] E. Derman and I. Kani. Riding on a smile. *Risk*, 7(2):32–39, 1994.

[16] B. Dumas, J. Fleming, and R.E. Whaley. Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6):2059–2106, 1998.

[17] B. Dupire. Pricing with smile. 2004.

[18] H. Ebens. Realized stock volatility. *Department of Economics, Johns Hopkins University*, 1999.

[19] J. Fan. A selective overview of nonparametric methods in finance. *Statistical Science*, 20(4):317–337, 2005.

[20] J. Fan, Y. Fan, and J. Jiang. Dynamic integration of time-and state-domain methods for volatility estimation. *Journal of the American Statistical Association*, 102(478):618–631, 2007.

[21] J. Fan and Y. Wang. Spot volatility estimation for high-frequency data. *Statistics and its Interface*, 1:279–288, 2008.

[22] D.P. Foster and D.B. Nelson. Continuous record asymptotics for rolling sample variance estimators. Technical report, National Bureau of Economic Research, 1994.

[23] J. Gatheral. *The volatility surface;A Practitioner's guide*. John Wiley & Sons, Inc, 2006.

[24] C.A.E. Goodhart and M. O'Hara. High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance*, 4(2-3):73–114, 1997.

[25] D. Lesniewski A. S. Hagan, P.S.Kumar and D. E. Woodward. Managing smile risk. *Wilmott Magazine*, 20:84–108, 2002.

[26] P.R. Hansen, J. Large, and A. Lunde. Moving average-based estimators of integrated variance. *Econometric Reviews*, 27(1-3):79–111, 2008.

[27] P.R. Hansen and A. Lunde. A forecast comparison of volatility models: does anything beat the garch (1,1)?. *Journal of applied econometrics*, 20:873–889, 2005.

[28] P.R. Hansen and A. Lunde. Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, 24(2):127–161, 2006.

[29] L. Harris. Stock price clustering and discreteness. *Review of Financial Studies*, 4(3):389–415, 1991.

[30] J.D. Hart. *Nonparametric smoothing and Lack of fit Test*. Springer Verlag, 1997.

[31] S.L. Heston. A closed-form solution for options with stochastic volatility with application to bond and currency options. *Review of Financial Studies,*, 6:327–343, 1993.

[32] A.L Hull, J.C White. The pricing of options on assets with stochastic volatility. *The Journal of finance*, 42:281–300, 1987.

[33] J. Jacod and P. Protter. Asymptotic error distributions for the euler method for stochastic differential equations. *The Annals of Probability*, 26(1):267–307, 1998.

[34] G Jiang. Nonparametric modeling of u.s. interest rate term structure dynamics and implications on the prices of derivative securities. *Journal of Financial and Quantitative Analysis*, 33(4):465–497, 1998.

[35] G. Jiang and J. Knight. Nonparametric pricing of interest rate derivative securities. *Econometric Theory*, 13(5):615–645, 1997.

[36] D. Kristensen. Nonparametric filtering of the realized spot volatility: a kernel-based approach. *Econometric Theory*, 26(01):60–93, 2010.

[37] A. Madhavan. Market microstructure: A survey. *Journal of Financial Markets*, 3(3):205–258, 2000.

[38] J.M. Maheu and T.H. McCurdy. Nonlinear features of realized fx volatility. *Review of Economics and Statistics*, 84(4):668–681, 2002.

[39] P. Malliavin and M.E. Mancino. Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 6(1):49–61, 2002.

[40] C. Mancini and R. Renò. Threshold estimation of markov models with jumps and interest rate modeling. *Journal of Econometrics*, 160(1):77–92, 2011.

[41] M. McAleer and M.C. Medeiros. Realized volatility: A review. *Econometric Reviews*, 27(3):10–45, 2008.

[42] N. Meddahi. A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, 17(5):479–508, 2002.

[43] R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.

[44] R.C. Merton. On estimating the expected return on the market* 1:: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361, 1980.

[45] P.A. Mykland and L. Zhang. Inference for volatility-type objects and implications for hedging. Technical report, Tech. rep., Carnegie-Mellon University, 2001.

[46] S. Ogawa and S. Sanfelici. An improved two-step regularization scheme for spot volatility estimation. *Economic Notes*, pages 105–132, 2011.

[47] P.C.B. Phillips, J.Y. Park, and Yale University. Cowles Foundation for Research in Economics. Nonstationary density estimation and kernel autoregression. *Cowles Foundation Discussion Paper*, 1998.

[48] P.E. Protter. *Stochastic integration and differential equations*, volume 21. Springer Verlag, 2004.

[49] R Renò. Nonparametric estimation of diffusion coefficient of stochastic volatility model. *Econometric theory*, 24(5):1174–1206, 2008.

[50] R. Renò, A. Roma, and S. Schaefer. A comparison of alternative nonparametric estimators of the short rate diffusion coefficient. *Economic Notes*, 35(3):227–252, 2008.

[51] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 4. Springer Verlag, 1998.

[52] R. Riccardo. *Volatility And Correlation*. 1999.

[53] M. Rubinstein. Implied binomial trees. *Journal of finance*, pages 771–818, 1994.

[54] J.R. Russell and F.M. Bandi. Microstructure noise, realized volatility, and optimal sampling. In *Econometric Society 2004 Latin American Meetings*. Econometric Society, 2004.

[55] S. Sanfelici. Calibration of a nonlinear feedback option pricing model. *Quantitative Finance*, 7(1):95–110, 2007.

[56] D.W. Scott. *Multivariate density estimation*, volume 139. Wiley Online Library, 1992.

[57] B.W. Silverman. *Density estimation for statistics and data analysis*, volume 26. Chapman & Hall/CRC, 1986.

[58] R. Stanton. A nonparametric model of term structure dynamics and the market price of interest rate risk. *The Journal of Finance*, 52(5):1973–2002, 1997.

[59] D. Tavella and C. Randall. *Pricing financial instruments: The finite difference method.* John Wiley & Sons, 2000.

[60] P. Wilmott, S. Howison, and J. Dewynne. *The mathematics of financial derivatives: a student introduction.* Cambridge Univ Pr, 1995.

[61] L. Zhang, P.A. Mykland, and Y. Ait-Sahalia. A tale of two time scales. *Journal of the American Statistical Association*, 100(472):1394–1411, 2005.

[62] B. Zhou. High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics*, 14(1):45–52, 1996.

[63] Y. ZU and P. BOSWIJK. Estimating realized spot volatility with noisy high frequency data. Technical report, Working Paper, 2009.