



DEPARTMENT OF ECONOMICS,  
MANAGEMENT AND STATISTICS  
UNIVERSITY OF MILAN – BICOCCA

DEMS WORKING PAPER SERIES

**Nonparametric Estimates of Option Prices  
Using Superhedging**

Gianluca Cassese

No. 293 – February 2015

Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa  
Università degli Studi di Milano - Bicocca  
<http://dems.unimib.it/>

# NONPARAMETRIC ESTIMATES OF OPTION PRICES USING SUPERHEDGING

GIANLUCA CASSESE

ABSTRACT. We propose a new nonparametric technique to estimate the CALL function based on the superhedging principle. Our approach does not require absence of arbitrage and easily accommodates bid/ask spreads and other market imperfections. We prove some optimal statistical properties of our estimates. As an application we first test the methodology on a simulated sample of option prices and then on the S&P 500 index options.

February, 16, 2015

JEL-codes: G12, G13

*Keywords and phrases:* Bid/Ask spreads, Implied risk-neutral measure, Non parametric regression.

## 1. INTRODUCTION

A classical exercise in the econometric analysis of financial markets is to estimate option prices and the risk neutral probability (or its density) which is implicit in them, as suggested by the famous works of Breeden and Litzenberger [5] and of Banz and Miller [3]. Of special importance in this exercise is the use of nonparametric techniques which have become popular in the last decades. Regardless of the econometric approach taken, a first step of crucial importance is purging the data from observations violating some no-arbitrage condition and thus conflicting with the conclusion that prices are the expected value of the asset discounted payoff computed with respect to the risk neutral probability, as assumed by Breeden and Litzenberger.

Recalcitrant observations may however originate from different sources other than just pure mispricing. There are first some microstructural issues, such as the bid/ask spread or other transaction costs, which are not considered in the risk neutral approach. The common practice is to assume that these components are uncorrelated with the fundamental value of the assets and may therefore be disposed of with simple transformations, such as computing mid prices. Nevertheless, a more accurate investigation reveals that market

---

In writing the empirical applications of this paper, I benefited from conversations with Patrick Gagliardini. I am also grateful to Alexandru Popescu for helping me with the MDR-CBOE Dataset. All errors are my responsibility.

makers often adjust the spread in response to pressure originating either from demand or supply, a fact that contributes to make the spread asymmetric and to induce a possibly significant correlation with the fundamental price. A second factor influencing market data is the existence of restrictions to trading, these too often overlooked in asset pricing models which typically assume the space of marketed claims to be linear. The extreme versions of these restrictions take the form of short selling prohibitions but even with less drastic market rules, the possibility of shorting assets is clearly delicate and subject to constraints such as the provision of appropriate margins the effect of which receives in general little attention – if any – in theoretical and applied work on options. In fact, violations of the lower bound for CALL options, a mispricing that may be exploited by shorting the underlying, are often more frequent than others associated with strategies involving positions which are easier to take. Some additional noise arises with strategies which prescribe to invest simultaneously on different markets, a characteristic which results not only in relevant fixed costs but also in the lack of trade synchronism thus inducing additional risks in the execution of arbitrage strategies. A further issue arises eventually in connection with the riskless asset appearing in virtually all financial models and definitely in all empirical exercises. The assumption of a riskless asset and its explicit identification in applied work appears more and more counterfactual as the interbank market turmoil in the years 2007-08 and the sovereign debt crisis following it have clearly demonstrated. Not only during periods of crisis it cannot be assumed that bonds are riskless but it is also hard to assert that the implicit risks are uncorrelated with the equity market.

Of course, the relevance of the preceding remarks would be much less disturbing if one could restrict attention to just the options market and consider only trading strategies which involve long positions. Unfortunately in this restricted perspective the fundamental premise of the Breeden Litzenberger exercise, i.e. that option prices are set on the basis of a risk neutral probability, cannot be assumed. Nevertheless, even if real option prices may fail to possess some fundamental properties, the superhedging price, conveniently computed will possibly satisfy them under mild conditions which involve only very simple and realistic trading strategies. The first step of our econometric approach consists of replacing original option prices with superhedging prices.

As a matter of fact this idea is not new. Aït-Sahalia and Duarte [1] have inaugurated a two stage approach to nonparametric estimation of option prices under shape restrictions which, in the first stage, prescribes to project market prices following a methodology adapted from Dykstra [8]. This procedure enforces the shape restrictions required. In the second stage, the prices so transformed are smoothed out using a local polynomial technique, a methodology which generalizes suitably the kernel by replacing constant functions with polynomials of

arbitrary but preassigned degree. It turns out, as will be shown in the body of the paper, that the projection technique suggested by Dykstra produces just the options superhedging price,  $q^0(k)$  – with  $k$  the strike price. The need for additional smoothing arises because the CALL function obtained,  $k \rightarrow q^0(k)$  is piecewise linear and therefore not informative enough for many a purpose, including the project to extract a probability density.

In a recent paper, [6], we have obtained a result which is particularly useful for the present purposes and that lays the ground for our nonparametric procedure. It states, roughly put, that it is possible to compute explicitly the superhedging price  $q^G(t)$  for a large class  $G$  of derivatives written on the same underlying and with payoff depending in a convex and decreasing way on a positive parameter,  $t$ . Not only, but for each such family  $G$  of derivatives it is possible to extract from the prices  $q^G(t)$  a probability measure,  $\nu^G$ , implicit in them in much the same way as suggested by Breeden and Litzenberger. A special case is of course the family  $G^0$  corresponding to plain options but in general the choice of the family  $G$  is open. Our suggestion is to consider new derivatives with a payoff which, while being perfectly smooth, approximates the option payoff uniformly. Moreover, the degree of approximation should be made sample dependent so to obtain good asymptotic properties. As long as these properties are guaranteed, any candidate  $G$  is suitable, more or less as the functional form of kernels is relatively unimportant in comparison with the choice of the bandwidth. We have found it easy to work with splines but it should be made clear from the outset that our use of splines has nothing to do with the econometric approaches based on this class of functions, such as those reviewed by Eubank [9] and successfully applied to options by Fengler and Hin [11]. Our problem in fact is not that of smoothing the option *prices* (or their implied volatility surface) but rather the option *payoff*. Splines simply turn out to be a conveniently tractable tool from a computational point of view; in addition, the smoothness/goodness of fit trade off may be tuned conveniently via a parameter acting as the bandwidth in kernel estimation. Other smoothing techniques may be employed, e.g. those based on some known probability density function. As a term of comparison, we shall briefly discuss the result obtained by smoothing options via the normal density rather than splines.

The econometric analysis of option prices has grown over the years to become almost a field of its own and has been masterly reviewed by Garcia, Ghysels and Renault in [13]. The interested reader may find in their work an exhaustive list of references that we shall not try to improve upon here, for reasons of brevity. The nonparametric approach has itself produced quite a number of important contributions that are worth discussing briefly with no pretension of completeness<sup>1</sup>. A forerunner of this stream of studies is the paper

---

<sup>1</sup>A nice review of nonparametric methods and their relevance for economics is in Yatchew [25].

by Jackwerth and Rubinstein [17] in which the parameters of a binomial tree are set so as to minimize the distance between binomial and actual option prices given a penalty for deviations from an initial *a priori* distribution. The assumption of absence of arbitrage opportunities is absolutely crucial here and is embodied in the tree parameters. Estimating the option price produces simultaneously an estimate of the risk neutral measure. Aït-Sahalia and Lo [2] estimate the CALL function using the Nadaraya-Watson kernel and recover the risk-neutral density by computing the second derivative. This methodology has optimal asymptotic properties but relatively poor performance in small samples, a fact common to many nonparametric methods and of special concern for options market. The curse of dimensionality problem, see [25, p. 675], requires to limit as much as possible the number of state variables. For example, the risk neutral density estimated in [2] via kernel smoothing is surely convergent to the true density but may fail even to be non negative in small samples [2, footnote 11, p. 508]. To circumvent this problem the authors propose a semi nonparametric approach in which the price is computed according to the Black and Scholes formula in which the volatility function is estimated nonparametrically. This choice has the clear advantage of guaranteeing the correct shape of the CALL function. Shape restrictions are easily accommodated in parametric modeling but are much more troublesome in the non parametric approach. Papers implementing the nonparametric methodology with shape restrictions are less numerous and more recent. Further to Aït-Sahalia and Duarte [1], another example of nonparametric techniques incorporating shape restrictions is the paper by Yatchew and Härdle [26] who follow a least squares approach in Sobolev spaces. There exist eventually several papers who adopt spline techniques to estimate option prices and one should mention Fengler [10], Fengler and Hin [11] and Yin, Wang and Qi [27].

We think that if our approach has any merit compared to the above references, then this lies in its clear financial interpretation. In fact we obtain the smoothness of the CALL function by pricing an appropriate derivative rather than performing some local averaging or implementing some other statistical technique<sup>2</sup>. Moreover, our method allows to take selection effects into full account, a fact not always clearly considered. Eventually the estimates produced are extremely tractable computationally speaking and have desirable convergence properties.

The paper is structured as follows. In section 2 we introduce the fundamental results obtained in [6] and needed in our exercise, together with the necessary notation. In section 3 we illustrate all details of our spline based technique and prove some of its properties. We

---

<sup>2</sup>A partial exception is the XMM methodology proposed by Gagliardini et al. [12], in which the GMM is applied with the additional constraint of reproducing a subset of given prices which are considered *a priori* to be correct.

also comment on the possibility of using given density functions. In section 4 we perform some numerical experiments on option prices generated by an a priori model. In particular we construct non parametric interval estimates in the presence of a realistic structure of noise. Eventually in section 5 we apply the proposed methodology to a sample of S&P option prices. All proofs are in the Appendix.

Let us mention in closing that, although as dictated by the literature on this topic we start assuming the existence of a given probability space,  $(\Omega, \mathcal{F}, P)$ , the results developed in [6] do not require this classical premise.

## 2. OPTION PRICING

Most of this section is adapted from [6]. Let the random variable  $X$  represent hereafter the payoff of a given underlying and  $K(X)$  the set of strike prices (including  $k = 0$ ) of all CALL options written on it. We assume  $P(X > 0) = 1$  and  $P(X < k) > 0$  for all  $k \in K(X) \setminus \{0\}$ <sup>3</sup>. The portfolio consisting of one option with strike price  $k \in K(X)$  and its price will be indicated by the symbols  $\theta(k)$  and  $q(k)$  respectively. Moreover, its payoff will be denoted with the symbol  $g_k^0(X) = (X - k)^+$ . As we only consider long positions,  $q(k)$  will actually be the *ask* price and we shall assume that it is positive. The set  $\Theta$  of admissible trading strategies is a convex set of portfolios formed by taking long positions in the set of traded options. The price and the payoff of  $\theta = \sum_{n=1}^N \alpha_n \theta(k_n) \in \Theta$  – so that  $\alpha_1, \dots, \alpha_N \geq 0$  and  $k_1, \dots, k_N \in K(X)$  – are denoted respectively by

$$(1) \quad q(\theta) = \sum_{n=1}^N \alpha_n q(k_n) \quad \text{and} \quad g_\theta^0(X) = \sum_{n=1}^N \alpha_n g_{k_n}^0(X)$$

The price function  $q$  does not include possible fixed trading costs. The apparent linear structure implicit in (1), notwithstanding the existence of bid/ask spreads (which many authors associate with subadditivity of prices, see [6]), is due to the restriction included in  $\Theta$  that only long positions may be assumed and the traditional anonymity of option trading. OTC trading of options follows, as is well known, different pricing schemes, often ad hoc, and our theory does not apply to these transactions<sup>4</sup>.

We define the superhedging price of an arbitrary random quantity  $h : \Omega \rightarrow \mathbb{R}$  as<sup>5</sup>

$$(2) \quad \pi(h) = \inf \left\{ \lambda q(\theta) : P(\lambda g_\theta^0(X)/(X \wedge 1) \geq h) = 1, \lambda > 0, \theta \in \Theta \right\}$$

<sup>3</sup>This corresponds to assuming that no quoted option is known to expire in the money with probability 1.

<sup>4</sup>More details on the assumptions behind the present construction are found in [6].

<sup>5</sup>In [6] the functional  $\pi$  is defined in much greater generality and no probability measure  $P$  is taken as exogenously given.

Write  $q^0(k) = \pi(g_k^0(X)/(X \wedge 1))$ . We say that an option with strike price  $k$  is priced efficiently if  $q(k) = q^0(k)$  and denote by  $K^0(X) = \{j_0, \dots, j_I\}$  the set of strike prices for which we have an efficient price. It is easily shown that necessarily  $j_0 = 0$ . Remark that the notion of efficiency adopted here is particularly poor as it only involves investments with long positions in CALL options and the underlying. Neither PUT options, Futures nor bonds are contemplated. This is desirable since the larger the set of derivatives involved the more likely is it that efficiency may fail, as in the case of the PUT/CALL parity.

We make the following

**Assumption 1.** *Let  $\bar{X} = \text{ess sup } X$  and assume that*

$$(3) \quad \bar{X} > j_I \quad \text{and} \quad \text{ess sup}(X \wedge z) = \bar{X} \wedge z \quad \text{for each } z \in \mathbb{R}$$

Denote by  $\Gamma$  the set of convex functions  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  with  $g(0) = 0$  and  $\lim_{n \rightarrow \infty} g(n)/n < \infty$ . Clearly, the function  $g_t^0$  introduced above and corresponding to the option payoff is an element of  $\Gamma$  for all  $t \geq 0$ : write  $G^0 = \{g_t^0 : t \geq 0\}$ . Of course, it is possible to write other derivatives on  $X$  possessing some of the properties of options. In the Appendix we prove the following:

**Theorem 1.** *For each  $f \in \Gamma$  there exists  $\theta(f) \in \Theta$  such that  $\pi(f(X)/(X \wedge 1)) = q(\theta(f))$  and*

$$(4) \quad q(\theta(f)) = [q(j_0), \dots, q(j_I)] \begin{bmatrix} (j_1 - j_0) & 0 & \dots & 0 \\ (j_2 - j_0) & (j_2 - j_1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ d_0(\bar{X}) & d_1(\bar{X}) & \dots & d_I(\bar{X}) \end{bmatrix}^{-1} \begin{bmatrix} f(j_1) \\ \vdots \\ f(j_I) \\ \hat{f}(\bar{X}) \end{bmatrix}$$

where  $d_i(\bar{X}) = (\bar{X} - j_i)$  and  $\hat{f}(\bar{X}) = f(\bar{X})$  if  $\bar{X} < \infty$  or else  $d_i(\bar{X}) = 1$  and  $\hat{f}(\bar{X}) = \lim_n f(n)/n$ . In particular, we can write

$$(5) \quad q(\theta(f)) = \sum_{i=0}^I w_i q(j_i) = \sum_{i=1}^I b_i f(j_i) + b_{I+1} \hat{f}(\bar{X})$$

Let  $D_i(h) = \frac{h(j_{i+1}) - h(j_i)}{j_{i+1} - j_i}$  when  $i = 0, \dots, I-1$  and  $D_{-1}(h) = D_I(h) = 0$ . Then, if  $\bar{X} = \infty$  we have

$$(6a) \quad w_I = \hat{f}(\bar{X}) - D_{I-1}(f), \quad b_{I+1} = q(j_I)$$

$$(6b) \quad w_i = D_i(f) - D_{i-1}(f) \quad \text{and} \quad b_{i+1} = D_{i+1}(q) - D_i(q) \quad i = 0, \dots, I-1$$

The next result is just a restatement of [6, Theorem 7].

**Theorem 2.** Let  $G = \{g_t : t \in \mathbb{R}_+\}$  be a collection of functions  $g_t \in \Gamma$  such that

$$(7) \quad \alpha g_s + (1 - \alpha)g_u \geq g_t \quad \text{whenever} \quad \alpha \in [0, 1], \quad s, t, u \in \mathbb{R}_+ \quad \text{and} \quad \alpha s + (1 - \alpha)u \leq t$$

Write  $q^G(t) = \pi(g_t(X)/(X \wedge 1))$ . There exists  $\beta^G \geq 0$  and  $\nu^G \in ca(\mathcal{B}(\mathbb{R}_+))$  such that

$$(8) \quad q^G(t) = \beta^G + \int_t^\infty \nu^G(x > z) dz \quad \text{for all } t \geq 0$$

Observe that, by standard rules,

$$(9) \quad \int_t^\infty \nu^G(x > z) dz = \int (x - t)^+ d\nu^G(x)$$

Thus (8), represents the price of the  $G$  derivatives as the sum of a bubble part and of their fundamental value. This view is substantiated by the observation that necessarily

$$(10) \quad \beta^G = \lim_{t \rightarrow \infty} q^G(t)$$

so that, upon choosing  $G = G^0$ , the term  $\beta^0$  represents the option price as the strike approaches infinity and contributes to explaining the overpricing of deeply out of the money CALL's often documented empirically in some form of the smile effect.

As remarked in [6], Breeden and Litzenberger formula applies, giving:

$$(11) \quad \nu^G(x > t) = - \left. \frac{dq^G(k)}{dk} \right|_{k=t} \quad \text{for all } t \geq 0$$

while in the model of Black and Scholes one has, as is well known,

$$(12) \quad \nu^{BS}(x > k) = e^{-rT} \Phi(d_2) \quad \text{with} \quad d_2 = \frac{\ln(S_0/k) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

so that  $\|\nu^{BS}\| = \exp(-rT)$ . Observe that in our setting, we do not have an *a priori* restriction on the norm of  $\nu^G$  because our market does not include any riskless bond.

### 3. ESTIMATING THE CALL FUNCTION

For what concerns the choice of  $G$ , it is natural to start considering the family  $G^0$ . By construction  $q^0(k)$  represents the efficient price of the corresponding option and so it coincides with  $q(k)$  if and only if  $k \in K^0(X)$ . More generally one sees that the function  $q^0$  is the highest among the positive, convex and decreasing curves passing through the knots  $\{(j, q(j)) : j \in K^0(X)\}$ . Ait-Sahalia and Duarte [1] obtain the same values in the first step of their approach. Since we have to make a choice for definiteness let's assume henceforth

**Assumption 2.**  $\bar{X} = \infty$ .



Replace  $g$  with  $g_k^0$  in Theorem 1. Then, according to (6), the cheapest way to superhedge the corresponding option is to invest

$$w_i^0(k) = \frac{j_{i+1} - k}{j_{i+1} - j_i} \quad \text{and} \quad w_{i+1}^0(k) = 1 - w_i^0(k)$$

in the options with strike prices  $j_i$  and  $j_{i+1}$  respectively, where  $j_{i+1} \geq k > j_i$ , and nothing in all other options; or else, if  $k > j_I$ , to buy one unit of the option with strike price  $j_I$ . The corresponding superhedging price would be

$$(13) \quad q^0(k) = \begin{cases} \frac{j_{i+1}-k}{j_{i+1}-j_i}q(j_i) + \frac{k-j_i}{j_{i+1}-j_i}q(j_{i+1}) & \text{if } j_i < k \leq j_{i+1} \\ q(j_I) & \text{if } j_I < k \end{cases}$$

Observe that the CALL function  $k \rightarrow q^0(k)$  satisfies: (i)  $q^0(j_i) = q(j_i)$  for  $i = 0, \dots, I$  and (ii) it is a straight line on each interval  $[j_i, j_{i+1}]$  for  $i = 0, \dots, I - 1$  and on  $[j_I, \infty)$ . Then it coincides with the projection obtained by following Dykstra's method, see [8, sec. 4.1]<sup>6</sup>.

From (13) we deduce easily that

$$(14) \quad \nu^0(x > k) = \begin{cases} \frac{q(j_i)-q(j_{i+1})}{j_{i+1}-j_i} & \text{if } j_i < k \leq j_{i+1} \\ 0 & \text{if } j_I < k \end{cases}$$

In other words, the implied set function  $\nu^0$  coincides – not surprisingly – with the derivative of the price function over the discrete set  $K^0(X)$  upon a change of sign.

Despite being a theoretically exact formula, the informational content of (14) is indeed quite poor empirically, as a consequence of the limited number of options which are priced efficiently by the market – the size of  $K^0(X)$ . In fact  $\nu^0(x > t)$  remains constant within two adjacent strike prices in  $K^0(X)$  or, equivalently, the efficient CALL function  $q^0(k)$  is piecewise linear. Adding the fictitious efficient prices  $q^0(k)$  to the original data set, thus, does not improve our knowledge of  $\nu$ .

Another way of putting it is saying that superhedging of options is an intrinsically trivial exercise, a conclusion to which contribute two distinct factors. First, the piecewise linear nature of the payoff function is such that superhedging never requires more than two efficient options which makes the corresponding price function locally linear, as in (13). On the other hand, the relative scarcity of available efficient prices makes the length of intervals on which the function is linear wider. If the size of the set  $K^0(X)$  is a constraint in this problem and essentially depends on the market structure, a quick look at (4) reveals that smoothness of  $q^G$  depends on the smoothness of the underlying family  $G$ . Given that obtaining smooth

---

<sup>6</sup>In fact, Ait-Shalia and Duarte have to adapt slightly the method of Dykstra since, assuming the lower bound for CALL options holds, the set unto which they project is not a convex cone but just a convex set. However, this additional restriction, for the reasons outlined above, does not apply here. See [1, p. 18 and Appendix A].

estimates is the goal of any econometric exercises, our first step is then to construct a family  $G^h$  of derivatives written on  $X$ , each having a payoff – denoted by  $g_k^h(X)$  – which is (i) conveniently close to the corresponding CALL option but (ii) twice continuously differentiable and such that (iii)  $G^h$  satisfies (7). These properties guarantee that the corresponding price,  $g^h(k)$ , will be a smooth estimate of the option prices and that it will admit the representation (8) in terms of an implicit pricing measure. The class  $G^h$  of derivatives will depend on a control parameter,  $h > 0$ , which acts in much the same way as the bandwidth parameter in kernel regression. We give now a more detailed description.

**3.1. The Spline Approach.** Let  $k, h > 0$  and  $N \in \mathbb{N}$  be given. Divide the interval  $[-h, h]$  into  $N$  intervals of equal length, with endpoints  $t_i^h = -h + 2ih/N$  and  $i = 0, \dots, N$ . Consider then the following functional (with  $D^2f$  denoting the second derivative of  $f$  and  $g_k^0$  the CALL payoff function with strike  $k$  as above):

$$(15) \quad I_k(h; g; \lambda) \equiv \sum_{i=0}^N [g_k^0(k + t_i^h) - g(k + t_i^h)]^2 + \lambda \int_{k-h}^{k+h} [D^2g(x)]^2 dx$$

and the program

$$(16) \quad I_k(h; \lambda) \equiv \min_{g \in \Gamma \cap \mathcal{C}^2} I_k(h; g; \lambda)$$

It is well known, see [9, Theorem 5.2], that a solution to this problem is given by a  $\mathcal{C}^2$  cubic spline which is linear outside of  $[k - h, k + h]$ . Based on the fact that the second derivative of a cubic spline is locally linear, the infinite dimensional problem (16) conveniently reduces to a  $2N$ -dimensional one. Turlach [24] developed a methodology to compute its solution under several shape restrictions such as (i)  $g, Dg, D^2g \geq 0$ , i.e. positivity, monotonicity and convexity. To these constraints we add the following: (ii)  $g(x) = Dg(x) = 0$  for  $x \leq k - h$ , (iii)  $g(k + h) = g_k^0(k + h)$  and (iv)  $Dg(x) = 1$  for  $x \geq k + h$ . Denote

$$(17) \quad \chi(k; h) = \{g \in \mathcal{C}^2(\mathbb{R}_+) : g \text{ cubic spline meeting the constraints (i) – (iv)}\}$$

The notation adopted is consistent with the choice of treating  $N$  as a fixed parameter and of focusing exclusively on properties which depend on  $h$ . Although we experimented several values, in what follows we will set  $N = 10$ . More importantly, we make the choice of  $\lambda$  endogenous by letting  $\lambda_h = (0.1h)^3$  – so that  $\lambda_{10} = 1$ . The existence of a constrained solution to (16) and the properties of such solution are proved in the next:

**Lemma 1.** *The problem*

$$(18) \quad I_k(h) \equiv \min \{I_k(h; g; \lambda_h) : g \in \chi(k; h)\}$$

admits one and only one solution,  $g_k^h \in \chi(k; h)$ , and this satisfies: (i)  $g_{k+m}^h(x) = g_k^h(x - m)$  for all  $k, m, x \geq 0$ , (ii)  $g_k^0 \leq g_k^{h'} \leq g_k^h$  whenever  $h' \leq h$  and (iii)  $\lim_{h \rightarrow 0} \sup_x (g_k^h - g_k^0)(x) = 0$ .

Thus if we set

$$(19) \quad G^h = \{g_k^h : k \geq 0\} \subset \Gamma$$

the family  $G^h$  satisfies (7). Actually, in the proof of Lemma 1 we obtain that for fixed  $k_0, h_0 > 0$

$$(20) \quad g_k^h(x) = g_{k_0}^{h_0} \left( k_0 + (x - k) \frac{h_0}{h} \right) \frac{h}{h_0} \quad \text{for every } k, h > 0$$

so that property (iii) of Lemma 1 follows easily from

$$(21) \quad 0 \leq q^h(k) - q^0(k) \leq \pi \left( \frac{g_k^h(X) - g_k^0(X)}{X \wedge 1} \right) \leq \frac{g_k^h(k)}{(k - h) \wedge 1} q(0) \leq \frac{g_k^{h_0}(k) q(0)}{(k - h_0) \wedge 1} \frac{h}{h_0}$$

Denote by  $\theta_k^h \in \Theta$  the portfolio  $\theta(g_k^h)$  involved in super replicating  $g_k^h$  and by  $\hat{g}_k^h$  its payoff.

The content of Lemma 1 is clearly illustrated in Figure 1, Panels A and B, where the function  $g_k^h$  and its derivative are plotted for different values of  $h$ .

[Figure 1 about here.]

In Panel C the target payoff  $g_k^h$  is plotted together with the payoff  $\hat{g}_k^h$  of the portfolio that super replicates it in order to understand the role of efficiency. By construction, superhedging  $g_k^h$  or  $\hat{g}_k^h$  are entirely equivalent exercises. Thus, the degree of smoothness involved depends not only on the shape of the notional payoff  $g_k^h$  but also on that of its market counterpart,  $\hat{g}_k^h$ , and particularly on the number of contracts involved in superhedging. This remark highlights the role of the bandwidth  $h$ . A small value of this parameter makes the width of the smoothing interval  $[k - h, k + h]$  narrow and diminishes, as a consequence, the distance between  $g_k^0$  and  $g_k^h$ , as shown in Lemma 1. At the same time the number of efficiently priced options with strike price included in such interval – i.e. the size of the set  $[k - h, k + h] \cap K^0(X)$  – is reduced, making the super hedging payoff  $\hat{g}_k^h$  less smooth. In fact, in Panel C – with 6 equally spaced efficient strike prices ranging from 95 to 105 – the payoff  $g_k^h$  is represented by a continuous, smooth line while  $\hat{g}_k^h$  is the piecewise linear curve of the same color dominating it. In the case  $h = 2$ , the super hedging portfolio only contains 2 options while for the case  $h = 10$  it contains all 6. It is clear from the picture that  $\hat{g}_k^h$  becomes less smooth as it approaches  $g_k^0$ , suggesting the need to balance these two countervailing effects.

To understand this point better, let  $M$  be the mesh of  $K(X) \setminus \{0\}$ , i.e.

$$(22) \quad M = \sup \{|k - k'| : k, k' \in K(X) \setminus \{0\}, k \neq k'\}$$

and  $M^0$  the mesh of  $K^0(X) \setminus \{0\}$ . Let also  $w_k^h(j)$  denote the quantity invested in option  $\theta(j)$ , with  $j \in K^0(X)$ , in order to super hedge  $g_k^h$ , obtained from (6). From the fact that  $g_k^h$  and  $g_k^0$  coincide outside of the interval  $[k-h, k+h]$  by (17) and from (6) we deduce the implications

$$(23a) \quad w_k^h(j_i) > 0 \quad \Rightarrow \quad j_{i+1} + h > k > j_{i-1} - h \quad \Rightarrow \quad |k - j_i| < h + M^0 \quad i = 1, \dots, I$$

$$(23b) \quad w_k^h(0) > 0 \quad \Rightarrow \quad j_1 > k - h$$

This suggests to set  $h = \delta M$  so to make the properties of the estimator sample dependent. In the sequel we will replace the superscript  $h$  with  $\delta$ . We observe that, as a consequence of (23), the number of efficient options employed in order to superhedge  $g_k^\delta$  is at most  $3 + 2\delta M/M^0$  so that it is desirable to fix  $\delta > M^0/M \geq 1$  in order to have at least 5 options available. In the following sections we will experiment values of  $\delta$  ranging from 1 to 10.

**3.2. Properties of the Estimator.** A different and more classical question is whether the proposed estimator converges to the *true* price function in the presence of disturbances. The classical formulation of this problem is

$$(24) \quad q(k) = F(k) + \varepsilon_k \quad k \in K(X)$$

where  $F$  is the model, and is thus assumed to be a positive, decreasing, convex,  $\mathcal{C}^2$  function, whereas the errors  $\varepsilon_i$  are identically and independently distributed with zero expectation.

Theorem 1 gives an explicit functional form to our estimator:

$$(25) \quad q^\delta(k) = \sum_{j \in K^0(X)} w_k^\delta(j) q(j) \quad t > 0$$

Observe that, by (37) and Assumption 2,

$$(26) \quad w_k^\delta(j) \geq 0 \quad \text{for all } j \in K^0(X) \quad \text{and} \quad \sum_{j \in K^0(X)} w_k^\delta(j) = 1$$

At first sight, then, our estimator appears as an exemplification of the local averaging approach which includes, as additional special cases, kernels and regressograms. For this class of estimators a well established theory demonstrates (see e.g. [25, p. 677]) that the *MSE* converges to 0 as the sample size diverges. However, in local average estimators weights are assumed to be uncorrelated with disturbances and this is of crucial importance in proving convergence. In our setting, instead, this property can in no way be assumed to hold. In fact the set  $K^0(X)$  of efficient option prices, hitherto treated as given, will in general strongly depend on prices and therefore on disturbances. Not only but the inclusion of a given price  $k$  in  $K^0(X)$  will depend not only on the corresponding noise but also on all other disturbances since a high value for  $\varepsilon_{k'}$  will make it more likely for  $q(k)$  to be efficient. The problem

discussed provides a clear exemplification of the selection bias studied by Heckman [15] and it is to some extent surprising that its pervasive role in empirical option pricing has not been fully recognized. From the classical work of Heckman we learn, in fact, that selection effects may result in biased and inconsistent estimates unless correcting for a term which proxies the expected value of disturbances conditional on selection. The work of Heckman has been extended to the nonparametric setting in a recent paper by Das et al. [7]. Our aim is that of giving sufficient conditions for the  $MSE$  to converge to 0 when taking the selection effect fully into account.

The first step is to treat  $K^0(X)$  as a random variable. Observe to this end that the inclusion  $k \in K^0(X)$  holds if and only if the  $\mathcal{F}$  measurable event  $A_k = \{q(k) \leq q^0(k)\}$  occurs. In order to take randomness into full account we replace the symbol  $K^0(X)$  with  $Z(\omega) \subset K(X)$ . Then, if  $z \subset K(X)$  we have that  $\{Z = z\} = \bigcap_{k \in z} A_k \in \mathcal{F}$ . We observe that although the value of  $w_k^\delta(j)$  depends crucially on  $Z$ , it is in fact deterministic once conditional on the event  $\{Z = z\}$ . It is thus natural to follow a two step procedure (similarly to Heckman and to Das et al.) by first conditioning all variables on the selection mechanism and then focusing on the unconditional properties. Denote by  $P_z(\cdot)$  the conditional expectation given the event  $\{Z = z\}$ , for  $z \subset K(X)$  and  $P_Z(\cdot) = \sum_{z \subset K(X)} P_z(\cdot) \mathbf{1}_{\{Z=z\}}$ . When needed we explicit  $z$  as  $\{0 = j_0^z < j_1^z < \dots < j_{j^z}^z\}$  and write

$$\tau_i = \sum_{z \subset K(X)} j_i^z \mathbf{1}_{\{Z=z\}}$$

Denote by  $M_z$  and  $M_Z$  the mesh of the sets  $z \setminus \{0\}$  and  $Z \setminus \{0\}$  respectively. The variable  $P_Z(\varepsilon)$  will be our correction term.

**Theorem 3.** *Let the function  $F : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  in (24) be decreasing and convex and assume that  $F \in \mathcal{C}^3$  with  $\sup_{x \in \mathbb{R}_+} |D^3 F(x)| < \alpha$ . Assume moreover that*

(27a)

$$E(\varepsilon_\tau \varepsilon_\sigma) = 0, \quad P_Z(\varepsilon_\tau) = P_z(\varepsilon_\sigma) \quad \text{and} \quad P_Z(\varepsilon_\sigma \varepsilon_\tau) = P_z(\varepsilon_{\sigma'} \varepsilon_{\tau'}) \quad \text{for all } \sigma, \tau, \sigma', \tau' \in Z$$

(27b)

$$E\left(\sup_{\tau \in Z} P_Z(|\varepsilon_\tau|)\right) < \infty \quad \text{and} \quad E\left(\sup_{\sigma, \tau \in Z} P_Z(|\varepsilon_\sigma \varepsilon_\tau|)\right) < \infty$$

(27c)

$$\lim_{M \rightarrow 0} E(M_Z^6) = 0 \quad \text{and} \quad \lim_{M \rightarrow 0} P(\tau_1 > 0) = 0$$

Then, for any compact interval  $I \subset \mathbb{R}_+ \setminus \{0\}$

(28)

$$\lim_{M \rightarrow 0} \sup_{t \in I} E\left(\left(q^\delta(t) - F(t)\right)^2\right) = 0$$

Condition (27a) is conceptually akin to the assumption made by Das et al. that the conditional expectation of the disturbance given selection depends only on the propensity

score, see [7, Assumptions 2.1, 2.2 and 2.3]. Condition (27b) is easily satisfied if, e.g., disturbances are uniformly distributed. The the most delicate property is definitely (27c) as it requires a strong form of convergence to 0 of the mesh of  $K^0(X)$ . The proof of Theorem 3 strongly relies on some properties of the smoothing spline  $g_k^\delta$  proved in Lemma 1, a fact emphasizing the role of splines.

**3.3. The density approach.** As mentioned in the introduction, splines are not the only possible choice for smoothing option payoff. We present in this paragraph an alternative based on some given probability density function,  $\phi$ , satisfying the properties

$$(29) \quad \int u\phi(u)du = 0 \quad \text{and} \quad \int |u|\phi(u)du < \infty$$

Denote by  $\Phi$  the distribution function associated with  $\phi$  and consider the following function:

$$(30) \quad g_k^h(x; \phi) = \Phi\left(\frac{x-k}{h}\right)(x-k) - h \int_{-\infty}^{(x-k)/h} u\phi(u)du$$

**Lemma 2.** *Let  $\phi$  be a probability density function satisfying (29). Then the function  $g_k^h(\cdot; \phi)$  defined in (30) satisfies properties (i), (ii) and (iii) of Lemma 1.*

The proof of the Lemma is rather clear given the explicit form of the derivatives of  $g_k^h(x; \phi)$ , i.e.

$$(31) \quad Dg_k^h(x; \phi) = \Phi\left(\frac{x-k}{h}\right) \quad \text{and} \quad D^2g_k^h(x; \phi) = \phi\left(\frac{x-k}{h}\right)\frac{1}{h}$$

in which it is implicit the inequality

$$0 \leq g_k^h(x; \phi) - g_k^0(x) \leq g_k^h(k; \phi) - g_k^0(k) \leq h \int |u|\phi(u)du$$

A special case of (30) is given by the standard normal distribution:

$$(32) \quad g_k^h(x; \mathcal{N}) = \Phi_{\mathcal{N}}\left(\frac{x-k}{h}\right)(x-k) + h\phi_{\mathcal{N}}\left(\frac{x-k}{h}\right)$$

that will be briefly considered in the applications that follow, as a term of comparison.

A noteworthy implication of (30), emerging clearly from (31) and (4), is the possibility to extract from the nonparametric prices  $q^h(k; \phi)$  an implicit risk neutral density in closed form, that is

$$(33) \quad d^h(x; \phi) = \sum_{i=1}^I b_i D^2g_{j_i}^h(x; \phi) = \sum_{i=1}^I b_i \phi\left(\frac{x-j_i}{h}\right)\frac{1}{h}$$

where the parameters  $b_i$  are the positive coefficients in (6)<sup>7</sup>. The implicit risk neutral density belongs thus to a preassigned family of density mixtures without actually having to make this assumption but rather as the consequence of the exact market pricing of a specific class

<sup>7</sup>In fact, when  $f = g_k^h(x; \phi)$  and  $\bar{X} = \infty$ , then  $\hat{f}(\bar{X}) = 1$ .

of derivatives,  $g_k^h(\phi)$ . Not only, but the mixing parameters are fixed by the market so that in order to estimate the implied risk neutral density one only has to specify the bandwidth  $h$ .

In the option pricing literature a lot of interest was raised by models which assume that the risk neutral density is a mixture of log normals. This modeling choice has been adopted, among others, by Ritchey [20], Melick and Thomas [19], Söderlind and Svensson [23] and Söderlind [22]. In [22], Söderlind lists among the advantages of this approach the closed form for option prices and the flexibility of this family of densities which easily accommodates for skewness and kurtosis of returns while avoiding the difficulties inherent in the nonparametric approach. Although this is clearly not the focus of our work, we mention however the relative ease of (33) in which the weights  $b_i$  are in fact fixed by the market and do not have to be estimated, which is typically a delicate step involving the EM algorithm. In the next section we confront the implicit density (33) with the true one when it is assumed that the returns follow a mixture of log normals to provide at least some evidence that our approach is a viable solution for those cases in which the assumption of a mixture of log normals is justified.

#### 4. EMPIRICAL APPLICATIONS: SIMULATION ANALYSIS

In this section we are going to test our methodology using as our benchmark the classical model of Black and Scholes, augmented for smile effects, using exactly the same values as in Ait-Sahalia and Duarte [1] for ease of comparison. The current underlying price is set at 1,365, maturity to 3 months, the interest rate at 4% and we assume no dividends. We consider 25 strike prices equally spaced between 1,000 and 1,700, so that  $M = 28$ . Volatility will be a linear, decreasing function of the strike, ranging from 40% to 20%. With these values, the option prices range from 378 to 0.9.

**4.1. Deterministic Analysis.** As a first step in our analysis we investigate how close are the quantities  $q^\delta(k)$ ,  $\sigma^\delta(k)$  and  $\nu^\delta(X > k)$  to the true values computed from the model, where  $\sigma^\delta(k)$  is the the level of volatility implicit in  $q^\delta(k)$  according to the Black and Scholes formula. We experiment three possible values of the smoothing parameter, namely  $\delta = 2, 5, 10$ . The results, in terms of deviations of estimated values from true ones, are plotted in Figure 2. Panel *A* represents the price gap, Panel *B* the volatility gap and Panel *C* the gap in terms of cumulative probability. For all choices of  $\delta$ , the distance between true and fictitious values appear to be rather small<sup>8</sup>. Deviations of the smooth price from the true one are indeed quite limited but contain a smile effect and increase for options out of the money. This

---

<sup>8</sup>Of course, the Black and Scholes implicit probability,  $\nu^{B\&S}$ , changes accordingly to incorporate the effect of the strike on volatility.

phenomenon, quite limited, soon disappears when we move from  $\delta = 10$  to  $\delta = 5$ . Deviations from the benchmark are more interesting when considering the implicit probability as the smooth functions lie below 0 up to 1,460 for all choices of  $\delta$ .

[Figure 2 about here.]

**4.2. Smoothness vs. Variance.** The picture changes significantly as we introduce noise into the original model. To avoid arbitrariness we still follow Ait-Sahalia and Duarte [1]. In particular we believe that noise has an implicit microstructural component that may be well captured by liquidity and the bid/ask spread. We proxy illiquidity of options via the factor  $1 + 5|\exp(r\tau)k/S_0 - 1|$  which is 1 for options exactly at the money and gets as high as 2.3 for deeply in and out of the money options. We fix the basis spread to 5% of the option price with a floor at 50 cents and a cap at 3 dollars. Noise is then modeled as a random variable which is uniformly distributed on an interval centered at the origin and with radius equal to the product of illiquidity and half of the spread.

In Figure 3 we draw the distance between the correct Black and Scholes price and its estimate  $q^\delta(k)$  when prices are affected by errors, in the way described above.

[Figure 3 about here.]

The picture shows that the higher the value of  $\delta$  the more smooth is the resulting estimate while, at the same time, goodness of fit decreases. This is a classical finding in nonparametric statistics but presents here some new feature. We observe first that non efficient option prices, corresponding to plain circles, have actually no impact on estimates. The corresponding superhedging prices are plotted as crosses (so that the superposition of a cross to a circle signifies that the corresponding strike is an element of  $K^0(X)$ ). 8 prices out of 24 are non efficient and, as can be seen, most of them are concentrated on the left-hand side corresponding to deeply ITM options. This is due to the combined effect of illiquidity and of the spread which, in our set-up as in the real world, magnifies noise for this segment of the options market. We remark that noise has an asymmetric effect. A positive shock, particularly in the presence of a strong microstructural multiplier, will make the corresponding price inefficient and thus has no impact on our estimates. If the shock is negative or small in magnitude then it is less likely to produce price inefficiency and will be included in the estimated values. Our method is subject to a limited underestimation error for those prices which are affected by errors in a more relevant way, as is the case here for options deeply ITM. Second, we notice that the higher is  $\delta$  the more upward shifted will be the fitting curve. This follows from the payoff  $g_k^\delta$  being larger. Including a constant term to the estimated CALL function would indeed reduce this problem as long as the constant make take on either sign. This addition would correspond however, in its financial counterpart,



to the possibility of taking long or short positions in the riskless asset, a modeling choice which, although completely standard in the literature, blatantly contrasts with our starting assumptions.

**4.3. Monte Carlo Simulation.** In order to construct interval estimates and evaluate the statistical aspects of our approach we run 5,000 Monte Carlo simulations of the error terms. Our model takes the form

$$(34) \quad q(k) + \varepsilon_k^s, \quad k \in K(X), \quad s = 1, 2, \dots, 5000$$

where the error terms are distributed as described above.

	dITM	ITM	ATM	OTM	dOTM	Total
Nr	6	5	5	4	5	25
mean	52.12	50.8	30.5	18.0	15.32	34.74
min	0.0	0.0	0.0	0.0	0.0	12.0
max	83.33	80.0	80.0	75.0	60.0	56.0

Table 1: Percentage of inefficient prices for each market segment.

The introduction of noise produces a number of arbitrage violations, as documented in the preceding Table 1. These violations amount on average to 34.74% of the sample but range up to 56% and are above 40% in the 20% of cases.

For any simulation  $s = 1, \dots, 5000$  we compute, via Theorem 1, the corresponding CALL function,  $q^{\delta,s}(k)$ , implied volatility,  $\sigma^{\delta,s}(k)$ , and the associated probability  $\nu^{\delta,s}$  for different values of  $h$ , although we only plot  $\delta = 5, 10$ . For each strike price  $k$ , we then compute the mean and the quantiles of the simulated sample obtaining for each quantity of interest maximum and minimum values, 90% and 95% confidence intervals, mean and standard deviation. The corresponding curves are plotted in Figure 4, for  $\delta = 5$ , and Figure 5, for  $\delta = 10$ .

We notice that, although in the worst possible scenarios the most deeply ITM options may be overpriced or underpriced by almost 5 \$ (when the price fixed by the model is however more than 300 \$), the mean pricing error is never larger than 2 \$ and, for reasonably liquid options just a few cents. Second, illiquid options are on average underpriced because of the effect highlighted before by which our method makes mainly use of prices affected by negative error terms. This may be considered as an asymmetric smile effect resulting, however, from microstructural factors.

[Figure 4 about here.]

[Figure 5 about here.]

It is noteworthy that for  $\delta = 5$  the true value always falls inside the confidence interval even at the 90% level for all three variables considered, suggesting that our approach produces quite reliable predictions. One may also notice that the two confidence bounds are not symmetric, especially for options deep in the money, reflecting the same distortion noted above. For the case  $\delta = 10$  the situation partly changes as the estimate of the option price becomes less precise and, in particular, the mean price exceeds the actual one by more than 50 cents for all options ATM or OTM. In addition, both lower confidence bounds are breached as frequently as 21.7% suggesting that the choice  $\delta = 10$  produces an increase in smoothness which results however in a significant pricing error. We investigated also the results relative to the intermediate values of  $\delta$  and we report the corresponding values of the mean squared error

$$(35) \quad MSE = \frac{1}{5,000 \times 24} \sum_{k \in K(X)} \sum_{s=1}^{5,000} (q^{\delta,s}(k) - q(k))^2$$

which we report in Table 2.

	$N = 4$	$N = 6$	$N = 8$	$N = 10$
$\delta = 1$	1.93	1.99	2.02	2.03
$\delta = 2$	1.78	1.87	1.92	1.95
$\delta = 5$	1.61	1.63	1.72	1.76
$\delta = 10$	6.17	2.06	1.61	1.58

Table 2. Values of the MSE for alternative parameter choices.

In the literature there is greater emphasis on the implied risk neutral density rather than on the implicit probability so that the focus is actually on the second derivative of the CALL function. We do not have a special interest for this quantity here, partly because the presumption that a density actually exists has no financial basis. In part, however, it is our choice to work with cubic splines that limits our ability to explore densities. Second derivatives in fact exist but are piecewise linear, making the candidate density function not particularly interesting for applications. To obtain smoothness of the implied risk neutral density one should perhaps adopt a different functional form than cubic splines such as splines of higher order (which are however much less tractable computationally speaking)

or as the normal option smoother described in (32)<sup>9</sup>. Another possibility would be to apply to the density obtained some local smoothing technique. In the Monte Carlo analysis performed here, however, smoothness of the density function arises upon averaging across all simulations. In Figure 6 we plot the estimated mean density function together with the one originated from model. The value for the MISE so obtained is  $4.402242e - 07$  and  $4.318549e - 07$ , for  $\delta = 5$  and  $\delta = 10$  respectively.

[Figure 6 about here.]

**4.4. A word on normal option smoothing.** We have performed the above analysis also via the normal options smoothing formula (32) rather than splines. The resulting values for the MSE as well as of confidence levels are less satisfactory than those reported above. We claim, however, that, in models assuming that the risk neutral density is a mixture of log-normals, then the density approach described in (32) produces remarkable results. To this end we have considered three normal densities with instantaneous parameters  $\mu = \{0.027, 0.033, 0.049\}$  and  $\sigma = \{0.3, 0.1, 0.4\}$  and weights  $\alpha = \{0.2, 0.3, 0.5\}$ , respectively. We have used the same values for strike prices, underlying, maturity and interest rate as above. Given these values<sup>10</sup>, we obtain risk neutral prices for options with strike prices ranging from 1,000 to 1,700. From such prices we extract the implied risk neutral density by using (32) after setting  $h = 20$ . We plot in Figure 7 the conditional risk-neutral densities over the interval 1,050 – 1,650. We stress that indeed the mixture of log-normals is rather fat tailed, as desired, but also that its nonparametric estimate is indeed very close to it. The corresponding value for the MISE amounts to  $3.34864e - 06$ .

[Figure 7 about here.]

## 5. EMPIRICAL APPLICATIONS: MARKET DATA

Eventually, we consider an application to market data by selecting an arbitrary trading day, 21<sup>st</sup> October 2010, on the S&P 500 options market<sup>11</sup>. We sample ask quotes at time intervals of one minute each and disregard quotes for which the reported ask size is below 100. On the subsample so obtained we have options quotes for 180 different strike prices – ranging from 50 to 2500 – and 12 possible maturities – from 22<sup>nd</sup> October 2010 to 22<sup>nd</sup>

<sup>9</sup>We have performed the Monte Carlo analysis described in this subsection also for option payoffs obtained via (32). Nevertheless the output we obtained, e.g. in terms of the MSE, is less satisfactory than the results illustrated here for cubic splines.

<sup>10</sup>In fact with the above values,  $\sum_{i=1}^r \alpha_i \mu_i = r$ .

<sup>11</sup>We make use of the quote prices provided by CBOE Market Data Retrieval (MDR). The dataset contains, among other things, information on bid and ask prices and volumes. Data are sampled at a frequency higher than 1 minute.

December 2012. We focus on options expiring in November 2010 and their quotes at 12:06, 12:39 and 13:03, around the market downturn of Figure 8.

[Figure 8 about here.]

At the three selected times and for the selected maturity there are 65, 53 and 67 efficiently quoted strike prices respectively out of 90, 75 and 99. We therefore have a relatively long cross section of strikes and an incidence of inefficient prices of 30% on average. The large number of available strikes is one of the advantages of working with quoted ask prices, as dictated by our model, rather than transaction prices. In the sample there is an overwhelming ratio of contracts ITM by 10% or more and virtually no OTM contract as reported in Table 3.

	Total	dITM	ITM	ATM	OTM	dOTM
Sample	23.972	45.53	27.15	27.22	0.09	0
12:06	90	39.77	35.23	25.00	0	0
12:39	75	52.31	32.31	15.38	0	0
13:03	99	38.37	31.39	30.23	0	0

Table 3: Contracts by Moneyness,  $S_t/K$ .

We select a subsample of strikes ranging from 670 to 1,255. At 12:39 the lesser number of strikes quoted corresponds to a larger maximum interval between consecutive strikes, i.e.  $M = 70$ , while at the other moments strikes do not differ by more than 25 and 30, respectively. Thus for the typical choice  $\delta = 5$  we expect to have a relatively poor performance at 12 : 39, due to the high value of  $M$ . In the following picture we plot for each time  $t$  the curves corresponding to the three distinct values of  $\delta = 2, 5, 10$ .

[Figure 9 about here.]

In fact we clearly see from Figure 9 that the performance of our estimates at 12 : 39 is quite poor due to the fact that there is just one quoted strike between 1075 and 1175. The price curve ends up being very smooth but overestimates actual prices by as much as 5\$ for the value  $\delta = 5$  while in the other selected instants the price gap never exceeds 50 cents for such a parameter choice.

Eventually we plot, for the value  $\delta = 5$ , the implied conditional risk neutral probability and density at 12 : 06 and 13 : 03, Figure 10, to capture the effect of the market downturn on  $\nu^\delta$  and  $d^\delta$ . As expected, the fall in the underlying price determines a more pessimistic view embodied in the implied risk neutral distribution. Of course, as explained above, the density is not a smooth quantity, due to our choice of working with cubic splines.

[Figure 10 about here.]

Eventually, we can use the above information to compute some measures of risk. In particular we consider a long position in a future contract expiring on November 2010, i.e. in a month. The future price is set according to the future/spot parity,  $F(t, T) = S_t \exp(r(T-t))$ . As a proxy for the riskless rate  $r$  we use the 1 month LIBOR rate on that date, quoted at 3.96%. We then compute the  $VaR$  and conditional  $VaR$  at a confidence level of 2% and 5% at the two instants of time. The values are reported in the following Table 4:

	2.5%	5%
$VaR$ 12:06	177.58	110.08
$VaR$ 13:03	163.94	123.95
$CVaR$ 12:06	183.00	134.30
$CVaR$ 13:03	171.26	149.65

Table 4: Measures of risk.

These last remarks suggest the importance to investigate the dependence of  $\nu$  on the current value of the underlying, although outside of our interests here. Another issue that would be important to address is the time evolution of the pricing measure.

## MATHEMATICAL APPENDIX

In this appendix we present some results which we used in the proofs of the main Theorems.

### RESULTS FROM SECTION 2

**Proof of Theorem 1.** The claim essentially follows from [6, Lemmas 7 and 8] in which, however, it is assumed that  $x < \infty$ . The proof given here is adapted from that one to cover the present setting. First of all, in superhedging a given claim, we can restrict to the options which are priced efficiently, i.e. whose strike is included in  $K^0(X)$ . Let  $j_0 = 0$  and write  $j_{I+1} = x$ , if  $x < \infty$  or else  $j_{I+1} > j_I$  arbitrarily.  $g_\theta^0(X) \geq f(X)$  for  $X = 0$ . Assume that  $g_\theta^0(j_i) < f(j_i) - \varepsilon$  for some  $i = 1, \dots, I+1$ . By continuity there exists then  $\delta > 0$  such that  $g_\theta^0(X) < f(X) - \varepsilon$  holds on the set  $\{j_i - \delta < X \leq j_i\}$ . However,  $P(j_i - \delta < X \leq j_i) > 0$  since otherwise  $\text{ess sup}(X \wedge j_i) \leq j_i - \delta$  in contrast with Assumption 1. Thus  $g_\theta^0(X) \geq f(X)$  a.s. implies that  $g_\theta^0(j_i) \geq f(j_i)$  for  $i = 0, \dots, I+1$ . If  $\bar{X} = \infty$  then by the fact that  $j_{I+1}$  was chosen arbitrarily we deduce  $g_\theta^0(x)/x \geq f(x)/x$ , i.e.  $1 \geq \hat{f}(\bar{X})$ .

Viceversa, if  $g_\theta^0(j_i) \geq f(j_i)$  for  $j = 0, \dots, I+1$ , then for each  $\omega \in \{X \leq \bar{X}\}$  and some choice of  $j_{I+1}$  there exists  $i(\omega) \in \{0, 1, \dots, I\}$  such that, for some  $0 \leq a \leq 1$ ,  $j_{i(\omega)} \leq X(\omega) \leq j_{i(\omega)+1}$

and thus

$$f(X(\omega)) \leq af(j_{i(\omega)}) + (1-a)f(j_{i(\omega)+1}) \leq ag_{\theta}^0(j_{i(\omega)}) + (1-a)g_{\theta}^0(j_{i(\omega)+1}) = g_{\theta}^0(X(\omega))$$

because  $f$  is convex and  $g_{\theta}^0$  is linear on each interval  $[j_i, j_{i+1}]$ . In other words,  $g_{\theta}^0(X) \geq f(X)$  outside of  $\{X > \bar{X}\}$  i.e.  $P$ -a.s. so that the inequality  $g_{\theta}^0(X) \geq f(X)$   $P$ -a.s. may be written in vector notation as

$$\mathbf{D}\mathbf{a} \geq \mathbf{f}$$

where  $\mathbf{a}$  is the vector of weights  $a_1, \dots, a_I$  such that  $\theta = \sum_{i=1}^I a_i \theta(j_i) \in \Theta$ ,  $\mathbf{D}$  is the  $(I+1) \times (I+1)$  matrix whose inverse appears in (4) and  $\mathbf{f} = [f(j_1), \dots, f(j_I), \hat{f}(\bar{X})]^T$ .

Define the vectors  $\mathbf{w}, \mathbf{b} \in \mathbb{R}^{I+1}$  implicitly by letting

$$(36) \quad b_I d_I = q(j_I) \quad \text{and} \quad b_I e_I + \sum_{i=n}^{I-1} b_i = \frac{q(j_n) - q(j_{n+1})}{j_{n+1} - j_n} \quad n = 0, \dots, I-1$$

with  $e_I = 1$  if  $\bar{X} < \infty$  or else  $e_I = 0$  and

$$(37) \quad \sum_{i=0}^n w_i = \frac{f(j_{n+1}) - f(j_n)}{j_{n+1} - j_n} \quad n = 0, \dots, I-1 \quad \text{and} \quad \sum_{i=0}^I w_i d_i = \hat{f}(\bar{X})$$

Clearly,  $w$  and  $b$  satisfy (6). The following properties are easily established by induction: (i)  $\mathbf{b} \geq 0$  (as  $j_0, \dots, j_I \in K^0(X)$ ), (ii)  $\mathbf{w} \geq 0$  (as  $f \in \Gamma$ ), (iii)  $\mathbf{b}^T \mathbf{D} = \mathbf{q}^T$  and (iv)  $\mathbf{w} = \mathbf{D}^{-1} \mathbf{f}$ . But then,

$$q(\theta(f)) = \min_{\{\mathbf{a} \in \mathbb{R}_+^{I+1} : \mathbf{D}\mathbf{a} \geq \mathbf{f}\}} \mathbf{q}^T \mathbf{a} = \min_{\{\mathbf{a} \in \mathbb{R}_+^{I+1} : \mathbf{D}\mathbf{a} \geq \mathbf{f}\}} \mathbf{b}^T \mathbf{D}\mathbf{a} \geq \mathbf{b}^T \mathbf{f} = \mathbf{q}^T \mathbf{w}$$

□

### RESULTS FROM SECTION 3

**Proof of Lemma 1.** By a result of Turlach [24, p. 85] the program (16) admits as its solution a cubic  $\mathcal{C}^2$  spline of the form

$$g_k^h(x) = \sum_{i=1}^N \mathbf{1}_{[k+t_i^h, k+t_{i+1}^h)}(x) \Phi_{k,i}^h(x - (k+t_i^h)) + \mathbf{1}_{[k+t_I^h, \infty)}(x)(x-k) \quad x \in \mathbb{R}_+$$

where  $t_1^h, \dots, t_{N+1}^h$  are as in the text and  $\Phi_{k,i}^h$  is a polynomial of degree 3 for  $i = 1, \dots, I$ . It is clear from the constraints imposed to (16) that indeed  $g_k^h \in \Gamma$ . Moreover, these same constraints imply that  $g_k^h(x) = g_k^0(x)$  when  $x \notin [k-h, k+h]$  while  $1 \geq Dg_k^h(x) \geq 0$  on  $[k-h, k+h]$ . Thus,  $g_k^h - g_k^0$  is increasing on  $(-\infty, k]$  and decreasing afterwards, so that  $\sup_x (g_k^h - g_k^0)(x) \leq (g_k^h - g_k^0)(k) = g_k^h(k)$ . Moreover, the (non empty) set of solutions is clearly

convex and the functional  $I_k(h; f)$  is strictly convex in  $f$  so that the solution is necessarily unique. Define

$$\bar{g}_k^h(x) = g_k^h(2k - x) + (x - k) \quad x \geq 0$$

and observe that  $\bar{g}_k^h \in \chi(k; h)$ . Moreover, one deduces from (15) that

$$\begin{aligned} I_k(h; \bar{g}_k^h) &= \sum_{i=1}^{N+1} [g_k^0(k + t_i^h) - \bar{g}_k^h(k + t_i^h)]^2 + (0.1h)^3 \int_{k-h}^{k+h} (D^2 \bar{g}_k^h(x))^2 dx \\ &= \sum_{t_i^h \leq 0} [\bar{g}_k^h(k + t_i^h)]^2 + \sum_{t_i^h > 0} [\bar{g}_k^h(k + t_i^h) - t_i^h]^2 + (0.1h)^3 \int_{k-h}^{k+h} (D^2 \bar{g}_k^h(x))^2 dx \\ &= \sum_{t_i^h \geq 0} [g_k^h(k + t_i^h) - t_i^h]^2 + \sum_{t_i^h < 0} [g_k^h(k + t_i^h)]^2 + (0.1h)^3 \int_{k-h}^{k+h} (D^2 g_k^h(x))^2 dx \\ &= I_k(h; g_k^h) \end{aligned}$$

However, since the solution is unique, we have the symmetry relation

$$(38) \quad g_k^h(x) = g_k^h(2k - x) + (x - k) \quad x \geq 0$$

Let  $f \in \chi(k; h)$ ,  $h > h'$  and define  $T : \chi(k; h) \rightarrow \chi(k; h')$  implicitly by letting

$$Tf(x) = f\left(k + (x - k)\frac{h}{h'}\right) \frac{h'}{h}$$

Observe that  $T$  is one to one and onto and that  $D^2 Tf(x) = h/h' D^2 f(k + (x - k)h/h')$ . Thus,

$$\begin{aligned} I_k(h'; Tf) &= \sum_{i=1}^{N+1} [g_k^0(k + t_i^{h'}) - Tf(k + t_i^{h'})]^2 + (0.1h')^3 \int_{k-h'}^{k+h'} (D^2 Tf(x))^2 dx \\ &= (h'/h)^2 \sum_{i=1}^{N+1} [g_k^0(k + t_i^h) - f(k + t_i^h)]^2 + (0.1h')^3 (h/h') \int_{k-h}^{k+h} (D^2 f(z))^2 dz \\ &= (h'/h)^2 I_k(h; f) \end{aligned}$$

Thus  $f$  solves the program (16) relatively to  $h$  if and only if  $Tf$  solves it relatively to  $h'$ . By uniqueness we conclude that (20) holds. Let  $0 < h' < h$  and observe that, if  $x \leq k$

$$g_k^{h'}(x) = g_k^h(k + (x - k)h/h')h'/h \leq g_k^h(x)h'/h \leq g_k^h(x)$$

a conclusion which extends to  $x > k$  by (38). This proves (ii). Given that  $0 \leq Dg_k^h(x) \leq 1$  we conclude that

$$0 \leq g_k^{h'}(x) - g_k^0(x) \leq g_k^{h'}(k) = \frac{h'}{h} g_k^h(k)$$

so that  $g_k^h$  decreases to  $g_k^0$  uniformly in  $x$ .

If  $y = k' - k \in \mathbb{R}$ , then writing  $t_{I+2}^h = \infty$  and

$$g_k^h(x) = \sum_{i=1}^{N+1} \mathbf{1}_{[k'+t_i^h, k'+t_{i+1}^h)}(x+y) \Phi_{k,i}^h(x+y - (k' + t_i^h)) \equiv \gamma_k^h(x+y; k')$$

It is obvious that  $\gamma_k^h(x+y; k') = \gamma_k^h(x; k)$  and that  $g_{k+y}^0(x+y) = g_k^0(x)$ . But then, for  $i = 1, \dots, N+1$ ,

$$g_{k'}^0(k' + t_i^h) - \gamma_k^h(k' + t_i^h; k') = g_k^0(k + t_i^h) - \gamma_k^h(k + t_i^h; k) = g_k^0(k + t_i^h) - g_k^h(k + t_i^h)$$

and that

$$\int_{k-h}^{k+h} [D^2 \gamma_k^h(x; k)]^2 dx = \int_{k'-h}^{k'+h} [D^2 \gamma_k^h(x+y; k')]^2 dx$$

Using the notation of (15), we conclude that

$$I_k(h) = I_k(h; \gamma_k^h(\cdot; k)) = I_{k+y}(h; \gamma_k^h(\cdot; k')) \geq I_{k'}(h) \quad k > 0, y \in \mathbb{R}$$

The same inequality holds after exchanging  $k$  for  $k'$  so that  $\gamma_k^h(\cdot; k')$  and  $g_{k'}^h$  are both  $\mathcal{C}^2$  splines of degree 3 solving (16) and thus coincide, by uniqueness. We conclude that  $g_{k+y}^h(x) = g_k^h(x-y)$  and thus, if  $0 \leq a \leq 1$  and  $k \leq ak_1 + (1-a)k_2$ ,

$$\begin{aligned} ag_{k_1}^h(x) + (1-a)g_{k_2}^h(x) &= ag_k^h((x + (k_1 - k)) + (1-a)g_k^h(x + k_2 - k)) \\ &\geq g_k^h(x + ak_1 + (1-a)k_2 - k) \\ &\geq g_k^h(x) \end{aligned}$$

proving (i). □

**Proof of Theorem 3.** Write

$$(39) \quad q^\delta(k) = F(k) + \sum_{\tau \in Z} w_k^\delta(\tau) [F(\tau) - F(k)] + \sum_{\tau \in Z} w_k^\delta(\tau) \varepsilon_\tau = F(k) + A_k^\delta(Z) + \sum_{\tau \in Z} w_k^\delta(\tau) \varepsilon_\tau$$

with  $A_k^\delta(Z) = \sum_{\tau \in Z} w_k^\delta(\tau) [F(\tau) - F(k)]$ . Then, by (27a)

$$\begin{aligned} P_Z\left(\left(q^\delta(k) - F(k)\right)^2\right) &= \sum_{\sigma, \tau \in Z} w_k^\delta(\sigma) w_k^\delta(\tau) P_Z\left(\left(F(\sigma) - F(k) + \varepsilon_\sigma\right)\left(F(\tau) - F(k) + \varepsilon_\tau\right)\right) \\ &= A_k^\delta(Z)^2 + [P_Z(\varepsilon_\sigma^2) - P_Z(\varepsilon_\sigma \varepsilon_\tau)] \sum_{\tau \in Z} w_k^\delta(\tau)^2 + P_Z(\varepsilon_\sigma \varepsilon_\tau) + 2P_Z(\varepsilon_\tau) A_k^\delta(Z) \end{aligned}$$

By (23a) we know that  $w_k^\delta(\tau) = 0$  when  $|k - \tau| > (1 + \delta)M_Z$  and  $\tau > 0$ . Using Taylor expansion with Lagrange remainder we get

$$\begin{aligned} |A_k^\delta(Z)| &\leq w_k^\delta(0) [F(0) - F(k)] + \sum_{\tau \in Z \setminus \{0\}} w_k^\delta(\tau) \left[ DF(k) |\tau - k| + \frac{(\tau - k)^2}{2} D^2 F(k) + \frac{|\tau - k|^3}{3!} D^3 F(x_\tau) \right] \\ &\leq F(0) \mathbf{1}_{\{\tau_1 > k - \delta M_Z\}} + |DF(k)| (\delta + 1) M_Z + D^2 F(k) (\delta + 1)^2 M_Z^2 + \frac{\alpha}{3!} (1 + \delta)^3 M_Z^3 \end{aligned}$$



where we made use of (23b). Under (27c) this implies

$$\lim_{M \rightarrow 0} E \left( A_k^\delta(Z)^2 + 2P_Z(\varepsilon_\tau) A_k^\delta(Z) \right) = \lim_{M \rightarrow 0} E \left( (A_k^\delta(Z)^2 + 2P_Z(\varepsilon_\tau) A_k^\delta(Z)) \mathbf{1}_{\{\tau_1 \leq k - \delta M_Z\}} \right) = 0$$

If  $i \geq 1$  we have  $w_k^\delta(\tau_i) = D_{i+1}(g_k^\delta) - D_i(g_k^\delta)$  (with  $\tau_{I+1} = \tau_I + 1$ , conventionally) so that

$$\begin{aligned} w_k^\delta(\tau_i) &= D_i(g_k^\delta) - D_{i-1}(g_k^\delta) \\ &\leq D g_k^\delta(\tau_{i+1}) - D g_k^\delta(\tau_{i-1}) && \text{(convexity)} \\ &= D^2 g_k^\delta(k)(\tau_{i+1} - \tau_{i-1}) + \frac{(\tau_{i+1} - k)^2}{2} D^3 g_k^\delta(x_{\tau_{i+1}}) - \frac{(k - \tau_{i-1})^2}{2} D^3 g_k^\delta(x_{\tau_{i-1}}) \\ &\leq 2D^2 g_k^\delta(k) M_Z + \alpha [1 + (1 + \delta)^2] M_Z^2 \end{aligned}$$

and thus

$$(40) \quad \sum_{\tau \in Z} w_k^\delta(\tau)^2 \leq \mathbf{1}_{\{\tau_1 > k - \delta M_Z\}} + 2D^2 g_k^\delta(k) M_Z + \alpha [1 + (1 + \delta)^2] M_Z^2$$

The claim follows from

$$\begin{aligned} \lim_{M \rightarrow 0} E \left( (g_k^\delta(k) - F(k))^2 \right) &= \lim_{M \rightarrow 0} E \left( [P_Z(\varepsilon_\sigma^2) - P_Z(\varepsilon_\sigma \varepsilon_\tau)] \sum_{\tau \in Z} w_k^\delta(\tau)^2 + P_Z(\varepsilon_\sigma \varepsilon_\tau) \right) \\ &\leq \lim_{M \rightarrow 0} E \left( 2 \sup_{\sigma, \tau} P_Z(\varepsilon_\sigma \varepsilon_\tau) \sum_{\tau \in Z} w_k^\delta(\tau)^2 \mathbf{1}_{\{\tau_1 \leq k - \delta M_Z\}} \right) + E(\varepsilon_\sigma \varepsilon_\tau) \\ &= 0 \end{aligned}$$

□

## REFERENCES

- [1] Y. Aït-Sahalia, J. Duarte (2003), *Nonparametric Option Pricing Under Shape Restrictions*, J. Econometrics **116**, 9-47.
- [2] Y. Aït-Sahalia, A. Lo (1998), *Nonparametric Estimation of State-Price-Densities Implicit in Financial Asset Prices*, J. Finance **53**, 499-547.
- [3] R. W. Banz, M. H. Miller (1978), *Prices for State-Contingent Claims: Some Estimates and Applications*, J. Business **51**, 653-672.
- [4] N. P. B. Bollen, T. Smith, R. E. Whaley (2004), *Modeling the bid/ask Spread: Measuring the Inventory-Holding Premium*, J. Financ. Econ. **72**, 97-141.
- [5] D. Breeden, R. Litzenberger (1978), *Prices of State-Contingent Claims Implicit in Option Prices*, J. Business **51**, 621-651.
- [6] G. Cassese (2014), *Asset Pricing in an Imperfect World*, mimeo, <http://arxiv.org/abs/1410.6408>.
- [7] M. Das, W. K. Newey, F. Vella (2003), *Estimation of Sample Selection Models*, Rev. Econ. Stud. **70**, 33-58.
- [8] R. L. Dykstra (1983), *An Algorithm for Restricted Least Squares Regression*, J. Amer. Stat. Ass. **78**, 837-842.

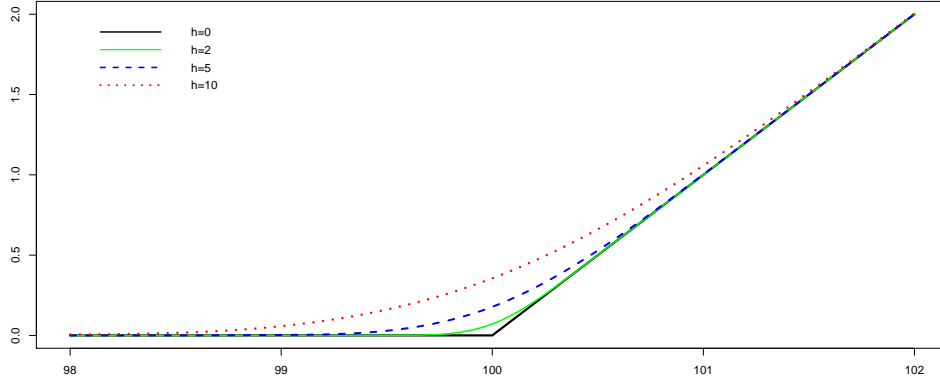
- [9] R. L. Eubank (1999), *Nonparametric Regression and Spline Smoothing*, Marcel Dekker, New York - Basel.
- [10] M. R. Fengler (2009), *Arbitrage-Free Smoothing of the Implied Volatility Surface*, *Quant. Finance* **9**, 417-428.
- [11] M. R. Fengler, L.-Y. Hin (2014), *Semi-nonparametric Estimation of the Call-Option Price Surface under Strike and Time-to-expiry No-arbitrage Constraints*, *J. Econometrics* **184**, 242-261.
- [12] P. Gagliardini, C. Gourieroux, E. Renault (2011) *Efficient Derivative Pricing by the Extended Method of Moments*, *Econometrica* **79**, 1181-1232.
- [13] R. Garcia, E. Ghysels, E. Renault (2010), *The Econometrics of Options Pricing*, in Y. Aït-Sahalia, L. P. Hansen (Eds.) *Handbook of Financial Econometrics*, vol. 1, 479-552, Amsterdam North-Holland.
- [14] J. Hasbrouck (2002), *Stalking the "Efficient Price" in Market Microstructure Specifications: an Overview*, *J. Financial Markets* **5**, 329-339.
- [15] J. J. Heckman (1979), *Sample Selection Bias as a Specification Error*, *Econometrica* **47**, 153-161.
- [16] R. D. Huang, H. R. Stoll (1997), *The Components of the bid/ask Spread: a General Approach*, *Rev. Financial Studies* **10**, 995-1034.
- [17] J. C. Jackwerth, M. E. Rubinstein (1996), *Recovering Probability Distributions from Option Prices*, *J. Finance* **51**, 1611-1631.
- [18] E. Mammen, C. Thomas-Agnan (1999), *Smoothing Splines and Shape Restrictions*, *Scand. J. Statist.* **26**, 239-252.
- [19] W. R. Melick, C. P. Thomas (1997), *Recovering Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis*, *J. Financ. Quant. Analysis* **32**, 91-115.
- [20] R. J. Ritchey (1990), *Call Option Valuation for Discrete Normal Mixtures*, *J. Financ. Res.* **13**, 285-296.
- [21] L. Rompolis, E. Tzavalis (2008), *Recovering Risk Neutral Densities from Option Prices: A New Approach*, *J. Financ. Quant. Analysis* **43**, 1037-1053.
- [22] P. Söderlind (2000), *Market Expectations in the UK before and after the ERM Crisis*, *Economica* **67**, 1-18.
- [23] P. Söderlind, L. Svensson (1997), *New Techniques to Extract Market Expectations from Financial Instruments*, *J. Monetary Econ.* **40**, 383-429.
- [24] B. A. Turlach (2005), *Shape Constrained Smoothing Using Smoothing Splines*, *Comp. Stat.* **20**, 81-103.
- [25] A. Yatchew (1998), *Nonparametric Regression Techniques in Economics*, *J. Econ. Lit.* **36**, 669-721.
- [26] A. Yatchew, W. Härdle (2006), *Nonparametric State Price Density Estimator Using Constrained Least Squares and Bootstrap*, *J. Econometrics* **133**, 579-599.
- [27] H. Yin, Y. Wang, L. Qi (2009), *Shape-Preserving Interpolation and Smoothing for Options Market Implied Volatility*, *J. Optim. Theory Appl.* **142**, 243-266.

UNIVERSITÀ MILANO BICOCCA

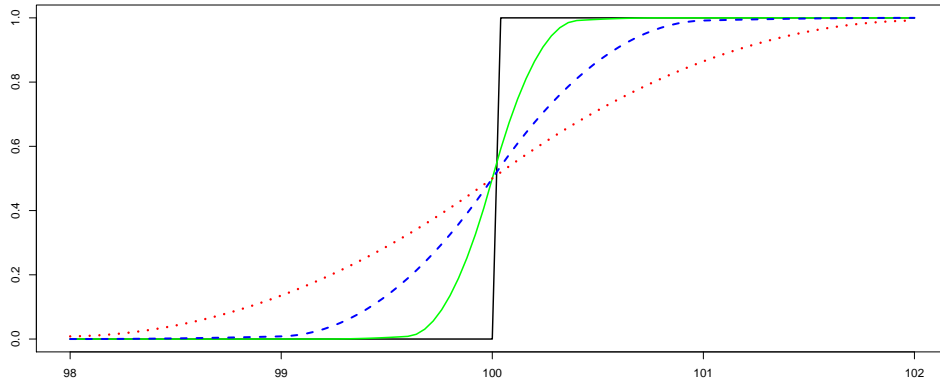
*E-mail address:* gianluca.cassese@unimib.it

*Current address:* Department of Economics, Statistics and Management Building U7, Room 2097, via Bicocca degli Arcimboldi 8, 20126 Milano - Italy

Panel A: Option Payoff  $g_k^h(X)$



Panel B: Option Payoff Derivative  $Dg_k^h(X)$



Panel C: Smooth Superhedging

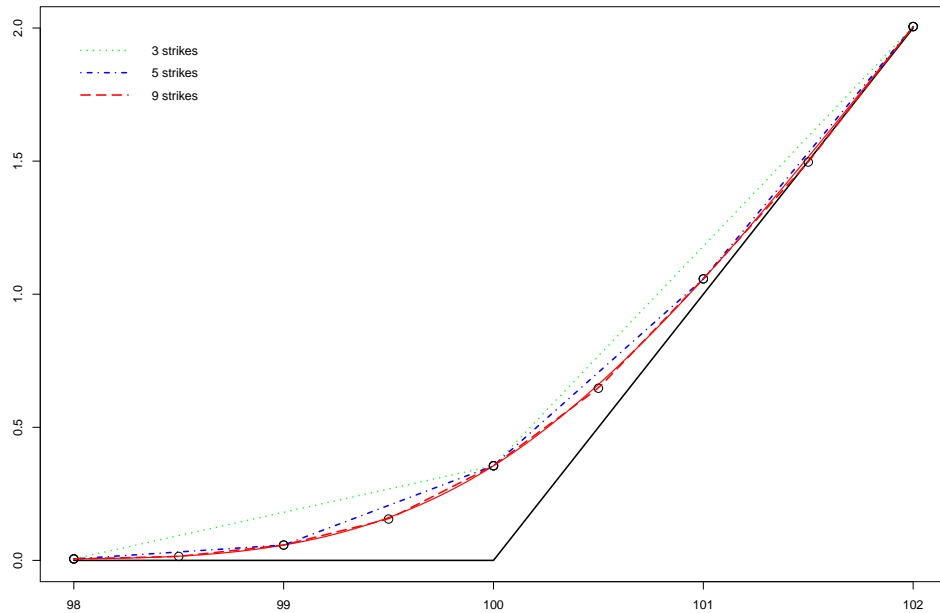
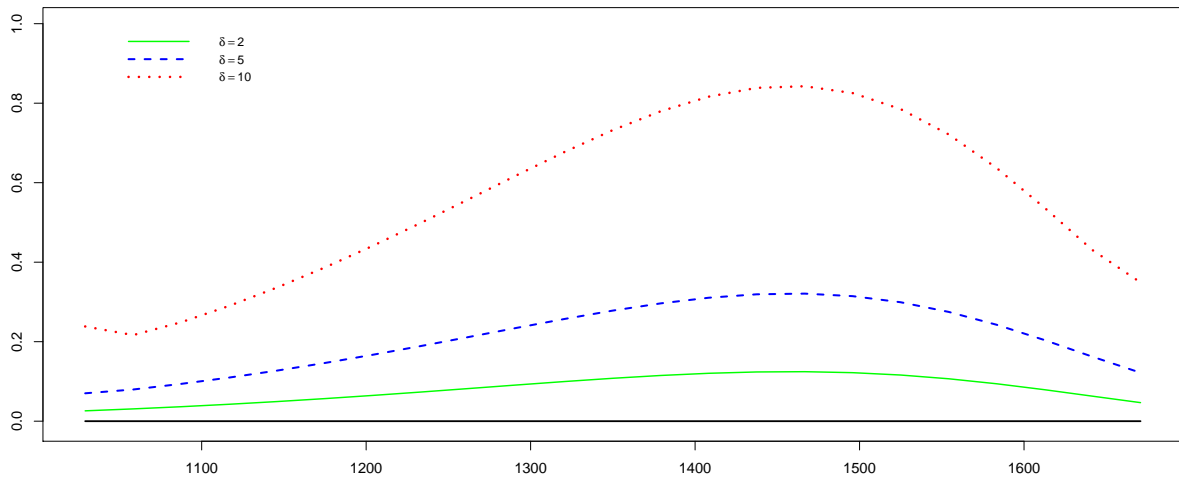


FIGURE 1. Smoothing Option Payoff by  $h$ .

Plot of  $g_k^h(x)$  (Panel A) and  $Dg_k^h(x)$  (Panel B) for  $k = 100$  and  $h = 2, 5, 10$ . In Panel C we draw  $g_k^{h=10}(x)$  (red dotted line) together with  $\hat{g}_k^{h=10}(X)$  (piecewise linear curves), the payoff of the portfolio superhedging it assuming to have 3, 5 or 9 equally spaced strike prices.

Panel A: Option Price  $q^{\delta M}(k)$ 

Panel B: Implied Volatility

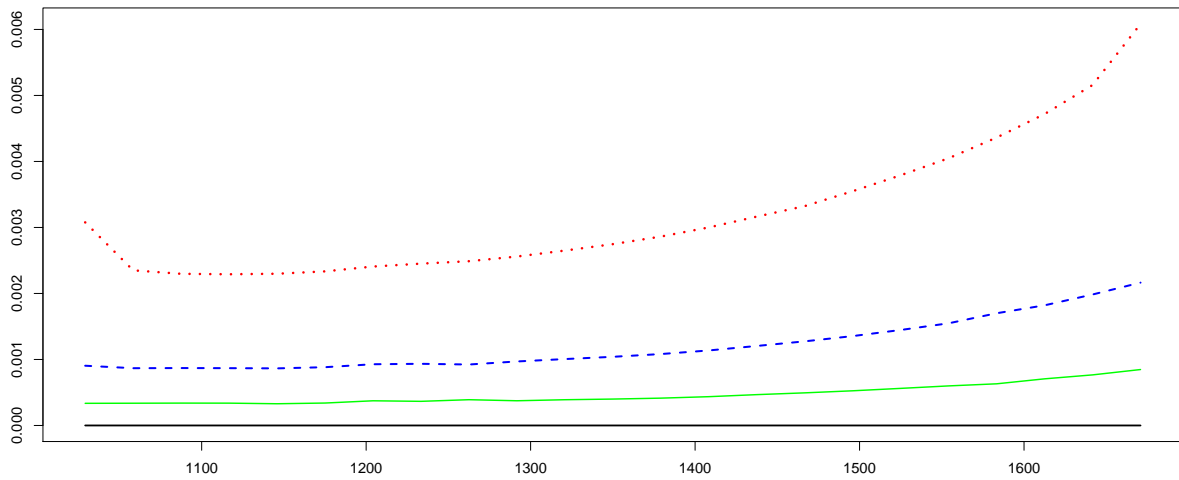
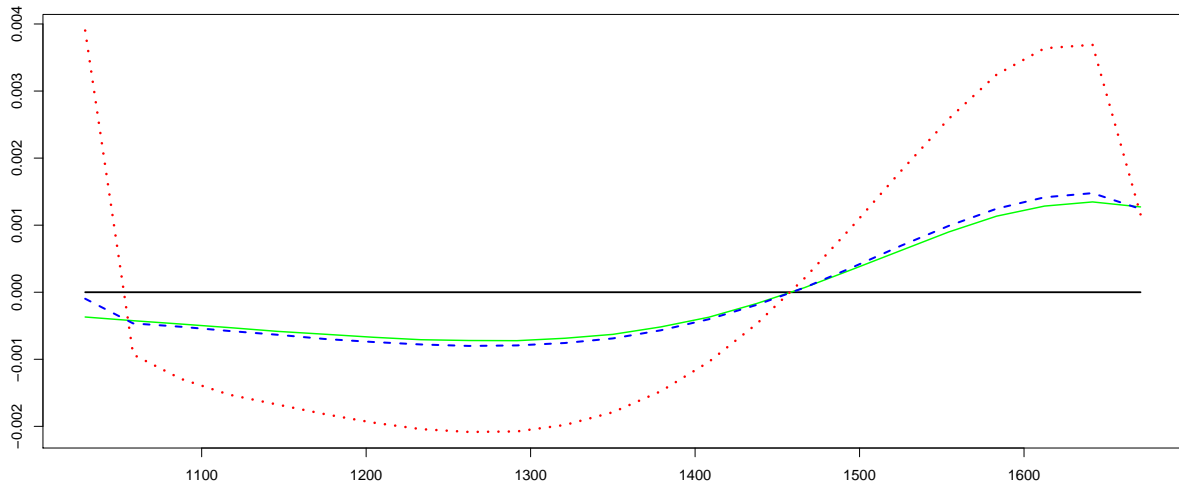
Panel C: Probability  $\nu^{\delta M}(X > k)$ 

FIGURE 2. Deviations from Black and Scholes with a Smile.

$\sigma = 0.4 - 0.2(k - 1000)/700$ ,  $\nu^{B\&S} = -\partial\sigma^{B\&S}/\partial K - \partial\sigma^{B\&S}/\partial\sigma d\sigma/dk$  and  $N = 10$ .

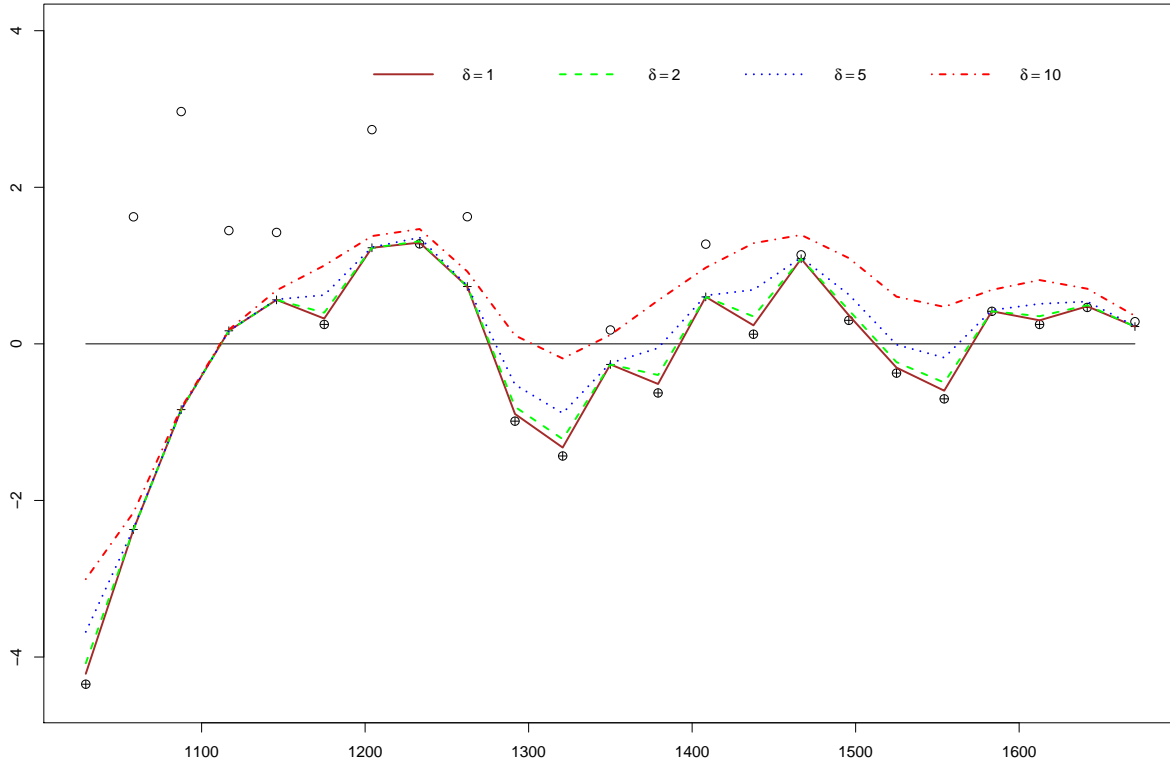
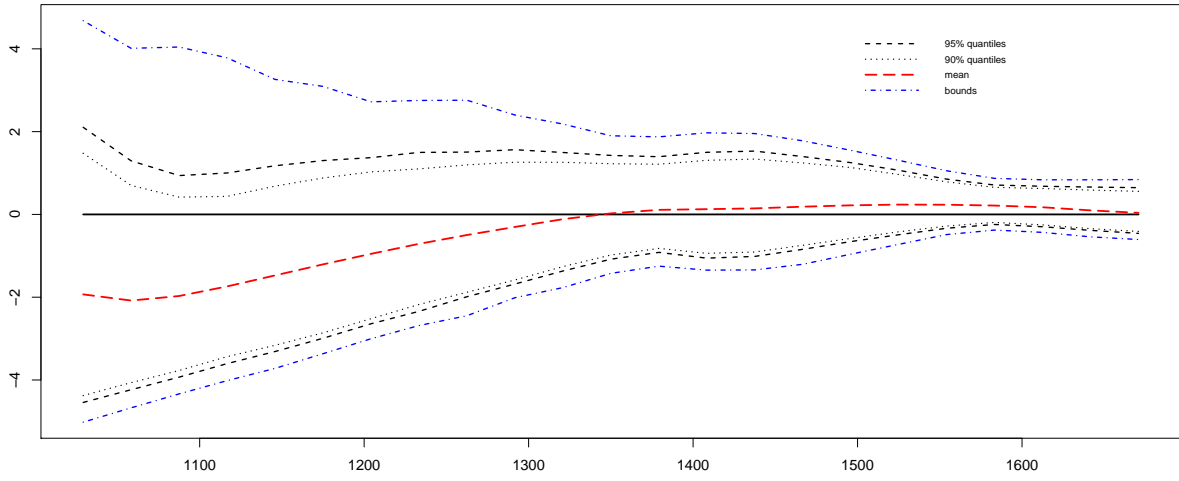


FIGURE 3. Smoothness vs. Variance.

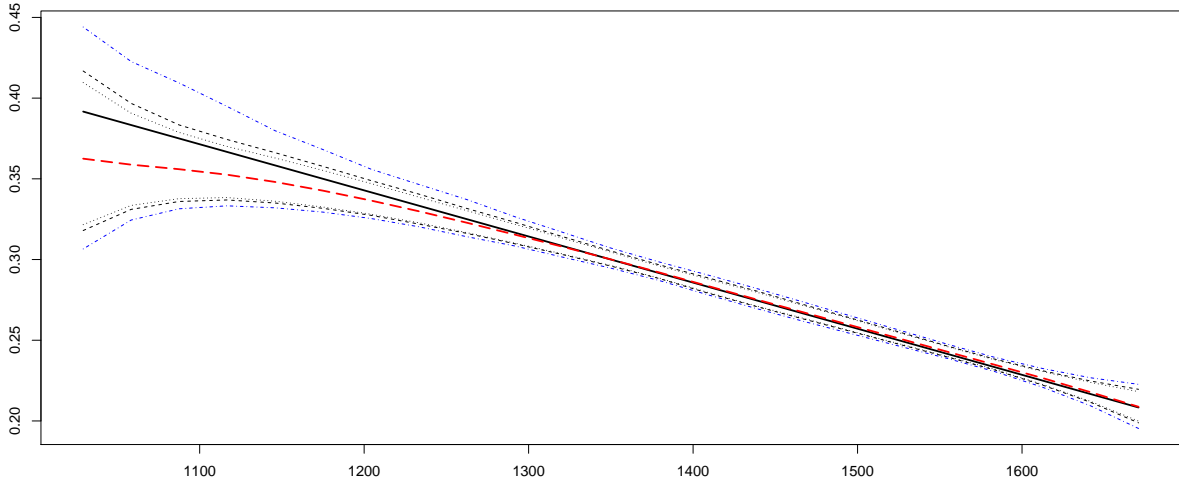
Estimating the model  $q(k) + \varepsilon_k$  with  $N = 10$  and  $\delta = 1, 2, 5, 10$ . Values are expressed in terms of their distance from the correct B&S price. Actual option prices are plotted as circles while the corresponding efficient prices are plotted as crosses. The two symbols are overwritten whenever the corresponding strike belongs to  $K^0(X)$ .

Figures

Panel A: Option Price



Panel B: Volatility



Panel C: Probability

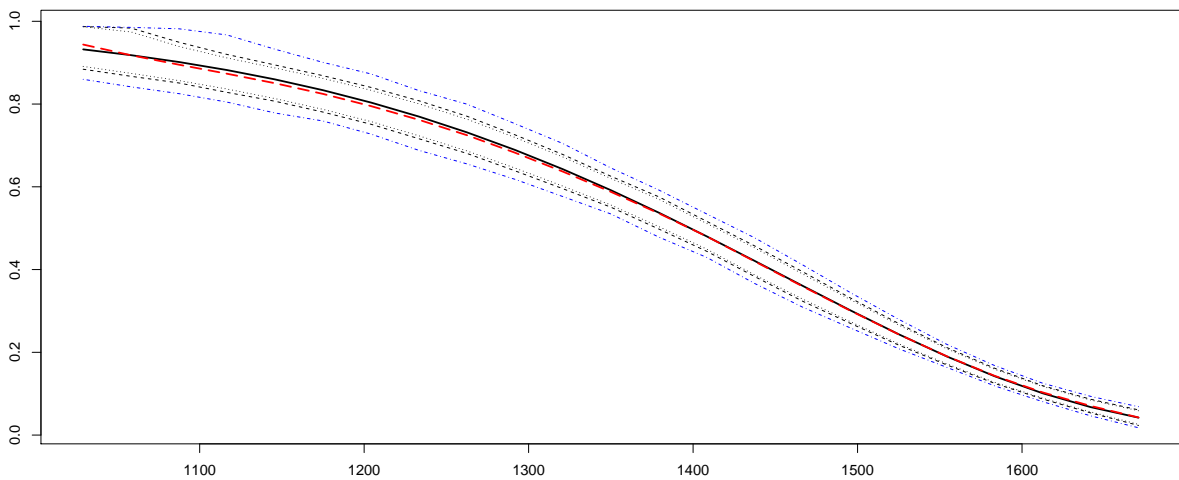
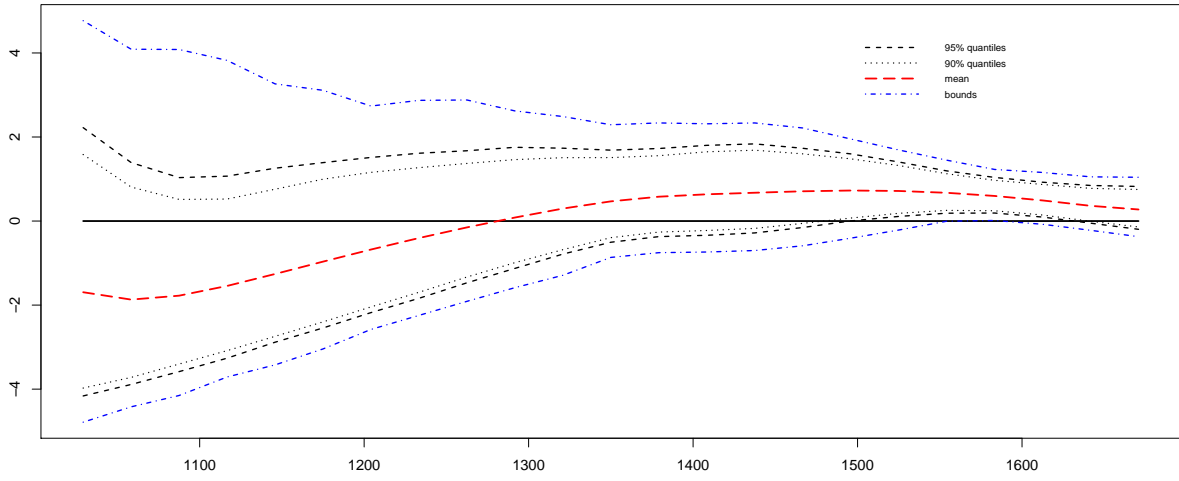


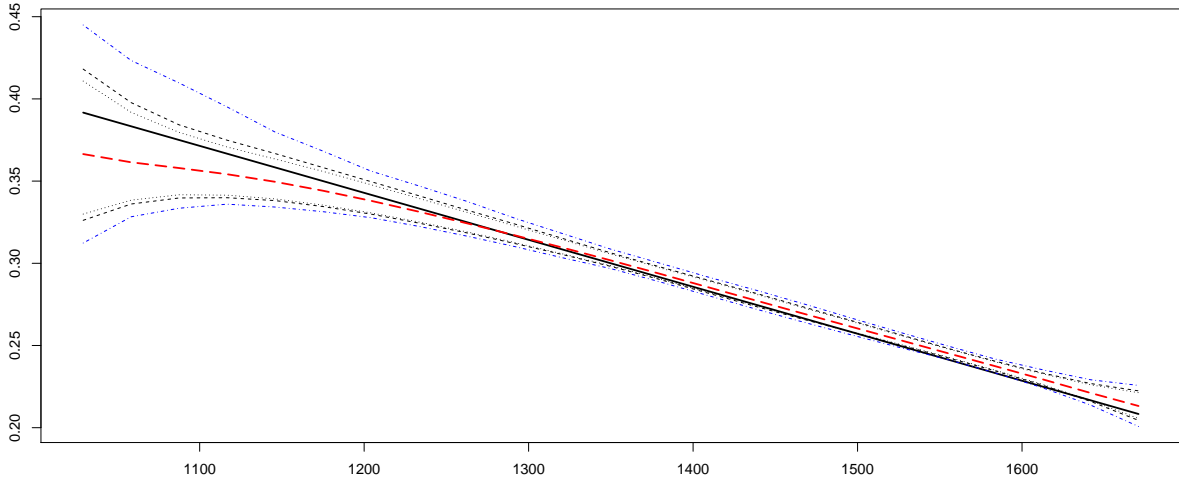
FIGURE 4. Simulated Simulated Confidence Intervals:  $N = 10$  and  $\delta = 5$ .

The confidence bands and mean were obtained after 5,000 Monte Carlo simulations.

Panel A: Option Price



Panel B: Volatility



Panel C: Probability

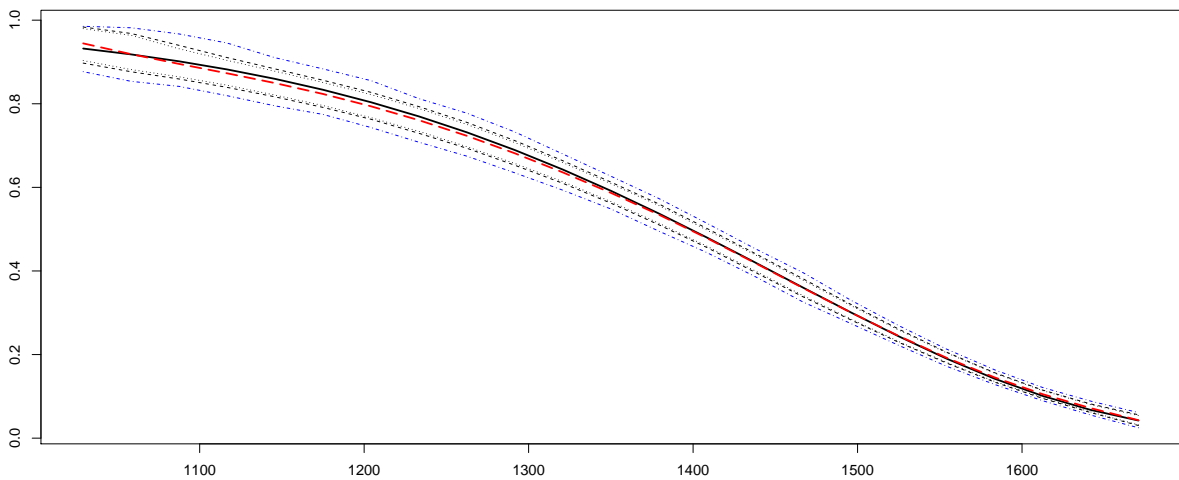
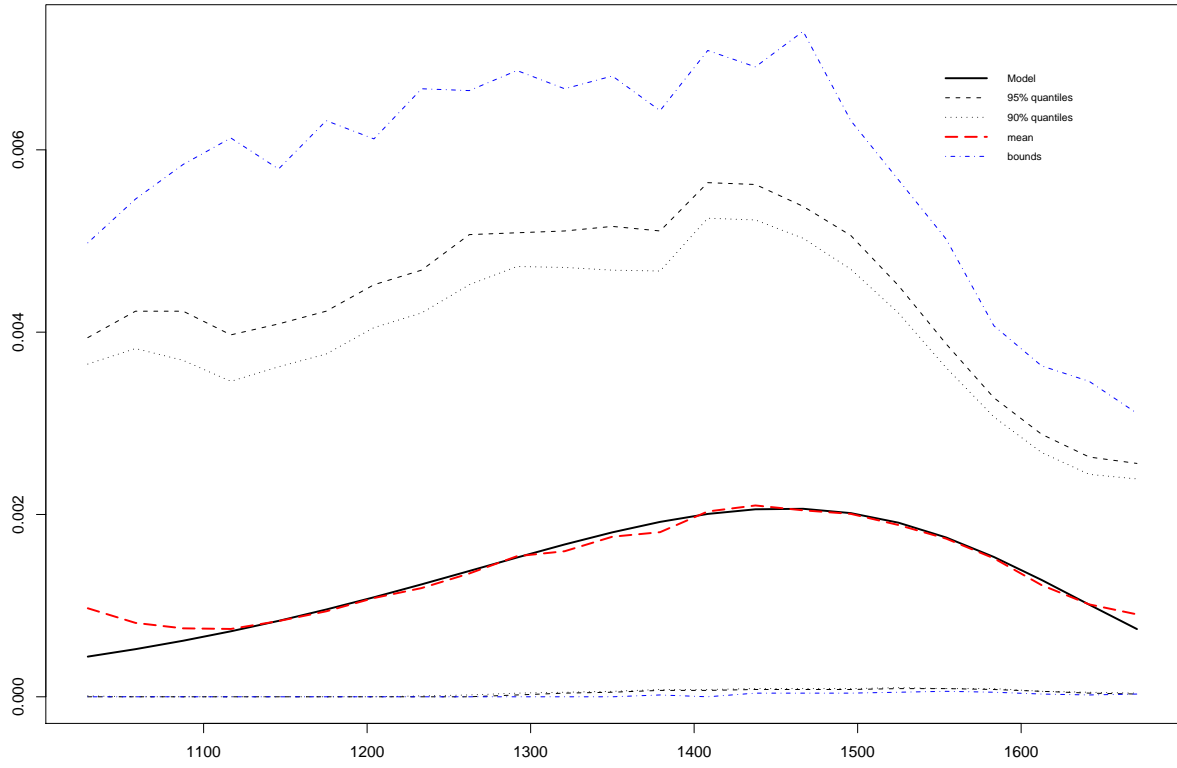
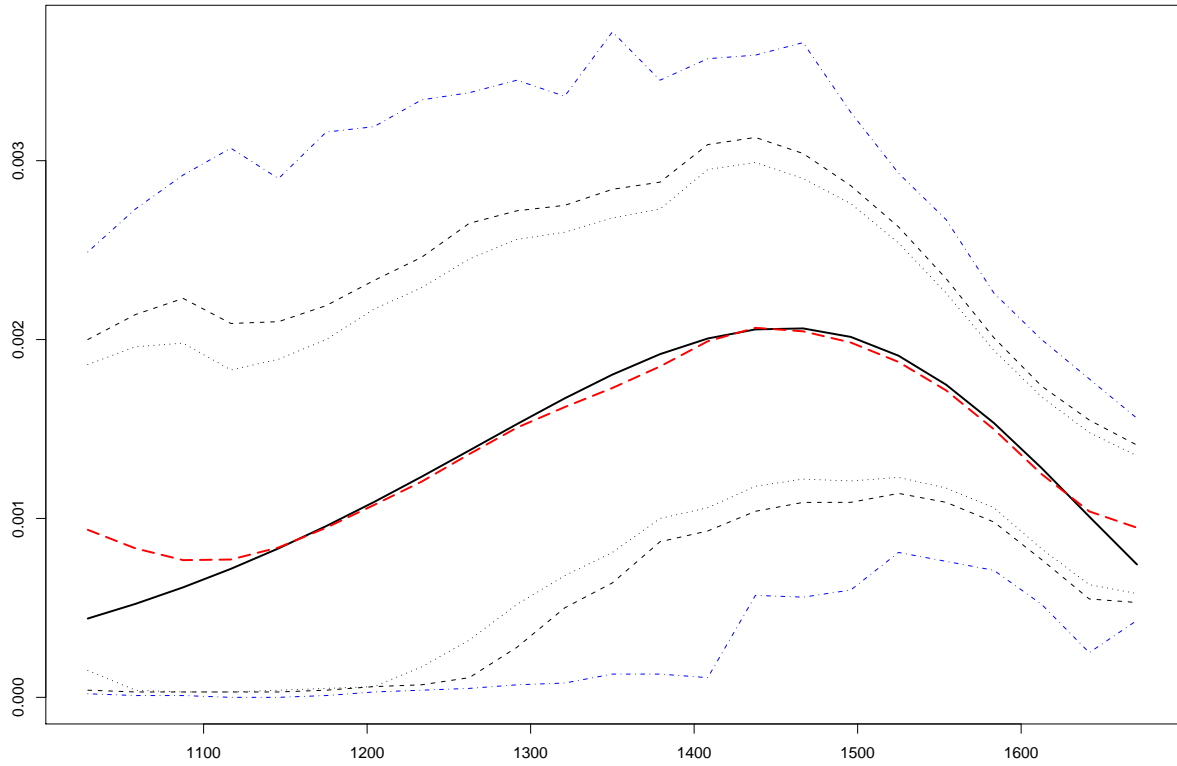


FIGURE 5. Simulated Confidence Intervals:  $N = 10$  and  $\delta = 10$ .

The confidence bands and mean were obtained after 5,000 Monte Carlo simulations.

## Figures

Panel A:  $\delta = 5$ Panel B:  $\delta = 10$ FIGURE 6. Implied risk-neutral average density:  $N = 10$  and  $\delta = 5, 10$ .



Risk neutral density

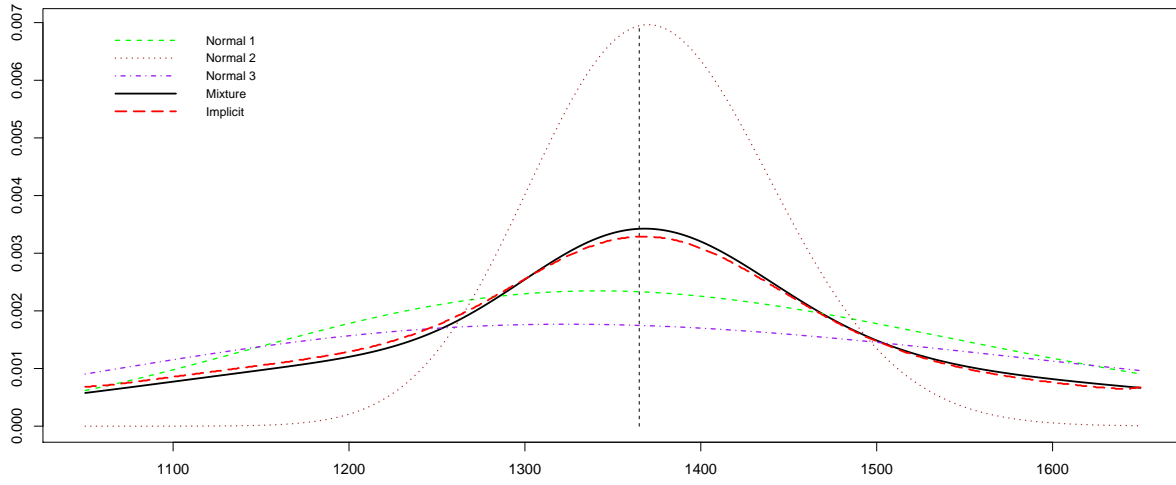


FIGURE 7. Risk-neutral densities for a mixture of log-normals.

All curves represent a corresponding conditional density over the interval 1050 – 1650. The dashed lines are the starting log-normal densities and the black solid line their mixture. The red line is the density extracted from the option prices.  $MISE = 3.34864e - 06$ .

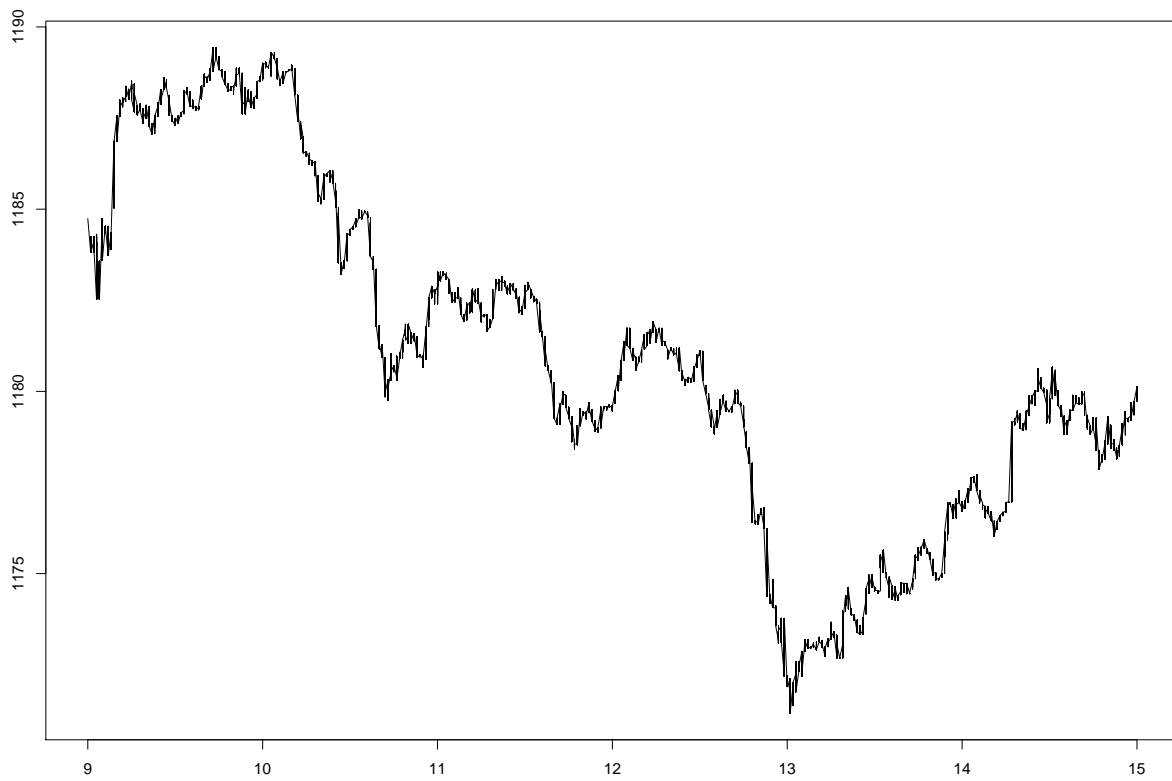


FIGURE 8. The S&P 500 Index on 21<sup>st</sup> October 2010

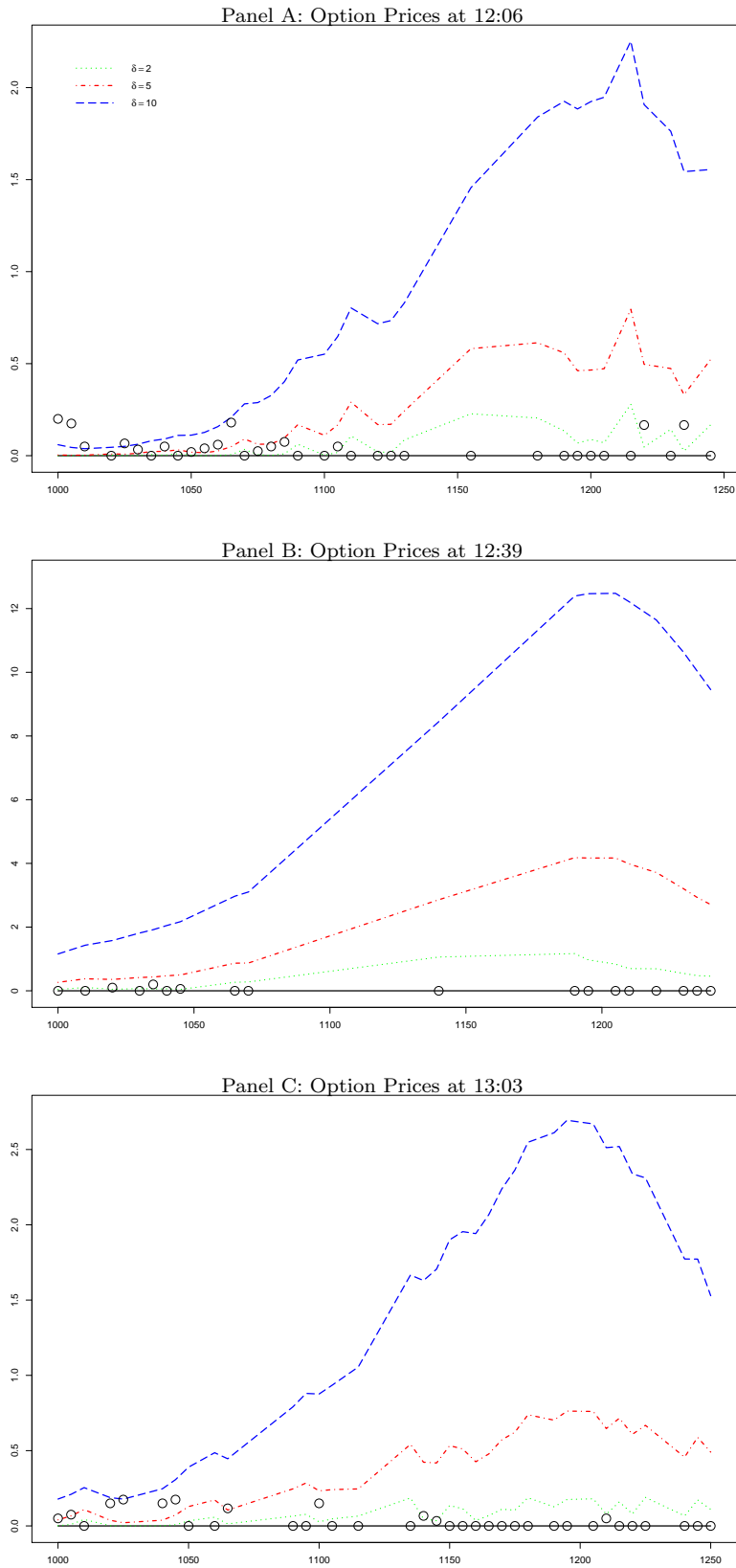


FIGURE 9. Actual and estimated option prices as differences with the efficient prices. Market prices are represented as circles.

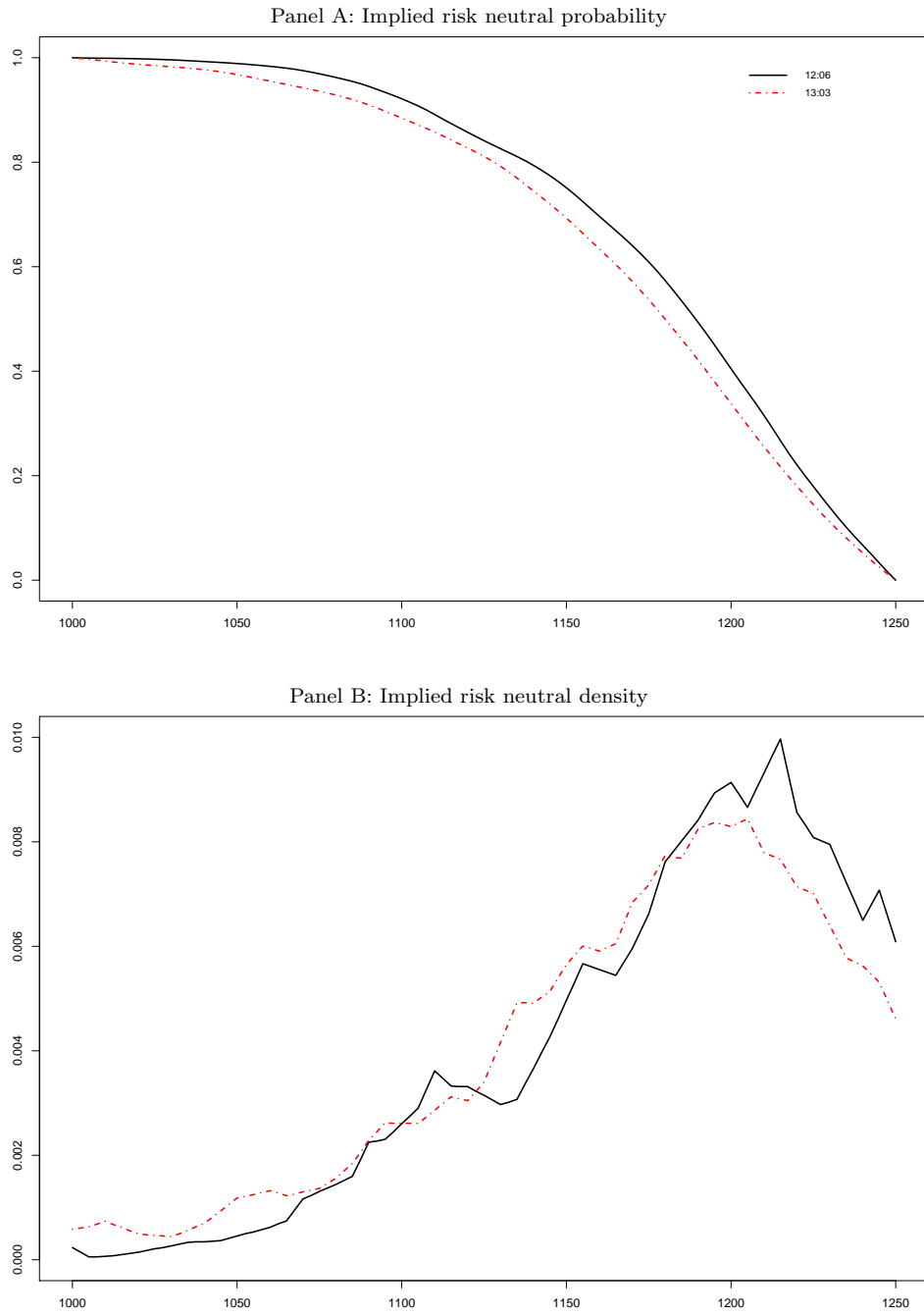


FIGURE 10. Conditional risk neutral distribution,  $\nu^\delta(X > x)$ , and density,  $d^\delta(x)$ , at 12:06 and 13:03 for the case  $\delta = 5$ .