Nonparametric State Price Density Estimation
Using Constrained Least Squares and the Bootstrap

Adonis Yatchew* and Wolfgang Härdle**

The economic theory of option pricing imposes constraints on the structure of call functions and state price densities. Except in a few polar cases, it does *not* prescribe functional forms. This paper proposes a nonparametric estimator of option pricing models which incorporates various restrictions within a single least squares procedure thus permitting investigation of a wide variety of model specifications and constraints. Among these we consider monotonicity and convexity of the call function and integration to one of the state price density. The procedure easily accommodates heteroskedasticity of the residuals. The bootstrap is used to produce confidence intervals for the call function and its first two derivatives. We apply the techniques to option pricing data on the DAX.

Keywords: option pricing, state price density estimation, nonparametric least squares, bootstrap inference, monotonicity, convexity

April 9, 2003.

*Department of Economics, University of Toronto. Support of the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

**Humboldt-Universität zu Berlin, Center for Applied Statistics and Economics. Support of SFB373 "Quantification und Simulation Ökonomischer Processe" Deutsche Forschungsgemeinschaft is gratefully acknowledged.

The authors are grateful to Christian Gourieroux and to two anonymous referees for helpful comments.

1. State Price Density Estimation

1.1 Parametric or Nonparametric?

Option price data have characteristics which are both nonparametric and parametric in nature. The economic theory of option pricing predicts that the price of a call option should be a monotone decreasing convex function of the strike price. It also predicts that the state price density (SPD) which is proportional to the second derivative of the call function, is a valid density function over future values of the underlying asset price, and hence must be non-negative and integrate to one. Except in a few polar cases, the theory does *not* prescribe specific functional forms. (Indeed the volatility smile is an example of a clear violation of the lognormal *parametric* specification implied by Black-Scholes.) All this points to a nonparametric approach to estimation of the call function and its derivatives.

On the other hand, multiple transactions are typically observed at a finite vector of strike prices. Thus, one could argue that the model for the option price – as a function of the strike price (other variables held constant) -- is intrinsically parametric. Indeed given sufficient data, one can obtain a good estimate of the call function by simply taking the mean transactions price at each strike price. Unfortunately, even with large data-sets, accurate estimation of the call function at a finite number of points does not assure good estimates of its first and second derivatives, should they exist. To incorporate smoothness and curvature properties, one can select a parametric family which is differentiable in the strike price, and impose constraints on coefficients. Such an approach, however, risks specification failures.

Fortunately, nonparametric regression provides a good reservoir of candidates for flexible estimation. Indeed, a number of authors have used nonparametric or semiparametric techniques in the estimation or testing of derivative asset models.

Among them are Aït-Sahalia (1996), Jackwerth and Rubinstein (1996), Ghysels et al. (1997), Aït-Sahalia and Lo (1998, 2000), Aït-Sahalia and Duarte (2000) , Broadie et al (2000a,b), Garcia and Gencay (2000), Aït-Sahalia, Bickel and Stoker (2001), Cont (2001), Cont and Fonseca (2002), Cont and Tankov (2002), Daglish (2002) and Härdle, Kleinow and Stahl (2002).

In earlier work, Yatchew and Bos (1997) showed how nonparametric least squares can easily incorporate a variety of constraints such as monotonicity, concavity, additive separability, homotheticity and other implications of economic theory. Their estimator uses least squares over sets of functions bounded in Sobolev norm. Such norms provide a simple means for imposing smoothness of derivatives of various order. There is a growing literature on the imposition and testing of curvature properties on nonparametric estimators. (See Wright and Wegman (1980), Schlee (1982), Friedman and Tibshirani (1984), Villalobas and Wahba (1987), Mukarjee (1988), Ramsay (1988), Robertson, Wright and Dykstra (1988), Kelly and Rice (1990), Mammen (1991), Goldman and Ruud (1992), Yatchew (1992), Mukarjee and Stern (1994), Bowman, Jones, and Gijbels (1998), Diack and Thomas-Agnan (1998), Ramsay (1998), Mammen and Thomas-Agnan (1999), Diack (2000), Gijbels et al.(2000), Hall and Heckman (2000), Hall and Huang (2001), Groeneboom, Jongbloed, and Wellner (2001), Juditsky and Nemirovski (2002), and Hall and Yatchew (2002).)

In the current paper, we combine shape restrictions with nonparametric regression to estimate the call price function within a *single* least squares procedure. Constraints include smoothness of various order derivatives, monotonicity and convexity of the call function and integration to one of the SPD. Confidence intervals and test procedures may be implemented using bootstrap methods. In addition to providing simulation results we apply the procedures to option data on the DAX index for the period January 4-15, 1999.

As an initial illustration of the benefits of smooth constrained estimation, particularly when estimating derivatives, we have generated 20 independent transactions prices at each of 25 strike prices. Details of the data generating mechanism are contained in Section 3 below. The top panel of Figure 1A depicts all 500 observations and the ‘true’ call function. As is typical in market data, the variance decreases as the option price declines. The second panel depicts the estimated call function obtained by taking the mean transactions price at each of the 25 strike prices. The bottom panel depicts our smooth constrained estimate. Both estimates lie close to the true function.

Insert Figure 1A

Figure 1B contains estimates of the first derivative. The upper panel depicts first-order divided differences of the point means, (these are the slopes of the lines joining the consecutive means in the middle panel of Figure 1A). By the mean value theorem, they should provide a reasonable estimate of the true first derivative near the point of approximation. But as can be seen, the estimate deteriorates rapidly as one moves to the left and the variance in transactions prices increases. The bottom panel depicts the first derivative of the proposed smooth constrained estimate which by comparison is close to the first derivative of the true call function.

Insert Figures 1B

Figure 1C illustrates estimates of the second derivative of the call function. The upper panel depicts second-order divided differences of the point means. (These are the slopes of the lines joining consecutive points in the top panel of Figure 1B.) The estimates gyrate wildly around the true second derivative. The lower panel depicts the second derivative of the smooth constrained estimate which tracks the true function reasonably well (note the change in scale of the vertical axis).

Insert Figure 1C

A number of practical advantages ensue from the procedures we propose. First, various combinations of constraints can be incorporated in a natural way within a single least squares procedure. Second, our 'smoothing' parameter has an intuitive interpretation since it measures the smoothness of the class of functions over which estimation is taking place by using a (Sobolev) norm. If one wants to impose smoothness on higher order derivatives, this can be done by a simple modification to the norm. Third, call functions and SPDs can be estimated on an hour-by-hour, day-by-day or 'moving window' basis, and changes in shape can be tracked and tested. Fourth, our procedures readily accommodate heteroskedasticity and time series structure in the residuals.

In the following, we consider two types of generating mechanisms for the "x" variable. In the first, x is drawn from a continuous distribution as would be the case if one were estimating the call function as a function of "moneyness". In the second, x is drawn from a discrete distribution at a finite set of strike prices as depicted in the upper panel of Figure 1A. The paper is organized as follows. The remainder of this section outlines the relevant financial theory and establishes notation. Section 2 outlines the estimator as well as inference procedures. Section 3 contains the results of simulations and estimation using DAX index options data. Section 4 contains our conclusions. Appendices contain proofs and derivations.

1.2 Financial Market Theory

Before proceeding, we briefly review some of the relevant financial theory. Implicit in the prices of traded financial assets are Arrow-Debreu prices or in a continuous setting, the state price density. These are elementary building blocks for understanding markets under uncertainty. The existence and characterization of SPDs has been studied by Black and Scholes (1973), Merton (1973), Rubinstein (1976) and Lucas (1978) amongst many others. Under the assumption of no-arbitrage, the SPD is usually called the risk neutral density because if one assumes that all investors are risk neutral, then the expected return on all assets must equal the risk free rate of interest. Cox and Ross (1976) showed that under this assumption Black-Scholes equation follows immediately. Other approaches have been proposed by Derman and Kani (1994) and Barle and Cakici (1998).

Let x be the strike price for a call option which will expire at time T. Let t be the current time, r the interest rate, τ the time to expiry and δ the dividend yield. Let P_t and P_T denote prices of the underlying asset at times t and T respectively. Then the call pricing function at time t is given by:

$$C\left(x,\tau,P_t,r,\delta\right) = e^{-r\tau}\int_0^\infty \max\left(P_T - x, 0\right) f^*\left(P_T | P_t,\tau,r,\delta\right) dP_T \qquad (1)$$

where the function f^* is the state price density. It assigns probabilities to various values of the asset at time of expiration given the current asset price, the time to expiry, the current risk-free interest rate and the corresponding dividend yield of the asset. As stated earlier, the call function is monotone decreasing and convex in x.

Breeden and Litzenberger (1978) show that the second derivative of the call pricing function with respect to the strike price is related to the state price density by:

$$f^*\left(P_T | P_t,\tau,r,\delta\right) = e^{r\tau}\frac{\partial^2 C\left(x,\tau,P_t,r,\delta\right)}{\partial x^2} \qquad (2)$$

We will focus on data over a sufficiently brief time span so that we may take the time to maturity, the underlying asset price, the interest rate and dividend yield to be roughly constant. Our objective will be to estimate the call function subject to monotonicity and convexity constraints and the constraint that the implied SPD is non-negative and integrates to a value not exceeding one.

We will use the following notational conventions. For an arbitrary vector z and matrices A, B we will use z_s, A_{st}, $[Az]_s$ and $[AB]_{st}$ to denote elements. Occasionally, we will need to refer to sub-matrices of a matrix. In this case we will adopt the notation $A[a{:}b, c{:}d]$ to refer to those elements which are in rows *a* through *b* and columns *c* through *d*. Given a function f, we will denote derivatives using bracketed superscripts, e.g., $f^{(1)}, f^{(2)}$.

2. Constrained Nonparametric Procedures

2.1 Nonparametric Least Squares

We begin with constrained nonparametric least squares estimation of a function of one variable on the interval $[a,b]$. Given data $(x_1,y_1),\ldots,(x_n,y_n)$, let $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$. (The x_i will be strike price or "moneyness" and y_i the option price.) With mild abuse of notation we will use x and y to denote the variable in question in addition to the vector of observations on that variable. Our model is given by:

$$y_i = m(x_i) + \varepsilon_i \qquad i = 1,\ldots,n \tag{3}$$

We will assume that the regression function m is four times differentiable, which in a nonparametric setting, will ensure consistent and smooth estimates of the function and its first and second derivatives. (Other orders of differentiation can readily be accommodated using the framework below.) We will assume that the x_i lie in the interval $[a,b]$. (For example, if the x variable is 'moneyness' then $[a,b]$ would typically be the interval [.8, 1.2].) For the time being assume the residuals ε_i are independent and heteroskedastic. Let $\underset{n\times n}{\Sigma}$ be the diagonal matrix with diagonal values $\sigma^2(x_1),\ldots,\sigma^2(x_n)$.

Let C^4 be the space of four times continuously differentiable scalar functions, i.e., $C^4 = \{ m:[a,b]\to\mathbb{R}^1 \mid m^{(i)} \in C^0, i = 0,1,..,4\}$ where C^0 is the set of continuous functions on $[a,b]$. On the space C^4 define the norm, $\| m \|_{\infty,4} = \sum_{i=0}^{4} max_x | m^{(i)} |$ in which case C^4 is a complete, normed, linear space, i.e., a Banach space. Consider the following inner product:

$$<m_1, m_2>_{Sob} = \sum_{i=0}^{4} \int_a^b m_1^{(i)} m_2^{(i)} dx \qquad (4)$$

with corresponding norm:

$$\| m \|_{Sob} = \left[\sum_{i=0}^{4} \int_a^b [m^{(i)}]^2 \, dx \right]^{1/2} \qquad (5)$$

and define the Sobolev space $\mathcal{H}^4$ as the completion of $\{m \in C^4\}$ with respect to $\|m\|_{Sob}$. We are interested in the following optimization problem:

$$\min_m \frac{1}{n} [y - m(x)]^\top \Sigma^{-1} [y - m(x)] \qquad \text{s.t.} \quad \|m\|^2_{Sob} \le L \qquad (6)$$

which imposes a smoothness condition with smoothing parameter L. By varying this parameter, we control the smoothness of the ball of functions over which estimation is taking place.

Using techniques well known in the spline function literature, it can be shown that the infinite dimensional optimization problem (6) can be replaced by a finite dimensional optimization problem as we outline below, (see e.g., Wahba (1990) or Yatchew and Bos (1997)).

Since the evaluation of functions at a specific point is a linear operator, by the Riesz Representation Theorem, given a point x_o there exists a function r_{x_o} in $\mathcal{H}^4$ called a *representor* such that $\langle r_{x_o}, h \rangle_{Sob} = h(x_o)$ for any $h \in \mathcal{H}^4$. Let $r_{x_1}, \ldots, r_{x_n}$ be the representors of function evaluation at $x_1,\ldots,x_n$ respectively. (Details of the calculation of representors are contained in Appendix B.) Let R be the $n \times n$ representor matrix whose columns (and rows) equal the representors evaluated at $x_1,\ldots,x_n$; that is, $R_{ij} = \langle r_{x_i}, r_{x_j} \rangle_{Sob} = r_{x_i}(x_j) = r_{x_j}(x_i)$. If one solves:

$$\min_c \frac{1}{n} [y - Rc]^\top \Sigma^{-1} [y - Rc] \qquad \text{s.t.} \qquad c^\top Rc \le L \tag{7}$$

where c is an $n \times 1$ vector, then the minimum value is equal to that obtained by solving (6). Furthermore, there exists a solution to (7) of the form:

$$\hat{m}(x) = \sum_{i=1}^{n} \hat{c}_i r_{x_i}(x) \tag{8}$$

where $\hat{c} = (\hat{c}_1, \ldots, \hat{c}_n)^\top$ solves (7). First and second derivatives may be estimated by differentiating (8):

$$\hat{m}^{(1)}(x) = \sum_{i=1}^{n} \hat{c}_i r_{x_i}^{(1)}(x) \tag{9}$$

and

$$\hat{m}^{(2)}(x) = \sum_{i=1}^{n} \hat{c}_i r_{x_i}^{(2)}(x) \tag{10}$$

Define $R^{(1)}$ to be the $n{\times}n$ matrix whose columns (and rows) are the first derivatives of the representors evaluated at $x_1,\ldots,x_n$. Define $R^{(2)}$ in a similar fashion. Then the estimates of the call function and its first two derivatives at the vector of observed strike prices are given by $\hat{m}(x) = R\hat{c}$, $\hat{m}^{(1)}(x) = R^{(1)}\hat{c}$ and $\hat{m}^{(2)}(x) = R^{(2)}\hat{c}$ where $x = (x_1,\ldots,x_n)$.

Proposition 1: Suppose one is given data $(x_1, y_1),\ldots,(x_n, y_n)$ where $y_i = m(x_i) + \varepsilon_i$, the ε_i are independently distributed with $Var(\varepsilon_i) = \sigma^2(x_i) = \sigma_i^2$ and $0 < \sigma^2(x) < K$ for some K. The x_i are i.i.d. with continuous density $p(x)$ on $[a,b]$ bounded away from zero. Let $\hat{m}$ satisfy (6). Then $\sup_{x\in[a,b]} \left|\hat{m}^{(s)} - m^{(s)}\right| \xrightarrow{P} 0$ for $s = 0,1,2$ and $1/n \sum \left(m(x_i) - \hat{m}(x_i)\right)^2 = O_P(n^{-8/9})$. ■

The result establishes the consistency of the estimator and its first two derivatives. It also establishes the rate of convergence of the estimator which equals the optimal rate for four-times differentiable nonparametric functions of one variable (see Stone (1980, 1982)). The rate of convergence will be useful for implementing residual regression tests of hypotheses on m. One can replace the true variances in equation (6) with consistent estimates (or even with ones).

2.2 Imposition of Constraints

Optimization problem (7) allow easy incorporation of various restrictions. Suppose one wants to impose the constraint that $\hat{m}$ is monotone decreasing at each x_i. Then one restricts the first derivative (9) to be negative at these points. To impose convexity, one can require the second derivative (10) to be positive. Then the quadratic optimization problem (7) can be supplemented with the monotonicity constraints:

$$R^{(1)}c \quad \leq \quad 0 \tag{11}$$

and the convexity constraints:

$$R^{(2)}c \quad \geq \quad 0 \tag{12}$$

Suppose we solve (7) subject to monotonicity and convexity constraints (11) and (12). Then the conclusions of Proposition 1 continue to hold as long as the true function m is *strictly* monotone and convex (see Yatchew and Bos (1997)). This is because, as sample size grows, and the estimate of the function and its first two derivatives approach their true counterparts, the smoothness constraint alone will be sufficient to ensure that the estimated function will be monotone and convex. That is, the monotonicity and convexity constraints will become non-binding.

Next we turn to imposing unimodality. Suppose the current asset price lies in the interval $\left[x_{i^*}, x_{i^*+1}\right]$ and that one wants to impose the constraint that the state price density is unimodal with the mode in this same interval. Since the SPD is (proportional to) the second derivative $m^{(2)}$, one needs to impose constraints on *its* derivative, that is on $m^{(3)}$. Define $R^{(3)}$ to be the $n \times n$ matrix whose columns (and rows) are the third derivatives of the representors evaluated at $x_1,...,x_n$. Then one imposes the constraints:

$$\underset{i^*\times n}{R^{(3)}}\left[1:i^*,1:n\right]\underset{n\times 1}{c} \geq 0 \qquad \underset{(n-i^*)\times n}{R^{(3)}}\left[i^*+1:n,\ 1:n\right]\underset{n\times 1}{c} \leq 0 \tag{13}$$

The first set of i^* inequalities ensures that the SPD has a positive derivative at strike prices below the current asset price, the remaining $n-i^*$ inequalities provide for a negative derivative at higher strike prices.

Finally, since the integral under the density cannot exceed one, we have:

$$e^{r\tau}\sum_{i=1}^{n} c_i\left[r_{x_i}^{(1)}(b) - r_{x_i}^{(1)}(a)\right] \leq 1 \tag{14}$$

where the exponential term reflects the proportionality factor relating the second derivative to the SPD as in equation (2).

2.3 Testing Monotonicity and Convexity

Suppose one wants to test monotonicity and convexity. The residual regression test considered by Fan and Li (1996) and Zheng (1996) can be adapted to produce a test of these properties. The basic idea underlying the procedure is that one takes the residuals from the "restricted regression" which imposes the constraints to be tested, then performs a kernel regression of these residuals on the explanatory variable x to see whether there is anything left to be explained. If so, then the null hypothesis is rejected.

Proposition 2: Suppose m is strictly monotone and convex and $\hat{m}$ is the smooth constrained estimator obtained by solving (7) subject to monotonicity and convexity constraints (11) and (12). Let $\lambda n \to \infty$, $\lambda^{1/2} n^{1-8/9} \to 0$. Let K be a kernel function (such as the normal, uniform or triangular kernel), and define

$$U = \frac{1}{n} \sum_i \left(y_i - \hat{m}(x_i)\right) \left[\frac{1}{\lambda n} \sum_{j \neq i} \left(y_j - \hat{m}(x_j)\right) K\left(\frac{x_j - x_i}{\lambda} \right) \right] \tag{15a}$$

then

$$\lambda^{1/2} n\, U \xrightarrow{D} N\left(0\ ,\ 2 \int \sigma^4(x) p^2(x) \int K^2(u)\right) \tag{15b}$$

Let the estimated variance of U be given by:

$$\hat{\sigma}_U^2 = \frac{2}{\lambda^2 n^4} \sum_i \sum_{j \neq i} \left(y_i - \hat{m}(x_i)\right)^2 \left(y_j - \hat{m}(x_j)\right)^2 K^2\left(\frac{x_j - x_i}{\lambda} \right) \tag{15c}$$

Then, $\lambda n^2 \hat{\sigma}_U^2 \xrightarrow{p} 2 \int \sigma^4(x) p^2(x) \int K^2(u)$. Hence $U / \hat{\sigma}_U \xrightarrow{D} N(0, 1)$. ■

The test described in Proposition 2, may be performed using the indicated asymptotic normal approximation. Alternatively, it may be implemented using the bootstrap as we describe below. It is consistent against non-monotone or non-convex alternatives.

2.4 Multiple Observations

As we indicated in the introduction, option price data often consist of multiple observations at a finite vector of strike prices. We will need to modify our set-up to incorporate this characteristic. Let $X = \left(X_1, \ldots, X_k\right)$ be the vector of k distinct strike prices. (In Figure 1A, there are k=25 distinct strike prices with 20 observations at each price so that n=500.) We will assume that the vector X is in increasing order. Let $\sigma^2\left(X_1\right), \ldots, \sigma^2\left(X_k\right)$ be the residual variances at each of the distinct strike prices. Let B be the $n \times k$ matrix such that:

$$\begin{aligned} B_{ij} &= 1 \quad \textit{if } x_i = X_j \\ &= 0 \quad \textit{otherwise} \end{aligned} \tag{16}$$

We may now rewrite (6) as

$$\min_m \ \frac{1}{n} \left[\underset{n\times 1}{y} - \underset{n\times k}{B} \ \underset{k\times 1}{m(X)} \right]^\top \underset{n\times n}{\Sigma^{-1}} \left[\underset{n\times 1}{y} - \underset{n\times k}{B} \ \underset{k\times 1}{m(X)} \right] \quad \text{s.t.} \quad \| m \|^2_{Sob} \le L \tag{17}$$

Noting that the representor matrix R is in this case $k \times k$, the analogue to (7) becomes:

$$\min_c \ \frac{1}{n} \left[y - BRc \right]^\top \Sigma^{-1} \left[y - BRc \right] \qquad \text{s.t.} \qquad c^\top Rc \le L \tag{18}$$

where c is a $k \times 1$ vector. Monotonicity and convexity constraints (11) and (12) can be added noting that $R^{(1)}$ and $R^{(2)}$ are now also $k \times k$ matrices.

Even if the number of distinct strike prices k does not increase, the call function can be estimated consistently at $X_1, \ldots, X_k$. However, as was pointed out by a referee, this does not assure that estimates of derivatives are estimated consistently. Indeed, no "nonparametric" estimator can consistently estimate derivatives at a point without an accumulation of observations in the neighborhood of the point, though of course a sufficient condition for consistency of the first two derivatives is that the function $m(x)$ is a linear combination of the representors $r_{X_1}, \ldots, r_{X_k}$.

Proposition 3: Suppose one is given data $(x_1, y_1), \ldots, (x_n, y_n)$ where $y_i = m(x_i) + \varepsilon_i$, the ε_i are independently distributed with $Var(\varepsilon_i) = \sigma_i^2$ and x_i are sampled from a discrete distribution whose support is $X_1, .., X_k$ with corresponding probabilities $\pi_1, .., \pi_k$. Suppose m lies strictly inside the ball of functions $\|m\|_{Sob}^2 < L$ and m is strictly monotone decreasing and strictly convex and is a linear combination of the representors $r_{X_1}, \ldots, r_{X_k}$. Let $\bar{y}(X) = \bar{y}(X_1, .., X_k)$ be the k-dimensional vector of average transactions prices at the k strike prices. Let $\hat{c}$ minimize (18) with the added constraints (11) and (12) and define $\hat{m}(X) = \hat{m}(X_1, .., X_k) = R\,\hat{c}$, $\hat{m}^{(1)}(X) = R^{(1)}\hat{c}$ and $\hat{m}^{(2)}(X) = R^{(2)}\hat{c}$. Let $\Omega / n = Var(\bar{y}(X))$ be the $k \times k$ diagonal matrix of variances of the point mean estimators, i.e., $\Omega_{jj} = \sigma^2(X_j) / \pi_j$. Then $\lim_{n \to \infty} Prob[\hat{m}(X) = \bar{y}(X)] = 1$, $\hat{m}(X) \xrightarrow{p} m(X)$, $\hat{m}^{(1)}(X) \xrightarrow{p} m^{(1)}(X)$ and $\hat{m}^{(2)}(X) \xrightarrow{p} m^{(2)}(X)$. Furthermore,

$$n^{1/2}(\hat{m}(X) - m(X)) \xrightarrow{D} N(0, \Omega) \tag{19}$$

$$\begin{aligned}
n^{1/2}(\hat{c} - c) &\xrightarrow{D} N(0, R^{-1}\Omega R^{-1}) \\
n^{1/2}(\hat{m}^{(1)}(X) - m^{(1)}(X)) &\xrightarrow{D} N(0, R^{(1)}R^{-1}\,\Omega\,R^{-1}R^{(1)}) \\
n^{1/2}(\hat{m}^{(2)}(X) - m^{(2)}(X)) &\xrightarrow{D} N(0, R^{(2)}R^{-1}\,\Omega\,R^{-1}R^{(2)}) \quad \blacksquare
\end{aligned} \tag{20}$$

Proposition 3 states that as data accumulate at each strike price, the inequalities implied by the smoothness, monotonicity and convexity constraints eventually become non-binding, the estimator becomes identical to the point mean estimator and the call function *m* is estimated consistently at the observed strike prices. Moreover, because the true call function is here assumed to be a linear combination of the representors at the observed strike price, the first and second derivatives are also estimated consistently.

Proposition 3 provides for asymptotic scalar and vector confidence regions of the call function, its first derivative and the SPD. For example, if one is interested in confidence intervals at strike price X_j, the asymptotic pivots are:

$$\tau^{(0)} = n^{1/2} \frac{\left(\hat{m}\left(X_j\right) - m\left(X_j\right)\right)}{\left[\hat{\Omega}\right]_{jj}^{1/2}} \tag{21a}$$

$$\tau^{(1)} = n^{1/2} \frac{\left(\hat{m}^{(1)}\left(X_j\right) - m^{(1)}\left(X_j\right)\right)}{\left[R^{(1)} R^{-1} \hat{\Omega} R^{-1} R^{(1)}\right]_{jj}^{1/2}} \tag{21b}$$

$$\tau^{(2)} = n^{1/2} \frac{\left(\hat{m}^{(2)}\left(X_j\right) - m^{(2)}\left(X_j\right)\right)}{\left[R^{(2)} R^{-1} \hat{\Omega} R^{-1} R^{(2)}\right]_{jj}^{1/2}} \tag{21c}$$

We note that the $\sigma^2\left(X_j\right)$ and the π_j may be estimated using the sample variance and sample proportion of observations at each distinct strike price.

2.5 Bootstrap Procedures

Percentile and percentile-*t* procedures are commonly used for constructing confidence intervals. The latter are often found to be more accurate when the statistic is an asymptotic pivot (see Hall (1992) for extensive arguments in support of this proposition). On the other hand, percentile methods might be better when the asymptotic approximation to the distribution of the pivot is poor as a result of small sample size or slow convergence.

Table 1 contains an algorithm for constructing percentile confidence intervals for the call function and its first two derivatives at X_j. As there are multiple observations at each strike price, we can accommodate heteroskedasticity by resampling from the estimated residuals at each strike price or we can use the wild bootstrap (see Wu (1986) or Härdle (1990, p.106-8, 247)). The procedures are applicable with the obvious modifications for a general confidence level α. Algorithms for constructing percentile-*t* confidence intervals may be constructed with modest additional effort. Table 2 summarizes the bootstrap algorithm for implementing the residual regression test in Proposition 2.

Table 1: Percentile Confidence Intervals for $\boldsymbol{m(X_j)}$, $\boldsymbol{m^{(1)}(X_j)}$and $\boldsymbol{m^{(2)}(X_j)}$
1. Calculate $\hat{c}$ and $\hat{m}$ by solving (18) subject to (11) and (12). Calculate the estimated residuals $\{\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n\}$. 2.a) Construct a bootstrap data-set $(x_i,y_1^B),\ldots,(x_n,y_n^B)$ where $y_i^B = \hat{m}(x_i) + \varepsilon_i^B$ and ε_i^B is obtained by sampling from $\{\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n\}$ using the wild bootstrap. b) Using the bootstrap data-set obtain $\hat{c}^B$ by solving (18) subject to (11) and (12). Calculate and save $\hat{m}^B(X_j)$, $\hat{m}^{(1)^B}(X_j)$ and $\hat{m}^{(2)^B}(X_j)$. 3. Repeat steps 2 multiple times. 4. To obtain a 95% point-wise confidence intervals for $m(X_j)$, $m^{(1)}(X_j)$ and $m^{(2)}(X_j)$ obtain .025 and .975 quantiles of the corresponding bootstrap estimates.

Table 2: Bootstrap Residual Regression Test of Monotonicity and Convexity
1. Calculate $\hat{c}$ and $\hat{m}$ by solving (18) subject to (11) and (12). Save the estimates of the regression function $\hat{m}(x_1),....,\hat{m}(x_n)$ and the residuals $\hat{\varepsilon}_1,....,\hat{\varepsilon}_n$. 2. Calculate U, $\hat{\sigma}_U$, and $U/\hat{\sigma}_U$ as in Proposition 2. 3. Sample using the wild bootstrap from $\hat{\varepsilon}_1,....,\hat{\varepsilon}_n$ to obtain $\hat{\varepsilon}_1^B,....,\hat{\varepsilon}_n^B$ and construct a bootstrap data set $(y_1^B,x_1),....,(y_n^B,x_n)$, where $y_i^B = \hat{m}(x_i) + \hat{\varepsilon}_i^B$. (b) Using the bootstrap data set, estimate the model under the null and calculate $\hat{\sigma}_U^B$, U^B, and $U^B/\hat{\sigma}_U{}^B$. (c) Repeat Steps (a) and (b) multiple times, each time saving the value of the standardized test statistic $U^B/\hat{\sigma}_U{}^B$. Define the bootstrap critical value for a 5 percent significance level test to be the 95th percentile of the $U^B/\hat{\sigma}_U{}^B$. 4. Compare $U/\hat{\sigma}_U$, the actual value of the statistic, with the bootstrap critical value.

3. Numerical Results

3.1 Simulations

In order to solve the various constrained optimization problems described in this paper, we used GAMS – General Algebraic Modeling System (see Brooke, Kendrick, and Meeraus 1992) which is a general package for solving a broad range of linear, nonlinear, integer, and other optimization problems subject to constraints.

In their simulations, Aït-Sahalia and Duarte (2000) calibrate their model using characteristics of the S&P options market. We calibrate our experiments using DAX options in January 1999 which expire in February of that year. At that time the DAX index was in the vicinity of 5000 (see Table 4 below). The 25 distinct strike prices range over the interval 4400 to 5600 in 50 unit increments. We set the short term interest rate r to 3.5% , the dividend yield δ to 2%, the time to maturity τ to .15 and the current price (value) of the index S to 5100. We assume the volatility smile function is linear in the strike price, i.e., $\sigma = .4 - .00025 * (x - 4400)$. Let $F = S\exp((r - \delta)\tau)$ be the forward price. Define $d_1 = \left(\log(F/x) + \sigma^2\tau/2\right)/\sigma\sqrt{\tau}$ and $d_2 = d_1 - \sigma\sqrt{\tau}$. Then the "true" call function is given by $m(x) = \exp(-r\tau)\left(F\,\Phi(d_1) - x\Phi(d_2)\right)$ where Φ is the standard normal cumulative distribution function. At each strike price, the residual standard deviation is set to 10% of the option price. For a given observation with strike price x_i, the "observed" option price is given by $y_i = m(x_i) + .1\,m(x_i)\varepsilon_i$ where the ε_i are i.i.d. standard normal.

We have already seen the 'true' call function, its first derivative and SPD plotted in Figures 1A,B, and C above. In each of the simulations below we assume there are 20 observations at each of the 25 strike prices for a total of 500 observations.

Figure 2A, B and C illustrate the impact of constraints on estimation. The 'unconstrained' estimator consists of the point means at each strike price. The 'smooth' estimator imposes only the Sobolev constraint as in equations (6) and (7) with the degree of smoothness identical to the true Sobolev norm of the underlying function which is the square root of 1.812. Monotonicity and convexity constraints are imposed using equations (11) and (12). Finally, we impose 'unimodality' constraints (13) which require the estimated SPD to be non-decreasing over the lowest five strike prices and non-increasing over the highest five. The purpose is to improve the estimator of the SPD at the boundaries. In each case the "90% point-wise intervals" contain the central 90% of estimates from 1000 replications. The "90% uniform intervals" are obtained by taking the central 99.6% of the estimates at each of the 25 distinct strike prices ($.996^{25} \cong .90$).

As may be seen in Figure 2A, the improvement in estimation of the call function resulting from adding constraints is barely discernible. Figure 2B illustrates the impacts on estimation of the first derivative. The benefits of adding shape constraints are clearly evident (note that the vertical scale narrows as one moves down the figure). The most dramatic impact of the constraints is on estimation of the SPD as may be seen in Figure 2C. Smoothness alone produces a very broad band of estimates, so much so that the true SPD looks quite flat. Adding monotonicity and convexity improves the estimates substantially, though they are quite imprecise at low strike prices. This is in part due to the much larger variance there. The 'unimodality' constraints alleviate this problem.

Table 3 summarizes the impact of imposing constraints on the MSE of the estimators of the call function, its first derivative and the SPD. The "unconstrained" estimator is obtained by taking point means, their first-order and second-order divided differences. Consistent with Figure 2, the average MSE of estimating the SPD falls dramatically – indeed by three orders of magnitude -- when smoothness, monotonicity, convexity and unimodality are imposed. Even supplementing the smoothness constraint with monotonicity, convexity and unimodality reduces the MSE of the SPD estimator by an order of magnitude.

Table 3: Average MSE

Model	Call Function	First Derivative	SPD
Unconstrained	200.02	.1219	1.49×10^{-4}
Smooth	25.68	.0052	1.26×10^{-6}
Smooth, Monotone, Convex	13.70	.0015	3.76×10^{-7}
Smooth, Monotone, Convex, Unimodal	13.26	.00097	2.00×10^{-7}

Next we turn to confidence interval estimation. With multiple observations at each strike price and under the assumptions of Proposition 3, asymptotic confidence intervals can be estimated using (21a,b,c). For bootstrap intervals we use the percentile method outlined in Table 1. In each case we performed 100 bootstrap draws. Figure 3 contrasts asymptotic and bootstrap confidence intervals for our data generating mechanism. (Recall there are 25 distinct strike prices with 20 observations at each one.) In the top panel which corresponds to the call function, the asymptotic confidence intervals (dashed lines) are somewhat broader than the bootstrap intervals (dotted lines). The middle and lower panels correspond to the first derivative and SPD estimates. In these two cases, confidence intervals based on the asymptotic approximation are extremely poor relative to the bootstrap intervals.

Insert Figure 3

Next we turn to the accuracy of the bootstrap intervals. We produced 500 samples and in each case obtained 100 bootstrap re-samples. The model which was estimated was the fully constrained version with smoothness, monotonicity, convexity and unimodality constraints. The smoothness constraint was set at the true norm. (Increasing the bound by 5% or 10% did not materially alter the conclusions below.) For the call function, the averages of the observed coverage frequencies across strike prices were .9795 for 99% nominal confidence intervals, .94 for 95%, .899 for 90% and .5185 for 50% intervals. For the first derivative, the averages of the observed coverage frequencies across strike prices were .9835 for 99% confidence intervals, .9545 for 95%, .9095 for 90% and .4935 for 50%. For the second derivative, the averages were .994 for 99% confidence intervals, .978 for 95%, .953 for 90% and .575 for 50%. Thus, percentile bootstrap confidence intervals for the SPD in the fully constrained model were, at least in these simulations, conservative.

Overall we found that while MSE improves and bootstrap confidence intervals narrow as one adds constraints to the 'smooth' model, bootstrap coverage accuracy deteriorates moderately.

In addition, we performed simulations in which the Sobolev smoothness parameter was estimated as the minimum of the cross-validation function. In this setting, the cross-validation function may be defined as:

$$CV(L) = \frac{1}{n} \sum_{i=1}^{n} \left[y_i - \hat{m}_{-i}(x_i)\right]^2 / \sigma^2\left(x_i\right) \tag{22}$$

where $\hat{m}_{-i}$ is obtained by solving (6) or (18) while omitting the i-th observation. We found that even with much smaller samples, e.g., with two observations at each of the 25 strike prices, the estimated cross-validation parameter was close to the true value.

In previous work, Li and Wang (1998), considered asymptotic and bootstrap versions of a residual regression test of a parametric null against a nonparametric alternative. They found that the actual mean of the asymptotically N(0,1) test statistic to be below zero and that the distribution was better approximated by the bootstrap. (For details, see Li and Wang (1998, page 155, Tables 1-2).) We performed simulations of the residual regression test of monotonicity and convexity in Proposition 2, implementing both the asymptotic and (wild) bootstrap versions of the procedure (see Table 2). We too found that the mean of the test statistic was below zero. For example, with 25 distinct strike prices and the data generating mechanism described at the beginning of this section, the mean and variance of the test statistic $U / \hat{\sigma}_U$ of Proposition 2 were -.27 and .97 respectively.

3.2 Applications to DAX Index Option Data

In this section we use the tools we have described to analyze data on DAX index options over the two week period January 4-15, 1999. Table 4 provides the closing daily values of the DAX over the period. During the first week, the index fluctuates in a range above 5200. In the early part of the second week it begins to decline and during the last three days of the sample period, the index remains below 5000.

Table 4: DAX Index, January 4-15, 1999

Mon Jan 4 5290	Tues Jan 5 5263	Wed Jan 6 5443	Thurs Jan 7 5346	Fri Jan 8 5371
Mon Jan 11 5267	Tues Jan 12 5196	Wed Jan 13 4982	Thurs Jan 14 4903	Fri Jan 15 4974

We restricted our estimates to call options which trade at 100 point intervals between 4500 and 5500 and expire at the end of February. Some trades do indeed take place outside this range of strike prices, but there are few of them. In each case, the smoothness parameter was selected using cross-validation. (For example, for January 4, the minimum of the cross-validation function was at L=1.3.)

We begin with data for January 4. The upper panel of Figure 4 illustrates the estimated SPD along with 90% point-wise bootstrap confidence intervals. In addition we have included uniform confidence bounds for the SPD by taking the central 99% of the estimates at each of the 11 distinct strike prices ($.99^{11} \cong .90$). With the DAX closing at 5290, one can expect that there is some probability mass beyond the 5500 level. Options at 5500 averaged 22Euro suggesting that the market assigned a positive probability to the DAX exceeding 5500 at time of expiration of the options. Indeed our estimated SPD integrates to about .8.

The lower panel of Figure 4 depicts the estimated SPD for January 14. By this time, the DAX had dropped to about 4900. The SPD is zero at both end points of our estimation range and constraint (14) which requires that the integral of the SPD not exceed one is binding and hence informative to the estimation process.

Insert Figure 4

Figure 5 plots the procession of daily estimated SPD's over the full period January 4-18, 1999, (January 9^{th} and 10^{th} were week-end days). During the first week market expectations as measured by the median of the SPD were that the DAX would be substantially over 5000 at time of expiry of the options. By the end of the second week, expectations had moved sharply lower, consistent with the DAX index value, and the SPDs become more highly concentrated, possibly a result of the fact that the expiry date was drawing closer.

Insert Figure 5

4. Conclusions

In this paper, we propose a nonparametric least squares estimator for modeling option prices and state price densities as a function of the strike price. The estimator readily incorporates various constraints such as monotonicity and convexity. We outline a "residual regression" type test of these properties. We also propose bootstrap procedures for constructing confidence intervals around the estimated call function and its derivatives.

Our methods can be readily extended in at least two ways. First, time series structure can be introduced into the residuals without complicating the optimization problem in equations (6) and (7). Second, given appropriate data, the call function can be estimated as a function of additional explanatory variables, such as the current asset price and the time to expiry. The nonparametric least squares problem in (7) remains as given, but the representor matrix R must now be calculated using representors which are functions of several variables. This can be readily accommodated using a generalization of the Sobolev inner product in (4). The resulting multivariate representors are simply products of the univariate representors used in this paper. The details of these kinds of multivariate nonparametric regression procedures may be found in Yatchew and Bos (1997).

References

Ait-Sahalia, Y. (1996): "Nonparametric Pricing of Interest Rate Derivative Securities", *Econometrica*, 64, 527-560.

Ait-Sahalia, Y., P. Bickel, P. T. Stoker (2001): "Goodness-of-Fit Tests for Kernel Regression with an Application to Option Implied Volatilities ", *Journal of Econometrics*, 105, 363-412.

Aït-Sahalia, Y. and A. Lo (1998): "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices", *Journal of Finance*, 53, 499-547.

Aït-Sahalia, Y. and A. Lo (2000): "Nonparametric Risk Management and Implied Risk Aversion", *Journal of Econometrics*, 94, 9-51.

Aït-Sahalia, Y. and J. Duarte (2000): "Nonparametric Option Pricing Under Shape Restrictions", manuscript, Princeton University.

Barle S. and N. Cakici (1998) "How to grow a smiling tree", *Journal Financial Engineering*, 7, 127 - 146,

Black, F. and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 637-659.

Bowman, A.W., M.C. Jones, and I. Gijbels (1998), "Testing Monotonicity of Regression," *Journal of Computational and Graphical Stati*stics, **7**, 489-500.

Breeden, D. and R. Litzenberger (1978): "Prices of State-Contingent Claims Implicit in Option Prices", *Journal of Business*, 51, 621-651.

Broadie, M. J. Detemple, E. Ghysels, and O. Torrés (2000): "Nonparametric Estimation of American Options' Exercise Boundaries and Call Prices", *Journal of Economic Dynamics and Control*, 24, 1829-57.

Broadie, M. J. Detemple, E. Ghysels, and O. Torrés (2000): "American Options With Stochastic Dividends and Volatility: A Nonparametric Investigation", *Journal of Econometrics*, 94, 53-92.

Brooke, A., D. Kendrick, and A. Meeraus (1992), *GAMS*, Redwood City, CA: Scientific Press.

Cont, R. (2001): "Empirical Properties of Asset Returns, Stylized Facts and Statistical Issues", *Quantitative Finance*, 1, 223-236.

Cont, R. and J. da Fonseca (2002): "Dynamics of Implied Volatility Surfaces", *Quantitative Finance*, 2, 45-60.

Cont, R. and P. Tankov (2002): " Calibration of Jump-Diffusion Option-Pricing Models: A Robust Non-parametric Approach", Centre de Mathematiques Appliquees, CNRS - Ecole Polytechnique, Palaiseau, France.

Cox, J. and S. Ross (1976): "The Valuation of Options for Alternative Stochastic Processes", *Journal of Financial Economics*, 3, 145-166.

Daglish, T. (2002): "SPI Options: A Thorough Comparison of Non-Parametric and Parametric Option Pricing", manuscript, Rotman School of Business, University of Toronto.

Derman, E. and I. Kani (1994): "Riding on the Smile", *Risk*, 7, 32-39.

Diack, C. (2000), "Sur la Convergence des Tests de Schlee et de Yatchew," *Canadian Journal of Statistics*, **28**, 653-668.

Diack, C. and C. Thomas-Agnan (1998), "A Nonparametric Test of the Non-Convexity of Regression," *Nonparametric Statistics*, **9**, 335-362.

Fan Y. and Q. Li (1996), "Consistent Model Specification Tests: Omitted Variables and Semiparametric Functional Forms," *Econometrica*, **64**, 865-890.

Friedman, J. and R. Tibshirani (1984), "The Monotone Smoothing of Scatterplots," *Technometrics*, **26**, 243-250.

Garcia, R. and R. Gencay (2000): "Pricing and Hedging Derivative Securities with Neural Networks and a Homogeneity Hint", *Journal of Econometrics*, 94, 93-115.

Gijbels, I., P. Hall, M.C. Jones, and I. Koch (2000), "Tests for Monotonicity of a Regression Mean With Guaranteed Level," *Biometrika*, **87**, 663-673.

Goldman, S. and P. Ruud (1992), "Nonparametric Multivariate Regression Subject to Monotonicity and Convexity Constraints," manuscript, University of California, Berkeley.

Groeneboom, P., G. Jongbloed, and J.A. Wellner (2001), “Estimation of a Convex Function: Characterizations and Asymptotic Theory,” *Annals of Statistics*, **29**, 1653-1698.

Ghysels, E., V. Patilea, E. Renault and O. Torres (1997): “Nonparametric Methods and Option Pricing”, Core Discussion Paper, 9775.

Hall, Peter, (1992): *The Bootstrap and Edgeworth Expansion*, Springer Verlag, New York

Hall, P. and N. Heckman (2000), “Testing for Monotonicity of a Regression Mean by Calibrating for Linear Functions,” *Annals of Statistics*, **28**, 20-39.

Hall, P. and L. Huang (2001), “Nonparametric Kernel Regression Subject to Monotonicity Constraints,” *Annals of Statistics*, **29**, 624-647.

Hall, P. and A. Yatchew (2002), “Unified Approach to Testing Functional Hypotheses in Semiparametric Contexts,” unpublished manuscript, Australian National University, School of Mathematical Sciences.

Härdle W. (1990): *Applied Nonparametric Regression*, Econometric Society Monograph Series, 19, Cambridge University Press.

Härdle, W., T. Kleinow and G. Stahl (2002), (eds.), Applied Quantitative Finance: Theory and Computational Tools, Springer, Berlin.

Jackwerth, J. And M. Rubinstein (1996): “Recovering Probability Distributions from Option Prices ”, *Journal of Finance*, 51, 1611-32.

Juditsky, A. and A. Nemirovski (2002), “On Nonparametric Tests of Positivity/ Monotonicity/ Convexity,” *Annals of Statistics*, **30**, 498-527.

Kelly C. and J. Rice (1990), “Monotone Smoothing with Application to Dose-Response Curves and the Assessment of Synergism,” *Biometrics*, **46**, 1071-1085.

Li, Q. and S. Wang. (1998): “A Simple Consistent Bootstrap Test for a Parametric Regression Function”, *Journal of Econometrics*, 87,145-165.

Lucas, R. (1978): “Asset Prices in an Exchange Economy”, *Econometrica*, 46, 1429-1446.

Mammen, E. (1991): "Estimating a Smooth Monotone Regression Function", *Annals of Statistics*, 19, 724-740.

Mammen, E. and C. Thomas-Agnan (1999), “Smoothing Splines and Shape Restrictions,” *Scandinavian Journal of Statistics*, **26**, 239-252.

Merton, R. (1973): “Rational Theory of Option Pricing”, *Bell Journal of Economics and Management Science*, 4, 141-183.

Mukarjee, Hari (1988): "Monotone Nonparametric Regression", *Annals of Statistics*, 16, 741-750.

Mukarjee, H. and S. Stern (1994), “Feasible Nonparametric Estimation of Multi-argument Monotone Functions,” *Journal of the American Statistical Association*, **89**, 77-80.

Ramsay, J. (1988), “Monotone Regression Splines in Action,” *Statistical Science*, **3**, 425-461.

Ramsay, J. O. (1998): “Estimating Smooth Monotone Functions”, *Journal of the Royal Statistical Society, Series B*, 60, 365-375.

Robertson, T., F.T. Wright and R.L. Dykstra (1988): *Order Restricted Statistical Inference*, New York: John Wiley and Sons.

Rubinstein, M. (1976): “The Valuation of Uncertain Income Streams and the Pricing of Options”, *Bell Journal of Economics*, 7, 407-425.

Schlee, W. (1982), “Nonparametric Tests of the Monotony and Convexity of Regression,” in. B.V. Gnedenko, M.L. Puri, and I. Vincze (eds.), *Nonparametric Statistical Inference*, **2**, 823-836, Amsterdam: North Holland.

Stone, C. (1980), “Optimal Rates of Convergence for Nonparametric Estimators,” *Annals of Statistics*, **8**, 1348-1360.

Stone, C. (1982), “Optimal Global Rates of Convergence for Nonparametric Regression,” *Annals of Statistics*, **10**, 1040-1053.

Villalobos, M. and G. Wahba (1987), “Inequality-Constrained Multivariate Smoothing Splines With Application to the Estimation of Posterior Probabilities,” *Journal of the American Statistical Association*, **82**, 239-248.

Wright, I. and E. Wegman (1980): "Isotonic, Convex and Related Splines", *Annals of Statistics*", 8, 1023-1035.

Wu, Chien-Fu Jeff (1986): "Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis" (with discussion), *Annals of Statistics*, 14, 1261-1350.

Yatchew, A. (1992), "Nonparametric Regression Model Tests Based on Least Squares," *Econometric Theory*, **8**, 435-451.

Yatchew, A.J. and L. Bos (1997) "Nonparametric Regression and Testing in Economic Models", *Journal of Quantitative Economics*, Vol. 13, 81-131.

Zheng, J. (1996), "A Consistent Test of Functional Form Via Nonparametric Estimation Techniques," *Journal of Econometrics*, **75**, 263-289.

Appendix A: Proof of Propositions

Proof of Proposition 1: By Yatchew and Bos (1997), Theorem 3.1.2. $\hat{m}$ converges to m in mean squared error at the indicated rate of convergence. By Theorem 2.3 in Yatchew and Bos (1997), all functions in the estimating set have derivatives up to order 3 uniformly bounded in sup-norm. Combining these two results ensures that $\hat{m}$, $\hat{m}^{(1)}$ and $\hat{m}^{(2)}$convergence in sup-norm. ■

Proof of Proposition 2: See Yatchew and Bos (1997) Theorem 5.1.1. ■

Proof of Proposition 3: first note that $R^{-1}\bar{y}(X)$ solves the unconstrained version of optimization problem (18):

$$\min_c \frac{1}{n}\left[y - BRc\right]^{\top} \Sigma^{-1}\left[y - BRc\right].$$

Now let $\hat{c}$ minimize (18) subject to (11) and (12). Since $m(x)$ is a linear combination of the representors at $X = \left(X_1,...,X_k\right)$, $\hat{m}(x) = \sum_{j=1}^{k}\hat{c}_j r_{x_j}(x)$ and its first two derivatives consistently estimate their true counterparts. Hence, in addition, $\hat{m}(X)^{\top} R^{-1}\hat{m}(X) \overset{P}{\to} m(X)^{\top} R^{-1} m(X) < L$ and as sample size increases, the smoothness constraint becomes non-binding in probability. Furthermore, because the first and second derivatives are estimated consistently, strict monotonicity and convexity implies that these constraints also become non-binding in probability. Thus $\lim_{n\to\infty} Prob\left[\hat{m}(X) = \bar{y}(X)\right] = 1$.

Using conventional central limit theorems, $n^{1/2}(\bar{y}(X) - m(X)) \overset{D}{\to} N(0,\Omega)$ and equation (20) follows immediately. ■

Appendix B: Calculation of Representors

Our objective is to show how to construct representors. Let $\mathcal{H}^d$ be a Sobolev space of functions from [0,1] to $\mathbb{R}^1$ with inner product $<f,g>_{Sob} = \int_0^1 \sum_{i=0}^{d} f^{(i)}(x)\, g^{(i)}(x) dx$. We choose the unit interval to simplify exposition, though the arguments apply directly to arbitrary finite intervals. We also note that in the main part of the paper, $d = 4$. The resulting representors of function evaluation consist of two functions spliced together, each of which is a linear combination of trigonometric functions. The formulae are derived using elementary methods, in particular integration by parts and the solution of a linear differential equation. For additional details, see Yatchew and Bos (1997).

We construct a function $r_a \in \mathcal{H}^d[0,1]$ such that $<f,r_a>_{Sob} = f(a)$ for all $f \in \mathcal{H}^d[0,1]$. This representor of function evaluation r_a will be of the form:

$$r_a(x) = \begin{cases} L_a(x) & 0 \le x \le a \\ R_a(x) & a \le x \le 1 \end{cases}$$

where L_a and R_a are both analytic functions. For r_a of this form to be an element of $\mathcal{H}^d[0,1]$, it suffices that $L_a^{(k)}(a) = R_a^{(k)}(a)$, $0 \le k \le d-1$. Now write:

$$f(a) = <r_a,f>_{Sob} = \int_0^a \sum_{k=0}^{d} L_a^{(k)}(x) f^{(k)}(x) dx + \int_a^1 \sum_{k=0}^{d} R_a^{(k)}(x) f^{(k)}(x) dx$$

We ask that this be true for all $f \in \mathcal{H}^d[0,1]$ but by density it suffices to demonstrate the result for all $f \in C^\infty[0,1]$, the set of infinitely differentiable functions on the unit interval. Hence assume that $f \in C^\infty[0,1]$. Thus, integrating by parts, we have:

$$\sum_{k=0}^{d} \int_0^a L_a^{(k)}(x)\, f^{(k)}(x)dx \;=\; \sum_{k=0}^{d} \left\{ \sum_{j=0}^{k-1} (-1)^j \, L_a^{(k+j)}(x)\, f^{(k-j-1)}(x) \Bigg|_0^a \right.$$

$$\left. + \; (-1)^k \int_0^a L_a^{(2k)}(x)\, f(x)dx \right\}$$

$$= \sum_{k=0}^{d} \sum_{j=0}^{k-1} (-1)^j \, L_a^{(k+j)}(x)\, f^{(k-j-1)}(x) \Bigg|_0^a \; + \; \int_0^a \left\{ \sum_{k=0}^{d} (-1)^k \, L_a^{(2k)}(x) \right\} f(x)dx$$

letting $i = k - j - 1$ in the first sum, this may be written as

$$\sum_{k=0}^{d} \int_0^a L_a^{(k)}(x)\, f^{(k)}(x)dx \;=\; \sum_{k=1}^{d} \sum_{i=0}^{k-1} (-1)^{k-i-1} \, L_a^{(2k-1-i)}(x)\, f^{(i)}(x) \Bigg|_0^a$$

$$+ \int_0^a \left\{ \sum_{k=0}^{d} (-1)^k \, L_a^{(2k)}(x) \right\} f(x)dx$$

$$= \sum_{i=0}^{d-1} \sum_{k=i+1}^{d} (-1)^{k-i-1} \, L_a^{(2k-1-i)}(x)\, f^{(i)}(x) \Bigg|_0^a \; + \; \int_0^a \left\{ \sum_{k=0}^{d} (-1)^k \, L_a^{(2k)}(x) \right\} f(x)dx$$

$$= \sum_{i=0}^{d-1} f^{(i)}(a) \left\{ \sum_{k=i+1}^{d} (-1)^{k-i-1} L_a^{(2k-1-i)}(a) \right\} - \sum_{i=0}^{d-1} f^{(i)}(0) \left\{ \sum_{k=i+1}^{d} (-1)^{k-i-1} L_a^{(2k-1-i)}\,(0) \right\}$$

$$+ \int_0^a \left\{ \sum_{k=0}^{d} (-1)^k \, L_a^{(2k)}(x) \right\} f(x)dx$$

Similarly, $\displaystyle \int_a^1 \sum_{k=0}^{d} R_a^{(k)}(x)\, f^{(k)}(x)dx$ may be written as

$$- \sum_{i=0}^{d-1} f^{(i)}(a) \left\{ \sum_{k=i+1}^{d} (-1)^{k-i-1} R_a^{(2k-1-i)}(a) \right\} + \sum_{i=0}^{d-1} f^{(i)}(1) \left\{ \sum_{k=i+1}^{d} (-1)^{k-i-1} R_a^{(2k-1-i)}(1) \right\}$$

$$+ \int_a^1 \left\{ \sum_{k=0}^{d} (-1)^k R_a^{(2k)}(x) \right\} f(x)dx .$$

Thus we require both L_a and R_a to be solutions of the constant coefficient differential equation

$$\sum_{k=0}^{m} (-1)^k u^{(2k)}(x) = 0 .$$

Boundary conditions are obtained by setting the coefficients of $f^{(i)}(a), 1 \le i \le d-1,$ $f^{(i)}(0),$ $0 \le i \le d-1$ and $f^{(i)}(1),$ $0 \le i \le d-1$ to zero and the coefficient of $f(a)$ to 1. That is,

$$\begin{aligned}
&\sum_{k=i+1}^{d} (-1)^{k-i-1} \left\{ L_a^{(2k-1-i)}(a) - R_a^{(2k-1-i)}(a) \right\} = 0 && 1 \le i \le d-1 \\
&\sum_{k=i+1}^{d} (-1)^{k-i-1} L_a^{(2k-1-i)}(0) = 0 && 0 \le i \le d-1 \\
&\sum_{k=i+1}^{d} (-1)^{k-i-1} R_a^{(2k-1-i)}(1) = 0 && 0 \le i \le d-1 \\
&\sum_{k=1}^{d} (-1)^{k-1} \left\{ L_a^{(2k-1)}(a) - R_a^{(2k-1)}(a) \right\} = 1 &&
\end{aligned}$$

Furthermore, for $\mathbf{r}_a \in \mathcal{H}^d[0,1]$, we require, $L_a^{(k)}(a) = R_a^{(k)}(a),\ 0 \le k \le d-1$. This results in $(d-1)+d+d+1+d = 4d$ boundary conditions. The general solution of the above differential equation is obtained by finding the roots of its characteristic polynomial $P(\lambda) = \sum_{k=0}^{d} (-1)^k \lambda^{2k}$. This is easily done by noting that $(1+\lambda^2)\ P(\lambda) = 1 + (-1)^d \lambda^{2d+2}$ and thus the characteristic roots are given by $\lambda_k = e^{i\theta_k}$, $\lambda_k \ne \pm i$, where

$$\theta_k = \begin{cases} \dfrac{(2k+1)\pi}{2d+2} & d \text{ even} \\[2ex] \dfrac{2k\pi}{2d+2} & d \text{ odd} \end{cases}$$

The general solution is given by the linear combination $\sum_k a_k \, e^{(Re(\lambda_k))x} \sin\left(Im(\lambda_k)\right)x$ where the sum is taken over *2d* linearly independent real solutions of the differential equation above.

Let $L_a(x) = \sum_{k=1}^{2d} a_k u_k(x)$ and $R_a(x) = \sum_{k=1}^{2d} b_k u_k(x)$ where the u_k, $1 \le k \le 2d$ are *2d* basis functions of the solution space of the differential equation. To show that r_a exists and is unique, we need only show that the boundary conditions uniquely determine the a_k and b_k. Since we have *4k* unknowns *(2d* a_k*'s and 2d* b_k*'s)* and *4d* boundary conditions, the boundary conditions constitute a square $4d \times 4d$ linear system in the a_k*'s* and b_k*'s*. Thus it suffices to show that the only solution of the associated homogenous system is the zero vector. Now suppose that $L_a^h(x)$ and $R_a^h(x)$ are the functions corresponding to the solutions of the homogeneous system (i.e. with the coefficient of *f(a)* in the boundary conditions set to *0* instead of *1*). Then, by exactly the same integration by parts, it follows that $< r_a^h, f >_{Sob} = 0$ for all $f \in C^\infty[0,1]$. Hence r_a^h, $L_a^h(x)$ and $R_a^h(x)$ are all identically zero and thus by the linear independence of the $u_k(x)$, so are the a_a and b_k.

Figure 1A: Data and Estimated Call Function

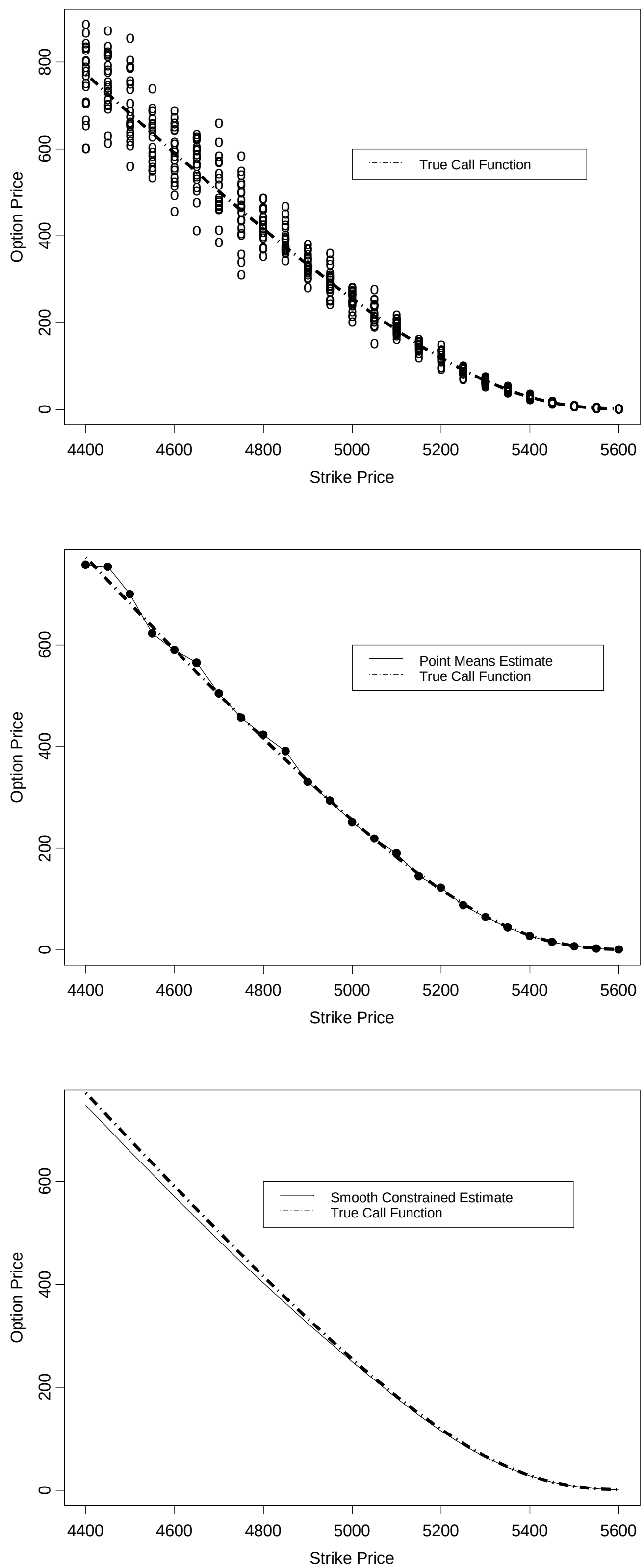
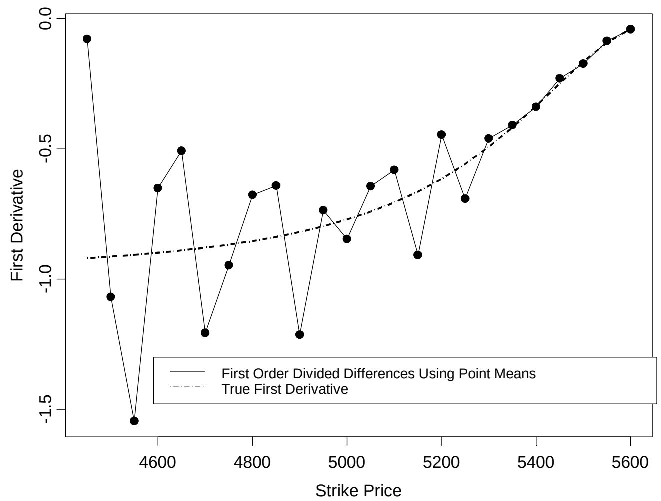


Figure 1B: Estimated First Derivative



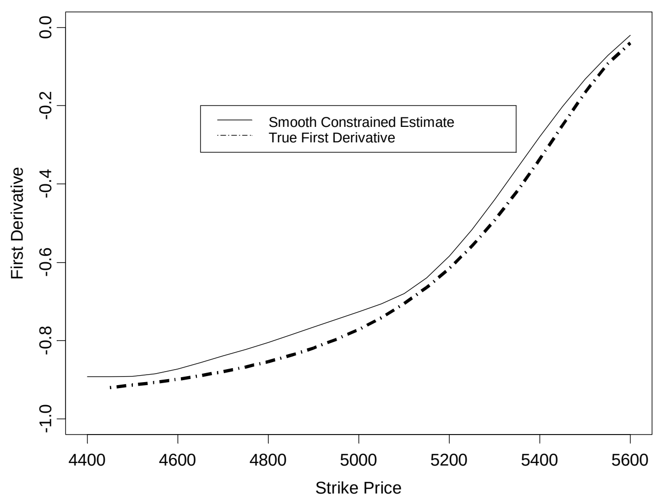
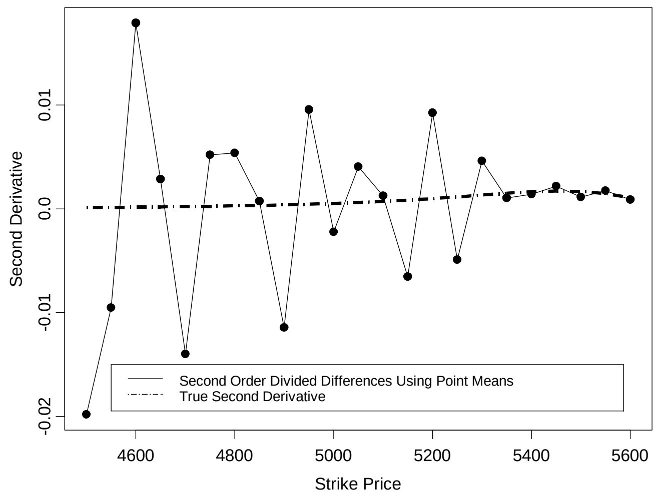


Figure 1C: Estimated SPDs



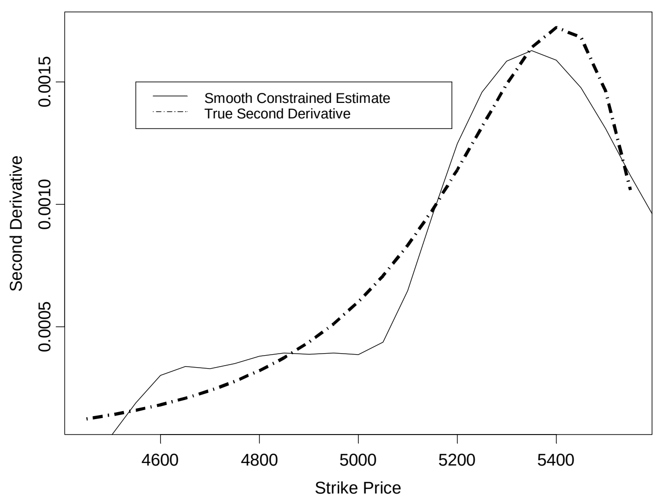


Figure 2A: Effects of Constraints on Call Function Estimates

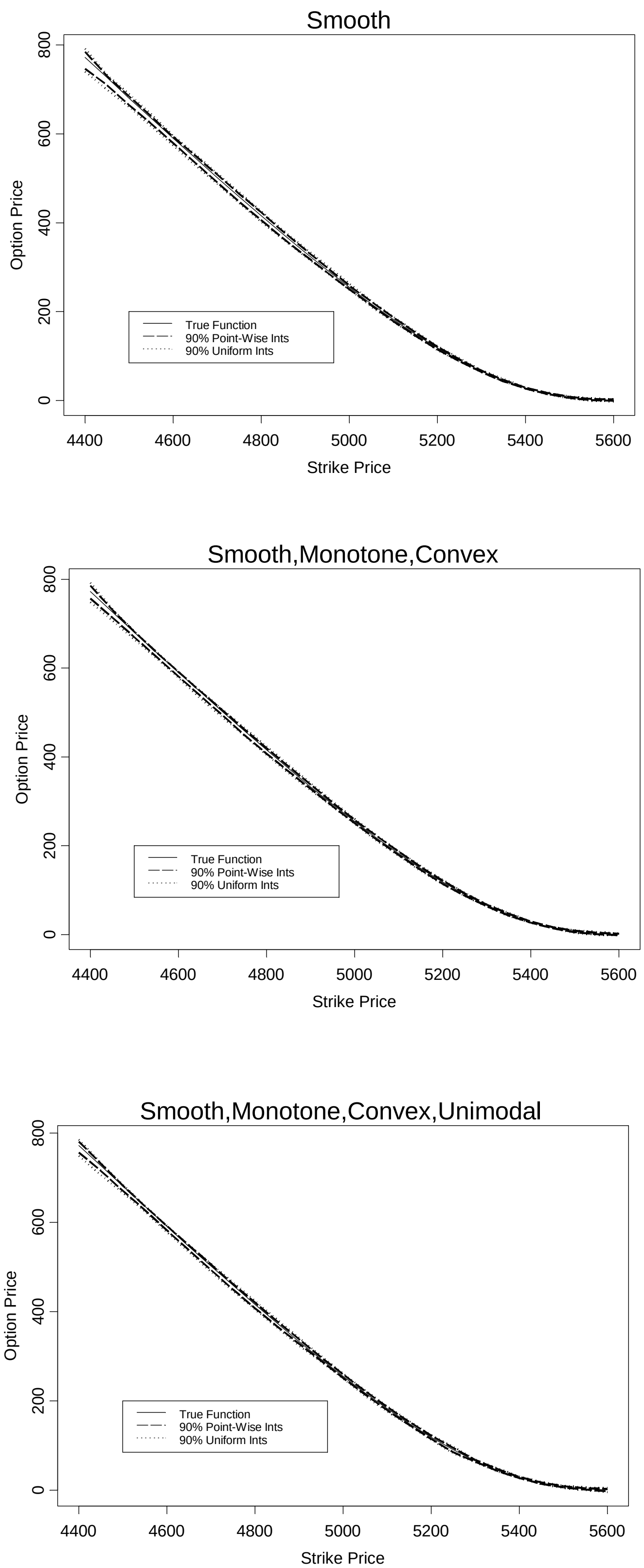


Figure 2B: Effects of Constraints on First Derivative Estimates

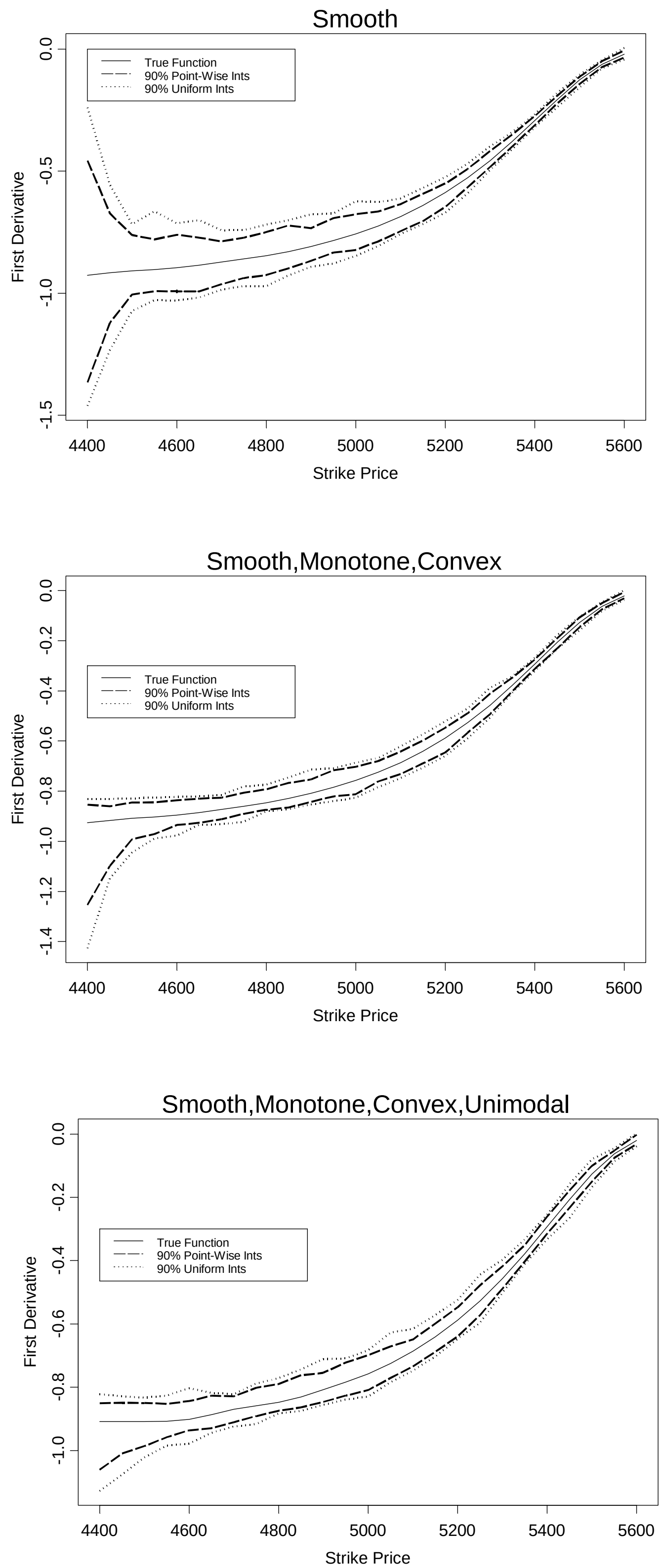


Figure 2C: Effects of Constraints on SPD Estimates

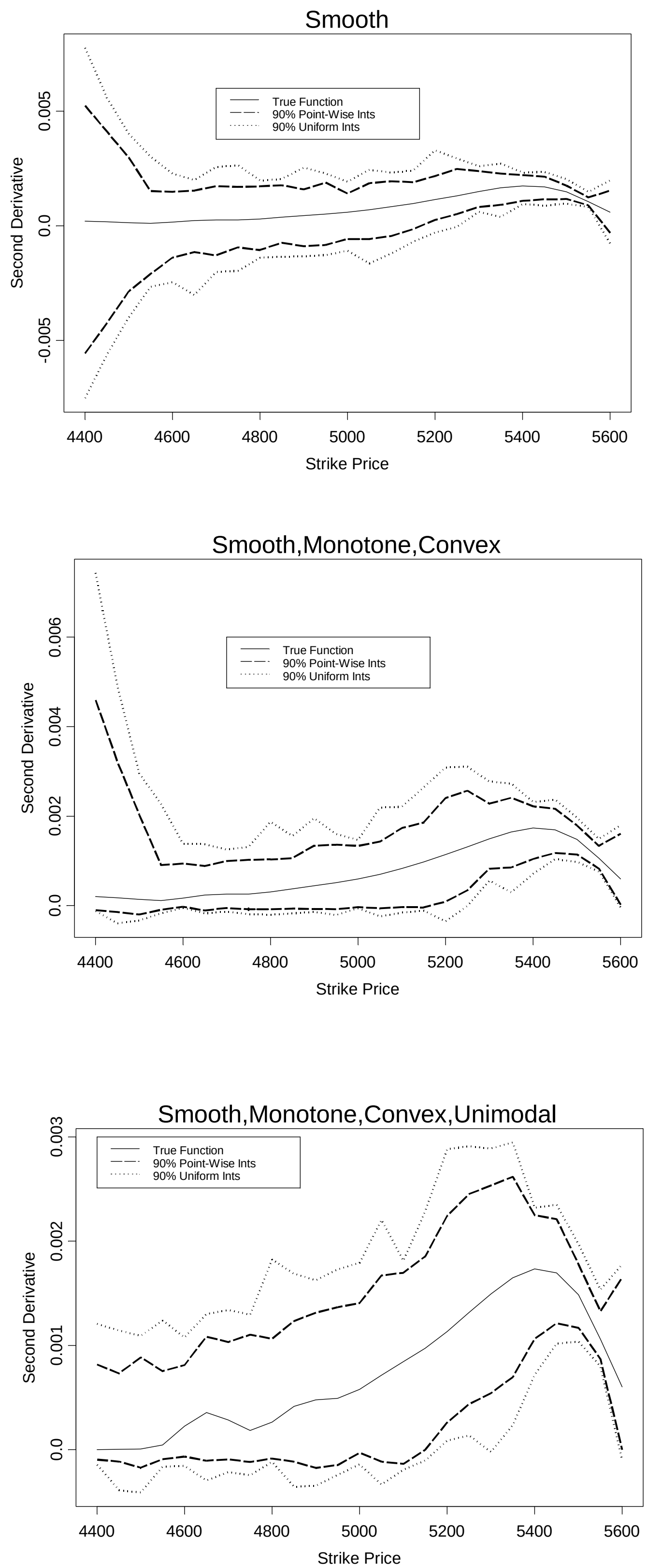


Figure 3: Asymptotic Vs. Bootstrap Confidence Intervals

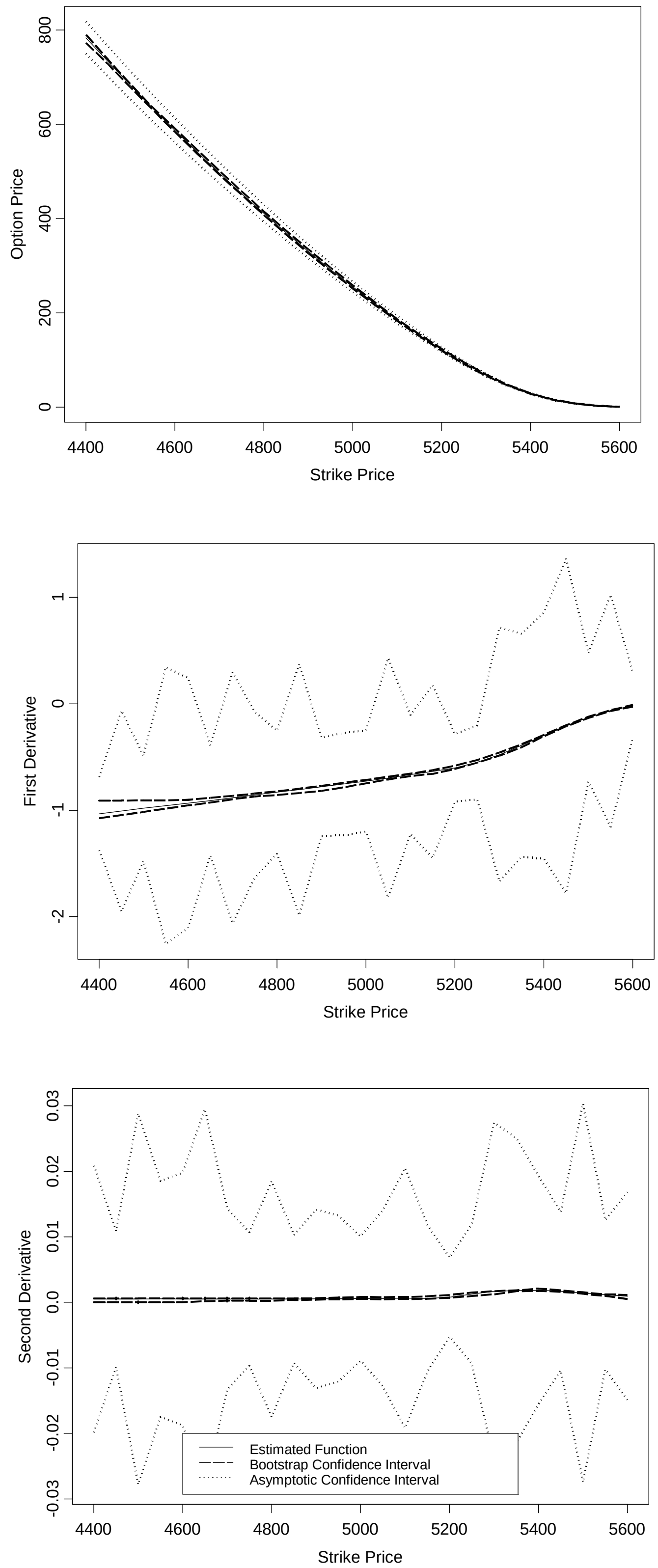
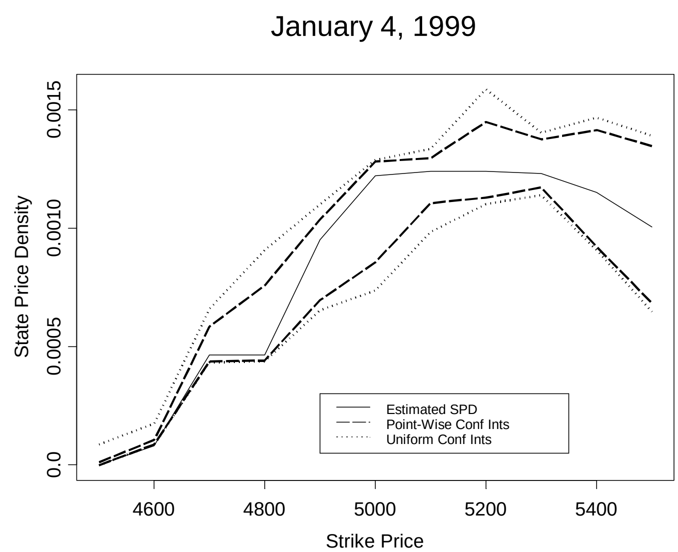


Figure 4: SPDs Using DAX Index Data



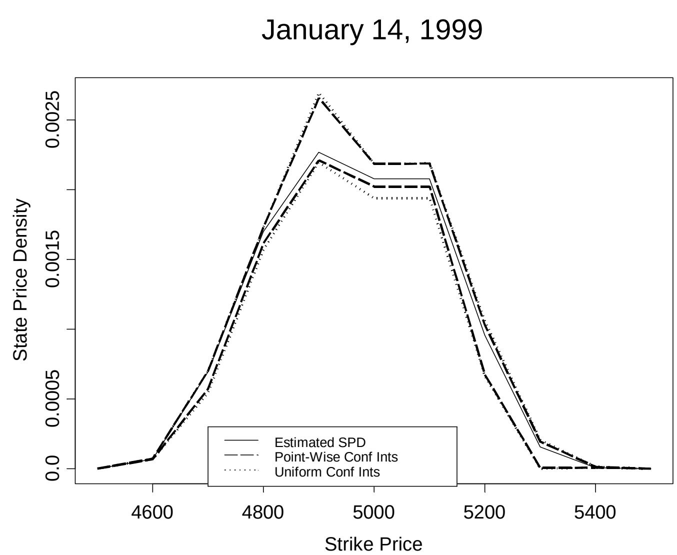


Figure 5: Evolution of State Price Densities Jan 4-8 and Jan 11-15, 1999, DAX Index Data

(Jan 9 and 10, 1999 were week-end days)