

RESEARCH

Open Access



Call option price function in Bernstein polynomial basis with no-arbitrage inequality constraints

Arindam Kundu¹, Sumit Kumar^{2*}, Nutan Kumar Tomar¹ and Shiv Kumar Gupta³

*Correspondence:

sumit.kumar@iimu.ac.in

²Faculty, Operations Management Quantitative Methods & Information System, Indian Institute of Management Udaipur, Udaipur, 313001, India

Full list of author information is available at the end of the article

Abstract

We propose an efficient method for the construction of an arbitrage-free call option price function from observed call price quotes. The no-arbitrage theory of option pricing places various shape constraints on the option price function. For each available maturity on a given trading day, the proposed method estimates an option price function of strike price using a Bernstein polynomial basis. Using the properties of this basis, we transform the constrained functional regression problem to the least-squares problem of finite dimension and derive the sufficiency conditions of no-arbitrage pricing to a set of linear constraints. The resultant linearly constrained least square minimization problem can easily be solved using an efficient quadratic programming algorithm. The proposed method is easy to use and constructs a smooth call price function which is arbitrage-free in the entire domain of the strike price with any finite number of observed call price quotes. We empirically test the proposed method on S&P 500 option price data and compare the results with the cubic spline smoothing method to see the applicability.

Keywords: call price function; no-arbitrage inequality constraints; constrained functional regression; Bernstein polynomial basis; quadratic programming

1 Introduction

The arbitrage-free option price function defined across strike price and estimated from the available quotes has been studied extensively by researchers and practitioners (see, e.g., [1–5]). This function contains precious information about the risk inherent in the underlying asset. Policy makers and investors estimate the state price density from option prices, which signifies the probability that market participants ascribe to the future asset price movements [6]. Practitioners frequently compute the implied volatility which sets the Black-Scholes model option value equal to the price of that option. The implied volatility is then used to obtain the price of other complex exotic options which are either not significantly traded, or for which the quotes are not recorded.

There are three essential difficulties in constructing an arbitrage-free option price function from the available bid-ask quotes. First, on a given trading day, the bid-ask quotes of an option contingent on an underlying asset are known only for a limited number of unevenly distributed strike prices. Second, high bid-ask spreads are observed for the options with the strike price far from the current underlying asset price which makes finding

the right price more challenging. Third, the observed quotes are not always free from an arbitrage opportunity. It is crucial to have an option price function which does not allow for any arbitrage, to avoid any mispricing.

An option price function is arbitrage-free in the strike domain if it is monotonic, convex, and satisfying certain bounds on itself and the first derivative [2, 7]. So, an arbitrage-free option price function needs to meet some shape restrictions although the actual functional form of the function is not known *a priori*. The only other information available is that it should fall ideally within the observed bid and ask quotes. Mathematically, estimating such a function is a generalized constrained functional regression problem.

Several approaches have been suggested for the construction of the arbitrage-free option price function. Bates [8] was first to use constrained cubic spline fitting to interpolate an arbitrage-free option price function from the observed option price transactions data. Kahalé [9], proposed a technique using the piecewise convex polynomial interpolation to approximate the call price function. However, the prerequisite for this algorithm is to prepare the data arbitrage-free which may lead to substantial loss of information.

Aït-Sahalia and Duarte [2] proposed a two-step method to estimate the arbitrage-free call price function. First, they use a constrained least-squares procedure which incorporate no-arbitrage shape constraints of monotonicity and convexity and then employ smoothing using local polynomials. In the same vein, Birke and Pilz [3] and Fan and Mancini [4] suggested alternative kernel regression estimators for the arbitrage-free call price function. The drawback of the kernel-based regression is that it is computationally intensive.

Fengler [5] proposed a procedure based on quadratic programming using the cubic spline smoothing on the call price data to get an arbitrage-free implied volatility. Along the same line, more spline-based models are described in the literature [10–13]. Unfortunately, selecting the optimal number and location of the knots for spline fitting under shape constraints leads to a highly nonlinear optimization problem [14].

In this study, we use a Bernstein polynomial basis [15] to estimate the arbitrage-free option price function. The choice of this basis is straightforward as the constraints of monotonicity, convexity, and other bounds on the function and its first derivative on the entire domain leads to a set of linear inequality constraints on unknown parameters. The advantage is that we need to solve a finite-dimensional least-squares problem with a few linear constraints only. Moreover, they provide a smooth estimate, which can be obtained as a unique solution of a quadratic programming problem, thereby making it computationally attractive and efficient.

The class of Bernstein polynomials is dense in the space of all continuous functions ($C^0[a, b]$) with any finite support with supremum norm and all of the derivatives possess the same convergence properties [16]. Furthermore, Chak *et al.* [17] proved that the univariate Bernstein polynomial estimator is consistent not only with the true function but also for its first and second derivatives. In the literature, the Bernstein polynomial regression has been employed with several shape constraints which include non-negativity, unimodality, monotonicity, convexity (or concavity), etc., either alone or in combination with each other (see, *e.g.*, [17–22]).

The organization of our paper is as follows: we present the no-arbitrage inequalities on the call price function in Section 2. Section 3 describes our proposed methodology to approximate the call price function using a Bernstein polynomial basis under the various inequality constraints arising from no-arbitrage conditions and derive the quadratic pro-

gramming formulation of the estimation problem. In Section 4, we examine the empirical applicability of our proposed model using S&P 500 option price data. We compare our computed results with the cubic spline smoothing method in Section 5. Finally, we conclude in Section 6.

2 No-arbitrage inequalities on the call price function

On a single underlying asset S_t , at current time t , the price of a European call option $C(K)$ with strike price K , at fixed expiry time T , time to maturity $\tau = T - t$, risk-free interest rate r , and deterministic dividend rate q , is given by

$$C(K) = e^{-r\tau} \int_0^\infty \max(S_T - K, 0)g(S_T) dS_T, \tag{1}$$

provided the probability density function $g(S_T)$ of state price S_T exists. The fundamental theorem of asset pricing ensures the existence of $g(S_T)$ under the assumption of no-arbitrage and is usually called the state price density (SPD) (see [23, 24]).

Differentiating (1) w.r.t. the strike price K we get

$$\frac{\partial C}{\partial K} = -e^{-r\tau} \int_K^\infty g(S_T) dS_T. \tag{2}$$

Since the state price density $g(S_T)$ is a probability density function, it is non-negative and integrable to one. Now,

$$g(S_T) \geq 0 \Rightarrow \frac{\partial C}{\partial K} \leq 0 \quad \text{and} \quad \int_0^\infty g(S_T) dS_T = 1 \Rightarrow \frac{\partial C}{\partial K} \geq -e^{-r\tau}.$$

So we have

$$-e^{-r\tau} \leq \frac{\partial C}{\partial K} \leq 0. \tag{3}$$

Thus, the call option price function needs to be monotonically non-increasing and further the slope of the function must be bounded below uniformly.

Differentiating (2) again w.r.t. the strike price K , we get the famous relation given by Breeden and Litzenberger [25],

$$\frac{\partial^2 C}{\partial K^2} = e^{-r\tau} g(K). \tag{4}$$

Equation (4) implies that the second derivative of the call price function is proportional to the state price density,

$$\Rightarrow \frac{\partial^2 C}{\partial K^2} \geq 0. \tag{5}$$

Hence, the call price function has to be globally convex as any local non-convexity of the call price function implies a negative state price density.

In addition, the no-arbitrage theory also imposes the following bounds on the call price function:

$$\max(e^{-q\tau} S_t - e^{-r\tau} K, 0) \leq C(K) \leq S_t e^{-q\tau}. \tag{6}$$

Therefore, an arbitrage-free call price function must satisfy the set of inequality constraints given by (3), (5), and (6).

3 Problem formulation

In this section, we describe the mathematical formulation to construct the arbitrage-free call price function $\hat{C} : [K_{\min}, K_{\max}] \rightarrow R$ using the given discrete set of observable data points. For a fixed expiry date, we assume that a finite sample of observations $\{(K_i, C_i) : K_i \in [K_{\min}, K_{\max}], i = 1 : n\}$ are available in the market, where the ordered pair (K_i, C_i) represents the price C_i corresponding to the European call option of strike price K_i contingent on a single underlying asset.

Let us consider the regression model,

$$C_i = f(K_i) + \epsilon_i, \quad i = 1, 2, \dots, n,$$

where the error term ϵ_i has zero mean with finite variance. The set of ordered pairs $\{(K_i, C_i)\}_{i=1}^n$ are assumed to be independent and identically distributed. Here, the unknown call price function $f(K)$ is assumed to belong to a class of smooth functions having some restrictions on their shapes and can be estimated by minimizing the empirical L_2 norm, *i.e.*,

$$\hat{C}(K) = \arg \min_{f(K) \in \mathfrak{S}_c} \frac{1}{n} \sum_{i=1}^n (f(K_i) - C_i)^2. \tag{7}$$

Here $\mathfrak{S}_c = \{f(K) \in C^2[K_{\min}, K_{\max}] : f(K) \text{ satisfies the inequality constraints given by (3), (5), and (6)}\}$.

Nevertheless, the constrained functional regression problem (7) has some basic mathematical challenges, such as

- only a finite sample of noisy and unevenly distributed observations are available;
- the actual functional form of the regression function f is not known *a priori*;
- the constraints arising from no-arbitrage conditions must be satisfied everywhere in the domain;
- the constraints are imposed not only on the function but also on its first and second derivative.

As the actual functional form of the regression function f is not known *a priori* we propose to approximate the regression function in a Bernstein polynomial basis. Using the properties of the Bernstein polynomial basis, we now transform the constrained functional regression problem (7) into a finite-dimensional least-squares problem with linear constraints.

3.1 Finite-dimensional problem in Bernstein polynomial basis

Let us apply a simple linear transformation $K \mapsto x$, where $x = \frac{K - K_{\min}}{K_{\max} - K_{\min}}$ to transform the domain $[K_{\min}, K_{\max}]$ to $[0, 1]$. Then it is easy to see that the inequality constraints on f given by (3), (5), and (6) transforms to the following conditions:

$$-(K_{\max} - K_{\min})e^{-r\tau} \leq \frac{\partial f}{\partial x} \leq 0, \tag{8}$$

$$\frac{\partial^2 f}{\partial x^2} \geq 0, \tag{9}$$

$$\max(e^{-qt} S_t - e^{-rt} ((K_{\max} - K_{\min})x + K_{\min}), 0) \leq f(x) \leq S_t e^{-qt}. \tag{10}$$

Now, for any continuous function f in $[0, 1]$, the approximating Bernstein polynomial of order N is given by

$$B_N(x; f) = \sum_{k=0}^N f(k/N) \binom{N}{k} x^k (1-x)^{N-k} = \sum_{k=0}^N \beta_k b_k(x, N), \tag{11}$$

where $\{b_k(x, N) = \binom{N}{k} x^k (1-x)^{N-k}, k = 0, 1, \dots, N\}$ forms a basis of the Bernstein polynomial of degree N and $\beta_N = \{\beta_k : \beta_k = f(\frac{k}{N}), k = 0, 1, \dots, N\}$ are the corresponding coefficients.

So, if we approximate $f(x)$ in the Bernstein polynomial basis of order N , the constrained functional regression problem (7) transforms into the following finite-dimensional problem of estimation of β_N :

$$\hat{C}(x) = \arg \min_{B_N(x; f) \in \mathfrak{S}_N} \frac{1}{n} \sum_{i=1}^n (B_N(x_i; f) - C_i)^2, \tag{12}$$

where

$$\mathfrak{S}_N = \{B_N(x; f) : B_N(x; f) \text{ satisfy the inequality constraints (8), (9), and (10)}\},$$

From the Weierstrass theorem, $B_N(x; f) \rightarrow f(x)$ uniformly over $[0, 1]$ as $N \rightarrow \infty$ [16]. Additionally, the derivatives of $B_N(x; f)$ also satisfy the existing bounds of the corresponding derivatives of f (see [26]) and possess the same convergence properties (see [16]). Moreover, using the properties of the Bernstein polynomial basis, the inequality constraints on $B_N(x; f)$ and its the derivatives can be transformed to the linear inequality constraints involving β_N only.

3.2 Constraints in Bernstein polynomial basis

Bernstein polynomial basis functions are non-negative on $x \in [0, 1]$. Also, they form a partition of unity, i.e., $\sum_{k=0}^N b_k(x, N) = 1$. The first and second derivatives of $B_N(x; f)$, $N \geq 2$ can be written

$$B'_N(x; f) = \sum_{k=0}^N \beta_k b'_k(x, N) = N \sum_{k=0}^{N-1} (\beta_{k+1} - \beta_k) b_k(x, N - 1), \tag{13}$$

$$B''_N(x; f) = \sum_{k=0}^N \beta_k b''_k(x, N) = N(N - 1) \sum_{k=0}^{N-2} (\beta_{k+2} - 2\beta_{k+1} + \beta_k) b_k(x, N - 2). \tag{14}$$

Since $b_k(x, N - 2)$ is positive in $[0, 1]$,

$$(\beta_{k+2} - 2\beta_{k+1} + \beta_k) \geq 0, \quad \forall k = 0 : N - 2 \quad \Rightarrow \quad B''_N(x; f) \geq 0, \quad \forall x \in [0, 1].$$

The above set of inequalities can also be written $\beta_1 - \beta_0 \leq \dots \leq \beta_N - \beta_{N-1}$. Then

$$(\beta_N - \beta_{N-1}) \leq 0 \tag{15}$$

$$\begin{aligned} &\Rightarrow \beta_1 - \beta_0 \leq \dots \leq \beta_N - \beta_{N-1} \leq 0 \\ &\Rightarrow B'_N(x;f) \leq 0, \quad \forall x \in [0,1]. \end{aligned}$$

Further, as $\sum_{k=0}^{N-1} b_k(x, N-1) = 1$,

$$\begin{aligned} \beta_1 - \beta_0 &\geq \frac{-(K_{\max} - K_{\min})e^{-r\tau}}{N} \tag{16} \\ &\Rightarrow \frac{-(K_{\max} - K_{\min})e^{-r\tau}}{N} \leq \beta_1 - \beta_0 \leq \dots \leq \beta_N - \beta_{N-1} \\ &\Rightarrow N \sum_{k=0}^{N-1} (\beta_{k+1} - \beta_k)b_k(x, N-1) \geq -(K_{\max} - K_{\min})e^{-r\tau} \sum_{k=0}^{N-1} b_k(x, N-1) \\ &\Rightarrow B'_N(x;f) \geq -(K_{\max} - K_{\min})e^{-r\tau}, \quad \forall x \in [0,1]. \end{aligned}$$

Furthermore,

$$\begin{aligned} \beta_1 - \beta_0 &\geq -(K_{\max} - K_{\min})e^{-r\tau}/N \\ &\Rightarrow \beta_1 \geq \beta_0 - (K_{\max} - K_{\min})e^{-r\tau}/N. \end{aligned}$$

So,

$$\begin{aligned} \beta_0 &\geq e^{-q\tau} S_t - e^{-r\tau} K_{\min} \\ &\Rightarrow \beta_1 \geq e^{-q\tau} S_t - e^{-r\tau} \left((K_{\max} - K_{\min})\frac{1}{N} + K_{\min} \right). \end{aligned}$$

Similarly, for $k = 2, 3, \dots, N$,

$$\beta_k \geq e^{-q\tau} S_t - e^{-r\tau} \left((K_{\max} - K_{\min})\frac{k}{N} + K_{\min} \right).$$

Then

$$B_N(x;f) \geq \sum_{k=0}^N \left(e^{-q\tau} S_t - e^{-r\tau} \left((K_{\max} - K_{\min})\frac{k}{N} + K_{\min} \right) \right) b_k(x, N).$$

Since the right hand side of the above inequality is the Bernstein polynomial representation of a linear function $e^{-q\tau} S_t - e^{-r\tau} ((K_{\max} - K_{\min})x + K_{\min})$.

Thus, we can always write

$$B_N(x;f) \geq e^{-q\tau} S_t - e^{-r\tau} ((K_{\max} - K_{\min})x + K_{\min}), \quad \forall x \in [0,1]. \tag{17}$$

Now,

$$B'_N(x;f) \leq 0 \Rightarrow B_N(0;f) \geq B_N(x;f) \geq B_N(1;f), \quad \forall x \in [0,1].$$

Moreover, we have $B_N(0;f) = \beta_0$, and $B_N(1;f) = \beta_N$.

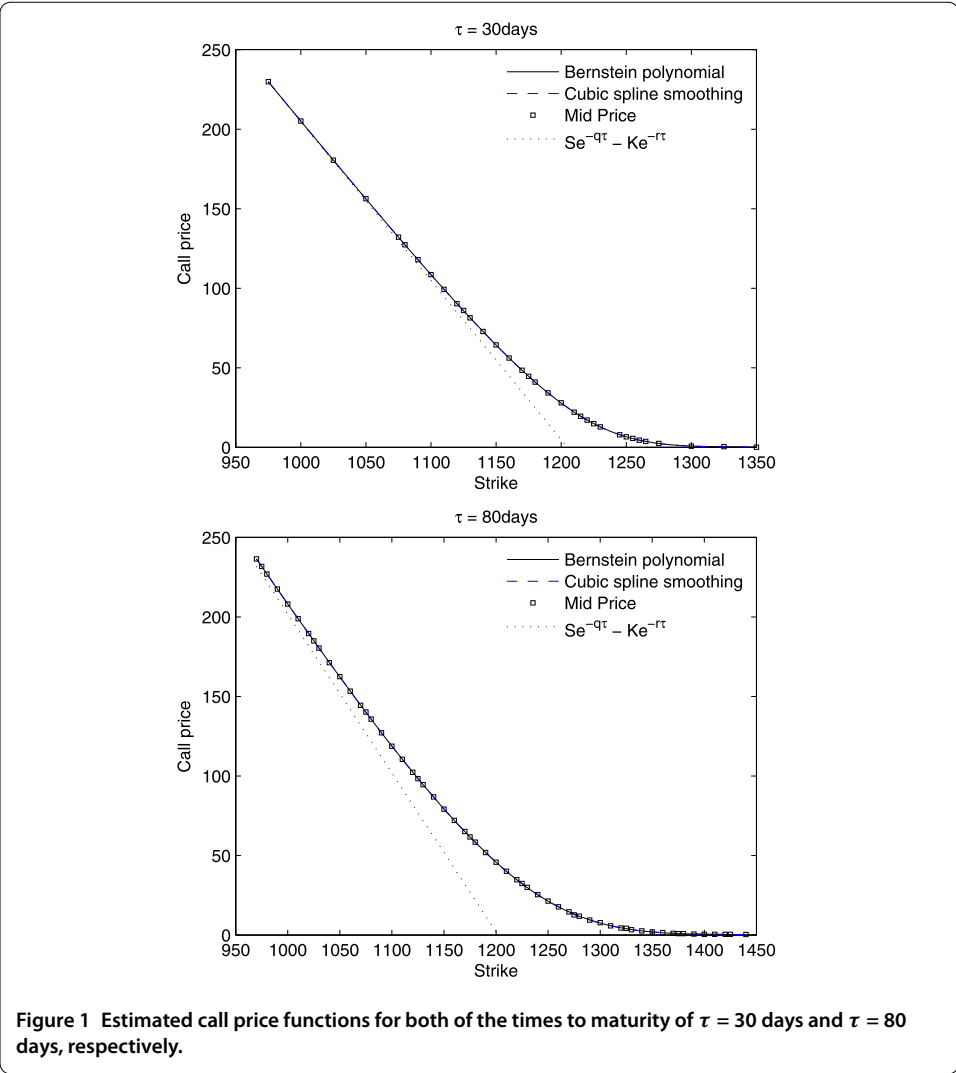


Figure 1 Estimated call price functions for both of the times to maturity of $\tau = 30$ days and $\tau = 80$ days, respectively.

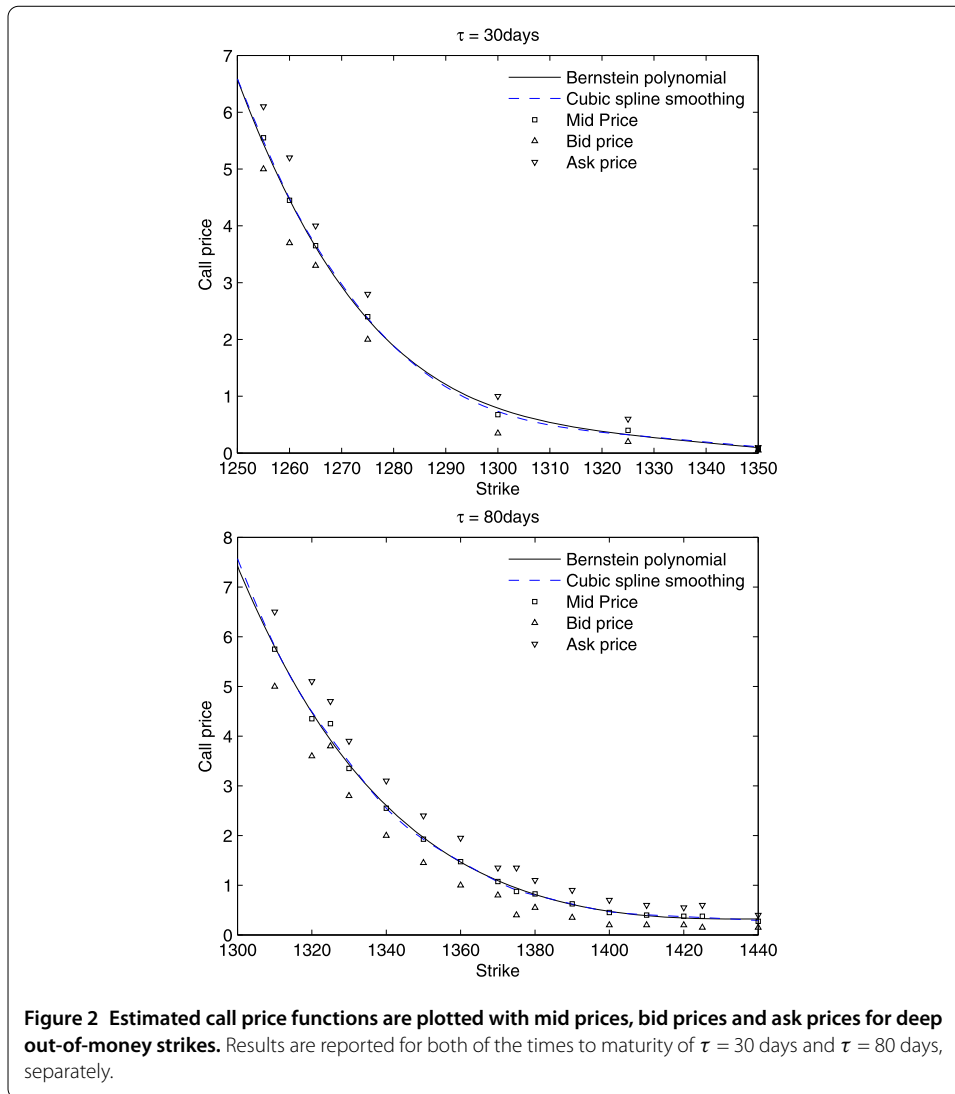
where $\mathbf{f}_N = [f_0, f_1, \dots, f_N]^T$ is a column vector of order $(N + 1)$, with the elements $f_i = 2 \sum_{j=1}^n C_j b_i(x_j, N)$, and the Hessian \mathbf{H}_N is a symmetric matrix of order $(N + 1)$ with the elements $h_{i,j}$ defined by

$$h_{i,j} = \begin{cases} \sum_{k=1}^n b_i^2(x_k, N); & \text{when } i = j, \\ \sum_{k=1}^n b_i(x_k, N) b_j(x_k, N); & \text{when } i \neq j. \end{cases}$$

For a given N , if the Hessian matrix \mathbf{H}_N is strictly positive definite, the quadratic programming problem (25) is well posed and can be solved in polynomial time [27]. Optimal N may be selected adaptively based on information criteria such as AIC used in [21] and defined by

$$AIC(N) = n \log(SSE/n) + 2edf, \tag{26}$$

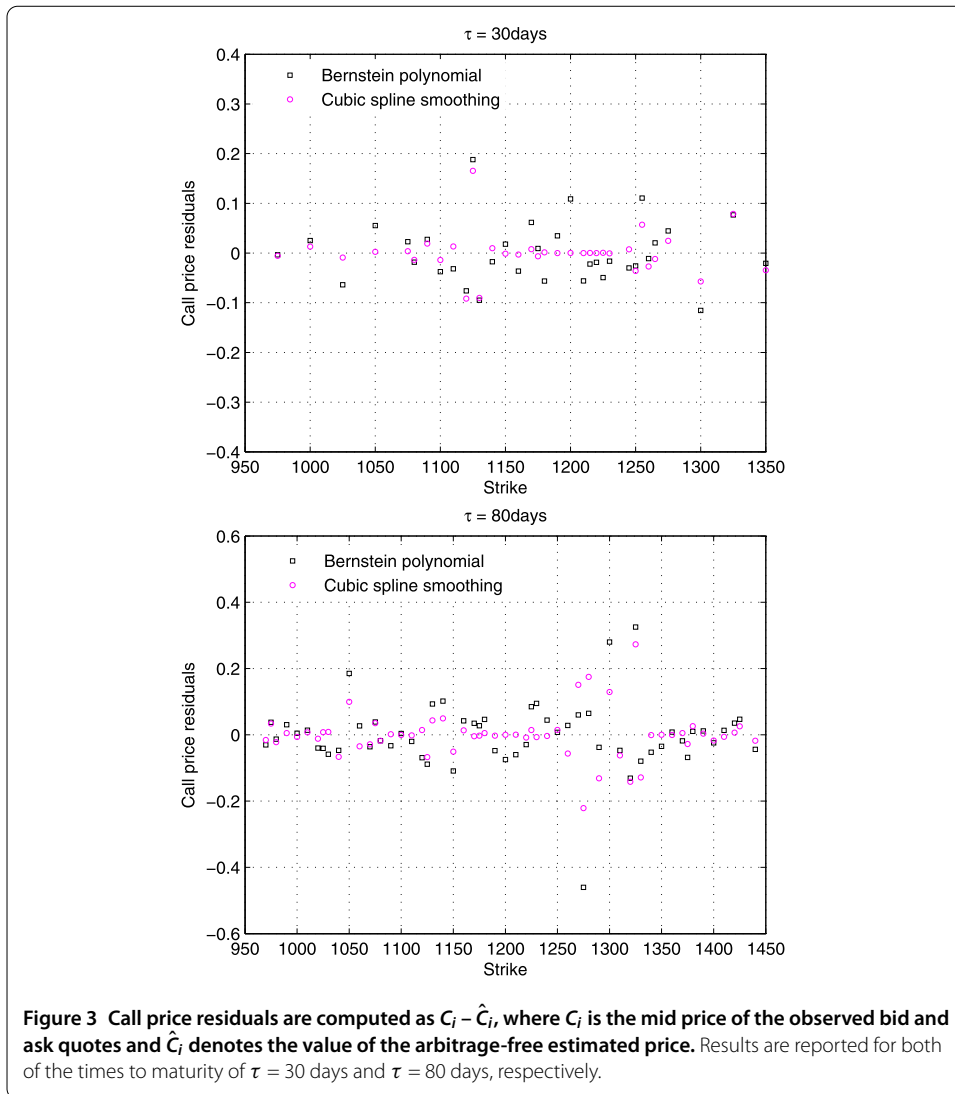
where $SSE = \sum_{i=1}^n (B_N(x_i; f) - C_i)^2$, and $edf = N + 1$ is the effective degree of freedom. From equation (26), we choose the optimal N as $N_{opt} = \arg \min_{N \in \mathbb{N}} AIC(N)$.



4 Empirical applications

In this paper, we test the empirical applicability of our proposed method on S&P 500 Index call option data. S&P 500 Index options belong the most liquid exchange traded options in the world and are well suited as a test case, since numerous empirical studies are performed on this data [2, 12]. We use the end of day quotes of the S&P 500 Index call option obtained from the Chicago Board of Options Exchange (CBOE). For each trading day, the quotes comprise last-bid and last-ask price for the option contingent on the S&P 500 Index with different strikes and maturities.

For the exposition, we choose the call options with dates of maturity (T) December 31, 2010 and February 19, 2011, which were recorded on $t =$ December 1, 2010. The S&P 500 Index closed at $S_t = 1,206.07$ on that day. As reported in [13], we consider an annualized risk-free rate r as 0.30% corresponding to the time to maturity $\tau = 30$ days, and 0.37% for $\tau = 80$ days. Also, we consider the implied forward $F_t^\tau = S_t e^{(r-q)\tau}$, as 1,204.85 and 1,201.95 for the respective maturities. Using the risk-free rates, we calculate the dividend $q = r - \frac{1}{\tau} \log\left(\frac{F_t^\tau}{S_t}\right)$, as 1.53% and 1.93% for the respective maturities. Since the call price can fluctuate around the median of the bid and ask price [28], we minimize the esti-



mation error from the mid price of the bid-ask quotes as suggested by Glaser and Heider [29]. To eliminate the error-prone observations and possible biases, we use a level 1 filter suggested by [30]. After filtering, the remaining number of observations are $n = 34$ for $\tau = 30$ days and $n = 57$ for $\tau = 80$ days.

We apply our proposed method to the data and solve the resultant quadratic programming problem by an *interior-point-convex* method using the *quadprog* function of the Matlab package. Using equation (26), for times to maturity τ equal to 30 days and 80 days, we get N_{opt} as 17 and 15, respectively.

To validate the computed results we compare the outcome with the popular cubic spline smoothing method for a single time to maturity as suggested by Fengler [5]. The natural cubic spline smoothing estimator can be obtained by minimizing the penalized sum of squares under the linear constraints given in Section 3.2 in [5],

$$\sum_{i=1}^n (C_i - f(K_i))^2 + \lambda \int_{K_{min}}^{K_{max}} (f''(K))^2 dK, \tag{27}$$

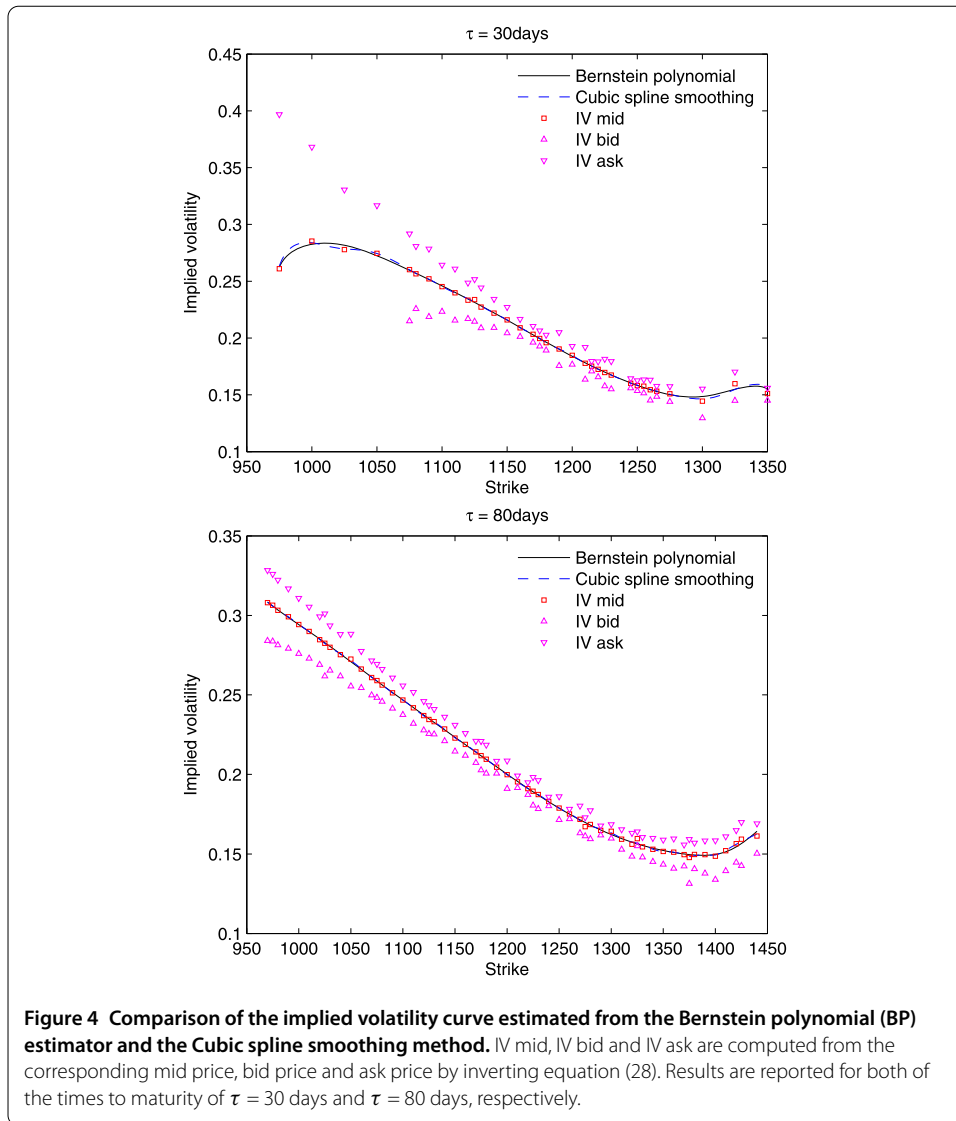
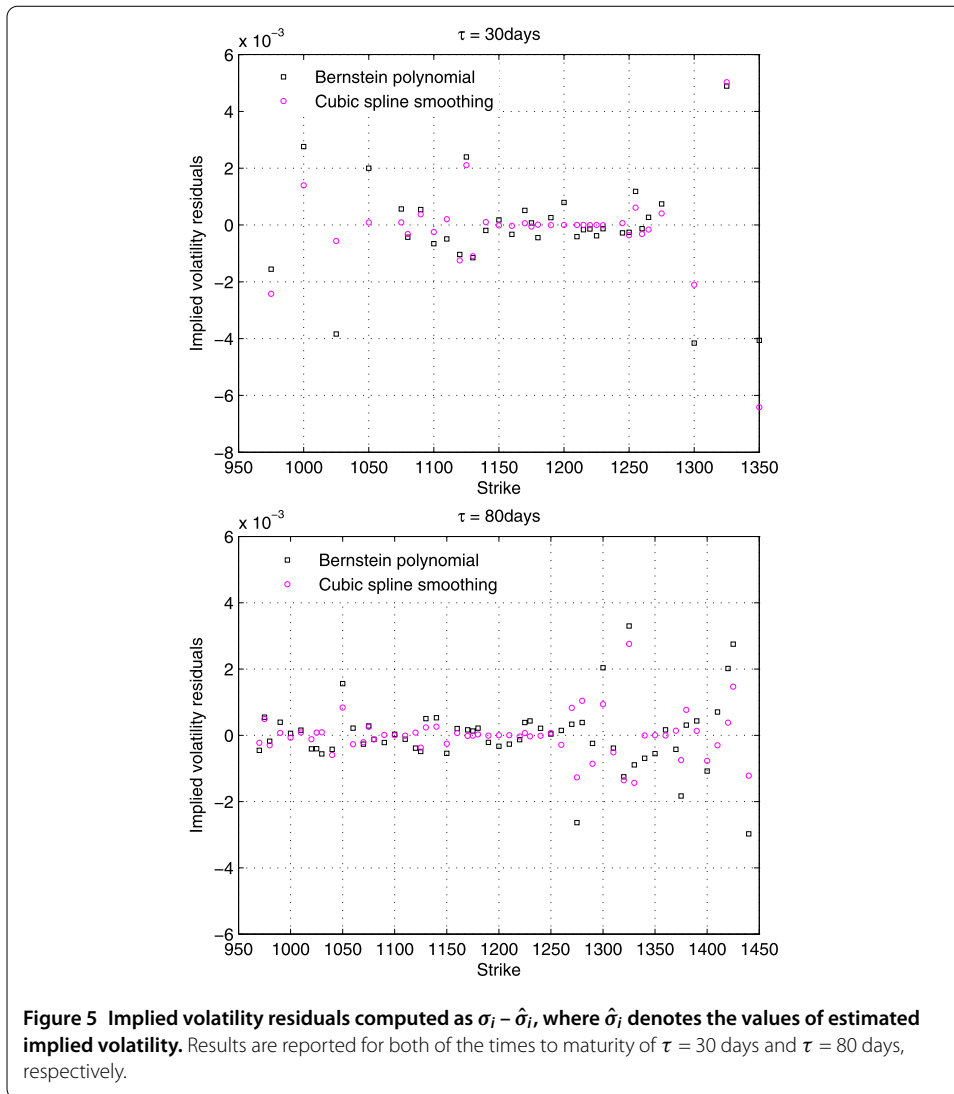


Figure 4 Comparison of the implied volatility curve estimated from the Bernstein polynomial (BP) estimator and the Cubic spline smoothing method. IV mid, IV bid and IV ask are computed from the corresponding mid price, bid price and ask price by inverting equation (28). Results are reported for both of the times to maturity of $\tau = 30$ days and $\tau = 80$ days, respectively.

where λ is a smoothing parameter. As the constraints on the first and second derivatives of the call price function act as a strong smoothing device, the choice of λ is of secondary importance, and so we fix λ as $1e-7$ as suggested in [5]. Finding an optimal number of knots and locations of these knots in the presence of constraints leads to a nonlinear optimization problem [14]. So, for the sake of simplicity, we fix the knots at the observation points.

Since practitioners regularly use the implied volatility to compare options prices across various strikes, expiries, and underlying assets, we also compare the implied volatility function corresponding to the estimated call price curves. In general, the implied volatilities are obtained by inverting the call option price function using the Black-Scholes equation [23] given by

$$C(S_t, t; K, \tau, r, q, \hat{\sigma}) = S_t e^{-q\tau} \Phi(d_1) - Ke^{-r\tau} \Phi(d_2), \tag{28}$$

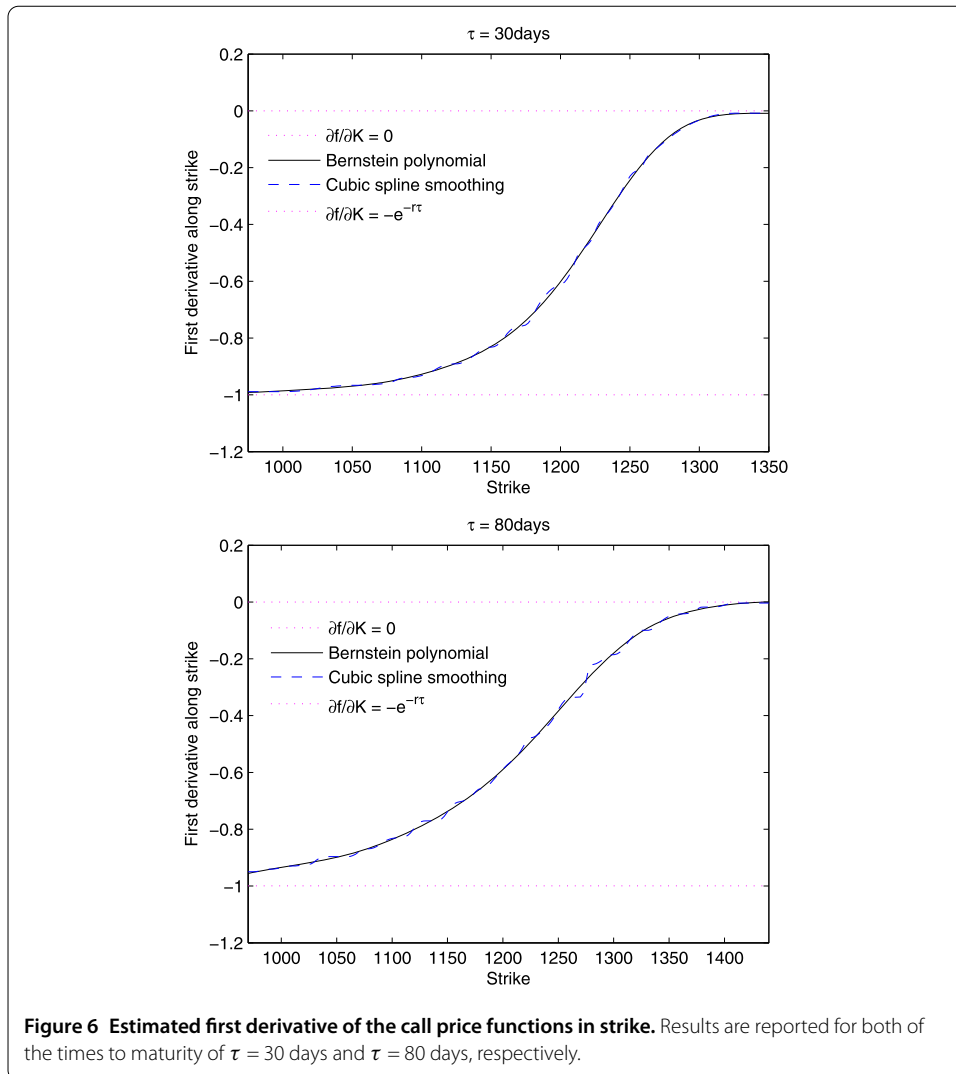


where $\hat{\sigma}$ is the implied volatility, Φ is the CDF of the standard normal distribution and $d_1 = \frac{\log(S_t/K) + (r - q + \hat{\sigma}^2/2)\tau}{\hat{\sigma}\sqrt{\tau}}$, and $d_2 = d_1 - \hat{\sigma}\sqrt{\tau}$. It is well known that the implied volatility curve will be arbitrage-free if it is deduced from the call price curve satisfying no-arbitrage conditions [9].

5 Results and discussion

The numerical results are computed and are compared with the cubic spline smoothing method to validate the performance. Figure 1 depicts the estimated arbitrage-free call price function for both of the maturities, which shows that our proposed estimator and CS method match well with the mid price of the observed bid-ask quotes. Also, Figure 1 ascertains that both of the estimators satisfy the price bound constraints defined in equation (10) globally.

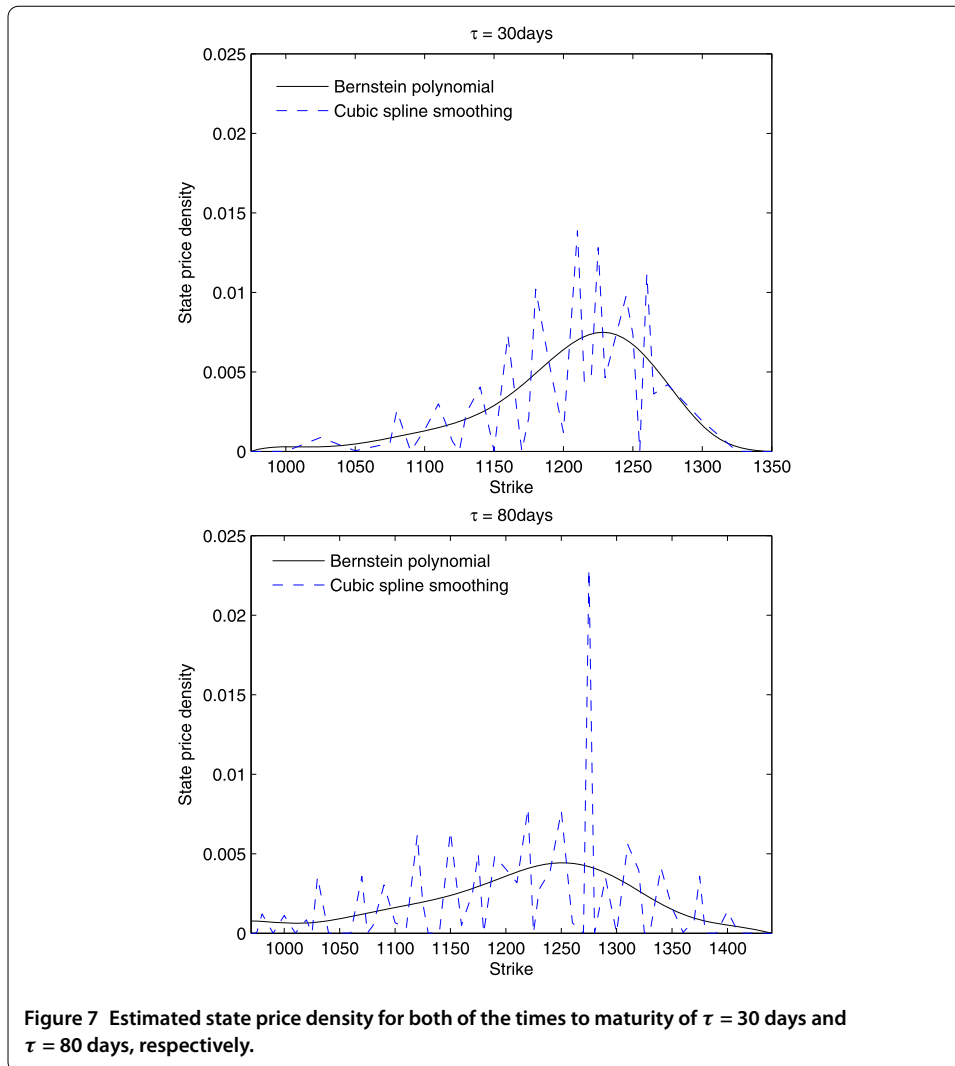
Figure 2 demonstrates that, even in the region of a deep out-of-the-money strike price (*i.e.*, $K \gg S_t$), the estimated curves from both of the methods lie between the bid and ask quotes. Estimates computed using the CS method are a bit closer to the mid price of the observed bid-ask quotes.



To see the closeness of the curve with the mid price, we present the price residuals in Figure 3. For both of the maturities, larger deviations are spotted for near-the-money options, which are consistent with the observation reported in [5].

Figure 4 displays the arbitrage-free implied volatility curve corresponding to the estimated call price function. We also show the computed implied volatility corresponding to the bid and ask price. It is observable that for the deep in-the-money strike price $K_{\text{deep}} (\ll S_t)$, the implied volatility for the corresponding bid price C_{bid} does not exist as $C_{\text{bid}} < (e^{-q\tau} S_t - e^{-r\tau} K_{\text{deep}})$, *i.e.*, the time value of the corresponding bid price is negative. Still, both of the methods produce a reasonable estimate of the implied volatility in that region. Again, we notice that the estimates from the CS method are a bit closer to the implied volatility corresponding to the mid price of the observed bid-ask quotes. On the other hand, our proposed method produce a bit smoother estimate.

Figure 5 plots pointwise implied volatility residuals. We observe larger deviations for in-the-money and out-of-the-money options where the first derivative of the implied volatility with respect to call price (known as the inverse vega) is sensitive.



To monitor the correct implementations of no-arbitrage inequalities, we compute the first derivative of the estimated call price function. It is obtained using equation (13) and transforming back to the strike domain. We display the derivative across strike in Figure 6, which shows that the estimated derivative from both of the methods is monotonically increasing and satisfy the bounds given by equation (3). However, unlike the cubic spline method our proposed method produces smooth estimates of the first derivative.

We also estimate the state price density function by computing the second derivative of the estimated call price function for both of the maturities. Figure 7 presents the estimated state price density function using both of the methods. The estimated state price density functions are positive in the entire domain, which shows that our proposed estimator satisfies the inequality arising from the convexity constraint. Since a relatively low penalty is added to the objective function to minimize equation (27), the contribution of the smoothness part becomes relatively negligible, which leads to a rough estimate of first and second derivatives of the cubic spline estimator. It is noticeable that unlike the CS method our proposed estimator produces a smooth state price density function.

Table 1 Root mean square error (RMSE), mean absolute error (MAE), and percentage relative mean error (PRME) of the estimated call option price function (Price) and the implied volatility function (IV) with respect to mid price quotes by the proposed Bernstein polynomial method (BP) and the cubic spline smoothing method (CS)

Time to maturity in days		RMSE		MAE		PRME	
		BP	CS	BP	CS	BP	CS
30	Price	0.0610	0.0427	0.0472	0.0237	-0.6694	-1.0065
	IV	0.0017	0.0016	0.0011	0.0008	-0.0632	-0.0891
80	Price	0.1008	0.0713	0.0635	0.0406	-0.0465	-0.0477
	IV	0.0010	0.0007	0.0007	0.0004	-0.0020	-0.0048

Results are reported for both of the times to maturity $\tau = 30$ days and $\tau = 80$ days.

Table 2 Number of constraints, number of iterations, and the average computational time taken by the optimization routine of the proposed Bernstein polynomial method (BP) and the cubic spline smoothing method (CS) for times to maturity of 30 days and 80 days, respectively

Time to maturity in days		No. of constraints (n)	Iterations	Average cpu time (in seconds)
30	BP	21	12	0.0090
	CS	69	12	0.0136
80	BP	19	12	0.0080
	CS	115	12	0.0165

Table 1 shows the comparative results in terms of the root mean square error (RMSE), the mean absolute error (MAE), and the percentage relative mean error (PRME) for the estimation of the call price function and the implied volatilities with respect to the mid price of the bid-ask quotes. It is apparent that both of the methods have an almost similar error behavior in terms of RMSE, MAE, and PRME measures.

Table 2 reports the computational details, *i.e.*, the number of constraints, number of iterations, and the average computational time taken by the optimization routine of the proposed Bernstein polynomial method (BP) and the cubic spline smoothing method (CS). The average is taken over 100 runs of the optimization routine starting from the second run. The number of constraints in the cubic spline method is equal to $2n + 1$, where n is the number of observations, while our proposed method requires only $N_{opt} + 4$ constraints, N_{opt} being the optimal order of the Bernstein polynomial basis and in general, $N_{opt} \ll n$. Although both of the methods take the same number of iterations, yet our proposed method consumes on average less computational time.

6 Conclusion

In this article, we propose an easy to use method for the construction of an arbitrage-free call option price function using a Bernstein polynomial basis. One of the most fundamental advantages here is that the estimation problem can be reduced to a quadratic programming problem with only a few linear constraints. The empirical results demonstrate that the proposed method is accurate and computationally efficient. For a limited number of bid-ask quotes, the method produces smooth estimates that satisfy all the inequality constraints imposed by no-arbitrage conditions on the entire domain. However, the spline-based model seems to provide a closer fit to the data, which may lead to overfitting and poor out-of-sample performance. These findings should motivate further research in the use of Bernstein polynomials for option pricing.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All the authors contributed equally and significantly in writing this paper. All authors read and approved the final manuscript.

Author details

¹Department of Mathematics, Indian Institute of Technology Patna, Bihta, Patna 801 103, India. ²Faculty, Operations Management Quantitative Methods & Information System, Indian Institute of Management Udaipur, Udaipur, 313001, India. ³Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee, 247667, India.

Acknowledgements

We thank anonymous referees for their helpful suggestions. The first author is grateful to the Ministry of Human Resource and Development, India for the financial support to carry out this work.

Received: 6 April 2016 Accepted: 31 May 2016 Published online: 10 June 2016

References

- Jackwerth, JC, Rubinstein, M: Recovering probability distributions from option prices. *J. Finance* **51**(5), 1611-1631 (1996)
- Ait-Sahalia, Y, Duarte, J: Nonparametric option pricing under shape restrictions. *J. Econom.* **116**(1), 9-47 (2003)
- Birke, M, Pilz, KF: Nonparametric option pricing with no-arbitrage constraints. *J. Financ. Econom.* **7**(2), 53-76 (2009)
- Fan, J, Mancini, L: Option pricing with model-guided nonparametric methods. *J. Am. Stat. Assoc.* **104**(488), 1351-1372 (2009)
- Fengler, MR: Arbitrage-free smoothing of the implied volatility surface. *Quant. Finance* **9**(4), 417-428 (2009)
- Bliss, RR, Panigirtzoglou, N: Testing the stability of implied probability density functions. *J. Bank. Finance* **26**(2), 381-422 (2002)
- Roper, M: Arbitrage free implied volatility surfaces. Preprint (2010)
- Bates, DS: The crash of '87: was it expected? The evidence from options markets. *J. Finance* **46**(3), 1009-1044 (1991)
- Kahalé, N: An arbitrage-free interpolation of volatilities. *Risk* **17**(5), 102-106 (2004)
- Laurini, MP: Imposing no-arbitrage conditions in implied volatilities using constrained smoothing splines. *Appl. Stoch. Models Bus. Ind.* **27**(6), 649-659 (2011)
- Orosi, G: Empirical performance of a spline-based implied volatility surface. *J. Deriv. Hedge Funds* **18**(4), 361-376 (2012)
- Orosi, G: Arbitrage-free call option surface construction using regression splines. *Appl. Stoch. Models Bus. Ind.* **31**(4), 515-527 (2014)
- Fengler, MR, Hin, L-Y: Semi-nonparametric estimation of the call-option price surface under strike and time-to-expiry no-arbitrage constraints. *J. Econom.* **184**(2), 242-261 (2015)
- Wang, X, Li, F: Isotonic smoothing spline regression. *J. Comput. Graph. Stat.* **17**(1), 21-37 (2008)
- Farouki, RT: The Bernstein polynomial basis: a centennial retrospective. *Comput. Aided Geom. Des.* **29**(6), 379-419 (2012)
- Lorentz, GG: Bernstein Polynomials. AMS Chelsea Publishing (2012)
- Chak, PM, Madras, N, Smith, B: Semi-nonparametric estimation with Bernstein polynomials. *Econ. Lett.* **89**(2), 153-156 (2005)
- Chang, I, Hsiung, CA, Wu, Y-J, Yang, C-C, et al.: Bayesian survival analysis using Bernstein polynomials. *Scand. J. Stat.* **32**(3), 447-466 (2005)
- Chang, I, Chien, L-C, Hsiung, CA, Wen, C-C, Wu, Y-J, et al.: Shape restricted regression with random Bernstein polynomials. In: *Complex Datasets and Inverse Problems: Tomography, Networks and Beyond*, pp. 187-202 (2007)
- Wang, J, Ghosh, SK: Shape restricted nonparametric regression with Bernstein polynomials. *Comput. Stat. Data Anal.* **56**(9), 2729-2741 (2012)
- Ding, J, Zhang, Z: Convex analysis in the semiparametric model with Bernstein polynomials. *J. Korean Stat. Soc.* **44**(1), 58-67 (2015)
- Turnbull, BC, Ghosh, SK: Unimodal density estimation using Bernstein polynomials. *Comput. Stat. Data Anal.* **72**, 13-29 (2014)
- Black, F, Scholes, M: The pricing of options and corporate liabilities. *J. Polit. Econ.* **81**(3), 637-654 (1973)
- Merton, RC: Theory of rational option pricing. *Bell J. Econ. Manag. Sci.* **4**(1), 141-183 (1973)
- Breeden, DT, Litzenberger, RH: Prices of state-contingent claims implicit in option prices. *J. Bus.* **51**(4), 621-651 (1978)
- Floater, MS: On the convergence of derivatives of Bernstein approximation. *J. Approx. Theory* **134**(1), 130-135 (2005)
- Floudas, CA, Visweswaran, V: Quadratic optimization. In: *Handbook of Global Optimization*, pp. 217-269 (1995)
- Hentschel, L: Errors in implied volatility estimation. *J. Financ. Quant. Anal.* **38**(4), 779-810 (2003)
- Glaser, J, Heider, P: Arbitrage-free approximation of call price surfaces and input data risk. *Quant. Finance* **12**(1), 61-73 (2012)
- Constantinides, GM, Jackwerth, JC, Savov, A: The puzzle of index option returns. *Rev. Asset Pricing Stud.* **3**(2), 229-257 (2013)