The Korean Journal of Applied Statistics (2011) DOI: http://dx.doi.org/10.5351/KJAS.2011.24.4.721
24(4), 721–733

Direct Nonparametric Estimation of State Price Density with Regularized Mixture

Yongho Jeon[1]

[1]Department of Applied Statistics, Yonsei University

(Received May 2011; accepted July 2011)

Abstract

We consider the state price densities that are implicit in financial asset prices. In the pricing of an option, the state price density is proportional to the second derivative of the option pricing function and this relationship together with no arbitrage principle imposes restrictions on the pricing function such as monotonicity and convexity. Since the state price density is a proper density function and most of the shape constraints are caused by this, we propose to estimate the state price density directly by specifying candidate densities in a flexible nonparametric way and applying methods of regularization under extra constraints. The problem is easy to solve and the resulting state price density estimates satisfy all the restrictions required by economic theory.

Keywords: State price density, European call option, shape constraints, gamma mixture density.

1. Introduction

The state price density(SPD) is the asset price density in the risk-neutral world where all investors are indifferent to risk and the expected return on all assets is the risk-free interest rate (Cox and Ross, 1976). The SPD also uniquely characterizes the equivalent martingale measure under which all asset prices discounted at the risk-free rate are martingale (Harrison and Kreps, 1979).

In the pricing of a call option, the SPD is directly related to the call pricing function and this relationship (together with the no arbitrage principle) imposes many restrictions including that the call pricing function should be a decreasing and convex function of the strike price. Let X be the strike price for a call option which will expire at time T, t the current time, $\tau = T - t$ the time to expiration, r the risk-free interest rate, and δ the dividend yield rate. Let S_t and S_T be the underlying asset price at time t and T, respectively. Assuming the existence of the SPD denoted by $f^*(S_T|S_t, \tau, r, \delta)$, the call pricing function of the European call option with payoff $(S_T - X)_+$ is given by:

$$C(S_t, X, \tau, r, \delta) = e^{-r\tau} \int_0^\infty (S_T - X)_+ f^*(S_T|S_t, \tau, r, \delta) dS_T, \tag{1.1}$$

[1]Assistant Professor, Department of Applied Statistics, Yonsei University, 262 Seongsanno, Seodaemun-gu, Seoul 120-749, Korea. E-mail: yhjeon@yonsei.ac.kr

where $(\nu)_+ = \max(\nu, 0)$. By taking derivative with respect to the strike price, we have

$$\frac{\partial C}{\partial X} = -e^{-r\tau} \int_X^\infty f^*(S_T) dS_T$$

and

$$f^*(S_T) = e^{r\tau} \frac{\partial^2 C}{\partial X^2}\bigg|_{X=S_T}$$

as observed by Breeden and Litzenberger (1978).

A proper density function should be nonnegative and integrate into unity, and the two constraints are respectively referred to as positivity and unity constraint. Since f^* is a proper density, the derivatives of the call pricing function C should satisfy

$$\begin{aligned} -e^{-r\tau} \le \frac{\partial C}{\partial X} &\le 0, \\ \frac{\partial^2 C}{\partial X^2} &\ge 0. \end{aligned} \tag{1.2}$$

The economic theory imposes no arbitrage bounds for C itself:

$$\max\left(0, S_t e^{-\delta\tau} - X e^{-r\tau}\right) \le C \le S_t e^{-\delta\tau}. \tag{1.3}$$

The no arbitrage principle places an additional constraint on the mean of SPD:

$$F_t = \int_0^\infty S_T f^*(S_T) dS_T = S_t \exp((r-\delta)\tau), \tag{1.4}$$

where F_t is the forward value at time t for delivery of the underlying asset at time T. With the condition (1.4), the constraint (1.3) is always satisfied as long as the SPD is a proper density. See Aït-Sahalia and Duarte (2003). Therefore the goal in this paper is to construct an estimator for the SPD, a valid density function over future values of the underlying asset which satisfies the constraint (1.4), with a good fit to the observed call option prices.

If asset prices follow geometric Brownian motion, the pricing function for a European call option is given by the Black-Scholes formula:

$$C_{BS}(F_t, X, \tau, r, \delta; \sigma) = e^{-r\tau}\{F_t \Phi(d_1) - X\Phi(d_2)\},$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, $F_t = S_t \exp((r-\delta)\tau)$ is the forward value, and $d_1 = (\log(F_t/X) + \sigma^2\tau/2)/(\sigma\sqrt{\tau})$ and $d_2 = d_1 - \sigma\sqrt{\tau}$. In this case the corresponding SPD is the lognormal density with mean $(r - \delta - \sigma^2/2)\tau$ and variance $\sigma^2\tau$ for $\log(S_T/S_t)$:

$$f^*(S_T) = \frac{1}{S_T\sqrt{2\pi\sigma^2\tau}} \exp\left[-\frac{\{\log(S_T/S_t) - (r-\delta-\sigma^2/2)\tau\}^2}{2\sigma^2\tau}\right].$$

This is a proper density function on $[0, \infty)$ and it satisfies the Equation (1.4).

For more flexible modeling, Aït-Sahalia and Lo (1998) considered the nonparametric regression problem

$$Y_i = C(S_i, X_i, \tau_i, r_i, \delta_i) + \epsilon_i \tag{1.5}$$

to estimate the call pricing function C and consequently the SPD f^*, based on historic data with option prices Y_i and the corresponding observations $(S_i, X_i, \tau_i, r_i, \delta_i)$ of the explanatory variable $(S_t, X, \tau, r, \delta)$, where ϵ_i are independent noises with mean zero. The sparseness of data in high dimensional space makes the general nonparametric regression difficult with a practical sample size unless any restrictions are further imposed on the model. So Aït-Sahalia and Lo (1998) reduced the dimensionality by specifying a semiparametric form using Black-Scholes formula:

$$C(S_t, X, \tau, r, \delta) = C_{BS}\left(F_t, X, \tau, r, \delta; \sigma\left(\frac{X}{F_t}, \tau\right)\right).$$

The call pricing function is given by the parametric Black-Scholes formula except that the volatility parameter σ is specified as a nonparametric function of the moneyness X/F_t and the time to expiration τ. Note that the resulting SPD with an arbitrary function σ may not be a proper density function and it may not satisfy the restriction (1.4) either.

For a nonparametric estimation of the SPD under the restrictions required, Aït-Sahalia and Duarte (2003) applied the local linear estimator to a transformed data instead of the original data to estimate $\partial C/\partial X$ and took a derivative on it to finally estimate $\partial^2 C/\partial X^2$ and the SPD. For a set of observation (x_i, y_i), $i = 1, \ldots, n$, ordered so that $x_1 \leq \cdots \leq x_n$, they found $m_1, \ldots, m_n$ to minimize

$$\sum_{i=1}^{n} (m_i - y_i)^2$$

subject to the slope and convexity constraints:

$$\begin{aligned} -e^{-r\tau} \leq \frac{m_{i+1} - m_i}{x_{i+1} - x_i} \leq 0, &\quad \text{for all } i = 1, \ldots, n-1, \\ \frac{m_{i+2} - m_{i+1}}{x_{i+2} - x_{i+1}} \geq \frac{m_{i+1} - m_i}{x_{i+1} - x_i}, &\quad \text{for all } i = 1, \ldots, n-2. \end{aligned}$$

Aït-Sahalia and Duarte (2003) showed that the estimated $\partial C/\partial X$ and $\partial^2 C/\partial X^2$ from the locally linear estimator with a log-concave kernel function based on the transformed data $m_1, \ldots, m_n$ satisfies the constraints (1.2). To further enforce the unity constraint for the SPD and the constraint (1.4), Aït-Sahalia and Duarte (2003) scaled and shifted the estimator of $\partial^2 C/\partial X^2$. All constraints are to be satisfied by adjusting the $\partial C/\partial X$ and C accordingly. However, it may not fit the data if the scale is not right.

For a constrained spline type approach, Yatchew and Härdle (2005) assumed that the X_i lie in a given interval $[a, b]$ and estimated the SPD by minimizing the squared error loss with a number of restrictions. The restrictions constrain the monotonicity and convexity of the call pricing function C to be satisfied at sample points and the integral of the SPD on $[a, b]$ not to exceed 1.

The difficulties in estimating the SPD using nonparametric smoothing methods are largely caused by the theory-motivated shape constraints. See Mammen *et al.* (2001) and references therein for a review of general constrained smoothing methods. Aforementioned nonparametric approaches estimate the call pricing function under shape constraints first to derive the corresponding SPD by taking its derivatives. These approaches tend to give large variances in the SPD estimation as a small change in the call price function estimator results in a big change in the corresponding SPD estimator.

For more stable but still flexible estimation, we propose to estimate the SPD directly by specifying the candidate densities in a flexible nonparametric way and finding the minimizer of a penalized empirical loss under some restrictions. The problem is easy to solve and the resulting SPD estimate is always a proper density function supported on $[0,\infty)$ which satisfies the constraint (1.4). Details of the proposed method are given in Section 2. We also discuss parameter tuning for the proposed method in Section 2. Simulation results are given in Section 3. In Section 4, we apply our method to option prices on the S&P 500 index. The article is concluded with a short discussion in Section 5.

2. Modeling

Suppose that (S_t, τ, r, δ) are given and we observe the strike prices X_i and the call option prices Y_i in the statistical model (1.5). Since our objective is to estimate the underlying SPD, which is a density function supported on $[0,\infty)$, we specify the SPD by a mixture

$$f^*(x) = c_1 K_1(x) + \cdots + c_q K_q(x) \tag{2.1}$$

of the *component densities* K_j's each of which is supported on $[0,\infty)$, where $c_j \geq 0$ and $\sum_j c_j = 1$ and the number of component q is prespecified. Let $\{\xi_1, \ldots, \xi_q\}$ be the location parameters, or knots, associated with component densities. This specification is a generalization of the kernel density estimation and automatically satisfies the positivity and unity constraints.

Our approach is to find the coefficients c_j minimizing a penalized empirical loss function over the class (2.1) under some restrictions. It is similar in spirit to the smoothing spline (Wahba, 1990) or penalized regression spline methods (Eilers and Marx, 1996; Ruppert *et al.*, 2003) in regression setting. In the smoothing spline, the observation itself or its subset is used for the knots. In the penalized regression spline, the knots different from the observation can be used such as equally spaced knots or equally spaced sample quantile knots.

The natural constraint in the mixture density (2.1) that the coefficients should be nonnegative and add up to one gives a novel form of regularization, and the shape of the constraint region which has many corners promotes the sparsity in the solution. With an SPD (2.1) specified with a large q in the beginning, most of the coefficients typically shrink to zero in the minimization procedure and only a small number of components remains in the solution.

The family of functions of the form (2.1) can be large enough for the flexibility required for nonparametric estimation. Since the support of the SPD is $[0,\infty)$, we may consider the lognormal components or the gamma components. With the lognormal components, the density of the logarithm of the prices is modeled as a mixture normal, which is known to be a rich family of distributions. The SPD estimation through the lognormal mixture is studied by Yuan (2009). Notice that (2.1) reduces to the famous Black-Sholes model when using one lognormal component. The gamma components with equally spaced location parameters are related to Tchebysheff-Laguerre functions which form an orthonormal system in the space $L^2(0,\infty)$. In this paper we focus on the mixture of the gamma components with the following form:

$$\hat{f}(x) = \sum_{j=1}^{q} c_j G_{\xi_j/b+1,b}(x) \tag{2.2}$$

with $c_j \geq 0$ and $\sum_{j=1}^{q} c_j = 1$, where $G_{a,b}(t)$ denotes the density of Gamma(a,b):

$$G_{a,b}(t) = \frac{t^{a-1} e^{-\frac{t}{b}}}{b^a \Gamma(a)}, \quad t > 0.$$

Note that each gamma component has its mode at the location parameter ξ_j, with the mean $\xi_j + b$ and the variance $(\xi_j + b)b$. Thus b can be understood as a smoothing parameter.

2.1. Constrained least squares

The main difference of our situation from the typical density estimation problem is that we do not observe a sample from the SPD which we want to estimate. Instead, we observe the call option prices that is directly related to the SPD through the relationships (1.1) and (1.5). Our approach is to minimize the weighted squared error loss

$$l_n(f^*) = \frac{1}{2}\sum_{i=1}^{n} w_i \left\{ Y_i - e^{-r\tau} \int_0^\infty (s - X_i)_+ f^*(s) ds \right\}^2, \tag{2.3}$$

under the no arbitrage constraint (1.4) on the mean of the SPD.

With f^* of the form (2.1), we have

$$e^{-r\tau} \int_0^\infty (s - X_i)_+ f^*(s) ds = e^{-r\tau} \sum_{j=1}^{q} c_j \{\mu_j^1(X_i) - \mu_j^0(X_i) X_i\},$$

where $\mu_j^k(z) = \int_z^\infty s^k K_j(s) ds$, $k = 0, 1$. In addition, the no arbitrage constraint (1.4) is equivalent to

$$\sum_{j=1}^{q} c_j \mu_j = S_t \exp((r - \delta)\tau), \tag{2.4}$$

where $\mu_j = \mu_j^1(0) = \int_0^\infty s K_j(s) ds$, since $\int_0^\infty s f^*(s) ds = \sum_{j=1}^{q} c_j \mu_j$.

Let $c = (c_1, \ldots, c_q)'$, $Y = (Y_1, \ldots, Y_n)'$, $W = \text{diag}(w_1, \ldots, w_n)$, and let D be the $n \times q$ matrix with (i, j) entry $d_{ij} = e^{-r\tau}\{\mu_j^1(X_i) - \mu_j^0(X_i) X_i\}$. Then (2.3) with f^* of the form (2.1) can be written as

$$\frac{1}{2} c'(D'WD)c - Y'WDc + \frac{1}{2} Y'WY. \tag{2.5}$$

We could minimize (2.5) subject to $\mathbf{1}'c = 1$, $c \geq 0$, and $\mu' c = S_t \exp((r - \delta)\tau)$, where $\mu = (\mu_1, \ldots, \mu_q)'$. More generally, we consider the l_2 penalized regularization and find c that minimizes

$$\frac{1}{2} c'(D'WD)c - Y'WDc + \frac{1}{2} Y'WY + \frac{\lambda}{2} c'c. \tag{2.6}$$

subject to $\mathbf{1}'c = 1$, $c \geq 0$, and $\mu' c = S_t \exp((r - \delta)\tau)$. The regularization parameter $\lambda \geq 0$ controls the balance between the weighted least square and the l_2 penalty. When $\lambda = 0$, we concentrate on the weighted least square. On the other hand, if $\lambda \to \infty$, the solution will be $c = (1/q, \ldots, 1/q)'$, since the quantity (2.6) is dominated by the l_2 penalty when λ is sufficiently large.

For the solution $\hat{c} = (\hat{c}_1, \ldots, \hat{c}_q)'$ to (2.6), the estimates for the SPD f^*, the call pricing function C, and the first derivative of the call pricing function C' at point x are respectively written as

$$\hat{f}^*(x) = \sum_j \hat{c}_j K_j(x),$$

$$\hat{C}(x) = e^{-r\tau} \int_0^\infty (s - x)_+ \hat{f}^*(s) ds = e^{-r\tau} \sum_j \hat{c}_j \{\mu_j^1(x) - \mu_j^0(x) x\},$$

$$\hat{C}'(x) = -e^{-r\tau} \int_z^\infty \hat{f}^*(s) ds = -e^{-r\tau} \sum_j \hat{c}_j \mu_j^0(x).$$

2.2. Tuning

In this section we discuss methods for estimation of tuning parameters. Among the most commonly used criteria for the function estimation is the mean integrated squared error(MISE):

$$\begin{aligned}\mathrm{MISE}_f &= E\int \left[\hat{f}(x) - f(x)\right]^2 dx \\ &= \int E\left\{\hat{f}(x) - E\left[\hat{f}(x)\right]\right\}^2 dx + \int \left\{E\left[\hat{f}(x)\right] - f(x)\right\}^2 dx \\ &= \int \mathrm{Var}\left[\hat{f}(x)\right] dx + \int \mathrm{Bias}^2\left[\hat{f}(x)\right] dx.\end{aligned}$$

The MISE is decomposed into the variance part and the bias part, and it can be viewed as a function of bandwidth b and regularization parameter λ. Typically, small b and λ reduce bias but increases variance, and large b and λ increases bias but reduces variance (variance-bias trade-off).

Since the observed data are call option prices, it seems appropriate for us to consider the MISE for the call pricing function $\mathrm{MISE}_C = E\int[\hat{C}(x) - C(x)]^2 dx$. For tuning in practice, we consider the cross-validation function(CV)

$$\mathrm{CV} = \frac{1}{n}\sum_{i=1}^{n} w_i \left[Y_i - \hat{C}_{-i}(X_i)\right]^2,$$

where $\hat{C}_{-i}$ is the estimate when omitting the i^{th} observation. This criterion would be justified better if Y_i follow $\mathcal{N}(C(X_i), \sigma^2/w_i)$ independently for given X_i, $i = 1, \ldots, n$.

Let DF be the degrees of freedom of the model for fixed b and λ. Then the generalized cross validation(GCV) may be considered as an approximation with simpler computation:

$$\mathrm{GCV} = \frac{\sum_{i=1}^{n} w_i \left(Y_i - \hat{Y}_i\right)^2}{(n - \mathrm{DF})^2}.$$

Alternative criteria for tuning parameters are Akaike(AIC) or Baysian(BIC) information criteria:

$$\begin{aligned}\mathrm{AIC} &= \sum_{i=1}^{n} w_i \left(Y_i - \hat{Y}_i\right)^2 \left(1 - \frac{n - 2\mathrm{DF}}{n - \mathrm{DF}}\right), \\ \mathrm{BIC} &= \sum_{i=1}^{n} w_i \left(Y_i - \hat{Y}_i\right)^2 \left(1 - \frac{n - \log n\,\mathrm{DF}}{n - \mathrm{DF}}\right).\end{aligned}$$

The degrees of freedom can be estimated as the sum of the sensitivities of the fitted values with respect to the observed values,

$$\widehat{\mathrm{DF}} = \sum_i \frac{\partial \hat{Y}_i}{\partial Y_i} = \mathrm{tr}\left(\frac{\partial \hat{Y}}{\partial Y}\right).$$

See Ye (1998) for more discussion on the degrees of freedom. Now let us consider the minimization problem (2.6). We ignore the constraint $\mu' c = S_t \exp((r - \delta)\tau)$ for the estimation of degrees of freedom as this constraint has minimal effect on the estimator in practice. Let $\mathcal{M} = \{\xi_1, \ldots, \xi_q\}$ and $\widehat{\mathcal{M}} = \{\xi_j : \hat{c}_j > 0\}$. The minimizer of (2.6) subject to $\mathbf{1}'c = 1$ and $c \geq 0$ is also the minimizer when we drop the positivity constraints $c_j \geq 0$ and concentrate on the coefficients corresponding to

$\widehat{\mathcal{M}}$. Without loss of generality, we assume that $\mathcal{M} = \widehat{\mathcal{M}}$ and drop the positivity constraints in the following discussion. Now the minimization problem can be equivalently written as

$$\frac{1}{2}c'(D'WD)c - Y'WDc + \frac{\lambda}{2}c'c - \tau\mathbf{1}'c,$$

using the Lagrange multiplier τ, and the solution is

$$\hat{c} = (D'WD + \lambda I)^{-1}(D'WY + \tau\mathbf{1}).$$

Letting $F = (D'WD + \lambda I)^{-1}$ and $\omega = (\mathbf{1}'F\mathbf{1})^{-1}$, we have

$$\tau = \omega(1 - \mathbf{1}'FD'WY)$$

from $\mathbf{1}'\hat{c} = 1$. Hence,

$$\hat{Y} = D\hat{c} = DFD'WY + \omega DF\mathbf{1} - \omega DF\mathbf{1}\mathbf{1}'FD'WY,$$

and we get $\widehat{DF}$ as follows:

$$\begin{aligned}
\mathrm{tr}\left(\frac{\partial \hat{Y}}{\partial Y}\right) &= \mathrm{tr}(WDFD') - \omega\mathrm{tr}(WDF\mathbf{1}\mathbf{1}'FD') \\
&= \mathrm{tr}(FD'WD) - \omega\mathbf{1}'FD'WDF\mathbf{1} \\
&= \mathrm{tr}(I - \lambda F) - \omega\mathbf{1}'(I - \lambda F)F\mathbf{1} \\
&= q - 1 - \lambda\ \mathrm{tr}(F) + \frac{\lambda(\mathbf{1}'F^2\mathbf{1})}{(\mathbf{1}'F\mathbf{1})}.
\end{aligned}$$

The equation $FD'WD = I - \lambda F$ comes from $(A + \lambda I)^{-1}A = I - \lambda(A + \lambda I)^{-1}$.

3. Simulation

In our simulation, we follow the settings in the simulation study of Aït-Sahalia and Duarte (2003) where the model is calibrated to match the basic features of S&P 500 index options market on a typical trading day in 1999. We set the current index price S_t at 1365, the short term interest rate at $r = 4.5\%$, the dividend yield at $\delta = 2.5\%$, and the time to maturity at $\tau = 0.119$. We assume the volatility smile is a linear function of the strike with volatility equal to 40% at the strike price 1000 and 20% at the strike price 1700, *i.e.*, $\sigma = .4 - .2(x - 1000)/700$. For the futures price $F_t = S_t \exp((r - \delta)\tau)$, define $d_1 = (\log(F_t/x) + \sigma^2\tau/2)/(\sigma\sqrt{\tau})$ and $d_2 = d_1 - \sigma\sqrt{\tau}$. Then the theoretical option price is given by $C(x) = \exp(-r\tau)(F_t\Phi(d_1) - x\Phi(d_2))$. We assume that we observe $n = 25$ option prices with strike prices equally spaced between 1000 and 1700.

The simulated option prices are determined by adding uniformly distributed random noise to the theoretical option prices. The amount of noise follows the second approach in Aït-Sahalia and Duarte (2003): the size of the noise range in percentage of the option value is a linear function of the strike with the range being $\pm 3\%$ of the option value for deep in the money options (at $x = 1000$) and $\pm 18\%$ for deep out of the money options (at $x = 1700$).

For each run, the call function, its first derivative, and the corresponding SPD are estimated by the proposed regularized mixture with the smoothing parameters tuned according to the AIC and the GCV. The strike prices are used as the knots. We assign the weights inversely proportional to the true call option prices. For each criterion, we consider two different ways of tuning: (i) fix $\lambda = 0$

Table 3.1. Mean and standard error of ISE for AIC and GCV tuning. Note that the + and − in the scale column indicate the power of 10 to multiply the result by.

ISE	Scale		b tuned, $\lambda = 0$	b and λ jointly tuned
C	+3	AIC	1.8247 (0.0244)	**1.6118** (0.0208)
		GCV	1.9727 (0.0250)	1.7583 (0.0242)
C'		AIC	0.1897 (0.0021)	**0.1375** (0.0018)
		GCV	0.2847 (0.0038)	0.2335 (0.0039)
f^*	−3	AIC	0.0496 (0.0006)	**0.0265** (0.0005)
		GCV	0.1218 (0.0025)	0.0954 (0.0025)

and tune the bandwidth b, and (ii) tune b and λ jointly. In Table 3.1, we report the mean and the standard error of the integrated squared errors(ISE)

$$\text{ISE} = \int_{800}^{1750} \left[\hat{f}^*(x) - f^*(x)\right]^2 dx$$

based on 5000 runs for the estimates with AIC and GCV tuning. The joint tuning results in a substantial improvement of the estimators in terms of the ISE, and the AIC performs superior to the GCV. Different ways in degrees of freedom estimation and weighting strategy may give different results. The BIC performs very similarly to the AIC and its results are omitted.

Figure 3.1 displays the average estimates and 95% pointwise confidence intervals for the call price function, its first derivative, and the SPD based on 5000 simulations when the smoothing parameters b and λ are jointly tuned according to AIC. The average estimates follow the theoretical curves well and the 95% confidence intervals are narrow around the theoretical values.

4. Real Example

We consider the S&P 500 index call options traded at the Chicago Board Options Exchange(CBOE). The daily closing prices of the call options that expire in August 2005 were obtained on weekdays between June 13, 2005 and July 15, 2005 except July 4 when the CBOE was closed. The closing spot prices of the S&P 500 index and the risk free interest rates (1- and 3-month Eurodollar Deposits, London) during the period are displayed in Figure 4.1.

We treat the average of closing bid and ask prices as raw call prices. The risk free interest rates are calculated by interpolating the daily 1- and 3-month Eurodollar Deposits (London). The dividend yield is implied through the put-call parity

$$P_t + S_t e^{-\delta\tau} = C_t + Xe^{-r\tau}$$

for the put-call pair at the money. We calculate the maturity as the number of days before expiration divided by 365.

The proposed method is applied to the data using the strike prices for the knots and the weights inversely proportional to the call option prices. The smoothing parameters b and λ are tuned jointly according to the AIC. Figure 4.2 shows the estimates of the call function, its first derivative, and the corresponding SPD during the week of June 13–June 17. The estimated functions satisfy all shape restrictions required. The mean of the estimated SPD is consistent with the forward value. Figure 4.3 shows the transition of the SPD during the whole period. The estimated SPD tracks the change of the forward value of the asset and it tends to get more concentrated at the forward value as the expiration approaches.

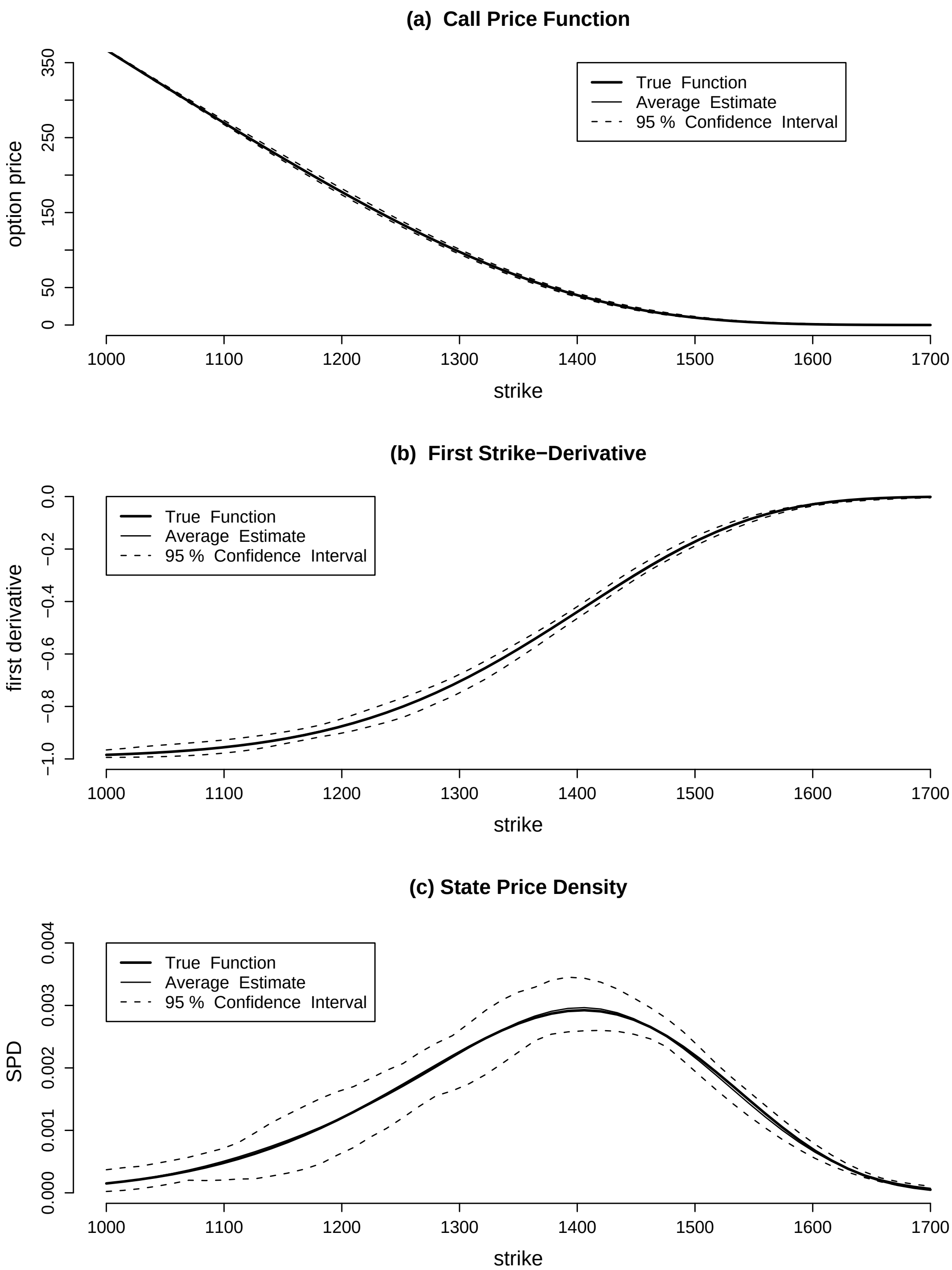

Figure 3.1. The average estimates and 95% pointwise confidence intervals for (a) the call price function, (b) its first derivative, and (c) the SPD, with jointly tuned smoothing parameters according to AIC.

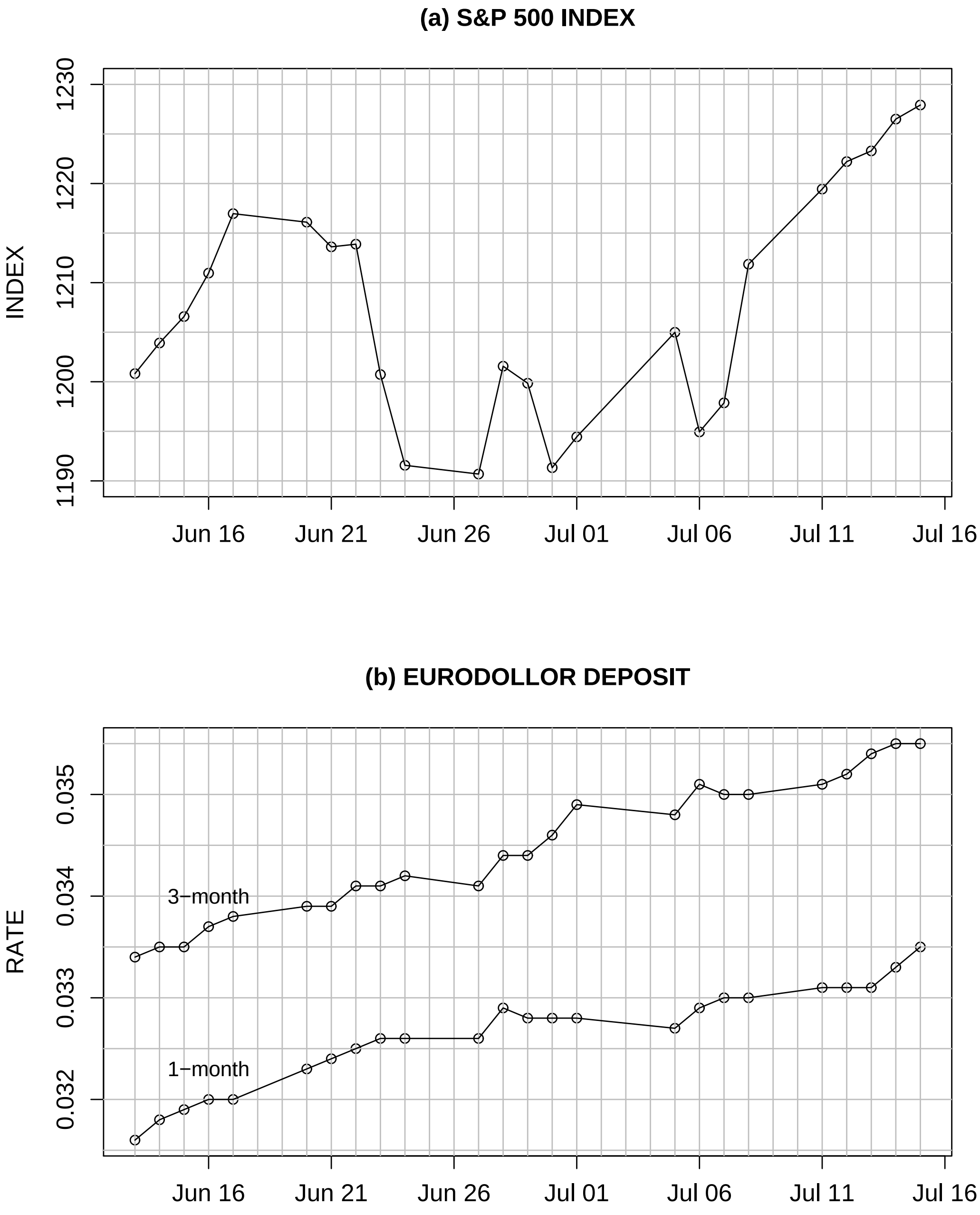

Figure 4.1. Closing spot prices of the S&P 500 index and Eurodollar Deposits between June 13, 2005 and July 15, 2005.

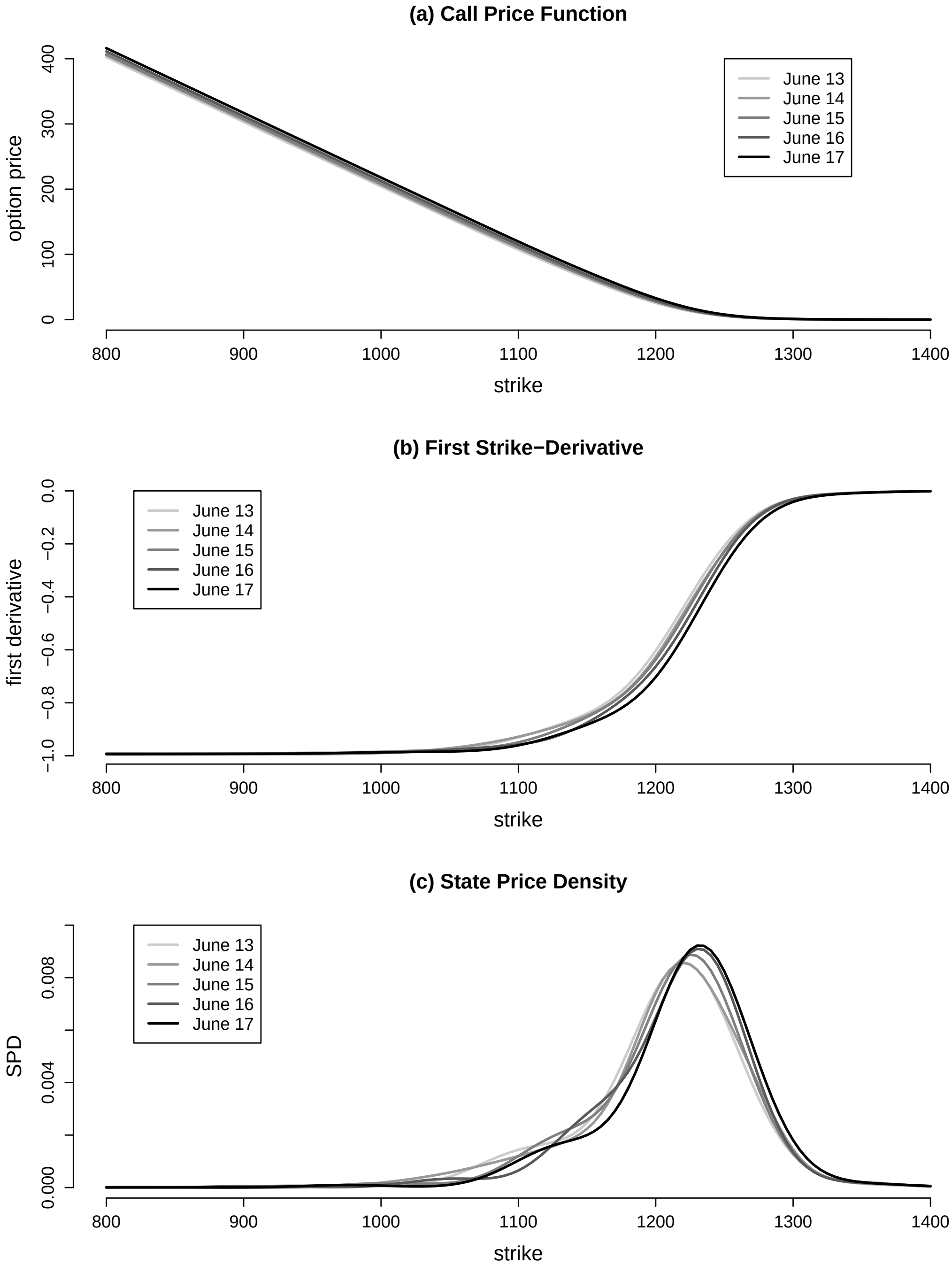

Figure 4.2. Estimated curves for (a) call price function, (b) its first derivative, and (c) the SPD in the week of June 13–June 17.

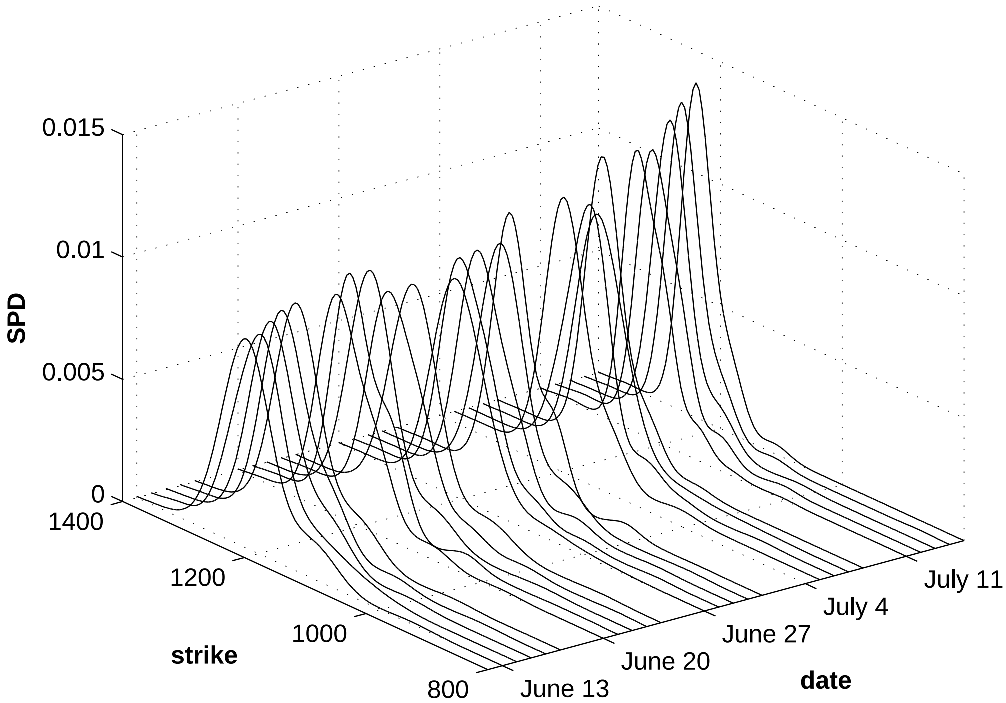

Figure 4.3. Transition of the SPD between June 13, 2005 and July 15, 2005.

5. Concluding Remarks

In this paper we proposed to estimate the SPD directly by specifying the candidate density in a flexible nonparametric way and finding the minimizer of the penalized squared error loss under no-arbitrage constraints. The estimator showed very good and stable performance in our simulation study with all restrictions required by economic theory satisfied. The method was successfully applied to call option prices on the S&P 500 index.

A new gamma kernel was proposed to avoid boundary bias in estimating the SPD supported on $[0,\infty)$. The proposed gamma kernel also can be used in the typical density estimation problems for densities supported on $[0,\infty)$. The theoretical properties of the gamma kernel density estimator of this type will require further investigation.

References

Aït-Sahalia, Y. and Duarte, J. (2003). Nonparametric option pricing under shape restrictions, *Journal of Econometrics*, **116**, 9–47.

Aït-Sahalia, Y. and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance*, **53**, 499–547.

Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices, *Journal of Business*, **51**, 621–651.

Cox, J. and Ross, S. A. (1976). The valuation of options for alternative stochastic processes, *Journal of Financial Economics*, **3**, 145–166.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties, *Statistical Science*, **11**, 89–121.

Harrison, J. M. and Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets, *Journal of Economic Theory*, **20**, 381–408.

Mammen, E., Marron, J. S., Turlach, B. A. and Wand, M. P. (2001). A general projection framework for constrained smoothing, *Statistical Science*, **16**, 232–248.

Ruppert, D., Wand, M. P. and Carroll, R. J. (2003). *Semiparametric Regression*, Cambridge University Press, Cambridge.

Wahba, G. (1990). *Spline Models for Observational Data*, SIAM, Philadelphia.

Yatchew, A. and Härdle, W. (2005). Dynamic nonparametric state price density estimation using constrained least-squares and the bootstrap, *Journal of Econometrics*, **133**, 579–599.

Ye, J. (1998). On measuring and correcting the effects of data mining and model selection, *Journal of the American Statistical Association*, **93**, 120–131.

Yuan, M. (2009). State price density estimation via nonparametric mixture, *The Annals of Applied Statistics*, **3**, 963–984.