Option Pricing with No-Arbitrage Constraints

Kay Frederik Pilz[1]

Quantitative Analysis
Sal. Oppenheim jr. & cie.
Frankfurt

Mathfinance Conference 2008

[1] Joint work with Melanie Birke, Ruhr-Universität Bochum, Germany

State Price Density

Breeden & Litzenberger (1978):

- In a complete market the continuity and linearity of the call price function implies the existence of a state price density

$$p^*(S_T) \equiv p^*(S_T \mid S_0, T, r_T, d_T),$$

where S_t denotes Spot at time t, T time to maturity and r_T, d_T interest rate and dividend yield, respectively.

State Price Density

Breeden & Litzenberger (1978):

- In a complete market the continuity and linearity of the call price function implies the existence of a state price density

$$p^*(S_T) \equiv p^*(S_T \mid S_0, T, r_T, d_T),$$

where S_t denotes Spot at time t, T time to maturity and r_T, d_T interest rate and dividend yield, respectively.

- The call price function is then given by

$$C(S_0, X, T, r_T, d_T) = e^{-r_T T} \int_0^{+\infty} (S_T - X)^+ \, p^*(S_T \mid S_0, T, r_T, d_T) \, dS_T,$$

where X is the strike of the option.

State Price Density

- Differentiating with respect to X yields

$$C'(X) \equiv \frac{\partial C(S_0, X, T, r_T, d_T)}{\partial X} = -e^{-r_T T} \int_X^{+\infty} p^*(S_T)\, dS_T .$$

State Price Density

- Differentiating with respect to X yields

$$C'(X) \equiv \frac{\partial C(S_0, X, T, r_T, d_T)}{\partial X} = -e^{-r_T T} \int_X^{+\infty} p^*(S_T)\, dS_T.$$

- Differentiating again with respect to X yields

$$C''(X) \equiv \frac{\partial^2 C(S_0, X, T, r_T, d_T)}{\partial X^2} = e^{-r_T T} p^*(X).$$

State Price Density

- Differentiating with respect to X yields

$$C'(X) \equiv \frac{\partial C(S_0, X, T, r_T, d_T)}{\partial X} = -e^{-r_T T} \int_X^{+\infty} p^*(S_T)\, dS_T.$$

- Differentiating again with respect to X yields

$$C''(X) \equiv \frac{\partial^2 C(S_0, X, T, r_T, d_T)}{\partial X^2} = e^{-r_T T} p^*(X).$$

From the first and second derivatives we get the constraints

$$\begin{array}{ll} \text{(I)} & -e^{-r_T T} \leq C'(X) \leq 0 \\ \text{(II)} & C''(X) \geq 0 \end{array}$$

State Price Density

No-arbitrage imposes bounds for the call price function,

$$C(X) \leq S_0 e^{-d_T T} \quad \text{and} \quad C(X) \geq \max(0, S_0 e^{-d_T T} - X e^{-r_T T}).$$

With (I) and (II) these bounds reduce to

$$\begin{array}{ll} \text{(III)} & C(0) = S_0 e^{-d_T T} \\ \text{(IV)} & \lim\limits_{X \to \infty} C(X) \geq 0. \end{array}$$

State Price Density

No-arbitrage imposes bounds for the call price function,

$$C(X) \leq S_0 e^{-d_T T} \quad \text{and} \quad C(X) \geq \max(0, S_0 e^{-d_T T} - X e^{-r_T T}).$$

With (I) and (II) these bounds reduce to

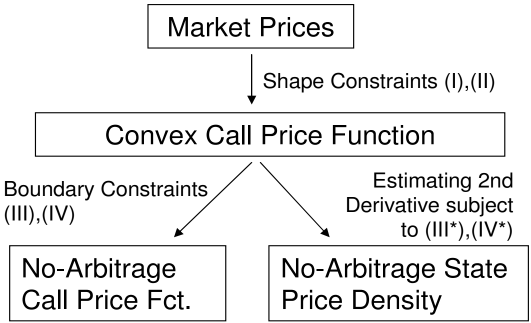
$$\begin{array}{ll} \text{(III)} & C(0) = S_0 e^{-d_T T} \\ \text{(IV)} & \lim_{X \to \infty} C(X) \geq 0. \end{array}$$

With $C''(X) = e^{-r_T T} p^*(S_T)$ one gets

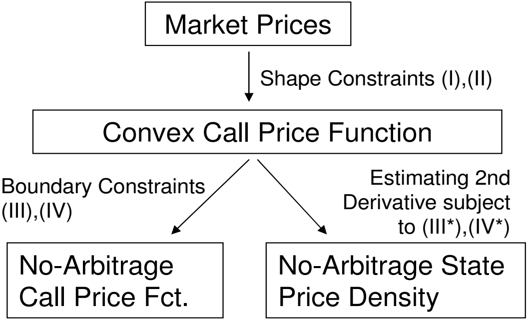
$$\begin{array}{ll} \text{(III}^*\text{)} & \int_0^{+\infty} C''(X)\, dX = e^{-r_T T} \\ \text{(IV}^*\text{)} & \int_0^{+\infty} X C''(X)\, dX = e^{-r_T T} F(0, T), \end{array}$$

where $F(0, T)$ is the instantaneous forward price with maturity T (see Aït-Sahalia & Duarte (2003)).

Constrained Estimation Procedure



Constrained Estimation Procedure



Practical limitation:
Option prices with strikes in an interval $[a, b] \subset [0, +\infty)$.

For theoretical considerations:
Without loss of generality $[a, b] = [0, 1]$.

Idea of Nonparametric Estimation

Let $\{X_i\}_{1 \leq i \leq n}$ be an i.i.d. sample of a r.v. with density f. The nonparametric density estimator is defined by

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^{n} K\Big(\frac{x - X_i}{h}\Big).$$

Idea of Nonparametric Estimation

Let $\{X_i\}_{1\leq i\leq n}$ be an i.i.d. sample of a r.v. with density f. The nonparametric density estimator is defined by

$$\hat{f}_n(x) = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x-X_i}{h}\Big).$$

Let $\{X_i, Y_i\}_{1\leq i\leq n}$ be an i.i.d. sample of a r.v. from the model

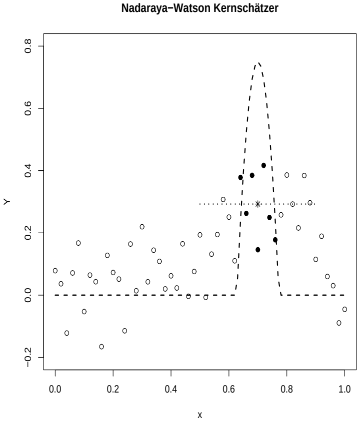
$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i,$$

(in particular with $\mathbb{E}[\varepsilon_i] = 0$ and $\mathrm{Var}(\varepsilon_i) = 1$), then the simplest non-parametric kernel regression estimator (Nadaraya-Watson type) is given by

$$\hat{m}_n(x) = \frac{\sum_{i=1}^{n} K\Big(\frac{x-X_i}{h}\Big) Y_i}{\sum_{i=1}^{n} K\Big(\frac{x-X_i}{h}\Big)}.$$

Idea of Nonparametric Regression

Interpretation of a local-polynomial estimation:



○, ●	:	Sample points with weight zero and non-zero, respectively
– – –	:	Kernel function ($x = 0.7$)
· · · · ·	:	local constant estimate

General Monotonization

$$\text{Increasing monotonicity of } \hat{C}', \ \hat{C}'(X) < 0 \text{ and } \hat{C}''(X) > 0 \text{ for all } X$$
$$\Longrightarrow \quad \text{convexity and decreasing monotonicity of } \hat{C}$$

General Monotonization

Increasing monotonicity of $\hat{C}'$, $\hat{C}'(X) < 0$ and $\hat{C}''(X) > 0$ for all X
$\Longrightarrow$ convexity and decreasing monotonicity of $\hat{C}$

HEURISTIC

$m : [0,1] \to \mathbb{R}$ strictly monotone increasing function.

- Realization $X_1, \ldots, X_n \sim \mathcal{U}[0,1]$ of uniform distributed r.v.

General Monotonization

Increasing monotonicity of $\hat{C}'$, $\hat{C}'(X) < 0$ and $\hat{C}''(X) > 0$ for all X
$\Longrightarrow$ convexity and decreasing monotonicity of $\hat{C}$

HEURISTIC

$m : [0,1] \to \mathbb{R}$ strictly monotone increasing function.

- Realization $X_1, \ldots, X_n \sim \mathcal{U}[0,1]$ of uniform distributed r.v.
- The density of the r.v. $m(X_i)$ at point x is given by

$$(m^{-1})'(x) \cdot I_{[m(0),m(1)]}(x)$$

The corresponding kernel density estimator is

$$\frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - m(X_i)}{h}\right)$$

General Monotonization

- Integration from $-\infty$ to $t \in [m(0), m(1)]$ yields $m^{-1}(t)$ for the true density and

$$\frac{1}{nh} \sum_{i=1}^{n} \int_{-\infty}^{t} K\left(\frac{u - m(X_i)}{h}\right) du$$

for the estimate.

General Monotonization

- Integration from $-\infty$ to $t \in [m(0), m(1)]$ yields $m^{-1}(t)$ for the true density and

$$\frac{1}{nh} \sum_{i=1}^{n} \int_{-\infty}^{t} K\left(\frac{u - m(X_i)}{h}\right) du$$

for the estimate.

- $m(X_i)$ is not observable. Estimate $m(X_i)$ by an unconstrained kernel estimator $\hat{m}(X_i)$:

$$\hat{m}_I^{-1}(t) = \frac{1}{Nh} \sum_{i=1}^{N} \int_{-\infty}^{t} K\left(\frac{u - \hat{m}(\frac{i}{N})}{h}\right) du$$

Here, the random uniform design X_i is replaced by N equidistant points $\frac{1}{N}, \ldots, 1$.

General Monotonization

GENERAL MONOTONIZATION PROCEDURE

Given a sample $\{X_i, Y_i\}_{1 \le i \le n}$ of the model

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i.$$

The strictly isotone estimator for the inverse regression function $m^{-1}(t)$ is given by

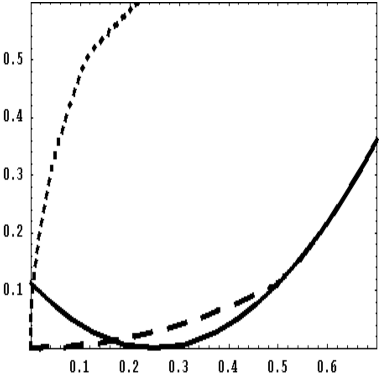
$$\hat{m}_I^{-1}(t) = \frac{1}{Nh_d} \sum_{i=1}^{N} \int_{-\infty}^{t} K_d\left(\frac{u - \hat{m}(\frac{i}{N})}{h_d}\right) du$$

The strictly isotone estimator $\hat{m}_I(x)$ for the originate function $m(x)$ results from $\hat{m}_I^{-1}(t)$ by numerical inversion.

General Monotonization

INTERPRETATION:

Monotonization as rearrangement



——— : true function

$m(x) = \frac{16}{9}(x - \frac{1}{4})^2$

······ : inverse monotone rearrangement

$$m^{*-1}(t) = \int_0^1 I\{m(x) \leq t\}\, dx = \begin{cases} \frac{3}{2}\sqrt{t} & \text{if } t \leq \frac{1}{9} \\ \frac{1}{4}(3\sqrt{t}+1) & \text{if } t > \frac{1}{9} \end{cases}$$

— — — : monotone rearrangement

$$m^{*}(x) = \inf\{t \;:\; m^{*-1}(t) > x\} = \begin{cases} \frac{4}{9}x^2 & \text{if } x \leq \frac{1}{2} \\ \frac{16}{9}(x-\frac{1}{4})^2 & \text{if } x > \frac{1}{2} \end{cases}$$

Call Price Function Estimation

To achieve the convexity for the estimate $\hat{C}$ we monotonize the estimate $\hat{C}'$.

- $\tilde{C}'(x)$ is the derivative of an unconstrained estimator $\tilde{C}(x)$ (e.g. the Nadaraya-Watson estimator), i.e. $\tilde{C}'(X) \equiv \frac{\partial}{\partial x}\tilde{C}(x)$.

Call Price Function Estimation

To achieve the convexity for the estimate $\hat{C}$ we monotonize the estimate $\hat{C}'$.

- $\tilde{C}'(x)$ is the derivative of an unconstrained estimator $\tilde{C}(x)$ (e.g. the Nadaraya-Watson estimator), i.e. $\tilde{C}'(X) \equiv \frac{\partial}{\partial x}\tilde{C}(x)$.
- Using the monotonization procedure we obtain an estimate for the inverse derivative of the call price function $(C')^{-1}(t)$,

$$\hat{\psi}(t) \equiv \frac{1}{h_d} \int_0^1 \int_{-\infty}^t K_d\left(\frac{u - \tilde{C}'(v)}{h_d}\right) du dv.$$

Call Price Function Estimation

To achieve the convexity for the estimate $\hat{C}$ we monotonize the estimate $\hat{C}'$.

- $\tilde{C}'(x)$ is the derivative of an unconstrained estimator $\tilde{C}(x)$ (e.g. the Nadaraya-Watson estimator), i.e. $\tilde{C}'(X) \equiv \frac{\partial}{\partial x}\tilde{C}(x)$.
- Using the monotonization procedure we obtain an estimate for the inverse derivative of the call price function $(C')^{-1}(t)$,

$$\hat{\psi}(t) \equiv \frac{1}{h_d} \int_0^1 \int_{-\infty}^t K_d\left(\frac{u - \tilde{C}'(v)}{h_d}\right) du dv.$$

- Inverting $\hat{\psi}$ yields $\hat{C}'(X) \equiv \hat{\psi}^{-1}(X)$.

Call Price Function Estimation

- Finally, integrating yields the estimator for the call price function, which is strictly convex, because $\hat{\psi}^{-1}$ is strictly increasing by construction,

$$\boxed{\hat{C}(X, u_0) \equiv \tilde{C}(u_0) + \int_{u_0}^{X} \hat{C}'(z)\, dz.}$$

Call Price Function Estimation

- Finally, integrating yields the estimator for the call price function, which is strictly convex, because $\hat{\psi}^{-1}$ is strictly increasing by construction,

$$\boxed{\hat{C}(X, u_0) \equiv \tilde{C}(u_0) + \int_{u_0}^{X} \hat{C}'(z)\, dz.}$$

Two further points have to be considered:

(1) Condition (I) claims $-e^{-r_T T} \leq C'(X) \leq 0$. Therefore we redefine $\hat{C}'$:

$$\begin{aligned} \hat{C}'(X) \quad \equiv \quad & \hat{\psi}^{-1}(X) \cdot I\{-e^{-r_T T} \leq \hat{\psi}^{-1}(X) \leq 0\} \\ & -e^{-r_T T} \cdot I\{\hat{\psi}^{-1}(X) < -e^{-r_T T}\} + 0 \cdot I\{0 < \hat{\psi}^{-1}(X)\} \end{aligned}$$

Call Price Function Estimation

(2) Choice of the initial point $u_0 \in [0, 1]$:

Call Price Function Estimation

(2) Choice of the initial point $u_0 \in [0, 1]$:
- Irrelevant from an asymptotic point of view (see Theorem 1)

Call Price Function Estimation

(2) Choice of the initial point $u_0 \in [0, 1]$:

- Irrelevant from an asymptotic point of view (see Theorem 1)
- For finite sample sizes it can be shown that an average over the initial points u_0 of the estimates $\hat{C}(X, u_0)$ is optimal regarding to a least squares criterion (see Theorem 2).

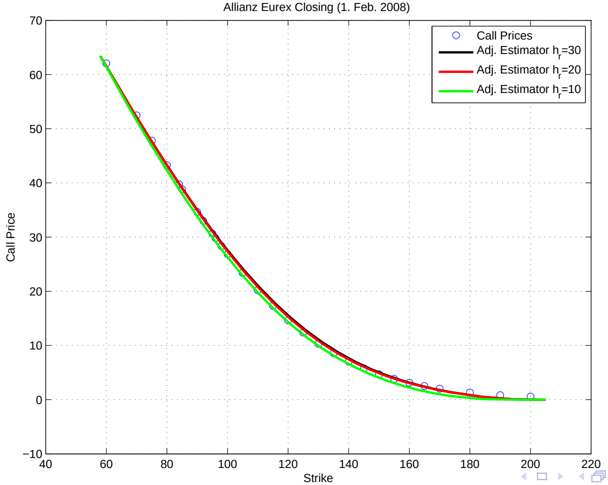
$$\hat{C}(X) \equiv \frac{1}{m}\sum_{j=1}^{m} \hat{C}(X, u_{0,j})$$

$$\equiv \frac{1}{m}\sum_{j=1}^{m}\left[\tilde{C}(u_{0,j}) + \int_{u_{0,j}}^{X} \hat{C}'(z)\,dz\right]$$

Example Call Price Function

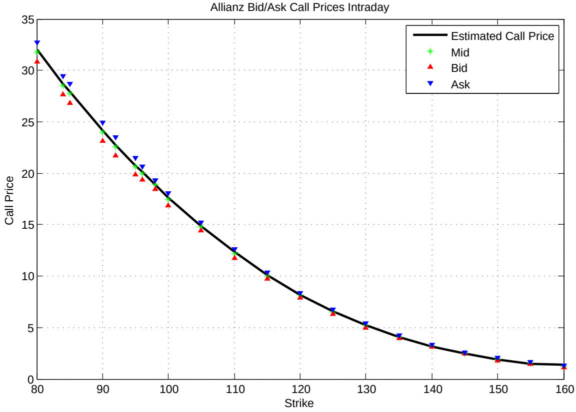
Eurex Closing Call Prices for Allianz from 1st February 2008

(Spot = 121.23, rate = 4.22%, time to maturity = 329 days)



Example Call Price Function

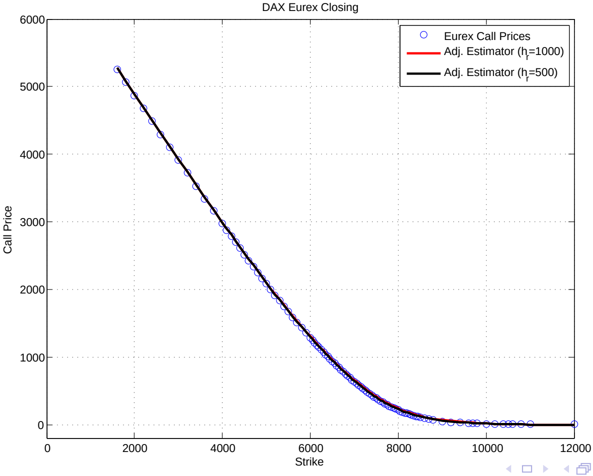
Eurex Call Prices for Allianz Intraday Snapshot from 13$^{\text{th}}$ March 2008
(Spot = 109.30, i-rate = 4.40%, div-yield = 6.66%, time to maturity = 280 days)



Example Call Price Function

Eurex Closing Call Prices for the DAX from 25$^{\text{th}}$ January 2008

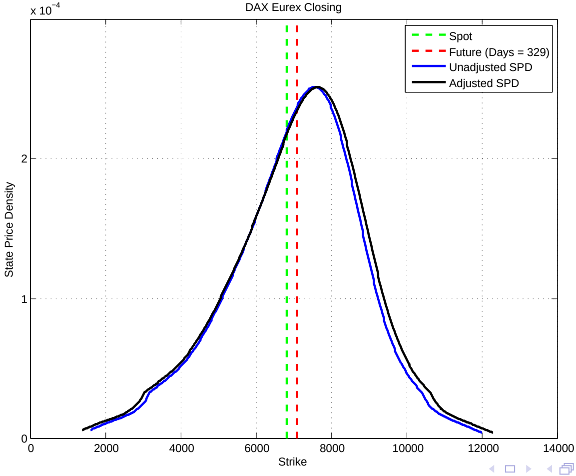
(Spot $= 6816.74$, rate $= 4.19\%$, time to maturity $= 322$ days)



Example State Price Density

Eurex Closing Call Prices for the DAX from 25$^{\text{th}}$ January 2008

(Spot = 6816.74, rate = 4.19%, time to maturity = 322 days)



Constraints (III) and (IV)

Case: Estimating call price function $C(X)$

Final modification to guarantee the constraints

$$\begin{array}{ll} \text{(III)} & C(X) \leq S_0 e^{-d_T T}, \\ \text{(IV)} & C(X) \geq \max(0, S_0 e^{-d_T T} - X e^{-r_T T}). \end{array}$$

Shifting the initial value by α and scaling the integrand by β leads to

$$\hat{C}(X, u_0) = \alpha + \tilde{C}(u_0) + \beta \cdot \int_{u_0}^{X} \hat{C}'(z)\, dz.$$

Constraints (III) and (IV)

Case: Estimating call price function $C(X)$

Final modification to guarantee the constraints

$$\begin{array}{ll} \text{(III)} & C(X) \leq S_0 e^{-d_T T}, \\ \text{(IV)} & C(X) \geq \max(0, S_0 e^{-d_T T} - X e^{-r_T T}). \end{array}$$

Shifting the initial value by α and scaling the integrand by β leads to

$$\hat{C}(X, u_0) = \alpha + \tilde{C}(u_0) + \beta \cdot \int_{u_0}^{X} \hat{C}'(z)\, dz.$$

The values for α and β are chosen such that the constraints

$$\hat{C}(0, u_0) \;=\; S_0 e^{-d_T T}, \qquad \hat{C}(1, u_0) \;\geq\; 0$$

are satisfied.

Constraints (III) and (IV)

Force the values α, β near to their „optimal" values 0 and 1, respectively by a loss function, e.g.

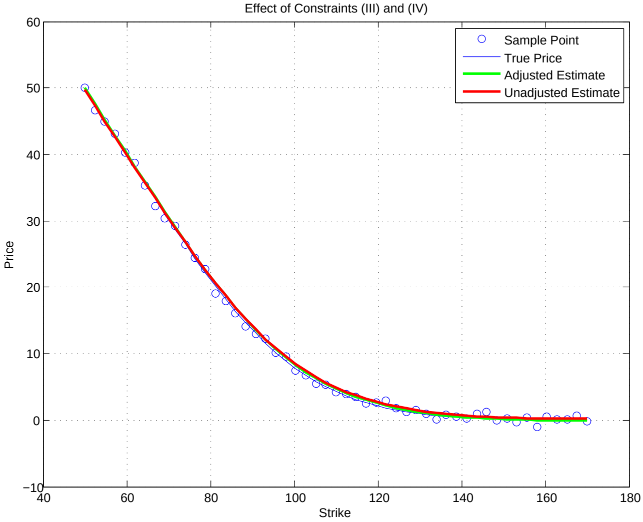
$$L(\alpha, \beta) = \alpha^2 + (1 - \beta)^2 \longrightarrow \text{min!}$$

Example:

Spot = 100, time to maturity = 1 year, interest = 0, dividend yield = 0, error std = 0.5
True values versus adjusted and unadjusted estimates:

Strike	true	adjusted	unadjusted
50.00	49.99	50.00	49.73
52.40	47.59	47.60	47.36
...	...	...	...
167.60	0.039	0.000	0.236
170.00	0.032	0.000	0.236

The corrections are typically very small:



Constraints (III) and (IV)

Case: Estimating SPD $p^*(X)$

The correction can be made as described in Ait-Sahalia and Duarte (2003).

- For condition (III*) the estimator $\hat{C}''(X)$ ist normalized by replacing

$$\hat{C}''(X) \quad \text{with} \quad \exp(-r_T T)\hat{C}''(X)\Big(\int_0^{+\infty} \hat{C}''(X)\Big)^{-1}$$

Constraints (III) and (IV)

Case: Estimating SPD $p^*(X)$

The correction can be made as described in Ait-Sahalia and Duarte (2003).

- For condition (III*) the estimator $\hat{C}''(X)$ ist normalized by replacing

$$\hat{C}''(X) \quad \text{with} \quad \exp(-r_T T)\hat{C}''(X)\Big(\int_0^{+\infty} \hat{C}''(X)\Big)^{-1}$$

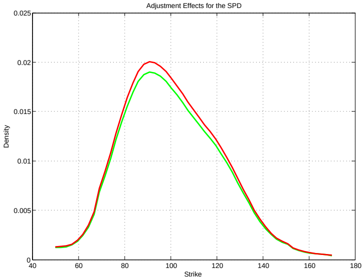
- For condition (IV*) a translation parameter z is computed, such that

$$\exp(r_T T) \int\limits_0^{+\infty} X\, \hat{C}''(X - z)\, dX = F_T(t)$$

is satisfied.

Constraints (III) and (IV)

Same Example as for the call price function.



The definite integrals

$$\int_{50}^{170} p^*(X)\, dX$$

are for the

unadjusted SPD	1.055
adjusted SPD	1.000

Asymptotic Results

<u>MODEL ASSUMPTIONS:</u>

$C(X)$ is a strictly convex function, and $\{X_i, Y_i\}_{1\le i\le n}$ is an i.i.d. sample of the regression model

$$Y_i = C(X_i) + \sigma(X_i)\varepsilon_i,$$

where

- $C(X) \in \mathcal{C}^3[0,1]$,
- $\{X_i\}_{1\le i\le n}$ are i.i.d. random variables, with
- density $f : [0,1] \to \mathbb{R}$, $f \in \mathcal{C}^3[0,1]$, f bounded away from zero
- $\{\varepsilon_i\}_{1\le i\le n}$ i.i.d. with $\mathbb{E}[\varepsilon_i] = 0$, $\mathrm{Var}(\varepsilon_\mathrm{i}) = 1$, independent from $\{X_i\}_{1\le i\le n}$, $\mathbb{E}[\varepsilon_i^4] < \infty$,
- $\sigma^2 : [0,1] \to \mathbb{R}^+$, σ continuous.

Asymptotic Results

Theorem (1)

Let the unconstrained estimator $\tilde{C}(X)$ be of Nadaraya-Watson type with kernel K_r of order three and bandwidth h_r.
Kernel K_d and bandwidth h_d refer to the monotonization procedure.

If the model assumptions are satisfied and the smoothing parameters satisfy the conditions

$$h_d, h_r \to 0, \quad nh_d, nh_r \to \infty,$$
$$h_d / h_r^{3/2} \to 0, \quad nh_r^7 = O(1), \quad (\log h_r^{-1})^{3/2} / nh_r^5 h_d = o(1),$$

then for every $X \in (0,1)$ with $C''(X) > 0$ and any $u_0 \in (0,1)$,

$$\sqrt{nh_r}(\hat{C}(X, u_0) - C(X) - b_n(X)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \gamma(X)),$$

Asymptotic Results

Theorem (1 Ctd.)

with asymptotic bias and variance given by

$$b_n(X) = h_r^3 \kappa_3(K_r) \frac{(Cf)^{(3)} - Cf^{(3)}}{f}(X) + o(h_r^3)$$
$$\gamma(X) = \int_{-1}^{1} K_r^2(y)\, dy \cdot \left(\frac{\sigma^2}{f}\right)(X),$$

where the constant $\kappa_3(K_r) = 1/3! \int_{-1}^{1} u^3 K_r(u)\, du$ *depends only on the kernel* K_r.

Choice of u_0

- If $\hat{C}'$ is strictly increasing, i.e. $\tilde{C}$ is strictly convex, then for $h_d \to 0$

$$\hat{C}(X, u_0) \longrightarrow \tilde{C}(X) \quad \text{for all } X, u_0 \in (0, 1).$$

Choice of u_0

- If $\hat{C}'$ is strictly increasing, i.e. $\tilde{C}$ is strictly convex, then for $h_d \to 0$

$$\hat{C}(X, u_0) \longrightarrow \tilde{C}(X) \quad \text{for all } X, u_0 \in (0,1).$$

- But if $\hat{C}'$ is not strictly increasing, i.e. $\tilde{C}$ is not strictly convex, each u_0 defines a different convex rearrangement.

Choice of u_0

- If $\hat{C}'$ is strictly increasing, i.e. $\tilde{C}$ is strictly convex, then for $h_d \to 0$

$$\hat{C}(X, u_0) \longrightarrow \tilde{C}(X) \quad \text{for all } X, u_0 \in (0, 1).$$

- But if $\hat{C}'$ is not strictly increasing, i.e. $\tilde{C}$ is not strictly convex, each u_0 defines a different convex rearrangement.

Define a loss criterion by using a L^2 norm,

$$L(u_0) \equiv \int_0^1 (C(X) - \hat{C}(X, u_0))^2 \, dX,$$

Choice of u_0

Theorem (2)

Let φ denote any function with $\varphi' = \hat{C}'$. A point $u_0^* \in (0,1)$ minimizes the function L if and only if

$$C(u_0^*) - \varphi(u_0^*) = \int_0^1 (C(u) - \varphi(u))\, du.$$

Moreover, for any u_0^* minimizing the above defined loss criterion we have

$$\hat{C}(X, u_0^*) = \int_0^1 \hat{C}(X, u)\, du.$$