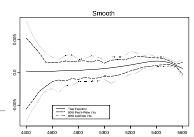
Dynamic Nonparametric State Price Density Estimation using Constrained Least Squares and the Bootstrap

Wolfgang Härdle

Institut für Statistik und Ökonometrie
CASE – Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin

Adonis Yatchew

Department of Economics
University of Toronto



State Price Densities

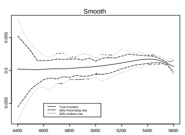
The State Price Density (SPD) describes the distribution of an asset S_t at future time T.

The SPD depends upon (t, T, r) and on volatility.

Using option prices, one may recover the SPD.

The SPD assigns probabilities to various possible values of the underlying asset at the option's expiration.

SPD differences may be used for trading strategies.



Geometric Brownian Motion (GBM)

$$\frac{dS_t}{S_t} = rdt + \sigma dZ_t,$$

yields Black Scholes (BS) SPD:

$$p_{BS}(S_t, S_T, r, \tau) = \frac{1}{S_T\sqrt{2\pi\sigma^2\tau}} \exp\left[-\frac{\left[\ln\left(\frac{S_T}{S_t}\right) - (r - \frac{\sigma^2}{2})\tau\right]^2}{2\sigma^2\tau} \right],$$

where $\tau = T - t$, S_t is the stock price at time t,
r interest rate, σ constant volatility.

Some concepts

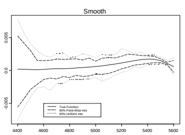
Let X be the strike price at time T, $p(S_t, S_T, r, \tau)$ the SPD.

Relation between the option prices and the SPD,

$$
\begin{aligned}
C(X,T) &= e^{-r\tau}\int_0^{+\infty}(S_T - X)^+ p(S_t, S_T, r, \tau)dS_T,\\
P(X,T) &= e^{-r\tau}\int_0^{+\infty}(X - S_T)^+ p(S_t, S_T, r, \tau)dS_T.
\end{aligned}
$$

Using the assumption of risk-neutrality

$$p(S_t, S_T, r, \tau) = p(S_T|S_t, r, \sigma).$$

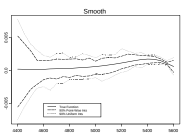


BS Formula

$$C(X,\tau) = S_t\Phi(d_1) - Xe^{-r\tau}\Phi(d_2)$$

$$d_1 = \frac{\ln(S_t/X) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} \qquad d_2 = d_1 - \sigma\sqrt{\tau}$$

Φ	CDF of the standard normal distribution	r	Interest rate
S_t	Asset price	$\tau = T - t$	Time to maturity
X	Strike price	σ	Constant volatility parameter

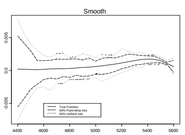


BS Formula

Example

Suppose $S_t = 100$, $X = 90$, $r = 3\%$, $\tau = 2$, and $\sigma = 10\%$.
Call price $C(X, \tau) = 16.018$ and put price $P(X, \tau) = 0.77663$.
Alternative calculation of $P(X, \tau)$ by the put-call-parity:

$$
\begin{aligned}
C(X, \tau) - P(X, \tau) &= S_t - Xe^{-r\tau}, \\
16.018 - 0.77663 &= 100 - 90e^{(-0.03 \cdot 2)}.
\end{aligned}
$$

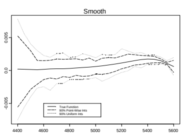


BS Implied Volatilities

However, σ is unknown! Hence define the volatility $\sigma_{imp}(X_j, \tau)$ *implied* by observed market option prices $C(X_j, \tau)$ as

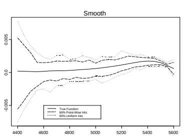
$$\sigma_{imp}: \quad C(X_j, \tau) - C_{BS}(X_j, \tau | S_t, r, \sigma_{imp}) = 0$$

This solution may be found using a Newton-Raphson or a bisection algorithm. It is unique as the BS formula is globally concave in σ.



Empirical Findings

- Implied volatility is not constant across time t.
- Implied volatility is not flat across strikes.
- Implied volatility is not flat across time to maturity.
- Implied volatility became asymmetric since the 1987 stock market crash.
- SPD is not lognormal and varies with t.
- SPD is different from density estimate based on history.



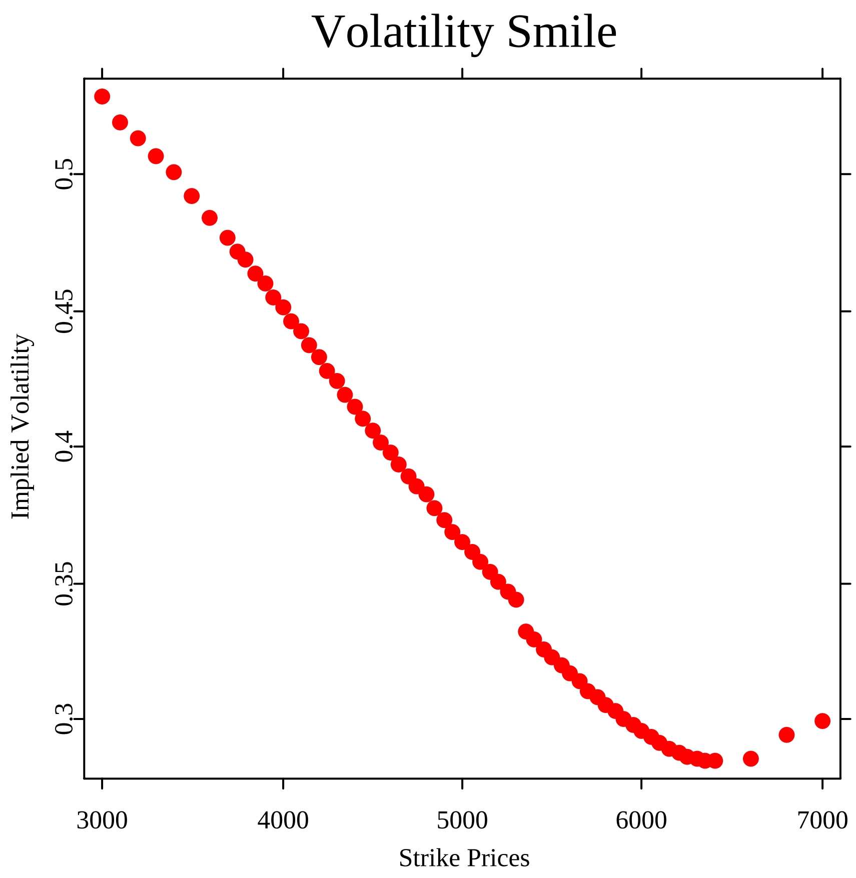
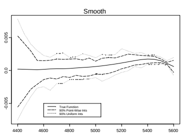


Figure 1: *Implied Volatility Smile: 6 months to expiry,* $t = 990104$, *ODAX*

XFGiv01.xpl



The SPD is proportional to the Arrow-Debreu price distribution at T,

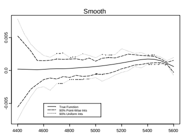
$$p(S_t, S_T, \tau) = e^{r\tau} \frac{\partial^2 C(X, \tau)}{\partial X^2}.$$

$C(\cdot, \tau)$ is monotonic decreasing and convex.

Theory does not prescribe the functional form of $C(\cdot, \tau)$.

For dynamic comparison, confidence intervals are necessary.

The bootstrap provides nonparametric confidence bounds that help in setting up trading strategies.



Notation

Call price function

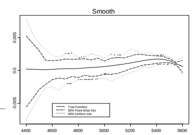
$$C(x,\tau) = m(x), \qquad (\tau \quad \text{fixed})$$

Observe several call prices at each distinct strike price X_j, $j = 1, \ldots, k$,

$$y_i = m(x_i) + \varepsilon_i, \qquad i = 1, \cdots, n$$

with error structure

$$\sigma^2(X_1), \sigma^2(X_2), \cdots, \sigma^2(X_k), \Sigma = \text{diag}\{\sigma^2(x_1), \sigma^2(x_2), \cdots, \sigma^2(x_n)\}$$



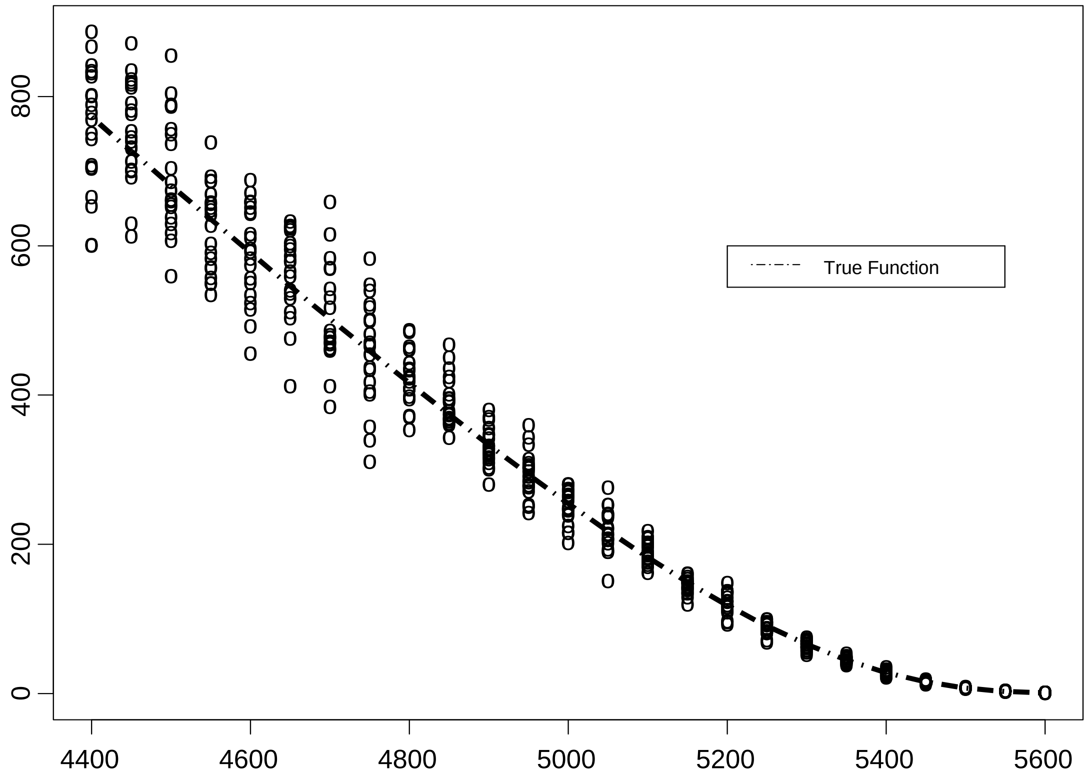
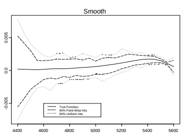


Figure 2: *Date and estimated call function,* $n = 500, k = 25, X_1 = 4400, X_2 = 4450, ..., ATM = 5290$*, on Jan 4, 1999.*



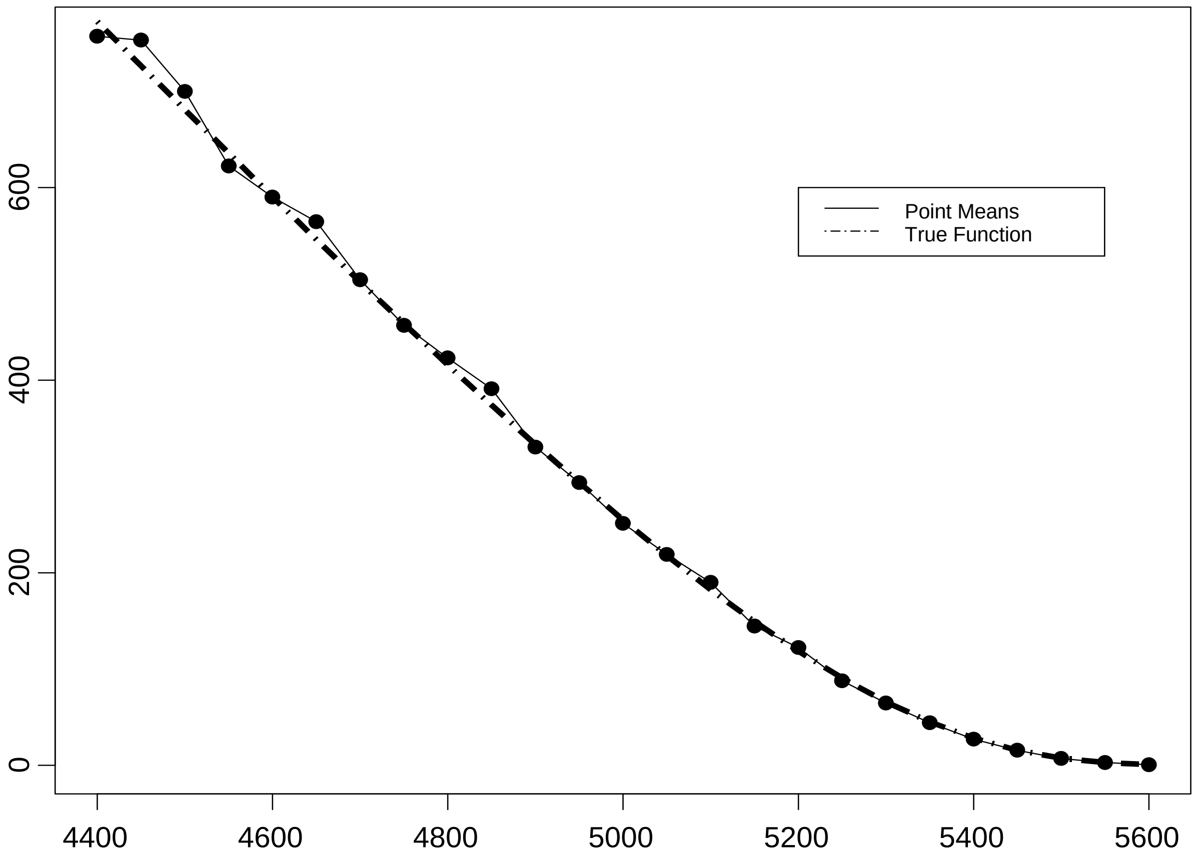
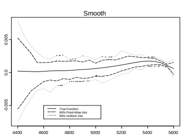


Figure 3: *Estimate of* $m(X_j)$ *using point means.*



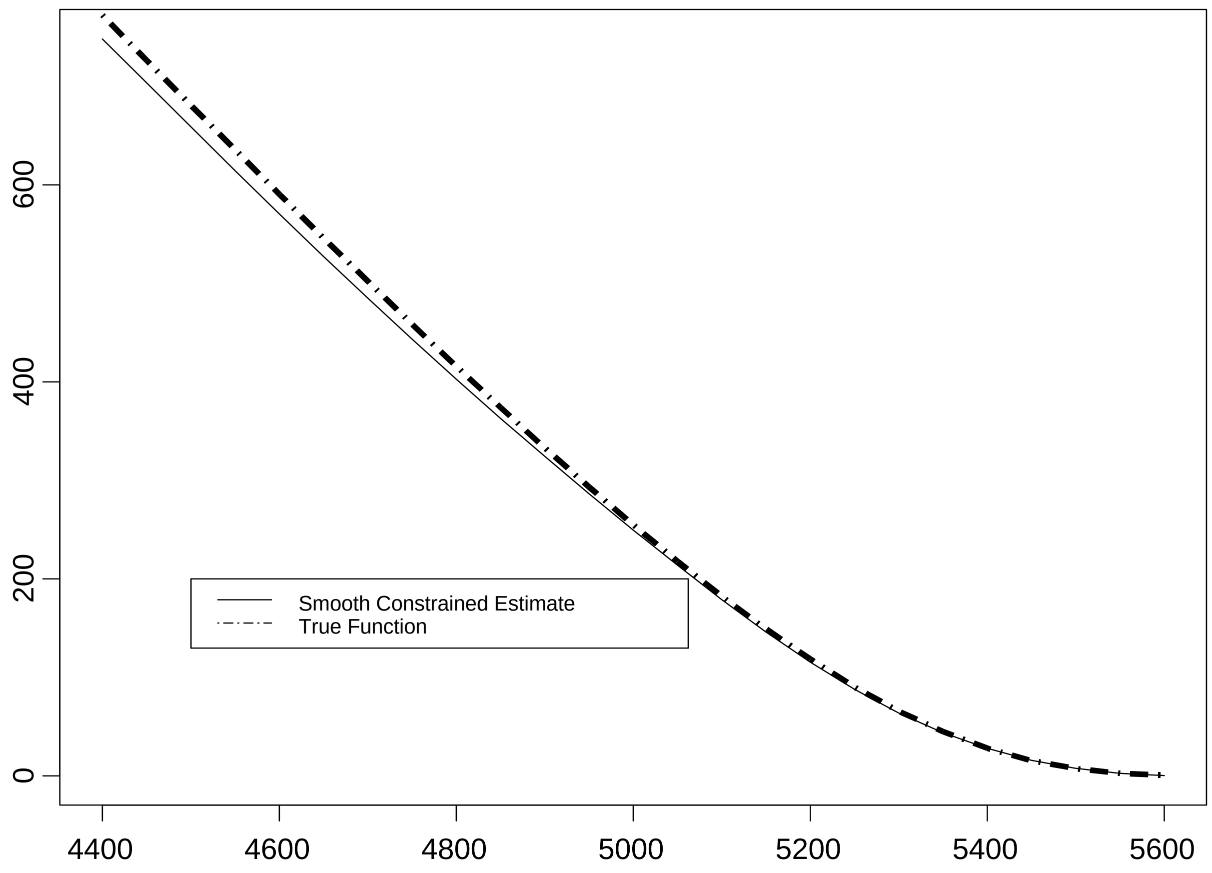
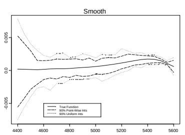


Figure 4: *Estimate of* $m(X_j)$ *constrained to be smooth, decreasing and convex.*



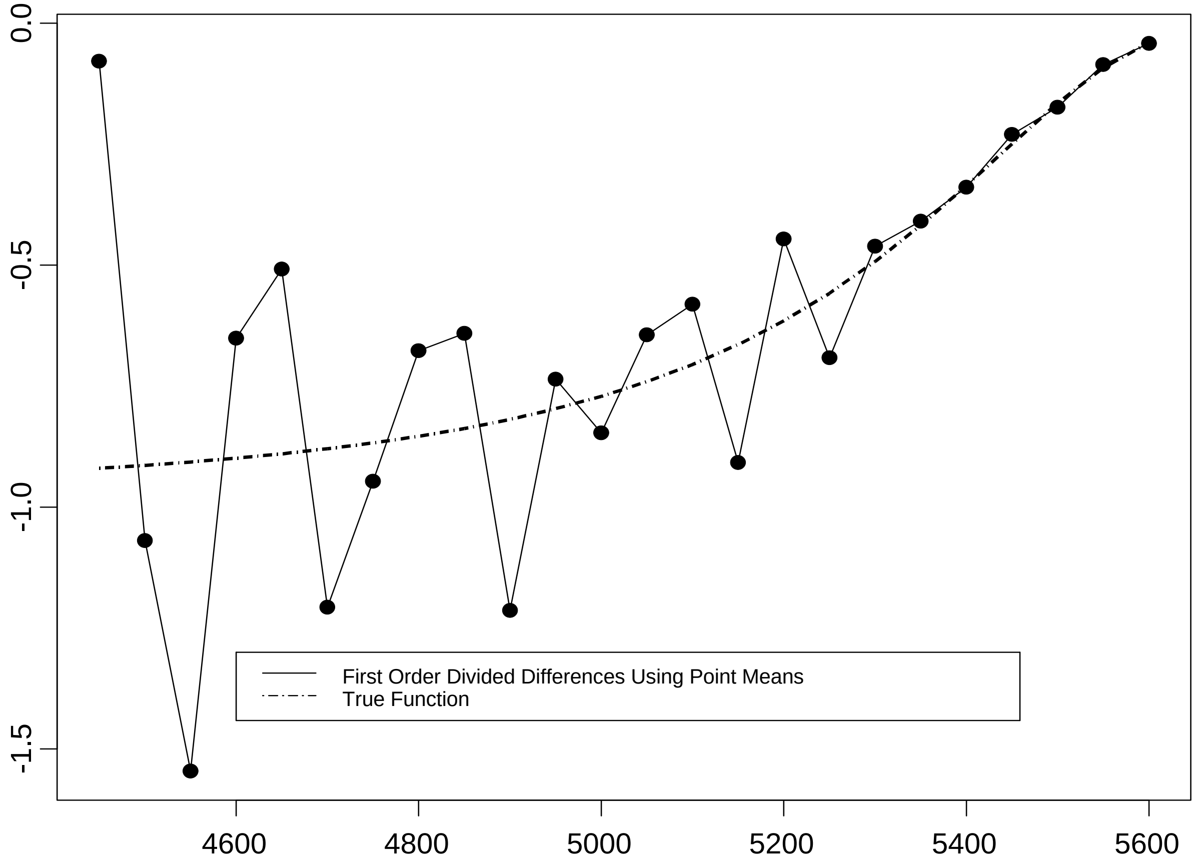
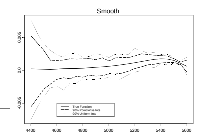


Figure 5: *First derivative $m^{(1)}(x)$ using divided differences of point means.*



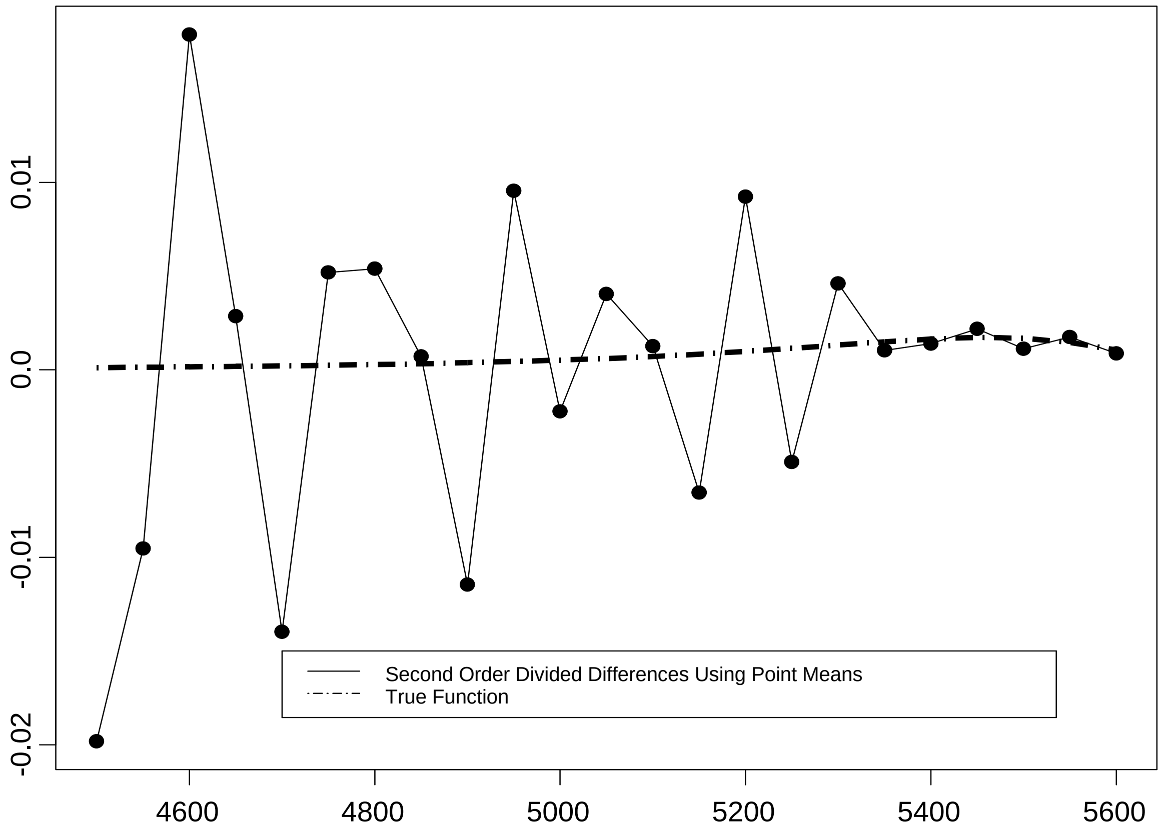


Figure 6: *Second derivative using divided differences of point means.*

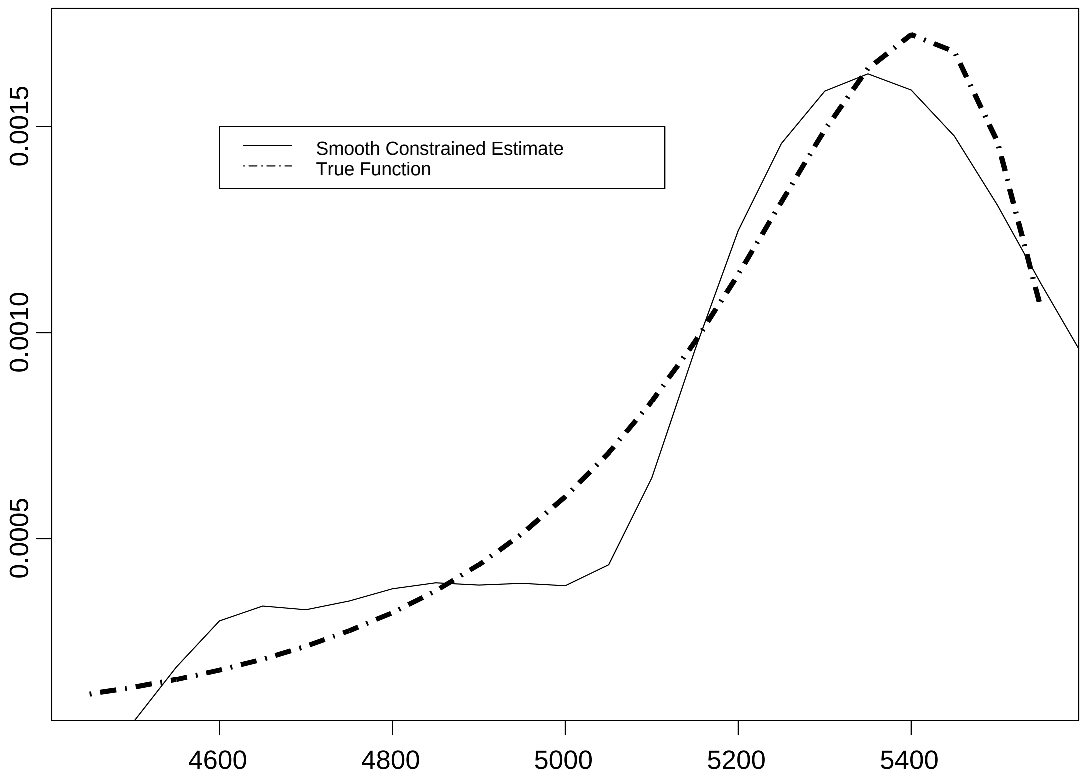
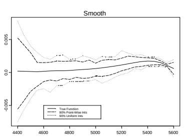
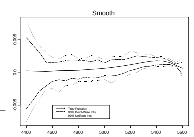


Figure 7: *SPD smooth restricted estimate (SRE)*



Overview

√ 1. Introduction and Motivation
2. Constrained nonparametric procedure
3. Monte Carlo Results
4. Application to the DAX
5. Trading rules



Constrained nonparametric procedure

$$\begin{aligned} y_i &= m(x_i) + \varepsilon_i, \qquad i = 1, \cdots, n, \\ m^{(l)}(x) &= l - \text{th} \quad \text{derivative.} \end{aligned}$$

Sobolev norm

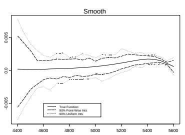
$$< m_1, m_2 >= \int m_1 m_2 + m_1^{(1)} m_2^{(1)} + m_1^{(2)} m_2^{(2)} + m_1^{(3)} m_2^{(3)} + m_1^{(4)} m_2^{(4)}$$

Estimation

$$\min_m n^{-1} \sum_{i=1}^{n} \{y_i - m(x_i)\}^2 \qquad s.t. \quad \|m\|^2 \leq L.$$

To accommodate multiple observations at each distinct strike price X_j, define

$$B = (B_{i,j}), \qquad B_{i,j} = \begin{cases} 1, & \text{if } x_i = X_j \\ 0, & \text{otherwise} \end{cases}$$



Estimation

$$\min_m n^{-1} \{y - Bm(X)\}^T \Sigma^{-1} \{y - Bm(X)\} \qquad s.t. \quad \|m\|^2 \leq L.$$

This is a finite dimensional problem!

Riesz representation

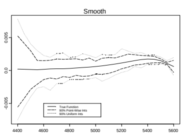
$$h(X_0) = < r_{X_0}, h >, \qquad r \quad \text{representor.}$$

Representor matrix

$$R = R_{i,j}, \quad R_{i,j} = < r_{X_i}, r_{X_j} > = r_{X_i}(X_j).$$

Equivalent problem

$$\min_c n^{-1}(y - BRc)^T \Sigma^{-1}(y - BRc), \qquad s.t. \quad c^T Rc \leq L.$$



Smooth Restricted Estimate(SRE)

$$\hat{m}^{(l)}(X) = \sum_{j=1}^{k} \hat{c}_j r_{X_j}^{(l)}(X),$$

with

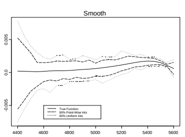
$$R_{i,j}^{(l)} = r_{X_i}^{(l)}(X_j).$$

Impose constraints:

monotonity: $R^{(1)}c \leq 0$

convexity: $R^{(2)}c \geq 0$

Other possible constraints: SPD unimodal with mode in some interval



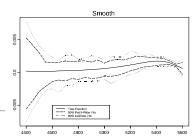
Proposition 1: Let $\overline{y}(X) = \overline{y}(X_1, \cdots, X_k)$ be the k dimensional vector of average transaction prices. Define $\hat{m}(X) = R\hat{c}$, then

$$\hat{m}(X) - \overline{y}(X) \quad \xrightarrow{P} \quad 0.$$

Proposition 2: Take $m^{(l)}(X) = R^{(l)}\hat{c}$ and define $\Omega/n = \mathrm{Var}\{\overline{y}(X)\}$. Then for $l = 0, 1, 2,$

$$\sqrt{n}(\hat{c} - c) \quad \xrightarrow{\mathcal{L}} \quad N\left(0, R^{-1}\Omega R^{-1}\right),$$

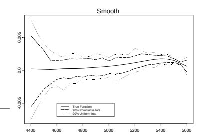
$$\sqrt{n}\left\{\hat{m}^{(l)}(X) - m^{(l)}(X)\right\} \quad \xrightarrow{\mathcal{L}} \quad N\left(0, R^{(l)}R^{-1}\Omega R^{-1}R^{(l)}\right).$$



Bootstrap procedure

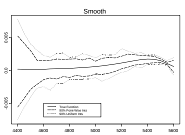
1. Calculate residuals $\hat{\varepsilon} = y - BR\hat{c}$
2. Use wild bootstrap to produce $y^* = BR\hat{c} + \varepsilon^*$
3. Recalculate $\hat{m}^{(l)*}(x)$ for the bootstrap sample
4. Approximate $\mathcal{L}\left\{\hat{m}^{(l)}(X) - m^{(l)}(X)\right\}$ by $\mathcal{L}\left\{\hat{m}^{(l)*}(X) - \hat{m}^{(l)}(X)\right\}$
5. Obtain pointwise and uniform confidence interval. For example

$$
\begin{aligned}
1 - \alpha &\cong P\left\{(\hat{m}^* - \hat{m}) \in [a^*, b^*]\right\} \\
&\cong P\left\{(\hat{m} - m) \in [a^*, b^*]\right\} \\
&= P\left\{\hat{m} - b^* \leq m \leq \hat{m} - a^*\right\}.
\end{aligned}
$$



Average MSE

model	Call Function	$m^{(1)}(x) * (10^{-4})$	SPD$(*10^{-7})$
Unconstrained	200	1219	1490
Smooth	26	52	12.6
Smooth, monotone, convex	14	15	3.8
Smooth, monotone, convex, unimodal	13	9.7	2



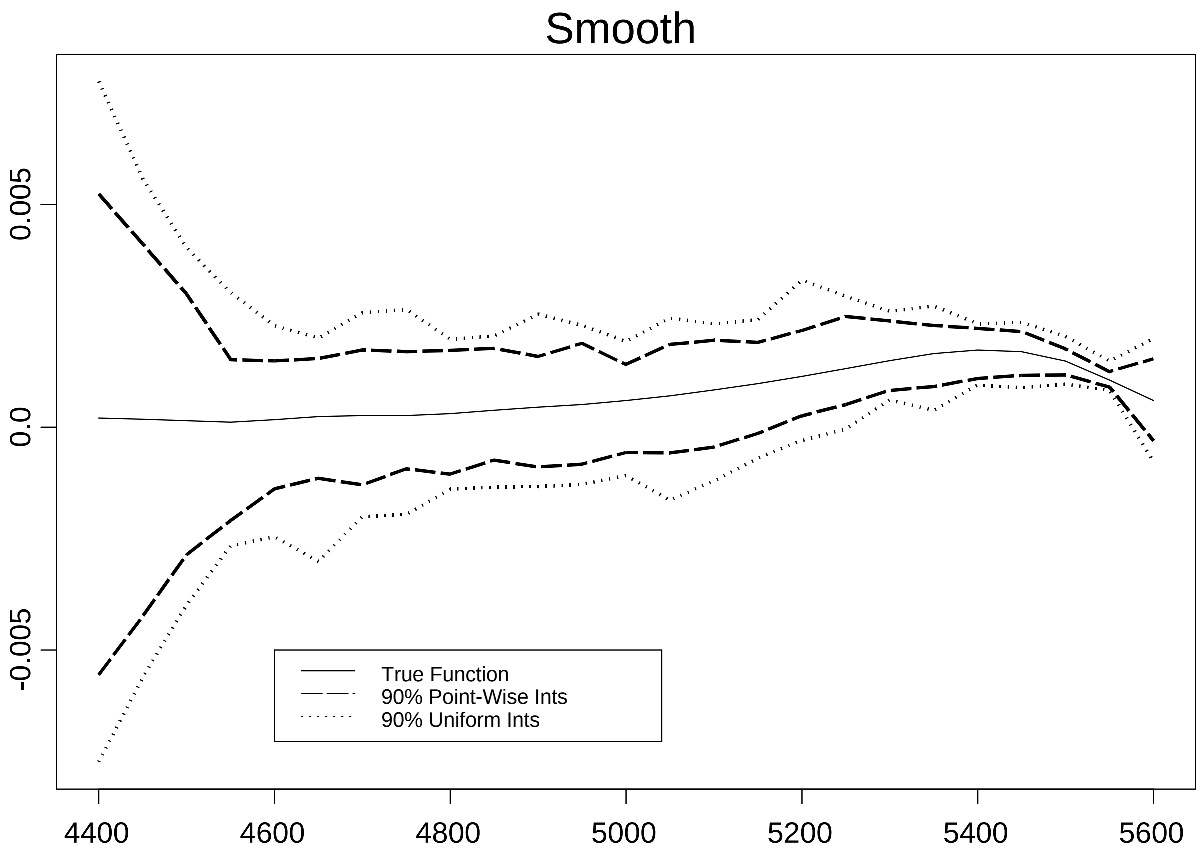


Figure 8: *Asymptotic CI for SRE SPD,* $L = 1.812 = \|m\|^2$.

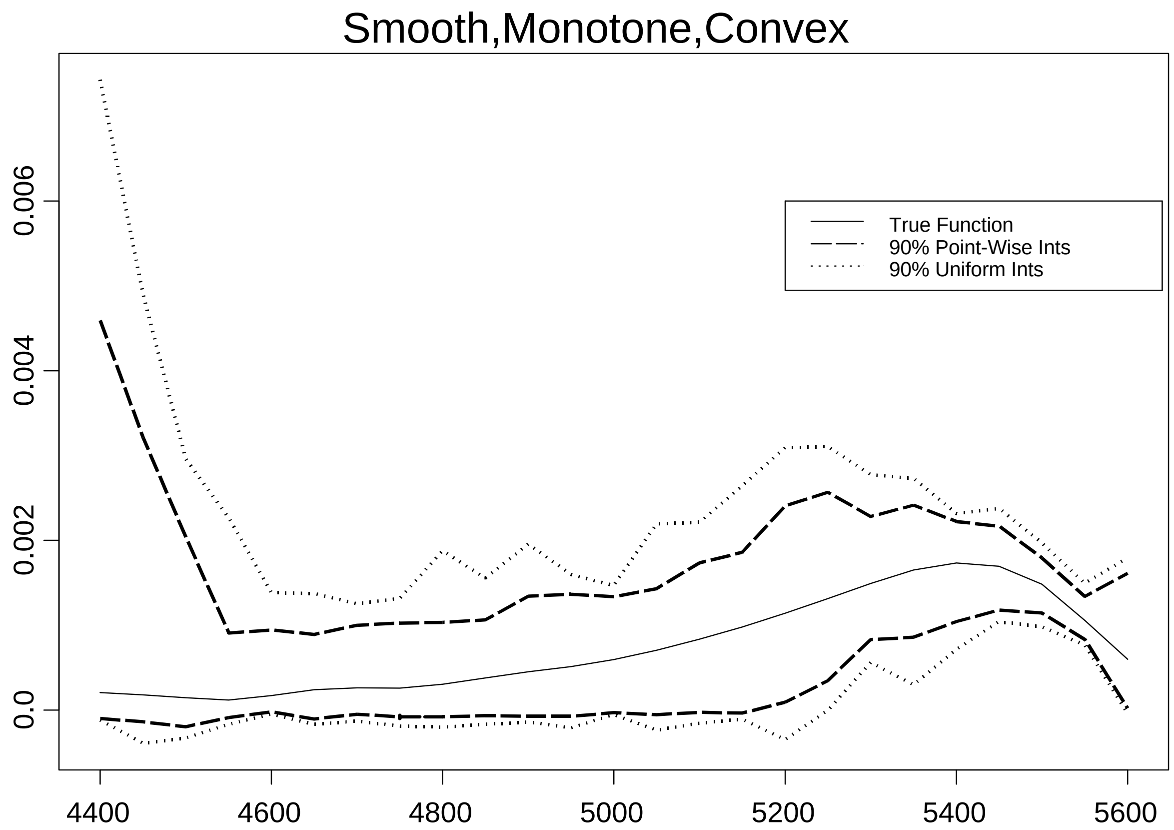
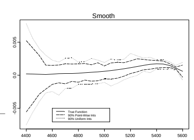


Figure 9: *Asymptotic CI for SRE SPD,* $L = 1.812$.



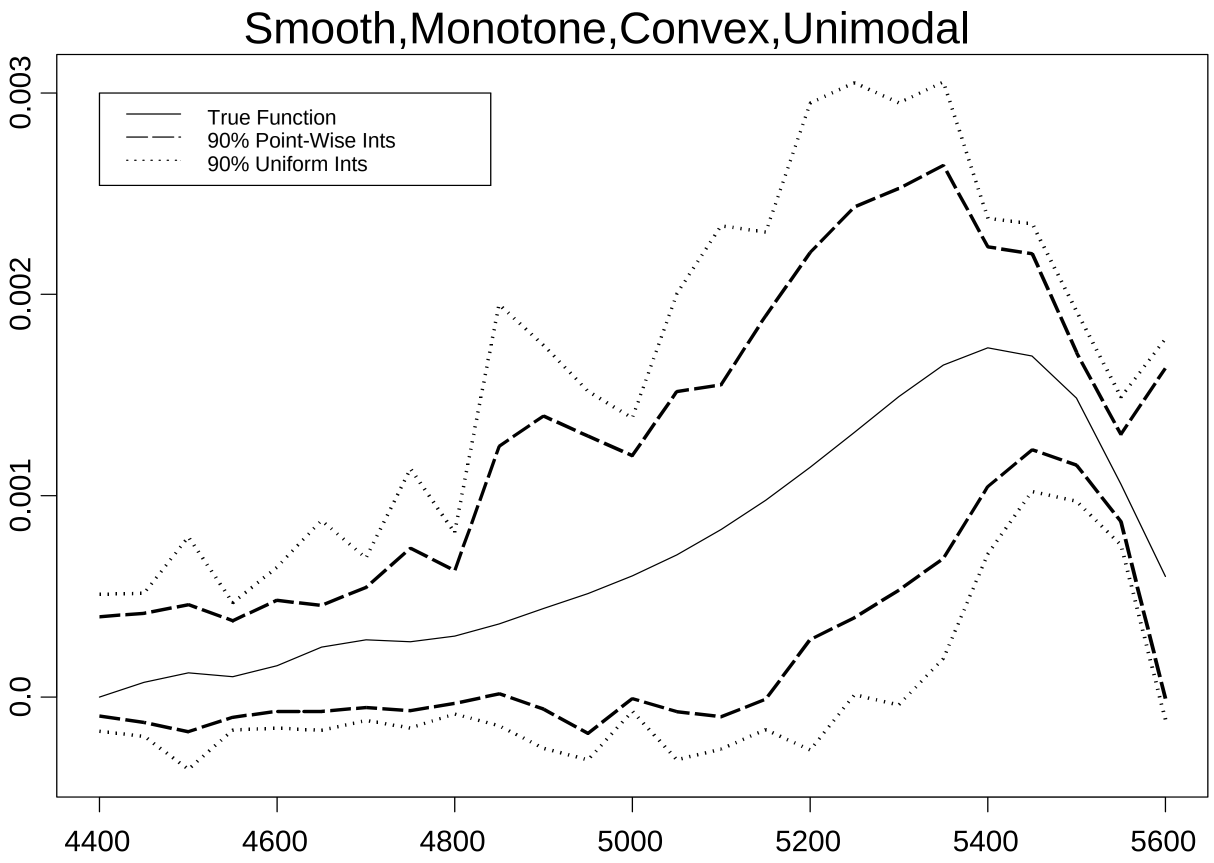
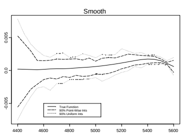


Figure 10: *Asymptotic CI for SRE SPD,* $L = 1.812$.



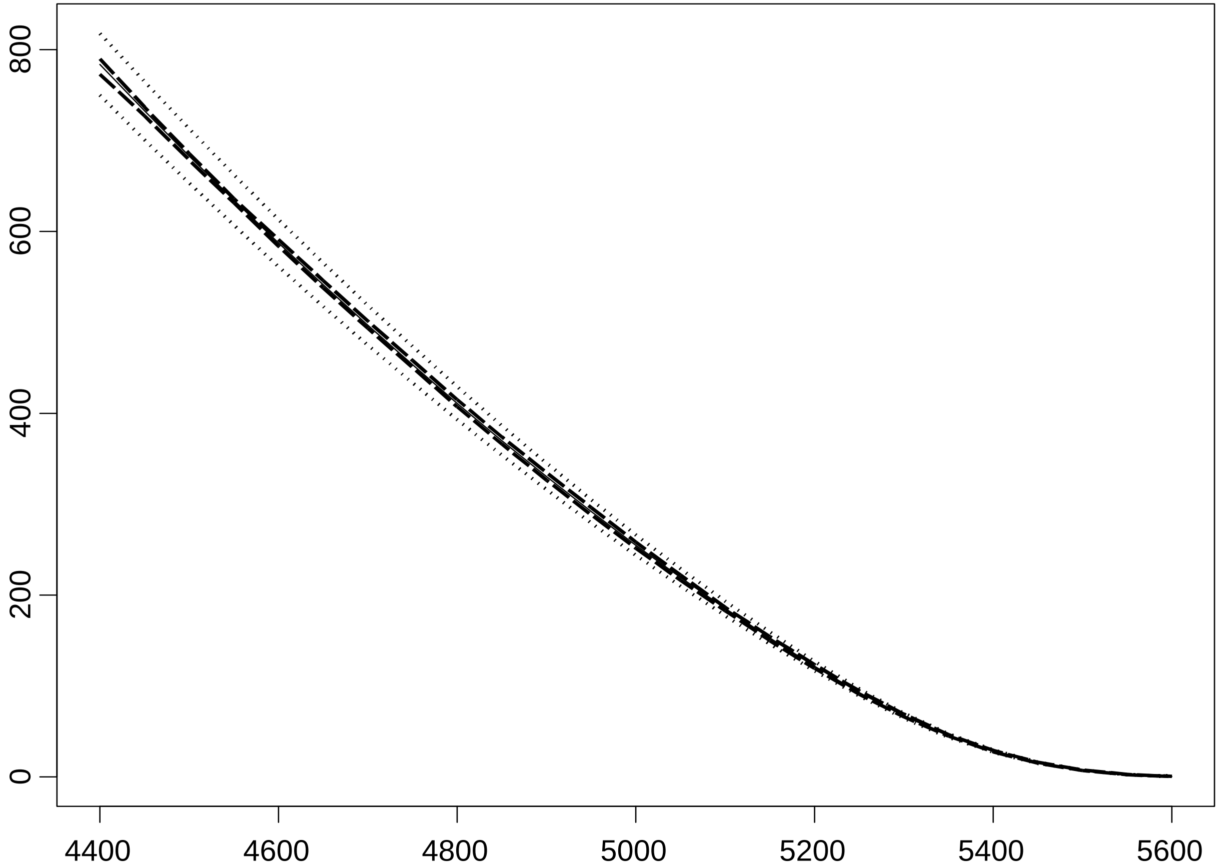
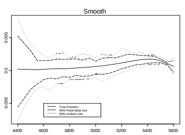


Figure 11: *Asymptotic vs. Bootstrap CI* $\hat{m}(x)$.



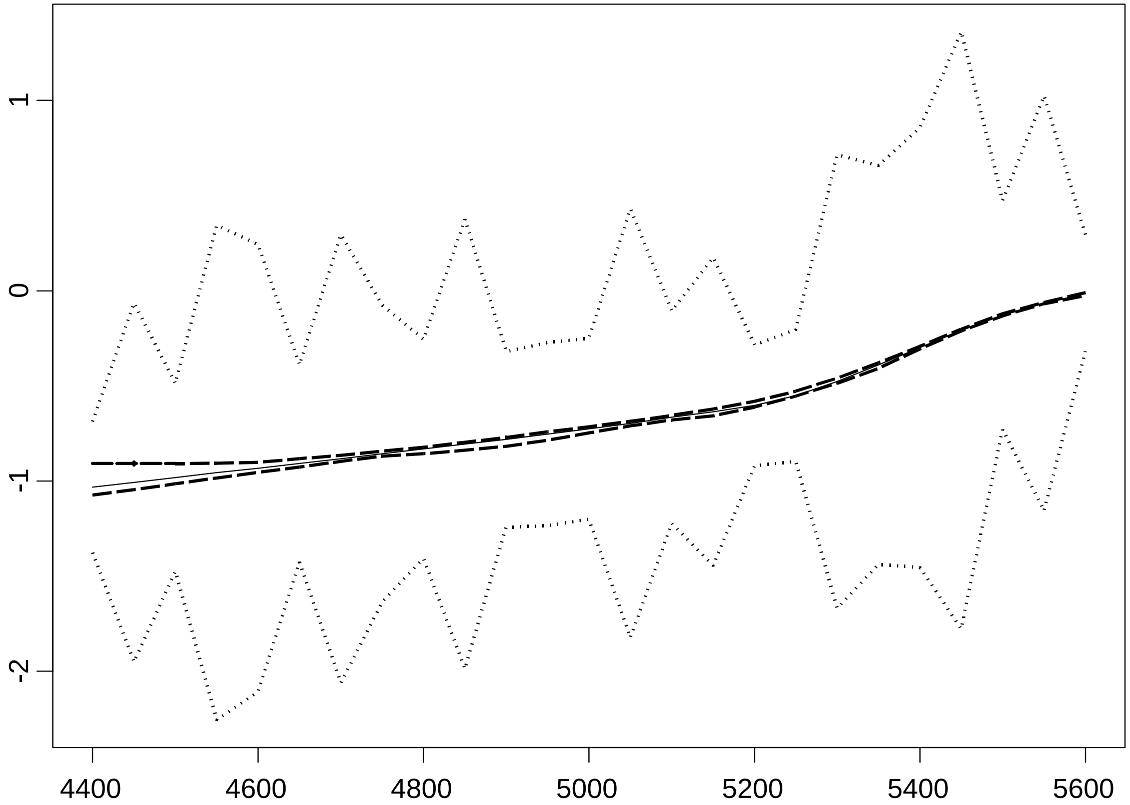


Figure 12: *Asymptotic vs. Bootstrap CI* $\hat{m}^{(1)}(x)$.

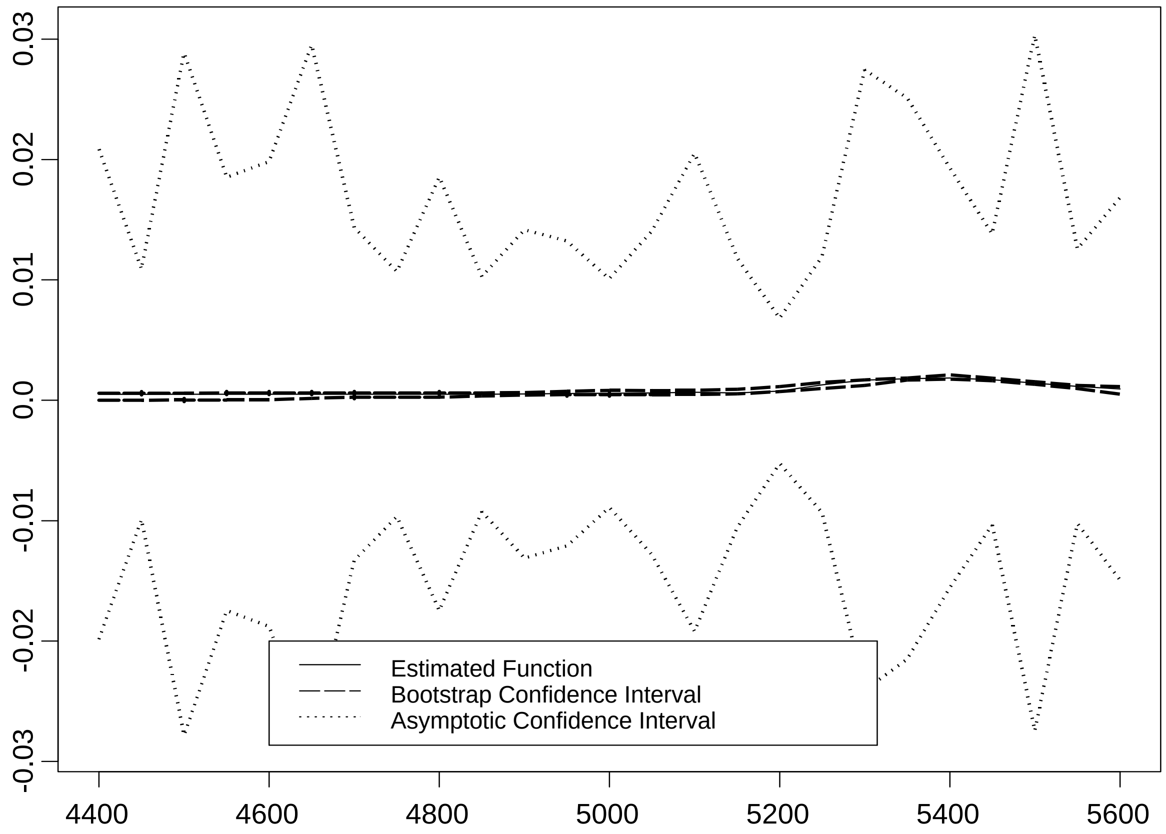
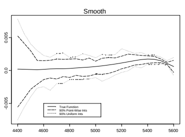


Figure 13: *Asymptotic vs. Bootstrap CI SPD.*



Accuracy of the Bootstrap Procedure

simulation repetitions: 500

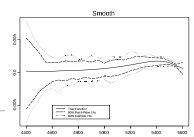
bootstrap resamples for each of these 500 samples : 100

calculate the coverage probability for CI

$1 - \alpha = 0.99, 0.95, 0.90, 0.5.$

Average coverage across strikes: for call and SPD

$1-\alpha$	average(call)	average(SPD)
0.99	0.979	0.994
0.95	0.94	0.978
0.90	0.899	0.953
0.50	0.5185	0.575



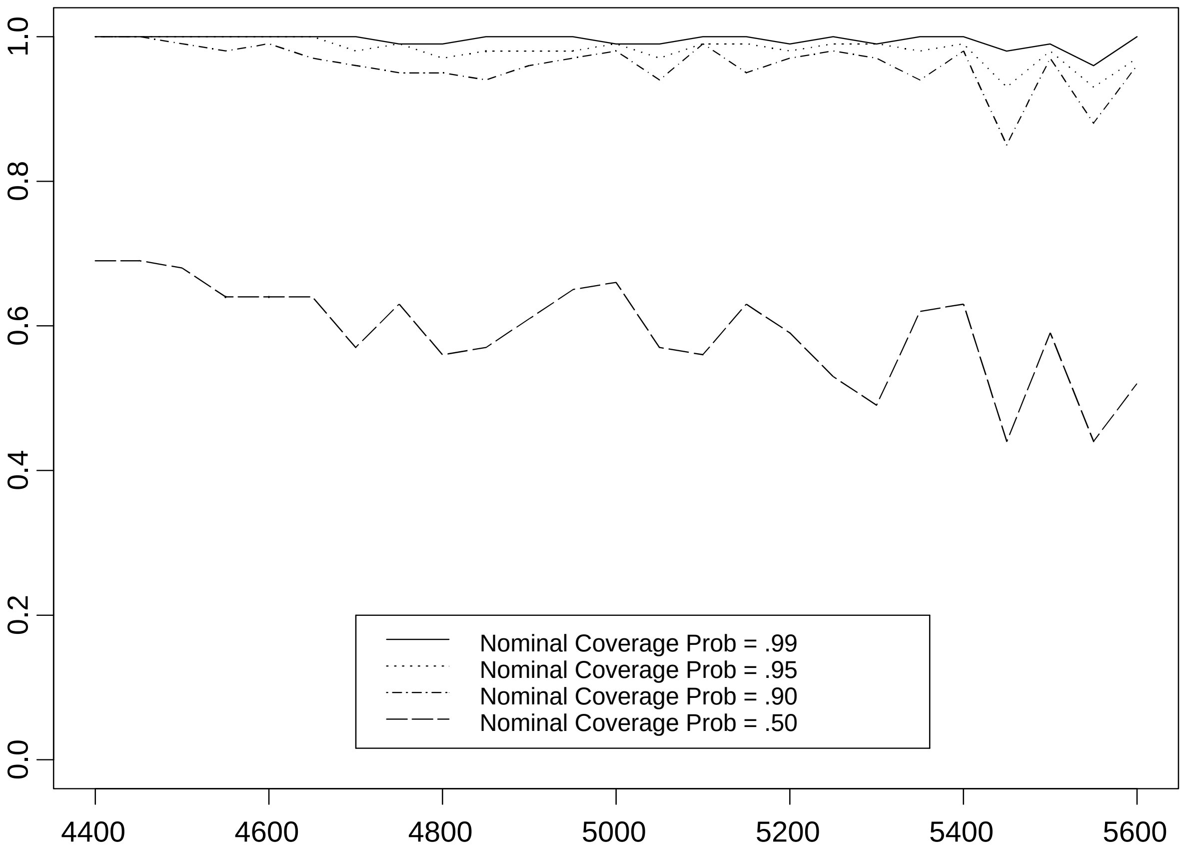


Figure 14: *Nominal coverage probability of bootstrap procedure for the SPD*

Application to DAX Data

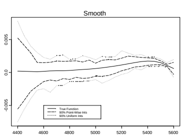
two week period, January 4-15, 1999

DAX fluctuates around 5200 during the first week

In the second week, DAX level below 5000

Estimates in $[4500, 5500]$ with 100 point intervals

Aim: Comparison of SPDs over time



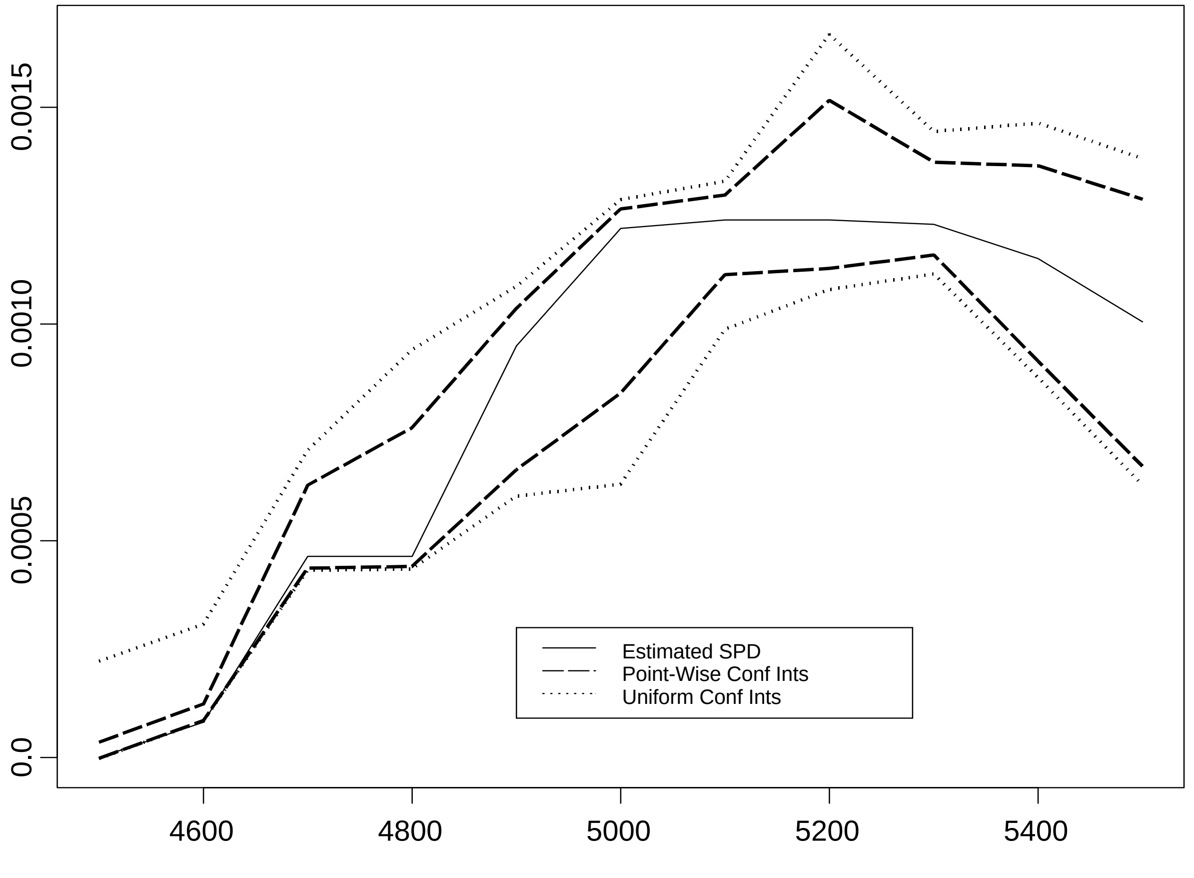


Figure 15: *SRE SPD with pointwise and uniform CI* 95%*, Jan 4, 1999*

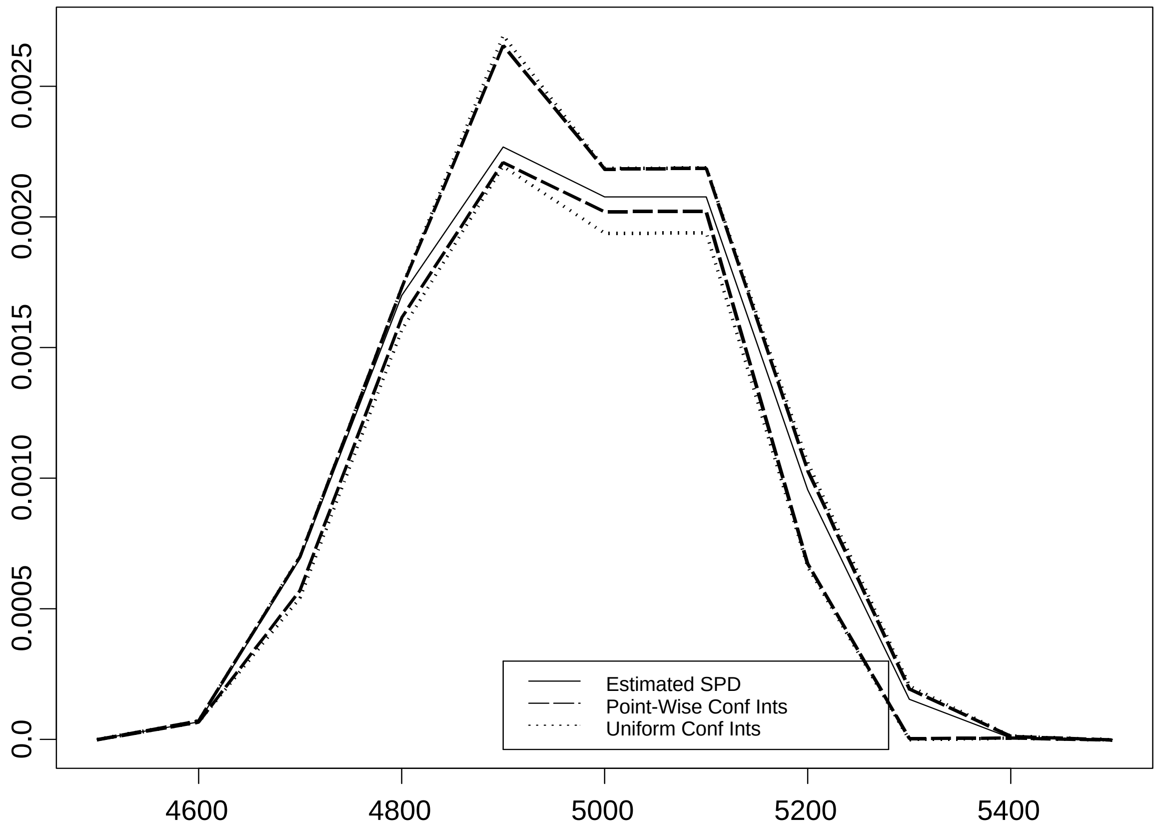
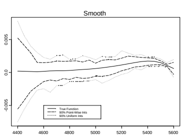


Figure 16: *SRE SPD with pointwise and uniform CI* 95%*, Jan 14, 1999*



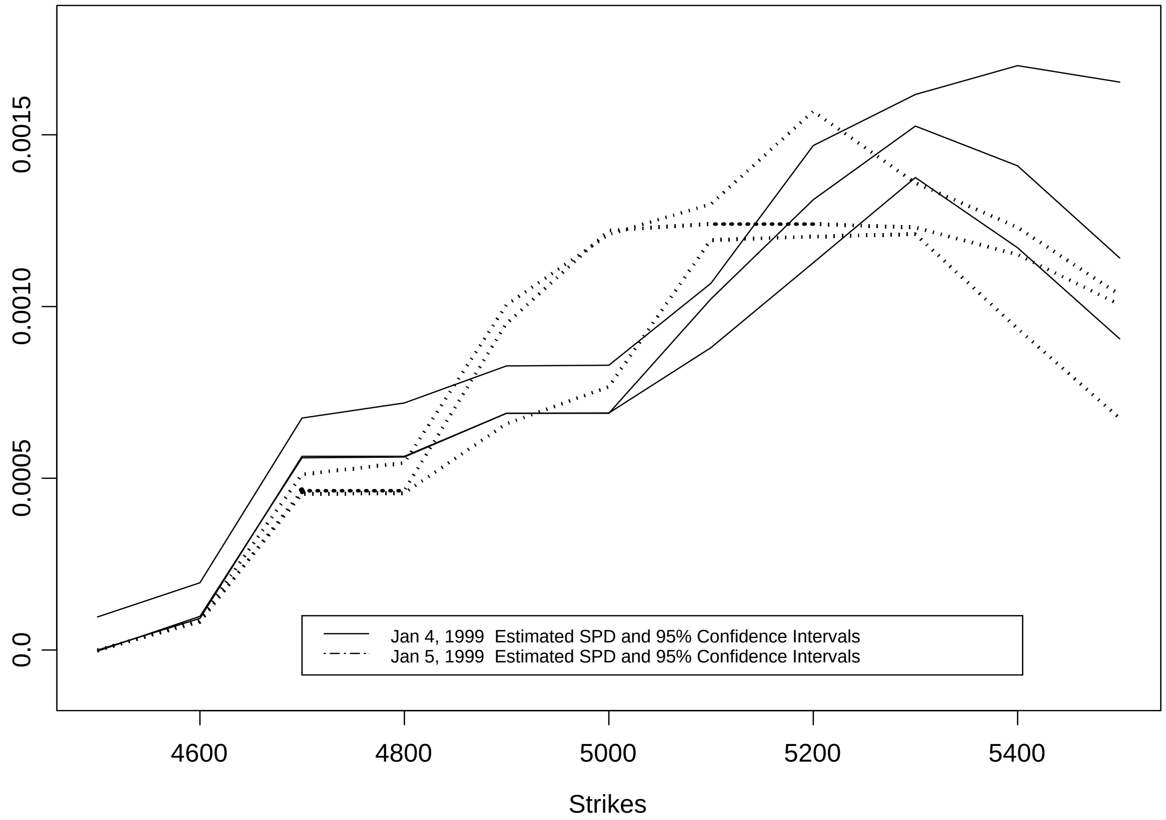
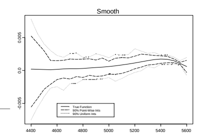


Figure 17: *SRE SPD with pointwise and uniform CI* 95%*, Jan 4 and 5, 1999*



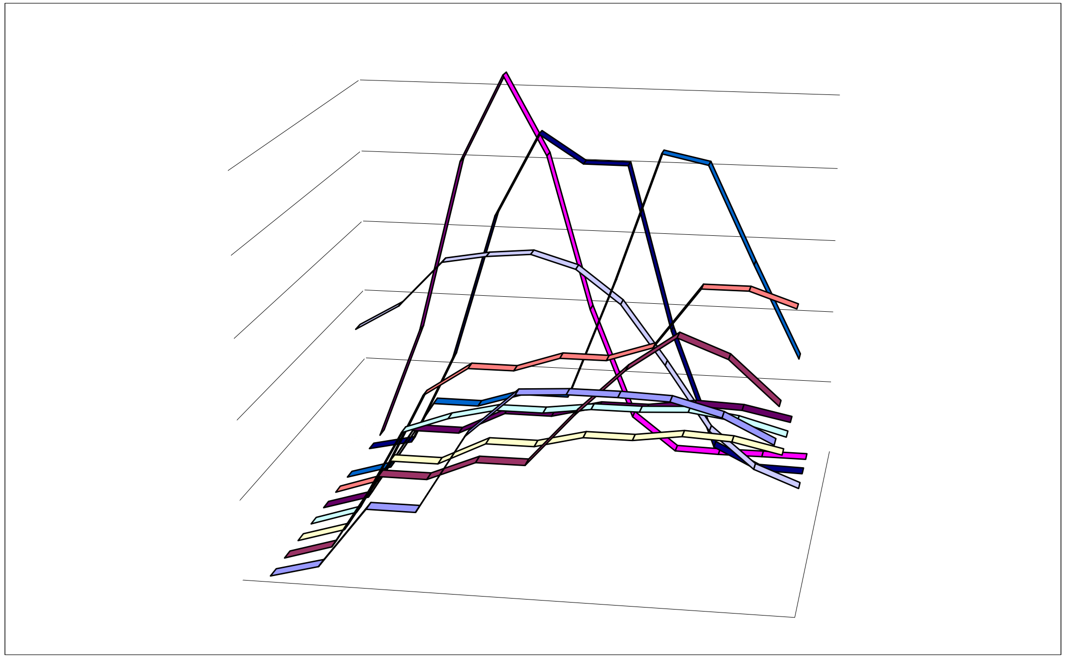
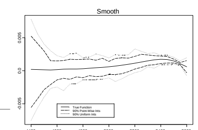


Figure 18: *Smooth constrained SPD estimate with pointwise and uniform CI* 95%, *Jan 4-15, 1999*

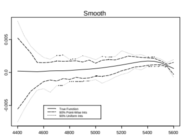


Trading rules

1. Calculate historical time series density g
   - Collect stock prices time series
   - Assume this time series is a sample path of the diffusion process

     $$\frac{dS_t}{S_t} = \mu_t dt + \sigma(S_t, t) dZ_t,$$

     where dZ_t is a Wiener process with mean zero and variance equal to dt.
   - Estimate diffusion function $\sigma(\cdot, \cdot)$ in the diffusion process model using historical time series
   - Estimate conditional density function $g = p(S_T | S_t, \hat{\mu}, \hat{\sigma})$ from Monte-Carlo simulated process



2. Calculate SRE SPD $\hat{m}^{(2)}(x)$

3. Compare the two densities

4. Set up skewness or kurtosis trade

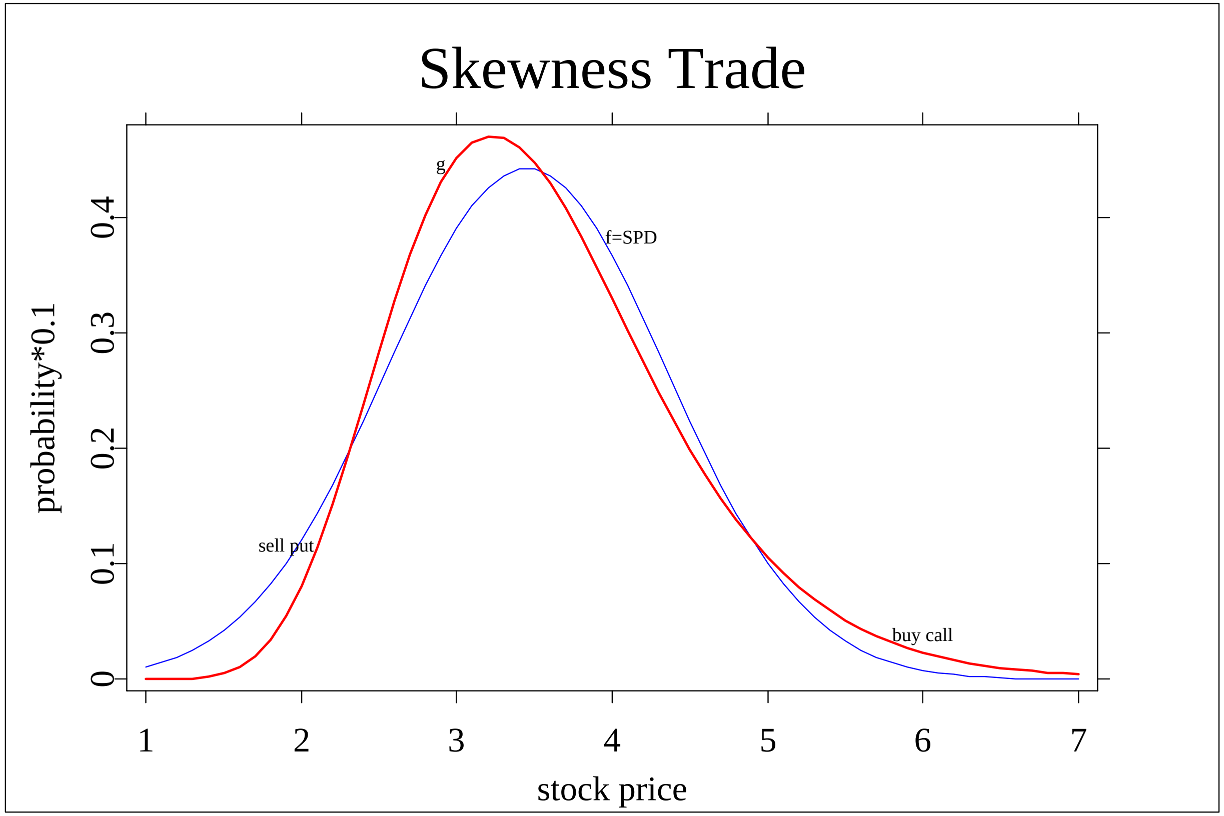
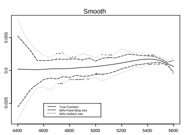


Figure 19: *Skewness Trade*, $skew(f) < skew(g)$

Trading Rules to exploit SPD differences			
Skewness			sell OTM put,
	$(S1)$	skew(f)$<$ skew(g)	buy OTM call
Trade	$(S2)$	skew(f) $>$ skew(g)	buy OTM put,
			sell OTM call
Kurtosis			sell far OTM and ATM ,
	$(K1)$	kurt(f)$>$ kurt(g)	buy near OTM options
Trade	$(K2)$	kurt(f) $<$ kurt(g)	buy far OTM and ATM,
			sell near OTM options

In this table, normal SPD is f and historical time series SPD is g. A far OTM call (put) is defined as one whose exercise price is 10% higher (lower) than the future price. A near OTM call (put) is defined as one whose exercise price is 5% higher (lower) but 10% lower(higher)than the future price.



More explanations about trading rules

when skew(f) $<$ skew(g), agents apparently assign a lower probability to high outcomes of the underlying than would be justified by the time series SPD (see Figure 19). Since for call options only the right 'tail' of the support determines the theoretical price the latter is smaller than the price implied by diffusion process using the time series SPD. That is we buy calls. The same reason applies to put options.

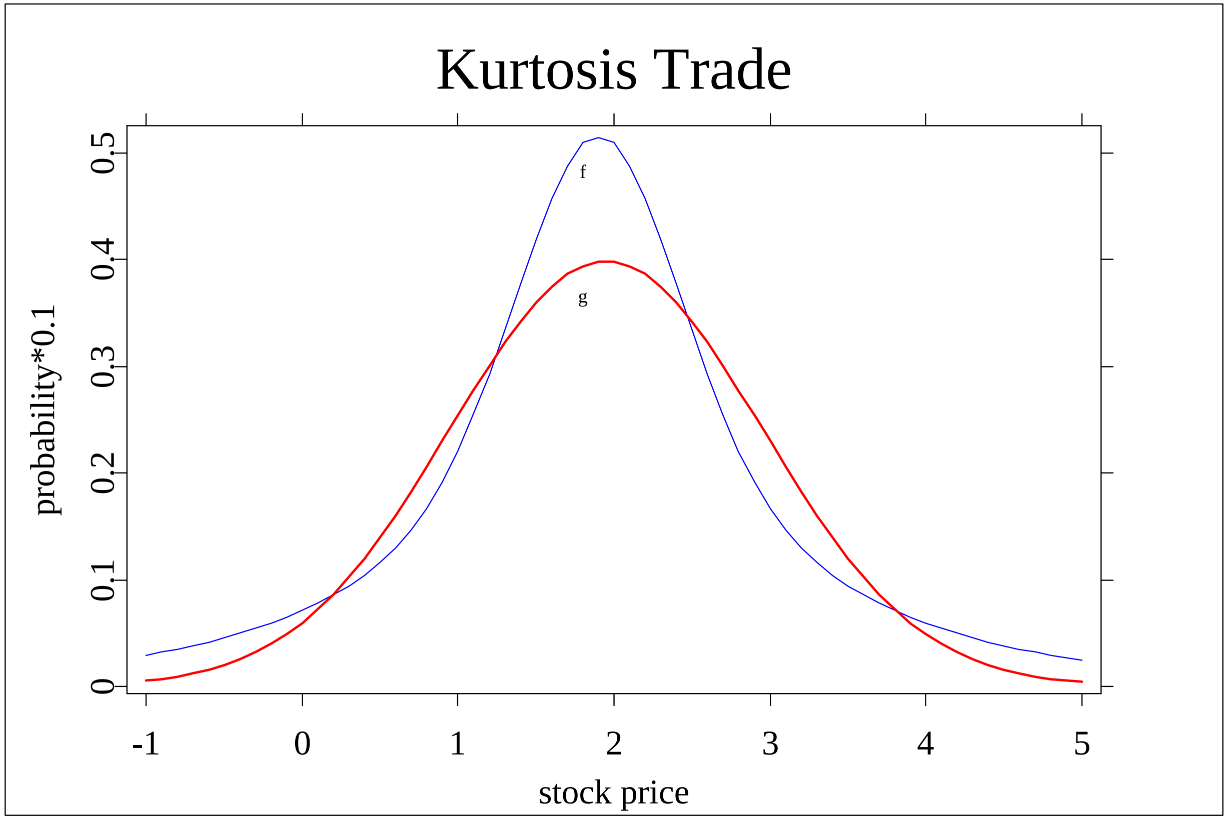


Figure 20: *Kurtosis Trade*, $kurt(f) > kurt(g)$