Härdle, Wolfgang Karl; Hlávka, Zdeněk

**Working Paper**

# Dynamics of state price densities

SFB 649 Discussion Paper, No. 2005,021

**Provided in Cooperation with:**
Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

*Suggested Citation:* Härdle, Wolfgang Karl; Hlávka, Zdeněk (2005) : Dynamics of state price densities, SFB 649 Discussion Paper, No. 2005,021, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:
https://hdl.handle.net/10419/25040

SFB 649 Discussion Paper 2005-021

# Dynamics of State Price Densities

Wolfgang Härdle*
Zdeněk Hlávka**

* CASE – Center for Applied Statistics and Economics,
Humboldt-Universität zu Berlin, Germany
** Charles University Prague, Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics,
Czech Republic

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

# Dynamics of State Price Densities

Wolfgang Härdle
CASE – Center for Applied Statistics and Economics
Wirtschaftswissenschaftliche Fakultät
Humboldt-Universität zu Berlin, Spandauer Str. 1, 10178 Berlin, Germany

Zdeněk Hlávka
Charles University Prague, Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics,
Sokolovská 83, 18675 Praha 8, Czech Republic

**Abstract**

State price densities (SPD) are an important element in applied quantitative finance. In a Black-Scholes model they are lognormal distributions with constant volatility parameter. In practice volatility changes and the distribution deviates from log-normality. We estimate SPDs using EUREX option data on the DAX index via a nonparametric estimator of the second derivative of the (European) call price function. The estimator is constrained so as to satisfy no-arbitrage constraints and it corrects for intraday covariance structure. Given a low dimensional representation of this SPD we study its dynamic for the years 1995–2003. We calculate a prediction corridor for the DAX for a 45 day forecast. The proposed algorithm is simple, it allows calculation of future volatility and can be applied to hedging exotic options.

*Key words and Phrases: option pricing, state price density estimation, nonlinear least squares, confidence intervals*

## 1 Introduction

The dynamics of option prices carries information on the changing (risk neutral) implied state price densities (SPD). Fitting SPDs over time provides useful insight into the behavior of the economic agents and the time inhomogeneity of the market. Knowledge of the SPD leads us also to pricing schemes for exotic options. Figure 1 displays the observed prices of European call options written on the DAX for the 16. January 1995 (abbreviated as 19950116). The left panel shows the ensemble of call option prices for different strikes and maturities as a free structure together with a smooth surface. The right panel shows the option prices only for the shortest maturity (4 days). In 1995, one observed every day such a point cloud with about 500 data points. In todays more liquid option markets this number has increased approximately 10 times. In our empirical study we will consider the time period up to 2003 thus also covering the more recent liquid option market.

The SPD bears important information on the behaviour and expectations of the market and is used for pricing and hedging (Fengler, Härdle, Mammen 2003). The most important application of SPD is that it allows

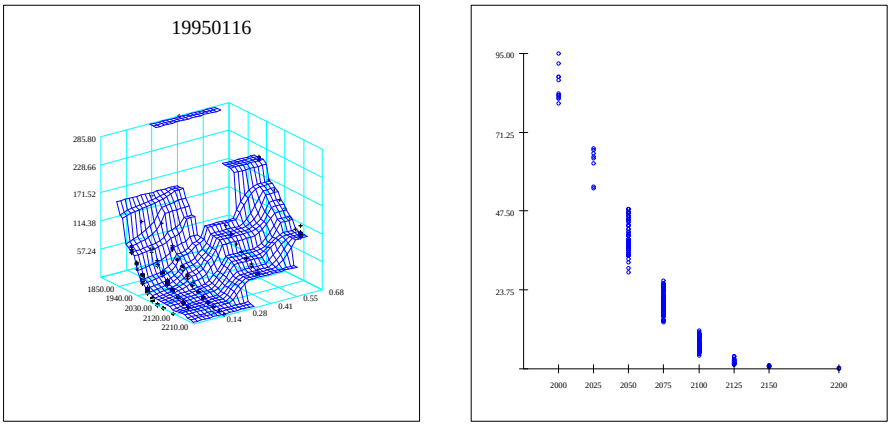


Figure 1: Option prices plotted against strike price and time to maturity with two-dimensional kernel regression surface (left) in January 1995 and the ensemble of the call option prices with shortest time to expiry against strike price (right) on 16. January 1995 (19950116). CASE financial data base MD*base.

to price options with complicated payoff functions simply by (numerical) integration of the payoff with respect to this density.

Prices $C_t(K,T)$ of European options with strike price $K$ observed at time $t$ and expiring at time $T$ allow to deduce the state price density in the following form (Breeden and Litzenberger 1978):

$$f(K) = \exp\{r(T-t)\}\frac{\partial^2 C_t(K,T)}{\partial K^2}. \tag{1}$$

Equation (1) is often used to estimate the state price density by the means of nonparametric regression.

Kernel smoothers were in this framework proposed and successfully applied by, e.g., Aït-Sahalia and Lo (1998), Aït-Sahalia and Lo (2000), Aït-Sahalia, Wang and Yared (2000) or Huynh, Kervella, and Zheng (2002). Aït-Sahalia and Duarte (2003) proposed a method for nonparametric estimation of the SPD under the appropriate constraints. Another sophisticated approach based on smoothing splines allowing to include the required constraints is described and applied on simulated data in Yatchew and Härdle (2005). An extensive overview of parametric and other estimation techniques can be found, e.g., in Jackwerth (1999). An application to option pricing is given in Buehler (2004). In these papers, the focus was on the smoothing technique rather than on a no-arbitrage argument. A crucial element of local volatility models, that are directly connected with the call price surface displayed in Figure 1, is the absence of arbitrage (Dupire 1994). Highly numerically efficient pricing algorithm, e.g., by Andersen and Brotherton-Ratcliffe (1997), rely heavily on no arbitrage properties. Kahalé (2004) proposed a procedure that requires solving a set of nonlinear equations with no guarantee of a unique solution. Moreover, for that algorithm the data feed alread is (unrealistically) expected to be arbitrage free (Fengler 2005). In addition, the covariance structure of the quoted option prices is rarely incorporated into the estimation procedure.

In this paper we develop a simple estimation technique in order to

construct SPD estimates satisfying the non-arbitrage constraints and to study the development of the estimated SPDs over time. The proposed technique involves contrained LS-estimation and enables us to construct confidence intervals for any future SPD. This, of course, is a vital feature for the trading floor where the derived (implied) volatility surfaces for different strikes and maturities are displayed for proper judgement of risk and return.

In the next two sections we construct an estimate of the SPD based on the observed call option prices satisfying all shape constraints given in Subsection 2.1 and we show that such estimate exists. In Section 4, we apply our estimation technique on option prices observed in year 1995. Some generalizations and natural models for covariance structure are proposed in Section 5. We will demonstrate that the true covariance structure of the option prices exhibits correlations depending both on the strike price and time of the trade. The dynamics of the estimated SPDs in years 1995–2003 is studied in Section 6.

# 2 Construction of the estimate

The fair price of a European call option with payoff $(S_T - K)_+ = \max(S_T - K, 0)$, with $S_T$ denoting the price of the stock at time $T$, $t$ the current time, $K$ the strike price, and $r$ the risk free interest rate, can be written as:

$$C_t(K,T) = \exp\{-r(T-t)\} \int_0^{+\infty} (S_T - K)_+ f(S_T) dS_T, \tag{2}$$

i.e., as the discounted expected value of the payoff with respect to the SPD $f(.)$. For the sake of simplicity of the following presentation, we assume in the rest of the paper that the discount factor $\exp\{-r(T-t)\} = 1$. In applications, this is achieved by correcting the observed option prices by the known risk free interest rate $r$ and the time to maturity $(T-t)$ in (2).

Our data set contains the observed option prices for various strike prices and maturities. Other variables are the interest rate, date, and time. We will analyze the option prices as a function of the strike price for fixed date and time to expiry. An example of such data set is displayed, for 16th January 1995 and for the shortest time to expiry, $\tau = T - t = 4$ days, on the right panel in Figure 1.

Let us denote the $i$-th observation of the strike price by $K_i$ and the corresponding option price, divided by the discount factor $\exp\{-r(T-t)\}$ from (2), by $C_i = C_{t,i}(K_i, T)$. In practice, one observes option prices repeatedly for a small number of distinct strike prices. Therefore, it is useful to adopt the following notation. Let $\mathcal{C} = (C_1, \ldots, C_n)^\top$ be the vector of the observed option prices at day $t$. The corresponding vector of the strike prices has the following structure:

$$\mathcal{K} = \begin{pmatrix} K_1 \\ K_2 \\ \vdots \\ K_n \end{pmatrix} = \begin{pmatrix} k_1 \mathbf{1}_{n_1} \\ k_2 \mathbf{1}_{n_2} \\ \vdots \\ k_p \mathbf{1}_{n_p} \end{pmatrix},$$

where $k_1 < k_2 < \cdots < k_p$, $n_j = \sum_{i=1}^n \boldsymbol{I}(K_i = k_j)$ with $\boldsymbol{I}(.)$ denoting the the indicator function and $\mathbf{1}_n$ a vector of ones of length $n$.

**EXAMPLE 1** *An example is given in Figure 1. In the left panel we display all European call prices on DAX traded on 16th January 1995 against the strike price and the maturity. In order to see the structure of the data set more clearly, we include also a kernel surface.*

*One can observe that options are traded only for few strike prices. Restricting ourselves only to the shortest maturity, we obtain the right panel of Figure 1. Here, at* $t = 19950116$, $n = 575$ *call options were observed on a grid of* $p = 8$ *strike prices* $k_1 = 2000, \ldots, k_8 = 2300$.

## 2.1 Assumptions and constraints

For fixed time $t$ and time to maturity $\tau = T - t$, the $i$-th observed option price (corresponding to strike price $K$) follows the model

$$C_{t,i}(K_i, T) = \mu(K_i) + \varepsilon_i, \tag{3}$$

where $\varepsilon_i$ are i.i.d. $N(0, \sigma^2)$ distributed variables. Heteroskedasticity can be incorporated in model (3) if we assume that the random errors $\varepsilon_i$ have $N(0, \sigma^2_{K_i})$ distribution. The assumptions on the distribution of random errors will be investigated in more detail in Subsection 4.3.

Harrison and Pliska (1981) characterized the absence of arbitrage by the existence of a unique risk neutral SPD $f(.)$. From formula 2 and properties of a probability density it follows that the function of the true conditional means $\mu(.)$ has to satisfy the following no-arbitrage constraints:

1'. it is positive,

2'. it is decreasing in $K$,

3'. it is convex,

4'. its second derivative exists and it is a density (i.e., nonnegative and it integrates to one).

Let us now have a look at functions satisfying Constraints 1'–4'.

**LEMMA 1** *Suppose that* $\mu$ *satisfies Constraints 1'–4'. Then we have for the first derivative,* $\mu^{(1)}(.)$, *that* $\lim_{x\to+\infty} \mu^{(1)}(x) = 0$ *and* $\lim_{x\to-\infty} \mu^{(1)}(x) = -1$.

**Proof:**
Constraint 4' implies that the first derivative, $\mu^{(1)}$, exists and that it is differentiable. $\lim_{x\to+\infty} \mu^{(1)}(x)$ exists since the function $\mu^{(1)}$ is increasing (Constraint 3') and bounded (Constraint 2'). Next, $\lim_{x\to\infty} \mu^{(1)}(x) = 0$ since a negative limit would violate Constraint 1' for large $x$ ($\mu^{(1)}(x)$ cannot be positive since $\mu(x)$ is decreasing). Finally, Constraint 4', $1 = \int_{-\infty}^{\infty} \mu^{(2)}(x)dx = \lim_{x\to+\infty} \mu^{(1)}(x) - \lim_{x\to-\infty} \mu^{(1)}(x)$, implies that $\lim_{x\to-\infty} \mu^{(1)}(x) = -1$. □

## 2.2 Existence and uniqueness

In this subsection we address the issue of existence and uniqueness of a regression function, $\hat{C}(.)$, satisfying the required assumptions and constraints.

In practice, we don't deal with a continuous function. Hence, we restate Constraints 1'–4' for discrete functions, defined only on a finite set of

distinct points, say $k_1 < \cdots < k_p$, in terms of their function values, $C(k_i)$, and their scaled first differences, $C^{(1)}_{k_i,k_j} = \{C(k_i) - C(k_j)\}/\{k_i - k_j\}$.

1. $C(k_i) > 0$, $i = 1, \ldots, p$,
2. $k_i < k_j$ implies that $C(k_i) \geq C(k_j)$,
3. $k_i < k_j < k_l$ implies that $-1 \leq C^{(1)}_{k_i,k_j} \leq C^{(1)}_{k_j,k_l} \leq 0$.

It is easy to see that Constraints 1–3 are discrete versions of Constraints 1'–4'.

From now on, similarly as in Robertson, Wright and Dykstra (1988), we think of the collection, $\mathcal{C}$, of functions satisfying Constraints 1–3 as a subset of a $p$-dimensional Euclidean space, where $p$ is the number of distinct $k_i$'s. The constrained regression, $\hat{C}$, is in this setting the closest point of $\mathcal{C}$ to the vector $C$ of the observed option prices with distances measured by the usual Euclidean distance

$$d(f, C) = (f - C)^\top (f - C) = \sum_{i=1}^{n} \{f(K_i) - C(K_i)\}^2. \tag{4}$$

From this point of view, the regression function, $\hat{C}$, consists only of the values of the function in the points $k_1, \ldots, k_p$. The first and second differences are used to approximate the first and the second derivatives, respectively.

We claim that the set, $\mathcal{C}$, of functions satisfying Constraints 1–3, has the following properties

1. $\mathcal{C}$ is closed in the topology induced by the metric given by Euclidean distance,
2. $\mathcal{C}$ is convex, i.e., if $f, g \in \mathcal{C}$ and $0 \leq a \leq 1$, then $af + (1 - a)g \in \mathcal{C}$.

**LEMMA 2** *Assume that $\hat{C} \in \mathcal{C}$ is the regression of $C(K_i)$, $i = 1, \ldots, n$, on $k_1 < \cdots < k_p$ under Constraints 1–3. If $a$ and $b$ are constants such that $a \leq C(K_i) \leq b$, $\forall i$, then $a - (k_p - k_1) \leq \hat{C}(k_i) \leq b + (k_p - k_1)$.*

**Proof:**
It is not possible that $\hat{C}(k_i)$ lies below $a$ or above $b$ for all $k_i$'s (otherwise we would get a better fit only by shifting $\hat{C}(k_i)$). The bounds now follow from Constraint 3. □

**THEOREM 1** *A regression, $\hat{C} = \arg\min_{f \in \mathcal{C}} d(f, C)$, satisfying Constraints 1–3 exists.*

**Proof:**
Lemma 2 implies that $\hat{C}$ belongs to a subset, $\mathcal{S}$, of $\mathcal{C}$ bounded below by $a - (k_p - k_1)$ and above by $b + (k_p - k_1)$. Thinking of the functions as of points in Euclidean space, it is clear that the continuous function $d(f, C)$ attains its minimum on the closed and bounded set $\mathcal{S}$. □

**REMARK 1** *Suppose that $\mathcal{C}$ is any convex set of functions on $\mathcal{X}$ and $C$ is a given function on $\mathcal{X}$. If $\hat{C} = \arg\min_{f \in \mathcal{C}} d(f, C)$ then for every $f \in \mathcal{C}$,*

$$\sum_{i=1}^{n} \{C(K_i) - \hat{C}(K_i)\}^\top \{\hat{C}(K_i) - f(K_i)\} \geq 0. \tag{5}$$

*There exists at most one function $\hat{C}$ satisfying (5).*

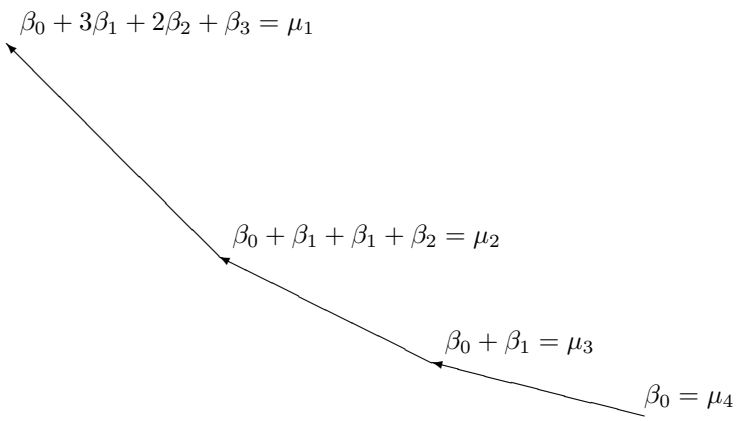


Figure 2: Illustration of the dummy variables for call options.

**Proof:**
See Robertson, Wright and Dykstra (1988, Theorem 1.3.1). □

**COROLLARY 1** *A regression, $\hat{C}$, satisfying Constraints 1–3 exists and it is unique.*

**Proof:**
It follows from Theorem 1 and Remark 1. □

## 2.3 Linear model

The configuration of data, under Constraints 1–3 of Subsection 2.1, can be easily described using regression models with constraints.

In the following, we fix the time $t$ and the expiry date $T$ and we omit these symbols from the notation. In Subsection 2.2 we have noted that the option prices are repeatedly observed for a small number $p$ of distinct strike prices. This situation is visible in our data, e.g., in the right hand panel in Figure 1 we have 575 observations observed at $p = 8$ distinct strike prices.

For simplicity of the following presentation we display the coefficients $\beta_i$ in the situation with only four distinct strike prices ($p = 4$) in Figure 2. Defining the expected values of the option prices given strike price, $\mu_j = \mu(k_j) = E\{C(k_j)\}$, we can write

$$\begin{aligned}
\mu_p &= \beta_0, \\
\mu_{p-1} &= \beta_0 + \beta_1, \\
\mu_{p-2} &= \beta_0 + 2\beta_1 + \beta_2, \\
\mu_{p-3} &= \beta_0 + 3\beta_1 + 2\beta_2 + \beta_3, \\
&\vdots \\
\mu_1 &= \beta_0 + (p-1)\beta_1 + (p-2)\beta_2 + \cdots + \beta_{p-1}.
\end{aligned}$$

Thus, we fit our data using coefficients $\beta_j$, $j = 1, \ldots, p$. The conditional means $\mu_i$, $i = 1, \ldots, p$ are replaced by the same number of parameters $\beta_j$,

$j = 0, \ldots, p-1$ which allow to impose the shape constraints in a more natural way.

The interpretation of the coefficients $\beta_j$ can be seen from Figure 2. $\beta_0$ is the mean option price at point 4. Constraint 1', Subsection 2.1, implies that it has to be positive. $\beta_1$ is the difference between the mean option prices at point 4 and point 3; Constraint 2' implies that it has to be positive. The next coefficient, $\beta_2$, approximates the change in first derivative in point 3 and it can be interpreted as an approximation of the second derivative in point 3. Constraint 3' implies that $\beta_2$ has to be positive. Similarly, $\beta_3$ is an estimate of the (positive) second derivative in point 2. Constraint 4' can be rewritten as $\beta_2 + \beta_3 \leq 1$.

In practice, we start with the construction of a design matrix which allows us to write the above model in the following linear form. For simplicity of presentation, we again set $p = 4$:

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}. \tag{6}$$

Ignoring the constraints on the coefficients would lead to a simple linear regression problem. Unfortunately, this approach does not have to lead, and usually does not, to interpretable and stable results.

Model (6) in the above form can be reasonably interpreted only if the observed strike prices are equidistant and if the distances between the neigbouring observed strike prices are equal to one. If we want to keep the interpretation of the parameters $\beta_j$ as the derivatives of the estimated function, we should use the design matrix

$$\Delta = \begin{pmatrix} 1 & \Delta_p^1 & \Delta_{p-1}^1 & \Delta_{p-2}^1 & \cdots & \Delta_3^1 & \Delta_2^1 \\ 1 & \Delta_p^2 & \Delta_{p-1}^2 & \Delta_{p-2}^2 & \cdots & \Delta_3^2 & 0 \\ \vdots & & & & & & \vdots \\ 1 & \Delta_p^{p-2} & \Delta_{p-1}^{p-2} & 0 & \cdots & 0 & 0 \\ 1 & \Delta_p^{p-1} & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{7}$$

where $\Delta_j^i = \max(k_j - k_i, 0)$ denotes the positive part of the distance between $k_i$ and $k_j$, the $i$-th and the $j$-th $(1 \leq i \leq j \leq p)$ sorted distinct observed values of the strike price.

The vector of conditional means $\mu$ can be written in terms of the parameters $\beta$ as follows

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \mu = \Delta\beta = \Delta \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix}. \tag{8}$$

The constraints on the conditional means $\mu_j$ can now be expressed as conditions on the parameters of the model (8). Namely, it suffices to request that $\beta_i > 0$, $i = 0, \ldots, p-1$ and that $\sum_{j=2}^{p-1} \beta_j \leq 1$.

The model for the option prices can now be written as

$$C(\mathcal{K}) = \mathcal{X}_\Delta \beta + \varepsilon, \tag{9}$$

where $\mathcal{X}_\Delta$ is the design matrix obtained by repeating each row of matrix $\Delta$ $n_i$-times, $i = 1, \ldots, p$.

# 3 Implementing the constraints

In order to impose Constraints 1–3 on parameters $\beta_i$, $i = 0, \ldots, p-1$, we propose the following reparametrization of the model in terms of parameters $\theta_j$, $j = 0, \ldots, p-1$:

$$\begin{aligned} \beta_0(\theta) &= \exp(\theta_0), \\ \beta_1(\theta) &= \exp(\theta_1), \\ &\vdots \\ \beta_{p-1}(\theta) &= \exp(\theta_{p-1}), \end{aligned}$$

under the constraint that $\sum_{j=2}^{p-1} \exp(\theta_j) < 1$. Clearly, the parameters $\beta_i(\theta)$ satisfy the constraints

$$\begin{aligned} \beta_i(\theta) &> 0, \quad i = 0, \ldots, p-1, \\ \sum_{j=2}^{p-1} \beta_j(\theta) &< 1. \end{aligned}$$

This means that the parameters $\beta_2(\theta), \ldots, \beta_{p-1}(\theta)$ can be considered as point estimates of the state price density (the estimates have to be positive and integrate to less than one). Furthermore, in view of Lemma 1, it is worthwhile to note that the parameters satisfy also

$$-\sum_{j=1}^{k} \beta_j \in (-1, 0), \quad \text{for } k = 1, \ldots, p-1.$$

The model (9) rewritten in terms of parameters $\theta_i$, $i = 0, \ldots, p$ is a nonlinear regression model which can be estimated using standard maximum likelihood methods. The main advantage of the maximum likelihood estimator (MLE) is that the asymptotic distribution is well known and that the asymptotic variance of the estimator can be approximated using numerical methods implemented in many statistical packages.

Using the data displayed in the right hand plot in Figure 1, we obtain the estimates displayed in Figure 3. The top plot displays the original data, the second plot shows the estimate of the first derivative, and the third plot shows the estimate of the second derivative, i.e., the state price density. Actually, all plots contain two curves, both obtained using model (9). The thick line is calculated using the parameters $\beta_i$ without constraints whereas the thin line uses the reparametrization $\beta_i(\xi)$ given in Subsection 3.1. In Figure 3, these two estimates coincide since the model maximizing the likelihood without constraints, by chance, fulfills the constraints $(\exists \xi : \beta_i = \beta_i(\xi), i = 0, \ldots, p-1)$ and hence it is clear that the same parameters maximize also the constrained likelihood.

The situation, in which the estimates with and without constraints differ, is displayed in Figure 4. Notice that the difference between the two regression curves is small whereas the difference between the estimates of the state price density (i.e., the second derivative of the curve) is surprisingly large. The unconstrained estimate shows very unstable behaviour on the left hand side of the plot. The constrained version behaves more reasonably. Very small differences between the fitted lines in the top plot in Figure 4 leads to huge differences in the estimates of second derivative.

We therefore conclude that small errors in the estimates of the curve may lead to large scale errors in the estimates of the first and second

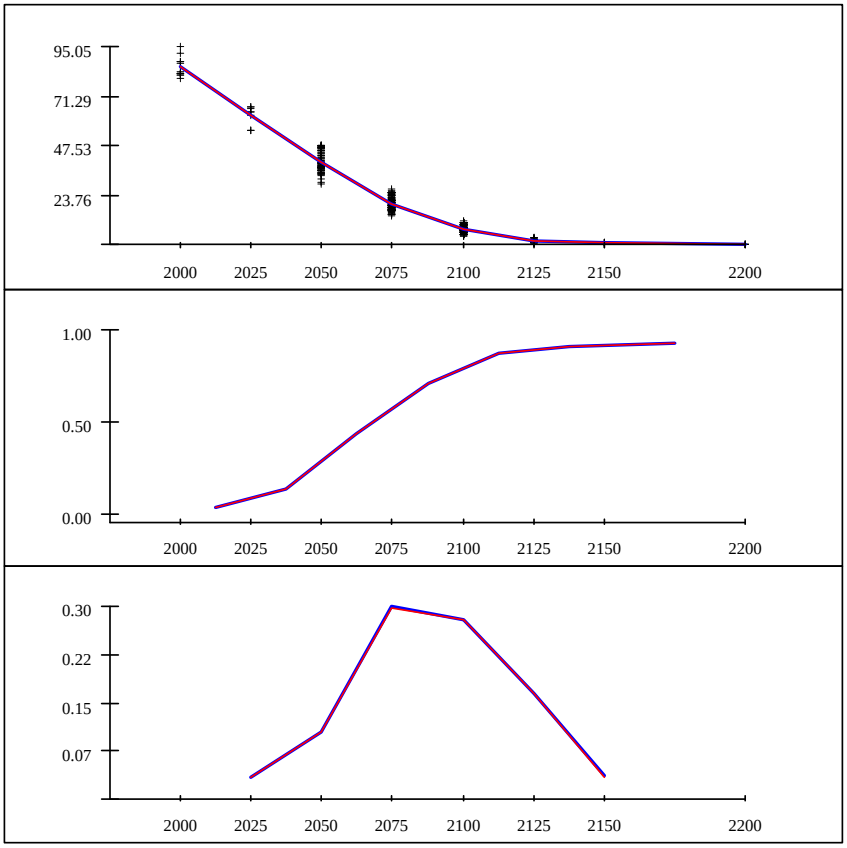


Figure 3: On 16th January 1995, the unconstrained estimate satisfies the constraints. Hence, it is equal to the constrained estimate. The top panel shows the original data with the fitted lines. The second and the third panel show the estimates of the first and second derivative, respectively.

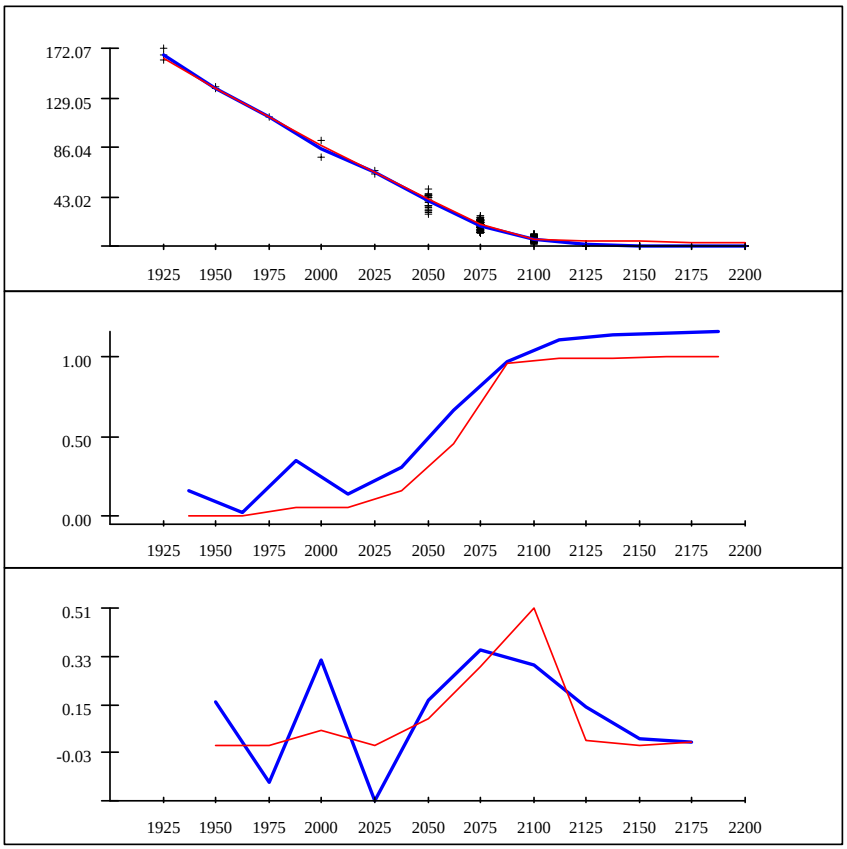


Figure 4: On 17th January 1995, the unconstrained estimate, displayed using the thicker and darker line, does not satisfy the constraints. The top panel shows the original data with and the two fitted lines. The estimates of the first derivative in the second panel look rather different. The constrained estimate of the second derivative in the bottom panel is clearly much more stable than the unconstrained estimate.

derivatives. The scale of this type of error seems to be limited by imposing the shape constraints given in Subsection 2.1.

## 3.1 Reparametrization

The following reparametrization of the model in terms of parameters $\xi_j$, $j = 0, \ldots, p$ is useful for calculating the estimates:

$$\begin{aligned} \beta_0(\xi) &= \exp(\xi_0), \\ \beta_1(\xi) &= \frac{\exp(\xi_1)}{\sum_{j=1}^{p} \exp(\xi_j)}, \\ &\vdots \\ \beta_{p-1}(\xi) &= \frac{\exp(\xi_{p-1})}{\sum_{j=1}^{p} \exp(\xi_j)}. \end{aligned}$$

The equality

$$\exp(\xi_p) \left\{ \sum_{j=1}^{p-1} \exp(\xi_j) \right\}^{-1} = 1 - \left\{ \sum_{j=1}^{p-1} \beta_j(\xi) \right\}^{-1}$$

shows the meaning of the parameter $\xi_p$. Setting this parameter to $-\infty$ would be the same as requiring that $\sum_{j=2}^{p-1} \beta_j(\xi) = 1$.

## 3.2 Inverse transformation of model parameters

For the numerical algorithm, it is useful to know how to calculate $\xi$s from given $\beta$s. This is needed, for example, to obtain reasonable starting points for the iterative procedure maximizing the likelihood.

**LEMMA 3** *Given $\beta = (\beta_1, \ldots, \beta_p)^\top$, where $\beta_p = 1 - \sum_{i=1}^{p-1} \beta_i$, the parameters $\xi = (\xi_1, \ldots, \xi_p)^\top$ satisfy the system of equations*

$$\left(\beta \mathbf{1}_p^\top - \mathbf{I}_p\right) \exp \xi^\top = \mathcal{A} \exp \xi^\top = 0, \tag{10}$$

*where $\mathbf{I}_p$ is the $(p \times p)$ identity matrix. Furthermore,*

$$\operatorname{rank} \mathcal{A} = p - 1. \tag{11}$$

*The system of equations (10) has infinitely many solutions which can be expressed as*

$$\exp(\xi) = \left(\mathcal{A}^- \mathcal{A} - \mathbf{I}_p\right) z, \tag{12}$$

*where $\mathcal{A}^-$ denotes the generalized inverse of $\mathcal{A}$ and where $z$ is an arbitrary vector in $\mathbb{R}^p$ such that the right hand side of (12) is positive.*

**Proof:**
Parts (10) and (11) follow from the definition of $\beta(\xi)$ and from simple algebra (notice that the sum of rows of $\mathcal{A}$ is equal to zero). Part (12) follows, e.g., from Anděl (1985, Theorem IV.18). □

It remains to choose the vector $z$ in (12) so that the solution of the system of equations (10) is positive.

**PROPOSITION 1** *The rank of the matrix $\mathcal{A}^-\mathcal{A} - \mathbf{I}_p$ is 1. Hence, any solution of the system of equations (10) is a multiple of the first column of the matrix $\mathcal{A}^-\mathcal{A} - \mathbf{I}_p$. The vector $z$ in (12) can be chosen, e.g., as $z = \pm \mathbf{1}_p$, where the sign is chosen so that the resulting solution is positive.*

**Proof:**
The definition of the generalized inverse is

$$\mathcal{A}\mathcal{A}^{-}\mathcal{A} - \mathcal{A} = \mathcal{A}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) = 0. \tag{13}$$

Lemma 3 says that $\operatorname{rank}\mathcal{A} = p - 1$. Hence, equation (13) implies that $\operatorname{rank}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) \leq 1$. Noticing that $\mathcal{A}^{-}\mathcal{A} \neq \mathbf{I}_p$ means that $\operatorname{rank}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) > 0$ and concludes the proof. □

## 3.3 The algorithm

The proposed algorithm consists of the following steps:

1. obtain a reasonable initial estimate $\hat{\beta}$, e.g., by running the Pool-Adjacent-Violators algorithm (Robertson, Wright and Dykstra 1988, Chapter 1) on the unconstrained least squares estimates of the first derivative of the curve,
2. transform the initial estimates $\hat{\beta}$ into the estimates $\hat{\xi}$ using the method described in Subsection 3.2,
3. minimize the nonlinear least squares as described in Subsection 3.1 using numerical methods.

An application of this simple algorithm on real data is given in the next section.

## 3.4 Asymptotic distribution

Assuming normality, the log-likelihood for the model (9) can be written as:

$$l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma) = -n \log \sigma - \frac{1}{2\sigma^2}\{\mathcal{C} - \mathcal{X}_\Delta\beta(\theta)\}^\top\{\mathcal{C} - \mathcal{X}_\Delta\beta(\theta)\}, \tag{14}$$

where $\mathcal{X}_\Delta$ is the design matrix given in (9). This normality assumption will be justified later by a residual analysis. The maximum likelihood estimator is defined as:

$$\hat{\theta} = \arg\max_\theta l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma). \tag{15}$$

Standard Maximum Likelihood theory (Serfling 1980) now says that the estimator $\hat{\theta}$, defined as:

$$\hat{\theta} = \arg\max_\theta l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma), \tag{16}$$

has asymptotically a $p$-dimensional normal distribution with mean $\theta$ and the variance given by the inverse of the Fisher information matrix:

$$\mathcal{F}_n^{-1} = \left\{-E\left(\frac{\partial^2}{\partial\theta\partial\theta^\top} l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma)\right)\right\}^{-1}. \tag{17}$$

More precisely,

$$n^{1/2}(\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} N_p(0, \mathcal{F}_n^{-1}). \tag{18}$$

The variance matrix of the asymptotic Normal distribution can be approximated using numerical maximization procedures.

In order to implement the described algorithm numerically, it is useful to express the contribution of the $i$-th row to the log-likelihood in the following form:

$$l_i(\theta) \quad = \quad -\log\sigma - \frac{1}{2\sigma^2} r_i^2, \tag{19}$$

where $\Delta_i^j = 0$ if $j \leq i$ and where $r_i$ denotes the $i$-th residual. The derivative of (19) with respect to the unknown parameters $\theta_0, \ldots, \theta_p$ is

$$\begin{aligned}
\frac{\partial l_i(\theta)}{\partial \theta_0} &= \frac{1}{\sigma^2} r_i \exp(\theta_0), \\
\frac{\partial l_i(\theta)}{\partial \theta_k} &= -\frac{1}{\sigma^2} r_i \left\{ -\Delta_p^i \frac{\exp(\theta_1)\exp(\theta_k)}{\left\{\sum_{j=1}^p \exp(\theta_j)\right\}^2} \right. \\
&\quad - \cdots + \Delta_{p-k+1}^i \frac{\exp(\theta_k)\left\{\sum_{j=1}^p \exp(\theta_j) - \exp(\theta_k)\right\}}{\left\{\sum_{j=1}^p \exp(\theta_j)\right\}^2} \\
&\quad \left. - \cdots - \Delta_2^i \frac{\exp(\theta_{p-1})\exp(\theta_k)}{\left\{\sum_{j=1}^p \exp(\theta_j)\right\}^2} \right\} \\
&= -\frac{1}{\sigma^2} r_i \beta_k(\theta) \left\{ \Delta_{p-k+1}^i - (\Delta_p^i, \ldots, \Delta_2^i) \begin{pmatrix} \beta_1(\theta) \\ \vdots \\ \beta_{p-1}(\theta) \end{pmatrix} \right\}.
\end{aligned}$$

The above expressions are useful especially for the software implementation of the numerical minimization in the second step of the algorithm 3.3.

## 3.5 Put-Call parity

The prices of put options can be easily included in our estimation technique by applying the Put-Call parity of the option prices. Each put option with price $P_t(K,T)$ corresponds to a call option with price

$$C_t(K,T) = P_t(K,T) + S_t - Ke^{-r(T-t)}.$$

In this way, the prices of the put options can be converted into the prices of call options and used in our model (Franke, Härdle, Hafner 2004). Statistically speaking, these additional observations will increase the precision of the SPD and will lead to more stable results. In Section 5, we will investigate also the covariance of the observed call and put option prices.

Another way to include the prices of the put options in our procedure is to fit the two curves separately using two sets of parameters under the assumptions of common SPD. The situation is displayed in Figure 5. The natural assumption that the same SPD drives both the put and call option prices leads to the following conditions on the coefficients $\alpha_i$ and $\beta_i$

$$\begin{aligned}
\alpha_i &= \beta_{p-i+1}, \quad \text{for } i = 2, \ldots, p-1 \\
\alpha_1 &= 1 - \sum_{i=1}^{p-1} \beta_i.
\end{aligned}$$

The problem of estimating regression functions under such linear equality constraints is solved, e.g., in Rao (1973).

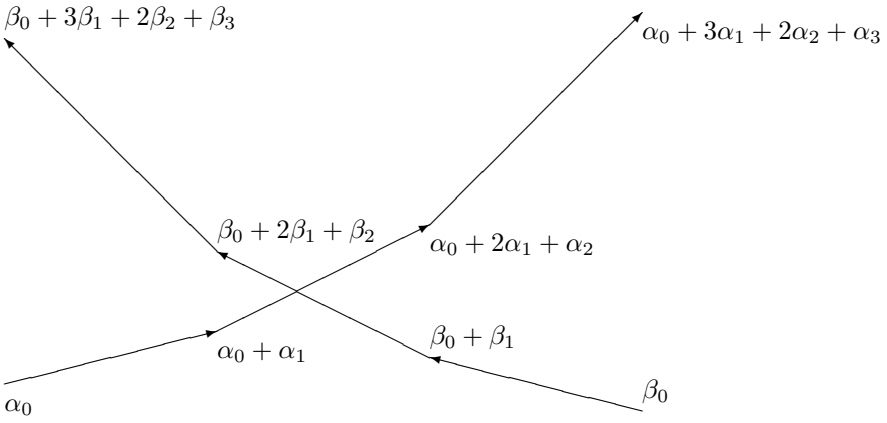


Figure 5: Illustration of the dummy variables for both call ($\beta$) and put ($\alpha$) options.

# 4 Application to DAX data

In order to illustrate the method, we apply it to DAX option prices on two consecutive days. These days (16th and 17th January 1995) were selected since they provide nice insight into the behaviour of the presented methods.

The observed option prices on one day (16th January 1995) are plotted on the left hand side of Figure 1 against maturity and strike price. The shape of dependency of the option price on the strike price can be nicely observed. For simplicity, in the following analyses we restrict ourselves only to data for fixed maturity as displayed on the right plot in Figure 1.

In Figures 3 and 4 we observe the difference between the unconstrained linear regression estimate and the constrained nonlinear regression estimate described in Section 3.1.

In Figure 3, the unconstrained model incidentally satisfies all conditions on the shape of the curve. Hence, the estimates of the curve itself (1st plot) its first derivative (2nd plot) and the SPD (3rd plot) coincide.

On 17th January, the situation becomes more interesting and it illustrates very clearly the advantages of the constrained estimator. In Figure 4, we plot the unconstrained and the constrained estimates using thick and thin line, respectively. Clearly, the difference between the fitted data (1st plot) is very small. However, this small difference in the first plot results in huge differences in the estimate of the first derivative (2nd plot) and especially in the estimate of the second derivate, the SPD, in the 3rd plot of Figure 4.

## 4.1 Interpretation of the estimates

The coefficients, $\hat{\beta}_{p-1}, \ldots, \hat{\beta}_2$, plotted in the bottom panels of Figures 3 and 4 can be described as estimates of the changes of the first derivative in that point. Since the first derivative of the curve corresponds to the integrated SPD, the coefficients $\hat{\beta}_{p-1}, \ldots, \hat{\beta}_2$ estimate probabilities associated with the corresponding strike price. We interpret the coefficients as a histogram-like estimator of the state price density. The asymptotic dis-

tribution of the estimator, based on Maximum Likelihood theory is given in Subsection 3.4.

## 4.2 Confidence intervals

We present two simple methods for calculating pointwise confidence intervals for the SPD. The description of the $x$-axis in Figures 6 and 7 shows the number of observations at each of the design points.

Notice that, in the unconstrained model, the estimates of the values of the SPD are just the parameters of the linear regression model. Hence, the confidence intervals for the parameters are also confidence intervals for the SPD. These confidence intervals for 16th and 17th January are displayed in the upper plots in Figures 6 and 7. The drawbacks of this method are clearly visible. In Figure 6, the lower bounds of the confidence intervals only asymptotically satisfy the condition of positivity. In Figure 7, we observe large variability on the left-hand side of the plot (the region with low number of observations). Again, some of the lower bounds are not positive.

Clearly, the confidence intervals based on the unconstrained model make sense only if the constraints are, by chance, satisfied. Even if this is the case, there is no guarantee that the lower bounds will be positive.

The lower panels in Figures 6 and 7 display confidence intervals conditional on the fact that $\sum_{i=1}^{p} \exp(\theta_i) < 1$. Using maximum likelihood theory, we calculate confidence intervals for the parameters $\theta$ (rescaled so that $\sum_{i=1}^{p} \exp(\theta_i) = 1$). Exponentiating the limits of these confidence intervals leads to valid confidence intervals for parameters $\beta$.

In Figure 6, both type of confidence intervals provide very similar results. The only difference is at the minimum and maximum value of the independent variable (strike price) where the unconstrained method provides negative lower bounds and the conditional method leads to very large upper bounds of the confidence intervals.

In Figure 7, we plot the confidence intervals for January 17th. Here, the unconstrained and the conditional methods lead to very different estimates. We can observe that the confidence intervals on the right hand side are much narrower for the conditional method. On the left hand side, both methods tend to provide confidence intervals that seem to be overly wide. For the conditional method, we observe that the length of the confidence intervals explodes when the estimated value of the SPD is very close to zero and, at the same time, the number of observation in that region (see the description of the $x$-axis) is small.

## 4.3 Residual analysis

The residuals on 17th January 1995 are plotted in Figure 8. The time of trade (in hours) is denoted by the plotting symbol. The circle, square and the star denote denote the trades carried out in the morning, midday and in the afternoon, respectively. The size of the symbols corresponds to the number of residuals lying in the respective areas.

The majority of the residuals correspond to the strike prices of 2075,- and 2100,-DM. The variance fo the residuals is very low on the right hand side of the plot and it rapidly increases when moving towards smaller maturities. On the left hand side of the plot, for maturities smaller than 2000, we have only very few observations and cannot judge the residual variability reliably.

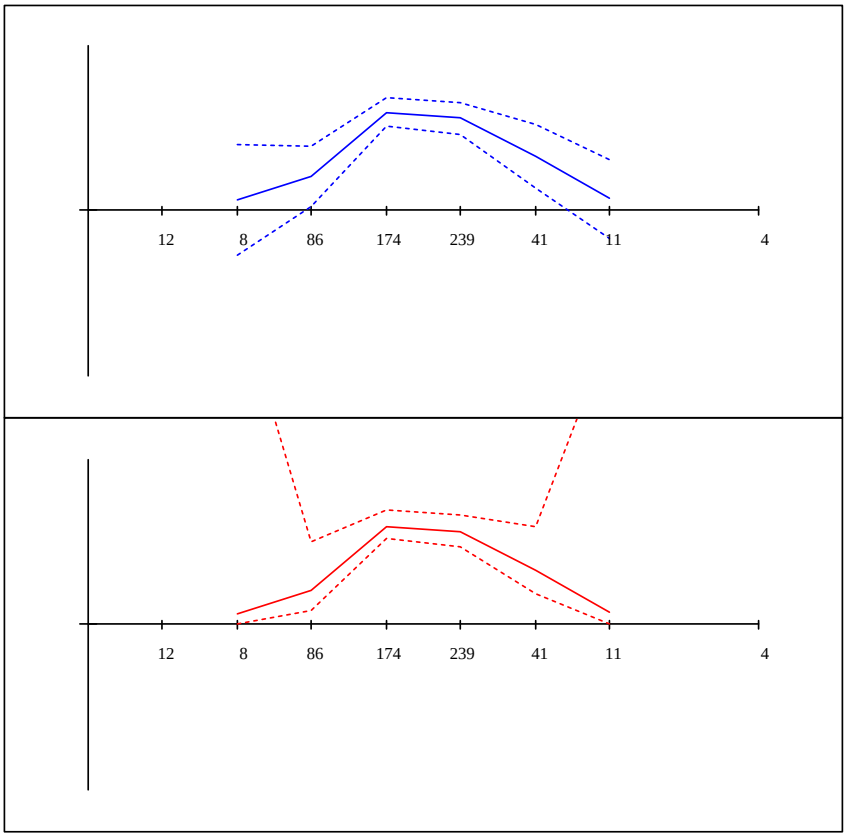


Figure 6: The unconstrained and constrained confidence intervals for SPD on 16th January 1995. The description on the $x$-axis shows the number of observations in each point.

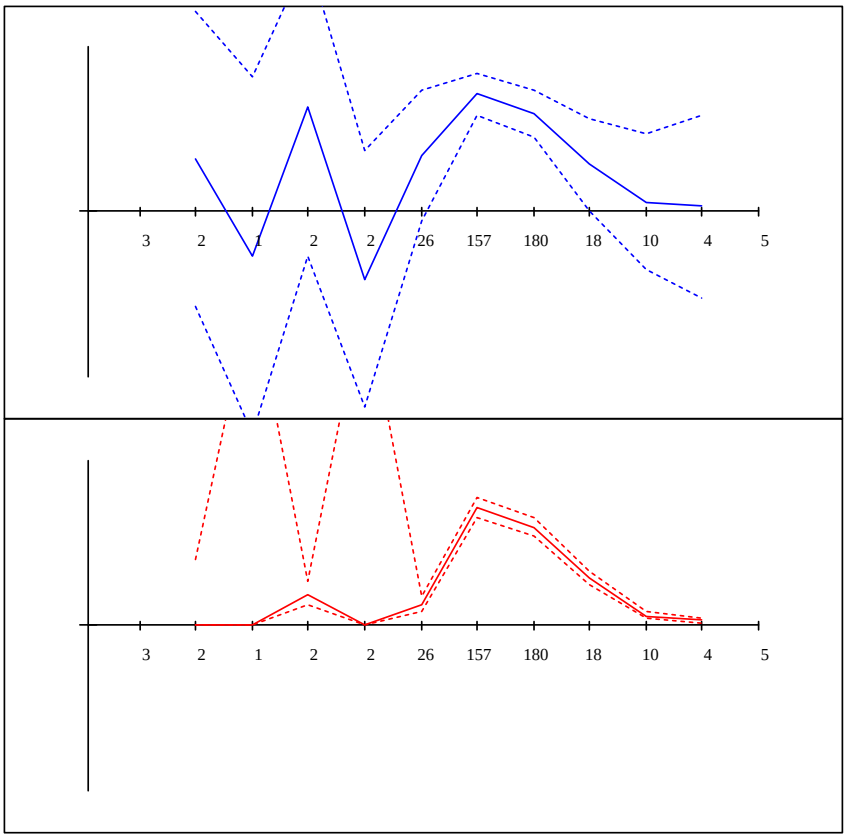


Figure 7: Confidence intervals for SPD on 17th January 1995. The description on the $x$-axis shows the number of observations in each point.

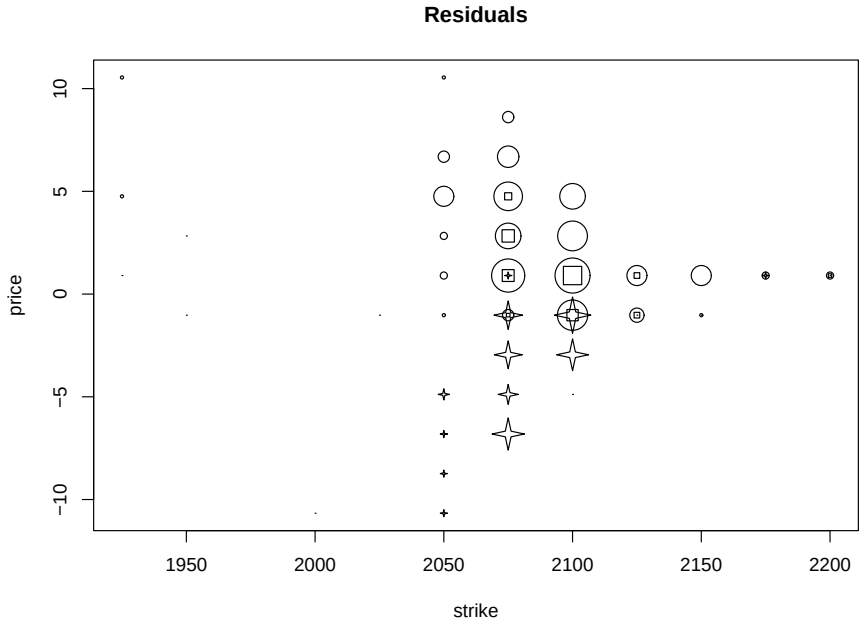


Figure 8: The time dependency and the heteroscedasticity of the residuals during one day (19950117). The circle, square and the star denote denote the trades carried out in the morning, midday and in the afternoon, respectively. Size of the symbols denotes number of residuals.

Apart of the obvious heteroscedasticity we observe also a very strong systematic movement in the SPD throughout the day: the circles, corresponding to the first third of the day, are positive and all stars, denoting the afternoon residuals, are negative. Similar patterns can be observed every day—residuals corresponding to the same time are having the same sign.

We conclude that the option prices naturally tend to follow the changes of the market during the day. A model, describing this behaviour, will be presented in the next Section 5.

# 5 Covariance structure

Up to now, we worked with the model

$$C_t(k_j) \quad = \quad \Delta_j \widetilde{\beta} + \varepsilon_t \tag{20}$$

or

$$\begin{aligned} C_t(k_j) &= \Delta_j \widetilde{\beta}_t + \varepsilon_t, \\ \widetilde{\beta}_t &= \widetilde{\beta}_{t-1}, \end{aligned} \tag{21}$$

where $t$ is the time, $\widetilde{\beta}$ denotes the column vector of the unknown parameters, and $\Delta_j$ denotes the $j$-th row of the matrix $\Delta$ defined in (7), i.e.,

$$\Delta_j = (1, \Delta_p^j, \Delta_{p-1}^j, \ldots, \Delta_{j+1}^j, \underbrace{0, \ldots, 0}_{(j-1)}).$$

However, the residual analysis in Section 4.3 clearly shows that it is not appropriate in this situation. Therefore, we will consider some generalizations that would lead to a better fit of the data set.

## 5.1 Heteroscedasticity

Assume that the $t$-th observation, corresponding to the $j$-th exercise price $k_j$, can be written as

$$C_t(k_j) \quad = \quad \Delta_j \widetilde{\beta}_t, \tag{22}$$

$$\widetilde{\beta}_t \quad = \quad \widetilde{\beta} + \varepsilon_t, \tag{23}$$

i.e., there is a random error $\varepsilon_t$ in the state price density $\widetilde{\beta}_t$. Clearly, the variance of the observed option prices $C$ is then

$$\operatorname{Var} C = \sigma^2 \operatorname{diag}(\mathcal{X}_\Delta \mathcal{X}_\Delta^\top), \tag{24}$$

where $\mathcal{X}_\Delta$ is the design matrix in which each row of the matrix $\Delta$ is repeated $n_j$ times, $j = 1, \ldots, p$.

**REMARK 2** *Assuming that the observations have the covariance structure (24), the least squares estimates do not change and*

$$\operatorname{Var} \hat{\beta} = \sigma^2 \{\mathcal{X}_\Delta^\top \operatorname{diag}(\mathcal{X}_\Delta \mathcal{X}_\Delta^\top)^{-1} \mathcal{X}_\Delta\}.$$

Another possible model for the heteroscedasticity would assume that the changes are multiplicative rather than additive.

$$\begin{aligned} C_t(k_j) &= \Delta_j \widetilde{\beta}_t \\ \log \widetilde{\beta}_t &= \log \widetilde{\beta} + \varepsilon_t \end{aligned}$$

This model leads to a variance of $C_t(k_j)$ that depends on the value of the SPD:

$$\mathrm{Var}\, C_t(k_j) = \sigma^2\{\beta_0^2+(\Delta_p^j)^2\beta_1^2+(\Delta_{p-1}^j)^2\beta_2^2+(\Delta_{p-2}^j)^2\beta_3^2+\cdots+(\Delta_{j+1}^j)^2\beta_j^2\}.$$

It is straightforward that Remark 2 applies also in this situation.

## 5.2 Covariance

Assume that there are random changes in the state price density coefficients $\widetilde{\beta}_t$ over time so that for equidistant time points $t$ we have

$$\begin{aligned} C_t(k_j) &= \Delta_j\widetilde{\beta}_t, \\ \widetilde{\beta}_t &= \widetilde{\beta}_{t-1}+\varepsilon_t, \end{aligned} \tag{25}$$

where, at fixed time $t$, $\widetilde{\beta}_t$ is the parameter vector and $\varepsilon_i$, $i = t, t-1, \ldots$ are iid random vectors having iid components with zero mean and variance $\sigma^2$. For nonequidistant time points, let $\delta_t$ denote the time between the $t$-th and $(t-1)$-st observation. The model is

$$\begin{aligned} C_t(k_j) &= \Delta_j\widetilde{\beta}_t, \\ \widetilde{\beta}_t &= \widetilde{\beta}_{t-1}+\delta_t^{1/2}\varepsilon_t \end{aligned} \tag{26}$$

and it leads to the covariance matrix with elements

$$\begin{aligned} \mathrm{Cov}\{C_{t-u}(k_j), C_{t-v}(k_i)\} &= \mathrm{Cov}(\Delta_j\widetilde{\beta}_{t-u}, \Delta_i\widetilde{\beta}_{t-v}) \\ &= \sigma^2\Delta_j\Delta_i^\top\sum_{l=1}^{\min(u,v)}\delta_{t+1-l}. \end{aligned} \tag{27}$$

The analogous model on the log scale is

$$\begin{aligned} C_t(k_j) &= \Delta_j\widetilde{\beta}_t, \\ \log\widetilde{\beta}_t &= \log\widetilde{\beta}_{t-1}+\delta_t^{1/2}\varepsilon_t. \end{aligned} \tag{28}$$

Applying Taylor expansion we obtain an approximation of the covariance structure as

$$\begin{aligned} \mathrm{Cov}\{C_{t-u}(k_j), C_{t-v}(k_i)\} &= \mathrm{Cov}(\Delta_j\widetilde{\beta}_{t-u}, \Delta_i\widetilde{\beta}_{t-v}) \\ &\doteq \sigma^2\Delta_j\widetilde{\beta}_t\widetilde{\beta}_t^\top\Delta_i^\top\sum_{l=1}^{\min(u,v)}\delta_{t+1-l}. \end{aligned} \tag{29}$$

## 5.3 Including put options

Similarly as the price of the call options, $C_t(k_j)$,we can write the model for the price of the put options, $P_t(k_j)$, as

$$P_t(k_j) = \Delta_j^P\widetilde{\alpha}+\delta_t\varepsilon_t \tag{30}$$

or

$$\begin{aligned} P_t(k_j) &= \Delta_j\widetilde{\alpha}_t+\delta_t^{1/2}\varepsilon_t, \\ \widetilde{\alpha}_t &= \widetilde{\alpha}_{t-1}, \end{aligned} \tag{31}$$

where $\widetilde{\alpha}_t$ denotes the column vector of the unknown parameters corresponding to time $t$ and $\Delta_j^P$ denotes the corresponding row of the design matrix, i.e.,

$$\Delta_j^P = (1, \Delta_j^1, \Delta_j^2, \ldots, \Delta_j^{j-1}, \underbrace{0, \ldots, 0}_{(p-j)}).$$

Using the relations between the $\alpha$ and $\beta$ parameters, $\alpha_i = \beta_{p-i+1}$, for $i = 2, \ldots, p-1$. After some simplifications we obtain

$$P_t(k_j) \quad = \quad \Delta_j^P \widetilde{\alpha} + \delta_t^{1/2} \varepsilon_t, \tag{32}$$

where $\widetilde{\alpha} = (\alpha_0, \alpha_1, \beta_{p-1}, \beta_{p-2}, \ldots, \beta_2)^\top$. In this way, we obtain a joint estimation strategy for both the call and put option prices:

$$\begin{aligned} C_t(k_j) &= \Delta_j \widetilde{\beta}_t, \\ P_t(k_j) &= \Delta_j^P \widetilde{\alpha}_t, \\ \begin{pmatrix} \widetilde{\beta}_t \\ \widetilde{\alpha}_t \end{pmatrix} &= \begin{pmatrix} \widetilde{\beta}_{t-1} \\ \widetilde{\alpha}_{t-1} \end{pmatrix} + \delta_t^{1/2} \varepsilon_t, \end{aligned} \tag{33}$$

which directly leads to the covariances

$$\begin{aligned} \mathrm{Cov}\{P_{t-u}(k_j), P_{t-v}(k_i)\} &= \mathrm{Cov}(\Delta_j^P \widetilde{\alpha}_{t-u}, \Delta_i^P \widetilde{\alpha}_{t-v}) \\ &= \sigma^2 \Delta_j^P (\Delta_i^P)^\top \sum_{l=1}^{\min(u,v)} \delta_{t+1-l}. \end{aligned} \tag{34}$$

and

$$\begin{aligned} &\mathrm{Cov}\{C_{t-u}(k_j), P_{t-v}(k_i)\} \\ &\quad = \mathrm{Cov}(\Delta_j \widetilde{\beta}_{t-u}, \Delta_i^P \widetilde{\alpha}_{t-v}) \\ &\quad = \sigma^2 \sum_{l=1}^{\min(u,v)} \delta_{t+1-l} \sum_{k=2}^{p-1} \Delta_{p+1-k}^j \Delta_i^{p+1-k}. \end{aligned} \tag{35}$$

## 5.4 Application

In Figure 9, we present the improvement achieved by considering both the call and put option prices and by applying the covariance structure (29,34,35) that takes into account the correlations of the observed option prices. In comparison with Figure 7, we observe the shorter length of the confidence intervals.

The residual plots in Figure 10 omits large residuals that correspond to the lowest and highest strike prices—the model doesn't seem to capture the variability in these points very well. All the remaining residuals plotted in the top panel of Figure 10 (using the same technique as in Figure 8) do not display any time dependency anymore. The lower panel of Figure 10 displays the histogram of the residuals. The distribution of the residuals seems to be symmetric and not too far from Normal distribution. However, the kurtosis of this distribution is too large and formal tests reject normality.

The proposed estimation technique can be also naturally used for predictions. In order to predict and construct confidence intervals for the SPD in time $t + \tau$ we enhance the model (33) by considering

$$\begin{pmatrix} \widetilde{\beta}_{t+\tau} \\ \widetilde{\alpha}_{t+\tau} \end{pmatrix} \quad = \quad \begin{pmatrix} \widetilde{\beta}_t \\ \widetilde{\alpha}_t \end{pmatrix} + \delta_{t,\tau}^{1/2} \varepsilon_{t,\tau}. \tag{36}$$

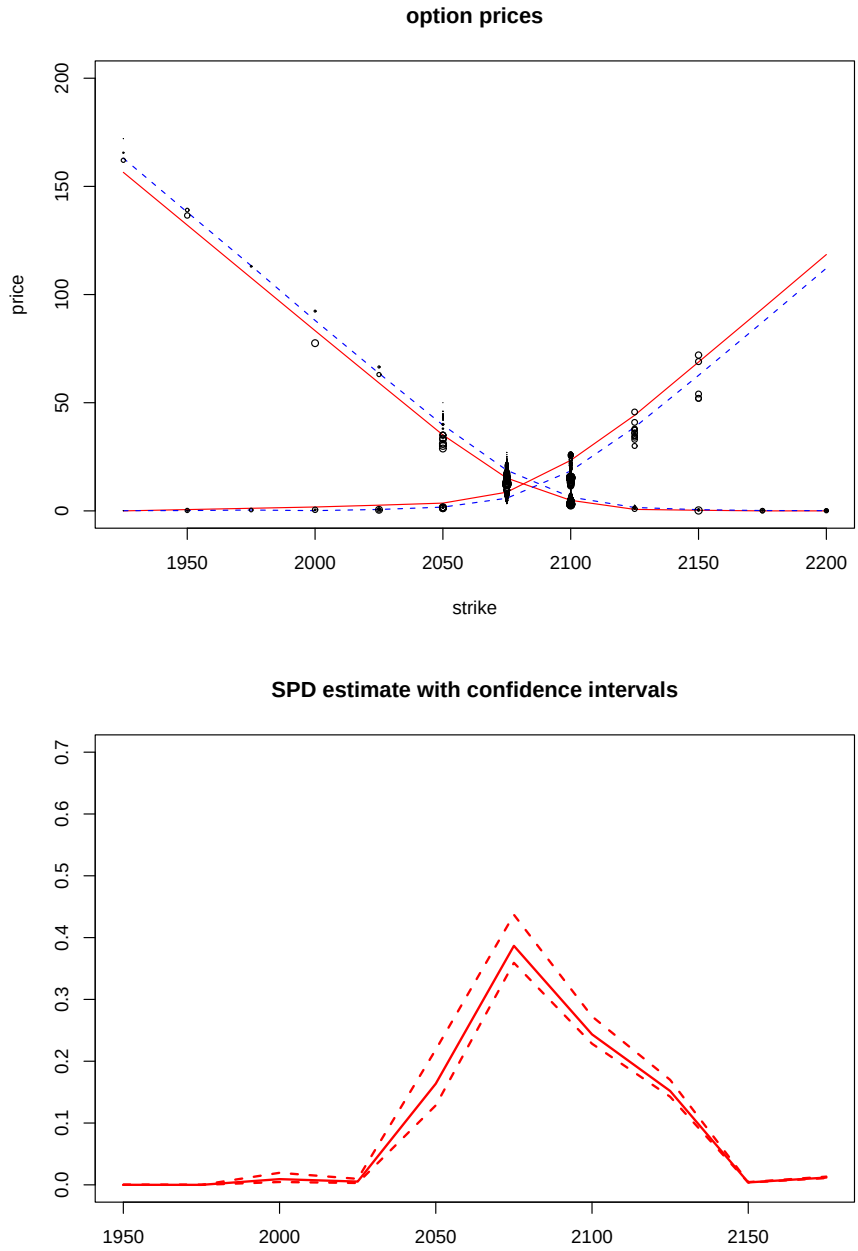


Figure 9: Estimate with confidence intervals and the residuals when using the covariance structure (29, 34, 35) on 17th January. The upper plot shows the observed option prices with the constrained (solid lines) and the unconstrained estimate (dashed lines). The size circles corresponds to the weight of the observations—the largest circles are the most recent observations. The lower plot shows the estimated SPD with the confidence intervals.

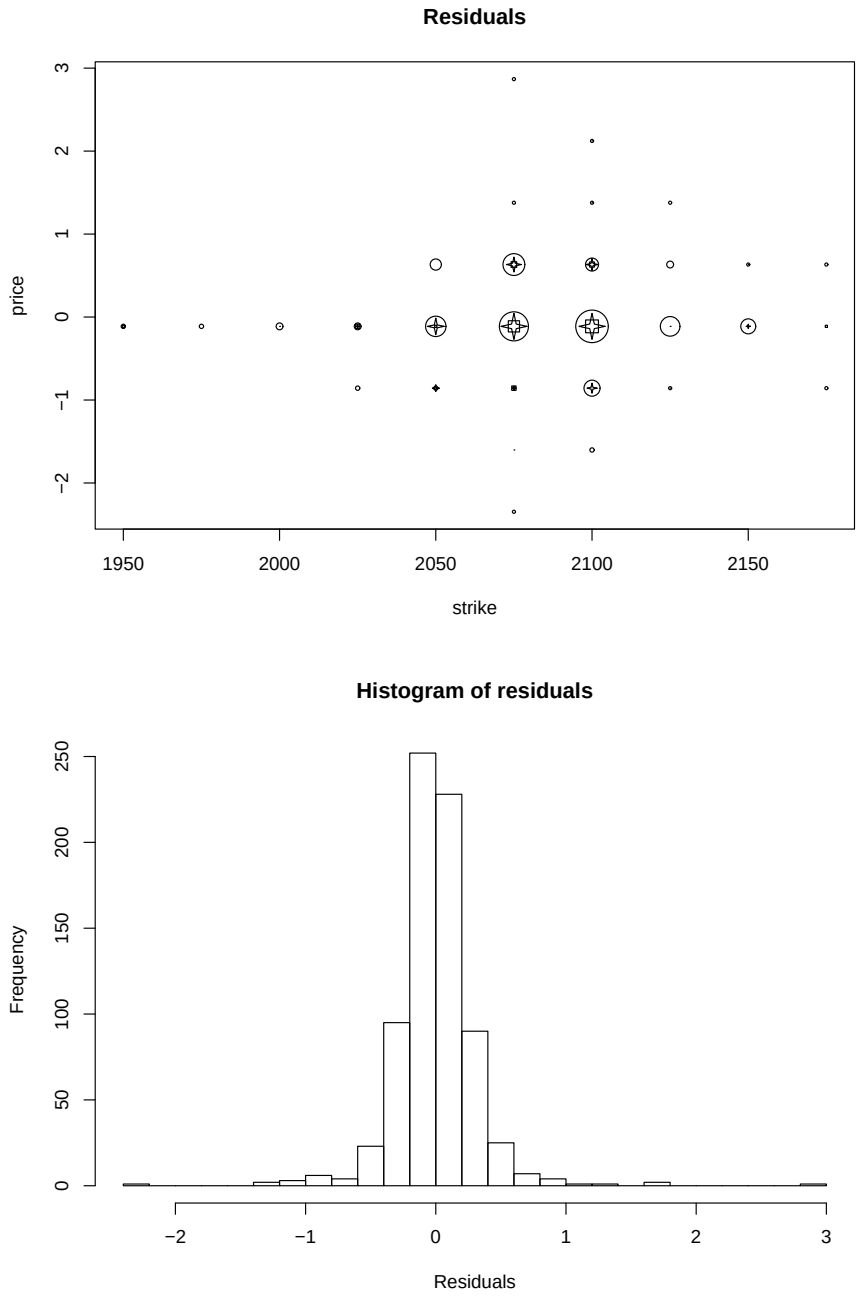


Figure 10: The development of the residuals resulting from model with the covariance structure (29, 34, 35) on 17th January during the day where circles, squares, and stars denote the residuals from morning, midday, and afternoon and a histogram of the residuals.

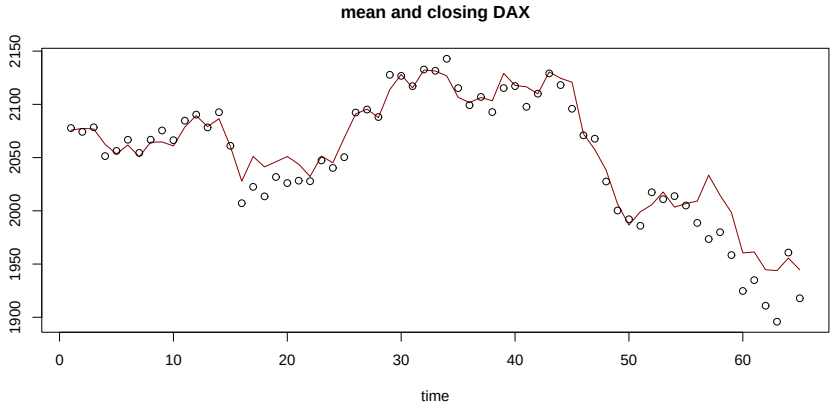


Figure 11: Daily development of the expected value of the SPD from January till March 1995. The circles denote the corresponding closing value of DAX.

It is straightforward that the estimation technique for the SPD in time $t+\tau$ does not change. The only difference is an increase in the covariances which now reflects the increase in uncertainty resulting from the random changes of the SPD between the time points $t$ and $t+\tau$.

# 6 Dynamics of SPD

In order to study the dynamics of SPD, we calculated basic moment characteristics of the estimated SPDs. The mean, variance, skewness, and kurtosis in the first Quarter of 1995 are plotted as lines in Figures 11–14. Note, that the SPDs in this period were always estimated using the options with shortest time to maturity. This means that the time to maturity is decreasing linearly in all plots, but it jumps up whenever the option with the shortest time to maturity expires. These jumps occured at days 16, 36, and 56 in all plots.

In Figure 11, the daily closing values of DAX are plotted as circles. We observe that the means of the SPD, displayed as the line, follow closely the value of the DAX index. In two periods, days 17–22 and days 56–63, the mean of the SPD lies above the DAX. It could be that the option market might have expected change of trend in these days.

In the second plot in Figure 12, we see that the variance of SPD decreases linearly as the option moves closer to its maturity. This observation suggests that SPD estimates calculated for neighbouring maturities can be linearly interpolated in order to obtain SPD estimate with arbitrary time to maturity. Such an estimate is important for making the SPD estimates comparable and for studying the development of the market expectations.

The skewness and the kurtosis in Figures 13 and 14 are for good reasons more volatile closely before the expiry of the option, but otherwise do not display any obvious pattern.

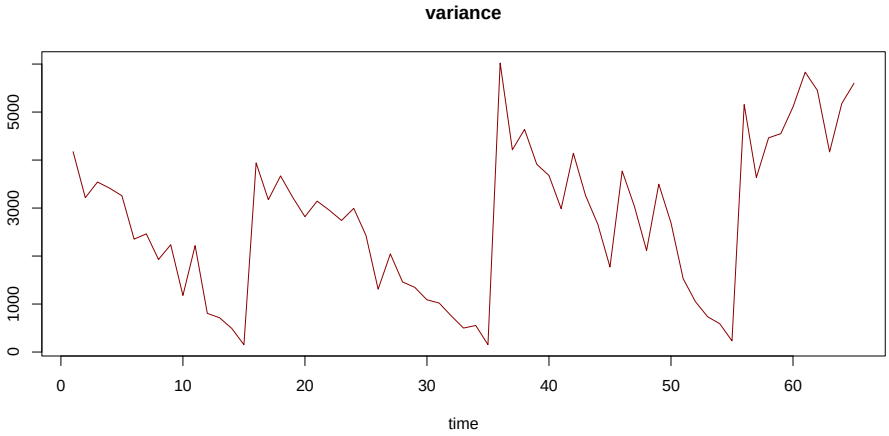


Figure 12: Daily development of the variance of the SPDs from January till March 1995.

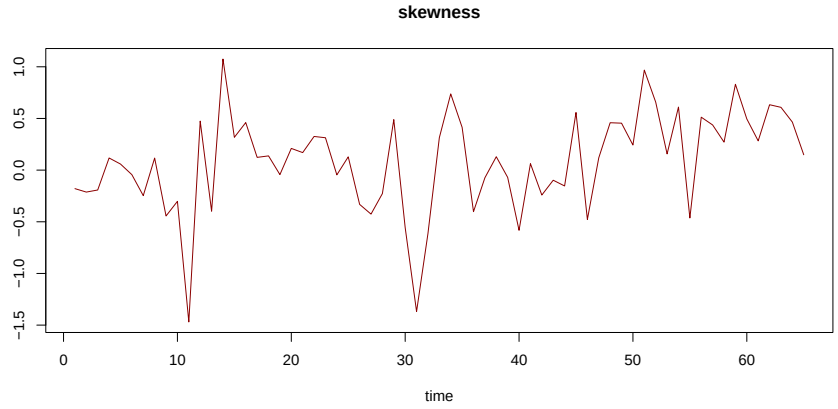


Figure 13: Daily development of the skewness of the SPDs from January till March 1995.

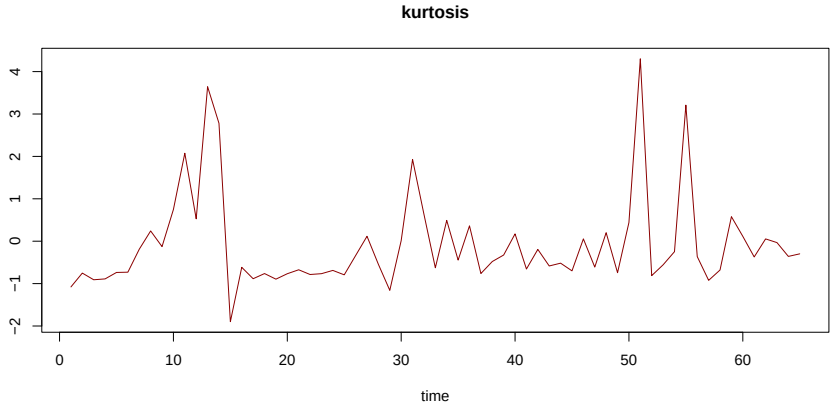


Figure 14: Development of the kurtosis of the SPDs from January till March 1995.

## 6.1 Estimate with the fixed time to expiry

The variances displayed in Figure 12 suggest that the variance of the SPD estimates changes approximately linearly in time when moving closer to the date of expiry.

Hence, from the estimates $f_{\tau_1}(.)$ and $f_{\tau_2}(.)$ of SPDs corresponding to the times of expiry $\tau_1$ and $\tau_2$, we construct an estimate $f_\tau(.)$ for any time of expiry $\tau \in (\tau_1, \tau_2)$ as

$$f_\tau(.) = \frac{(\tau_2 - \tau)^{1/2} f_{\tau_1}(.) + (\tau - \tau_1)^{1/2} f_{\tau_2}(.)}{(\tau_2 - \tau_1)^{1/2}}. \tag{37}$$

In this way, the variance, $V_\tau$, of the centered SPD with time to expiry equal to $\tau$ can be expressed as

$$\begin{aligned} V_\tau &= \int x^2 f_\tau(x) dx \\ &= \int x^2 \frac{(\tau_2 - \tau)^{1/2} f_{\tau_1}(.) + (\tau - \tau_1)^{1/2} f_{\tau_2}(.)}{(\tau_2 - \tau_1)^{1/2}} dx \\ &= \frac{(\tau_2 - \tau) V_{\tau_1}(.) + (\tau - \tau_1) V_{\tau_2}(.)}{\tau_2 - \tau_1}. \end{aligned}$$

We argue that such an estimate is reasonable since we observed in Figure 12 that the SPD variances change linearly in time. The centering of the SPDs does not present any additional difficulty since all SPD estimates calculated at time $t$ should have the same expected value.

The resulting estimates for the time of expiry $\tau = 45$ days for the first trading days in years 1995, 2001, and 2003 are given in Figure 15. We observe very large changes in the market expectations during this period. In 1995, the expectations (45 days ahead) were concentrated just above 2000. Six years later, in 2001, the expectations of the market look very different. The center of the histogram is now three times larger than in 1995. Also the variance of the SPD is much larger. In 2003, the prices are only about one half of the prices in 2001, but the variance of the distribution remains very high.

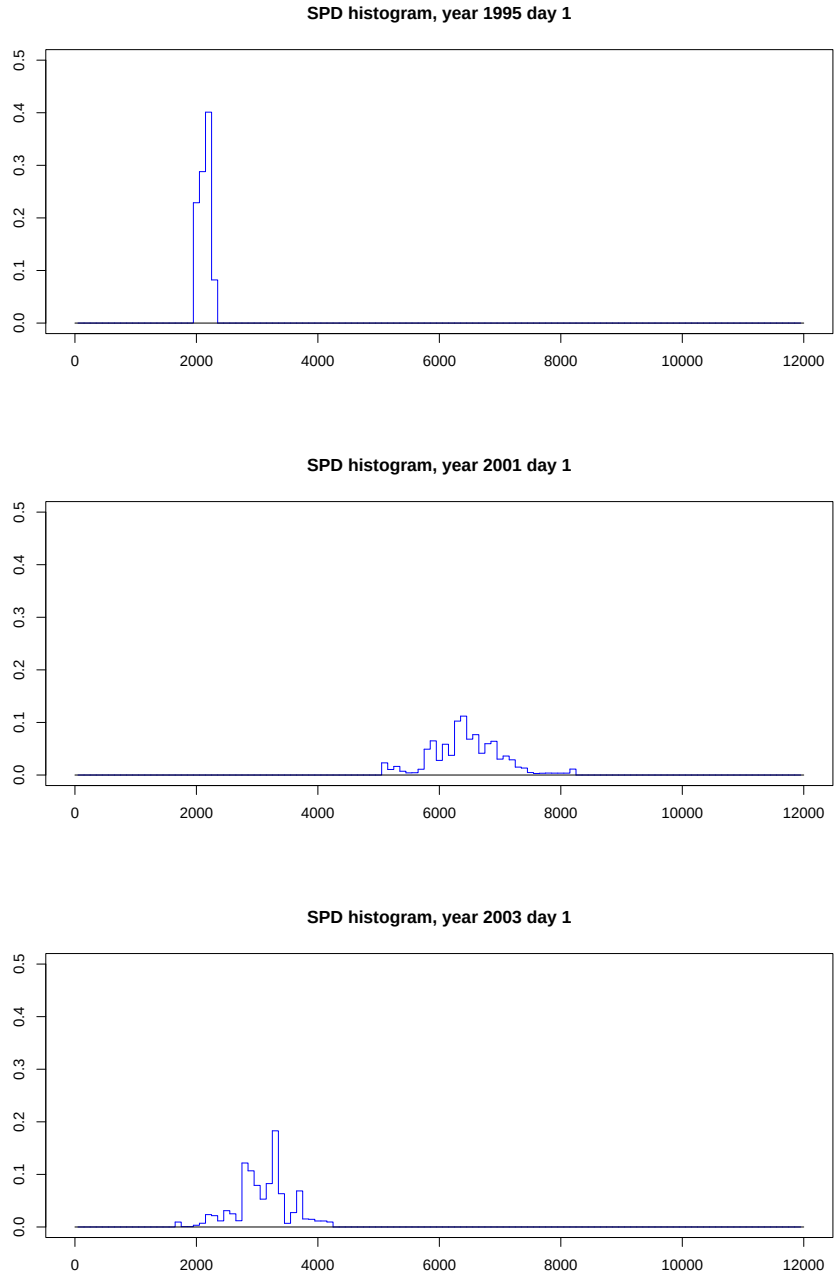


Figure 15: Histograms of the SPD estimates on the first trading day in years 1995, 2001, and 2003.

## 6.2 Verification of the market's expectations

The SPD can be interpreted as the market's expectation of the behaviour of the value of the DAX in 45 days. Hence, it is interesting to use our data set to verify how these expectations compare with reality. In Figures 16–18, we plot two types of prediction intervals based on the SPD together with the true future value of the DAX. The red dashed line is calculated as the mean of the SPD plus and minus two times its standard deviation. The green dotted line displays the 2.5% and 97.5% quantiles of the SPD. Both intervals should cover the future value of the DAX, displayed as a full black line, with probability 0.95.

Figures 16–18 suggest that the method works well and that the DAX mostly stays well within our expectations. The DAX was sometimes rising faster than the market expected from 1995 till mid 1998. After a fast decrease in the second half of 1998, the market increases again till the beginning of year 2000. Since then, the market decreases. However, the changes stay mostly within or very close to the bounds predicted by our SPD estimates. The only exception is a large shock observed in September 2001 caused by the terrorist attack on the World Trade Center.

The plots in Figures 16–18 have different scales which makes the comparison of different years more difficult. Therefore, we display the predictions for all years in Figure 19. Here, we can see the huge changes in the market between 1995 and 2003. The value of DAX was in 2003 only slightly higher than in 1995, but the volatility was much larger.

## 6.3 Evaluation of the quality of the forecasts

The quality of the forecasts can be evaluated by comparing the true future observation with its predicted distribution (the SPD). Craig, Glatzer, Keller and Scheicher (2003) evaluate the forecasting performance of German stock option densities using the probability integral transformed observations $z_{h,t}$, where $t$ denotes the time and $h$ the forecasting horizon. More precisely, we define

$$z_{h,t} = \int_{-\infty}^{X_{t+h}} \widehat{f}_{h,t}(u)du,$$

where $\widehat{f}_{h,t}(.)$ denotes our estimate of the SPD $h$ days ahead at time $t$ and $X_{t+h}$ is the future observation. In other words, $z_{h,t}$ is the probability value of $X_{t+h}$ with respect to $\widehat{f}_{h,t}(.)$. Clearly, the $z_{h,t}$ should be uniformly $U(0,1)$ distributed if the estimated SPD $\widehat{f}_{h,t}(.)$ is equal to the true density of $X_{t+h}$. In Figure 20, we display the histograms of $z_{h,t}$'s for each year. Clearly, in the ideal case, the histograms should not be too far from Uniform $U(0,1)$ distribution. In our data, we observe that the histograms look quite different than we would expect. Especially in years 1995–1999, the DAX was moving mainly in the upper quantiles of the predicted SPD.

In order to account for the overlapping forecasting periods, we calculate the confidence limits for the empirical distribution function

$$\widehat{F}(u) = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{I}(z_{h,t} \leq u)$$

of $z_{h,t}$'s that take into account the autocorrelation structure.

$$\widehat{\mathrm{Var}}(\widehat{F}(u)) = \frac{1}{T}\left\{\widehat{\gamma}_u(0) + 2\sum_{j=1}^{h}\left(1-\frac{j}{T}\right)\widehat{\gamma}_u(j)\right\}, \qquad (38)$$


Figure 16: SPD based predictions in years 1995 up to 1997.


Figure 17: SPD based predictions in years 1998 up to 2000.

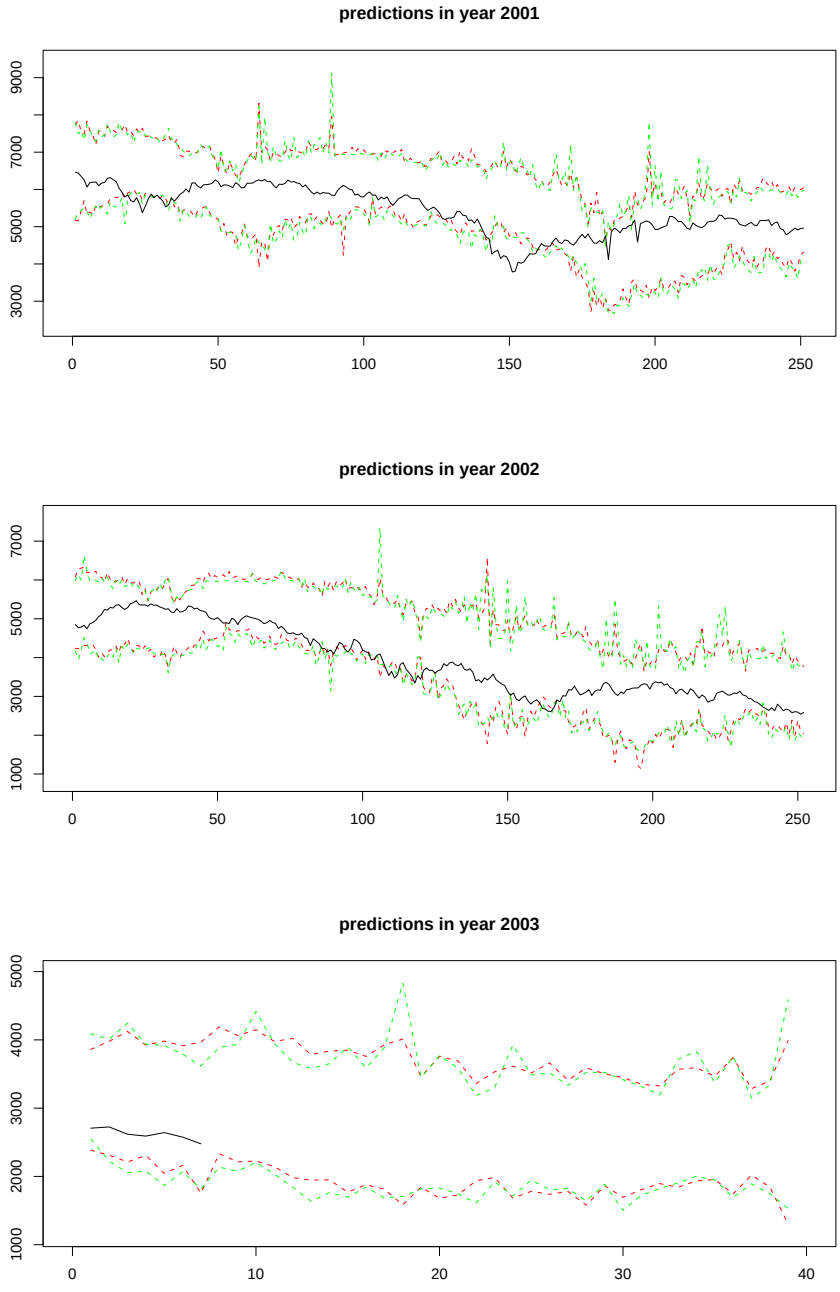


Figure 18: SPD based predictions in years 2001 up to 2003.

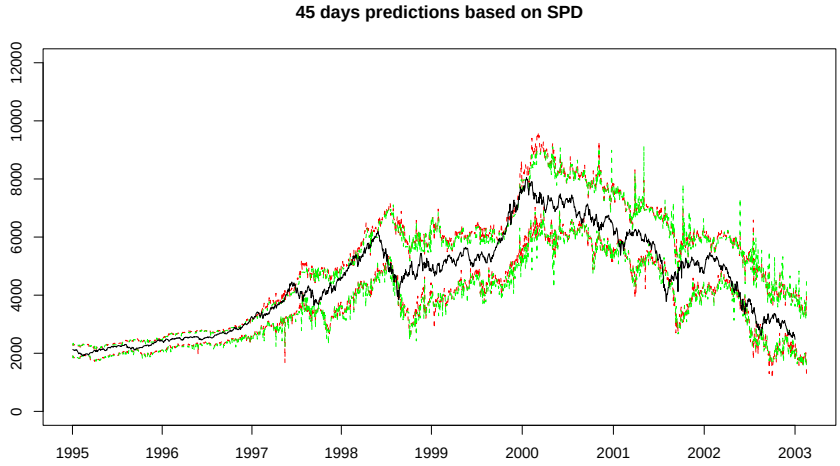


Figure 19: SPD based predictions from January 1995 till March 2003.

where $\gamma_u(j)$ is the sample autocovariance of order $j$:

$$\gamma_u(j) = \frac{1}{T}\sum_{t=j+1}^{T}\left\{\boldsymbol{I}(z_{h,t} \le u) - \widehat{F}(u)\right\}\left\{\boldsymbol{I}(z_{h,t-j} \le u) - \widehat{F}(u)\right\}.$$

The empirical distribution functions $\widehat{F}(.)$ are plotted separately for years 1995–2002 in Figure 21. The distribution function of $U(0,1)$ and the limits following from (38) are displayed as dotted lines. The year 2003 was not included since our dataset contains only two months of the year 2003 which did not leave enough observations to confirm the forecasts.

In years 1995–1997, the market was growing much faster than the SPDs were predicting. In year 1996, it never happened that the DAX fell bellow the 40% quantile of the forecast distribution and there were only few days when this value was bellow 60%. The situation in years 1998 and 1999 was less extreme even though the fast growth of DAX continued. The forecast distribution given by the SPD estimate $\widehat{f}_{t,h}(.)$ does not differ significantly from the true distribution of $X_{t+h}$ in years 2000–2002. Thus, the DAX was growing much faster than the option market expected in years 1995–1999 and it was falling according to the SPD forecasts in years 2000–2002.

# 7 Conclusion

We have proposed a simple linear model for the arbitrage free estimation of the SPD. Our procedure takes care of the daily changing covaariance structure and involves—using the Put-Call parity—both types of European options. We analyze the moment dynamics of the SPD from 1995–2003 and find by Maximum Likelihood theory confidence intervals for the (future) SPD. An application to DAX EUREX data for the years 1995–2003 produces a forecast prediction corridor that is almost perfectly

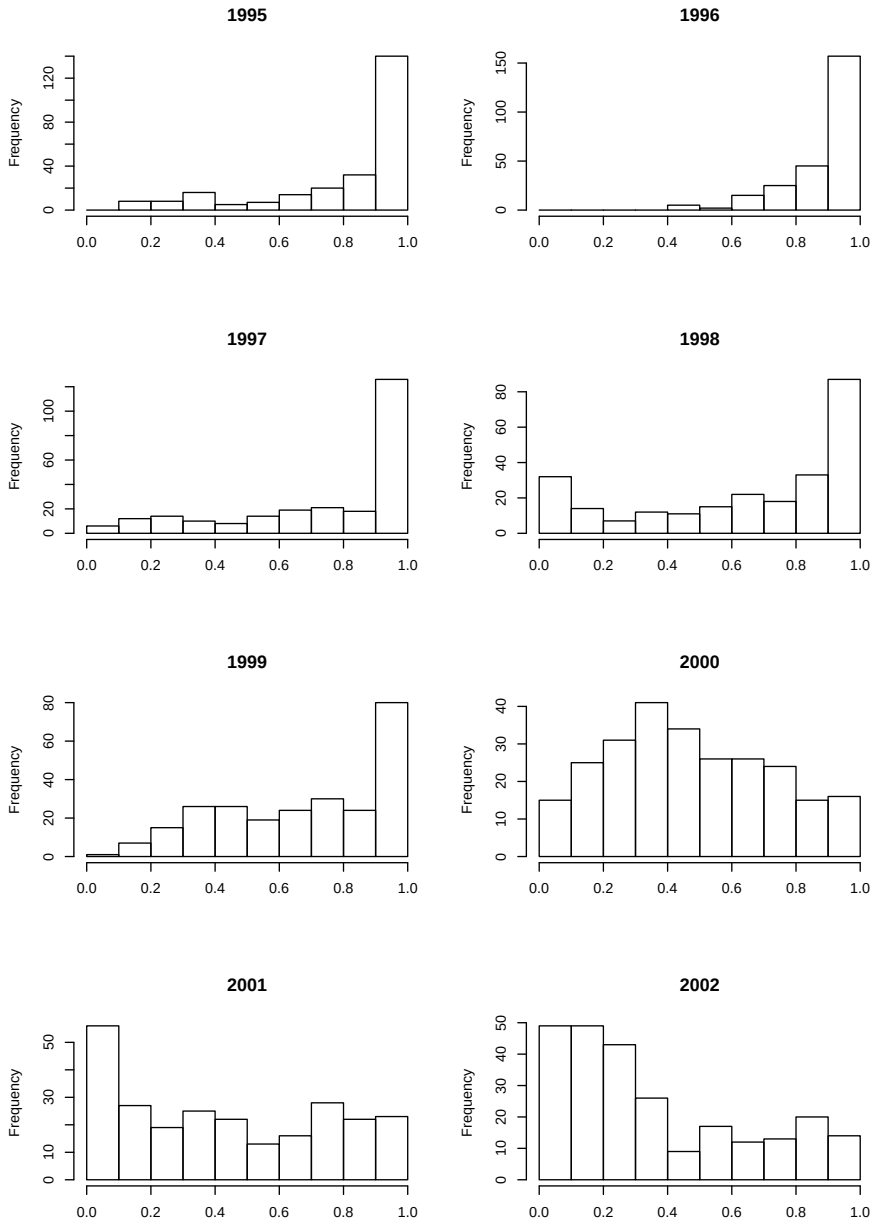


Figure 20: Histograms.

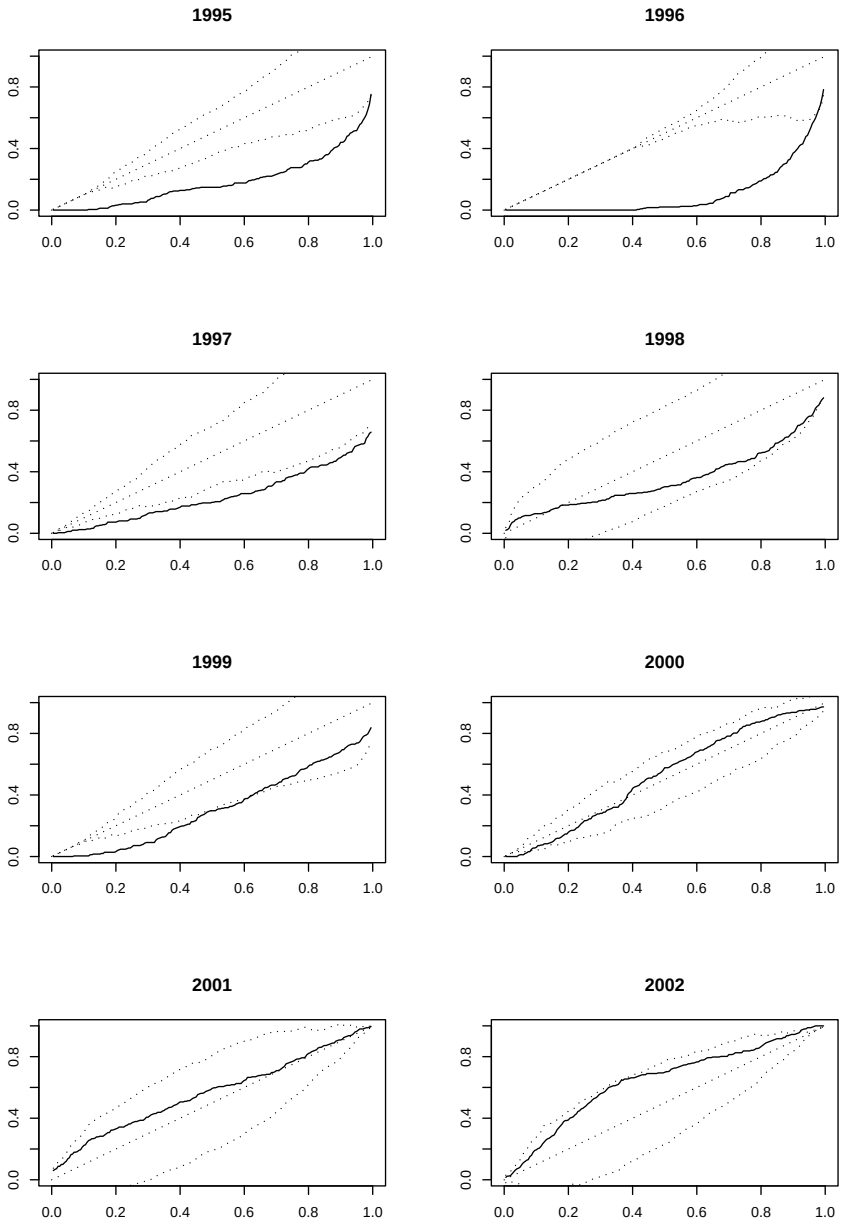


Figure 21: Integral transformation.

covering the future DAX index value. The only exception occured in Septenber 2001. The proposed technique enables us not only to price exotic options byt also to measure the risk and volatility ahead of us.

# 8 Acknowledgements

The authors' work on this project was supported by Deutsche Forschungsgemeinschaft, SFB 649 "Ökonomisches Risiko", by MSM 0021620839, and by MŠMT 1K04018.

# References

Aït-Sahalia, Y. & Duarte, J. (2003). Nonparametric Option Pricing under Shape Restrictions, *Journal of Econometrics*, **116**, 9–47.

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

Aït-Sahalia, Y. & Lo, A.W. (2000). Nonparametric risk management and implied risk aversion, *Journal of Econometrics* **94**, 9–51.

Aït-Sahalia, Y., Wang, Y. & Yared, F. (2000). Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset?, *Journal of Econometrics* **102**, 67–110.

Anděl, J. (1985). *Mathematical Statistics (in Czech)*, SNTL/Alfa, Prague.

Andersen, L.B.G. & Brotherton-Ratcliffe, R. (1997). The equity option volatility smile: An implicit finite-difference approach, *Journal of Computational Finance* **1**(2), 5–37.

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

Buehler, H. (2004). Expensive martingales, *unpublished manuscript*.

Craig, B., Glatzer, E., Keller, J. & Scheicher, M. (2003). The forecasting performance of German stock option densities, Discussion paper 17/2003, Series 1: Studies of the Economics Research Centre, Deutsche Bundesbank.

Dupire, B. (1994). Pricing with a smile, *RISK* **7**(1), 18–20.

Fengler, M.R. (2005). Arbitrage-free smoothing of the implied volatility surface, *unpublished manuscript*.

Fengler, M.R., Härdle, W. & Mammen, E. (2003). Implied volatility string dynamics, Discussion paper 54/2003, Sonderforschungsbereich 373, Humboldt-Universität zu Berlin.

Franke, J., Härdle, W. & Hafner, C. (2004). *Statistics of Financial Markets, An Introduction*, Springer Verlag, Heidelberg.

Huynh, K., Kervella, P. & Zheng, J. (2002). Estimating State-Price Densities with Nonparametric Regression. In *Härdle, Kleinow, Stahl: Applied Quantitative Finance*, Springer Verlag, Heidelberg.

Harison, J. & Pliska, S. (1981). Martingale and stochastic integral in the theory of continuous trading, *Stochastic Processes and their Applications* **11**, 215–260.

Jackwerth, J.C. (1999). Option-implied risk-neutral distributions and implied binomial trees: a literature review, *Journal of Derivatives* **7**, 66–82.

Kahalé, N. (2004). An arbitrage-free interpolation of volatilities, *RISK* **17**(5), 102–106.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, Wiley, New York.

Robertson, T., Wright, F.T. & Dykstra, R.L. (1988). *Order Restricted Statistical Inference*, Wiley, Chichester.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

Yatchew, A. & Härdle, W. (2005). Nonparametric state price density estimation using constrained least squares and the bootstrap, *Journal of Econometrics*, in print.

# SFB 649 Discussion Paper Series

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Nonparametric Risk Management with Generalized Hyperbolic Distributions" by Ying Chen, Wolfgang Härdle and Seok-Oh Jeong, January 2005.
002 "Selecting Comparables for the Valuation of the European Firms" by Ingolf Dittmann and Christian Weiner, February 2005.
003 "Competitive Risk Sharing Contracts with One-sided Commitment" by Dirk Krueger and Harald Uhlig, February 2005.
004 "Value-at-Risk Calculations with Time Varying Copulae" by Enzo Giacomini and Wolfgang Härdle, February 2005.
005 "An Optimal Stopping Problem in a Diffusion-type Model with Delay" by Pavel V. Gapeev and Markus Reiß, February 2005.
006 "Conditional and Dynamic Convex Risk Measures" by Kai Detlefsen and Giacomo Scandolo, February 2005.
007 "Implied Trinomial Trees" by Pavel Čížek and Karel Komorád, February 2005.
008 "Stable Distributions" by Szymon Borak, Wolfgang Härdle and Rafal Weron, February 2005.
009 "Predicting Bankruptcy with Support Vector Machines" by Wolfgang Härdle, Rouslan A. Moro and Dorothea Schäfer, February 2005.
010 "Working with the XQC" by Wolfgang Härdle and Heiko Lehmann, February 2005.
011 "FFT Based Option Pricing" by Szymon Borak, Kai Detlefsen and Wolfgang Härdle, February 2005.
012 "Common Functional Implied Volatility Analysis" by Michal Benko and Wolfgang Härdle, February 2005.
013 "Nonparametric Productivity Analysis" by Wolfgang Härdle and Seok-Oh Jeong, March 2005.
014 "Are Eastern European Countries Catching Up? Time Series Evidence for Czech Republic, Hungary, and Poland" by Ralf Brüggemann and Carsten Trenkler, March 2005.
015 "Robust Estimation of Dimension Reduction Space" by Pavel Čížek and Wolfgang Härdle, March 2005.
016 "Common Functional Component Modelling" by Alois Kneip and Michal Benko, March 2005.
017 "A Two State Model for Noise-induced Resonance in Bistable Systems with Delay" by Markus Fischer and Peter Imkeller, March 2005.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**
**http://sfb649.wiwi.hu-berlin.de**

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

018 "Yxilon – a Modular Open-source Statistical Programming Language" by Sigbert Klinke, Uwe Ziegenhagen and Yuval Guri, March 2005.
019 "Arbitrage-free Smoothing of the Implied Volatility Surface" by Matthias R. Fengler, March 2005.
020 "A Dynamic Semiparametric Factor Model for Implied Volatility String Dynamics" by Matthias R. Fengler, Wolfgang Härdle and Enno Mammen, March 2005.
021 "Dynamics of State Price Densities" by Wolfgang Härdle and Zdeněk Hlávka, March 2005.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**
**http://sfb649.wiwi.hu-berlin.de**

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".