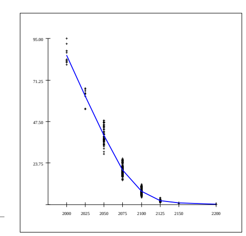
Arbitrage free state price density dynamics

Wolfgang Härdle
C.A.S.E.
Humboldt-Universität zu Berlin

Zdeněk Hlávka
Charles University in Prague
Faculty of Mathematics and Physics

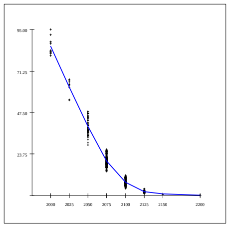
SFB 649, MSM 0021620839, MŠMT 1K04018

S_t	K	r	τ	P_t	Put=0	M	IV_t	time	volume	date
2079.23083	1925.00000	0.05153	0.05000	1.50000	0.00000	0.91667	0.20430	36369.67000	100.00000	19950102
2079.23083	1925.00000	0.05153	0.05000	1.50000	0.00000	0.91667	0.20430	38483.60000	100.00000	19950102
2079.23083	1925.00000	0.05153	0.05000	1.50000	0.00000	0.91667	0.20430	38794.47000	100.00000	19950102
2079.72548	1925.00000	0.05153	0.05000	1.40000	0.00000	0.91645	0.20217	40751.65000	100.00000	19950102
2079.23083	1925.00000	0.05153	0.05000	1.30000	0.00000	0.91667	0.19896	44073.10000	100.00000	19950102
2078.73618	1925.00000	0.05153	0.05000	1.80000	0.00000	0.91688	0.21110	45841.68000	20.00000	19950102
2083.68272	1925.00000	0.05153	0.05000	1.40000	0.00000	0.91471	0.20605	54200.51000	30.00000	19950102
2083.18806	1925.00000	0.05153	0.05000	1.30000	0.00000	0.91492	0.20280	54609.59000	90.00000	19950102
2081.20945	1925.00000	0.05153	0.05000	1.50000	0.00000	0.91579	0.20626	57224.43000	1.00000	19950102
2081.20945	1925.00000	0.05153	0.05000	1.20000	0.00000	0.91579	0.19801	57224.88000	1.00000	19950102
2079.23083	1925.00000	0.05153	0.05000	1.20000	0.00000	0.91667	0.19611	57555.44000	49.00000	19950102
2079.23083	1950.00000	0.05153	0.05000	2.60000	0.00000	0.92857	0.19862	45867.00000	40.00000	19950102
2074.28429	1975.00000	0.05153	0.05000	4.80000	0.00000	0.94272	0.19150	36827.09000	2.00000	19950102
2076.75756	1975.00000	0.05153	0.05000	4.30000	0.00000	0.94160	0.18869	44673.96000	30.00000	19950102
2075.27360	1975.00000	0.05153	0.05000	4.30000	0.00000	0.94227	0.18685	46724.45000	50.00000	19950102
2083.68272	1975.00000	0.05153	0.05000	4.30000	0.00000	0.93847	0.19716	54543.59000	10.00000	19950102
2078.24152	2000.00000	0.05153	0.05000	7.50000	0.00000	0.95283	0.18810	36450.55000	100.00000	19950102
2079.23083	2000.00000	0.05153	0.05000	7.20000	0.00000	0.95238	0.18682	38786.55000	100.00000	19950102
2082.19876	2000.00000	0.05153	0.05000	7.00000	0.00000	0.95102	0.18913	42075.81000	10.00000	19950102
2082.19876	2000.00000	0.05153	0.05000	6.30000	0.00000	0.95102	0.18258	42077.89000	50.00000	19950102
2077.74687	2000.00000	0.05153	0.05000	7.00000	0.00000	0.95306	0.18294	44219.41000	40.00000	19950102
2076.75756	2000.00000	0.05153	0.05000	7.20000	0.00000	0.95352	0.18334	44601.38000	100.00000	19950102
2074.77894	2000.00000	0.05153	0.05000	7.50000	0.00000	0.95443	0.18318	44856.34000	100.00000	19950102
2075.27360	2000.00000	0.05153	0.05000	7.50000	0.00000	0.95420	0.18389	46747.75000	3.00000	19950102
2084.17737	2000.00000	0.05153	0.05000	6.00000	0.00000	0.95012	0.18235	54276.01000	75.00000	19950102
2082.19876	2000.00000	0.05153	0.05000	6.40000	0.00000	0.95102	0.18353	54887.14000	100.00000	19950102
2089.12392	2025.00000	0.05153	0.05000	9.50000	0.00000	0.95972	0.18170	34442.91000	10.00000	19950102



Data:

1. S_t underlying (corrected)
2. K strike
3. r interest rate
4. τ time to maturity
5. P_t option price
6. I (Call=yes)
7. M moneyness
8. σ_{IV} implied volatility
9. time (in sec. after midnight)
10. volume
11. date (yyyymmdd)



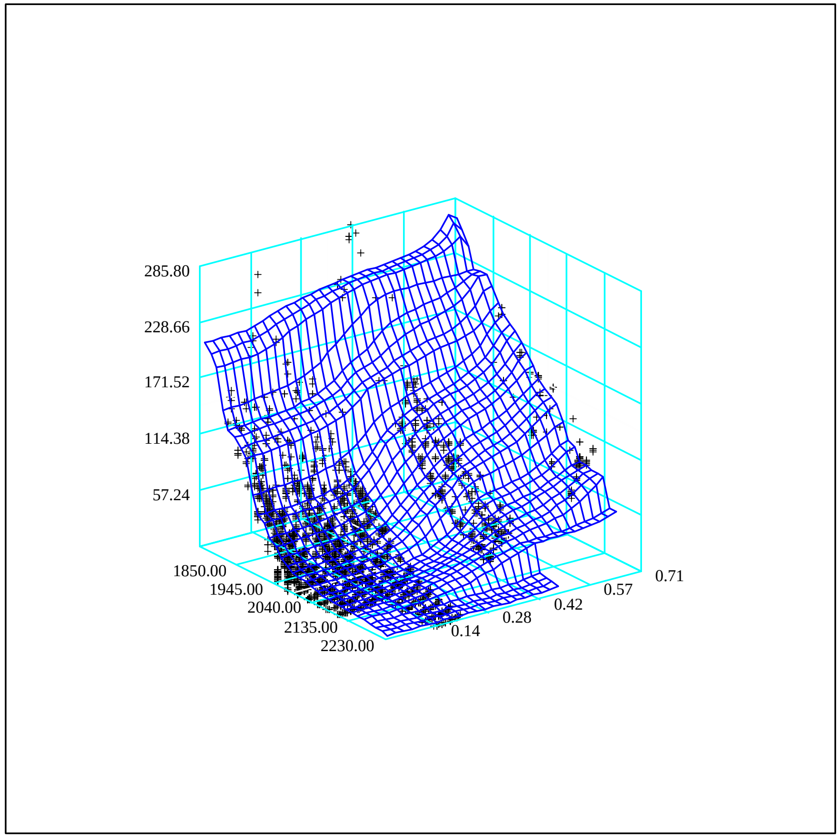
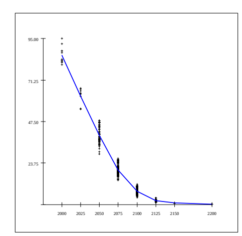


Figure 1: Option price as a function of time to maturity and moneyness.



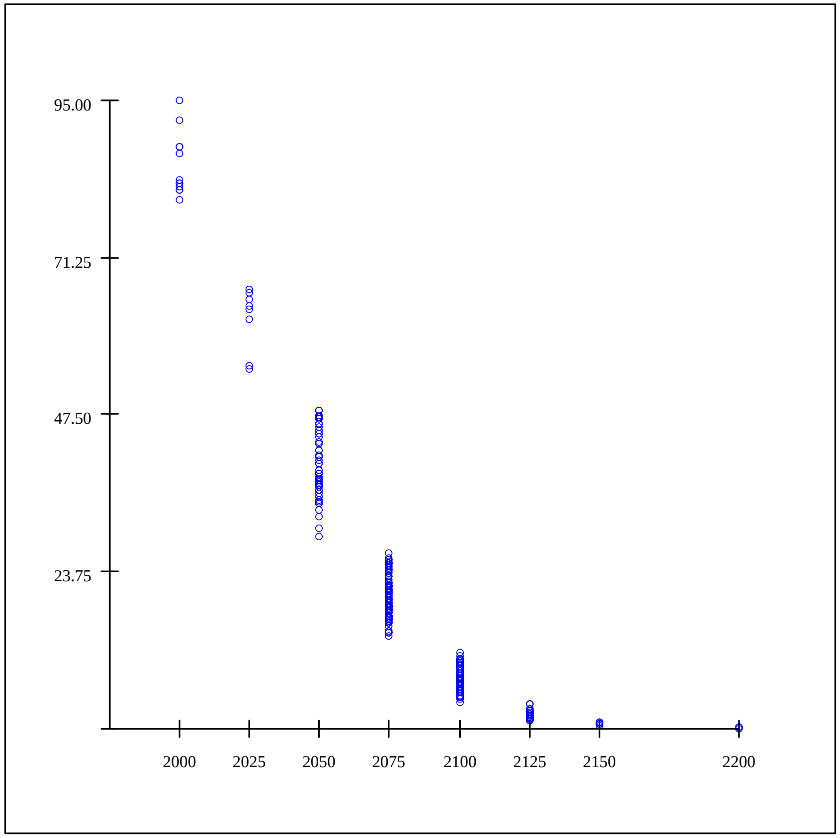
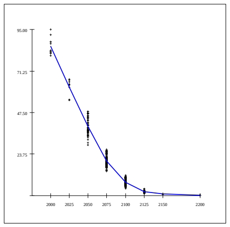


Figure 2: Option price as a function of moneyness for fixed time to maturity.



The fair price of an option with payoff $(S_T - K)_+ = \max(S_T - K, 0)$ can be written as

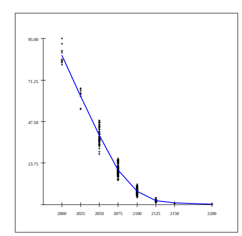
$$C_t(K,T) = \exp\{-r_{t,T-t}(T-t)\} \int_0^{+\infty} (S_T - K)_+ f(S_T) dS_T. \qquad (1)$$

The state price density $f(S_T)$ of the prices at time T can be expressed as:

$$f_{\text{SPD}}(K) = \exp\{r_{t,T-t}(T-t)\} \frac{\partial^2 C_t(K,T)}{\partial K^2}. \qquad (2)$$

Equation (2) is often used to estimate the SPD.

See, e.g., Aït-Sahalia and Duarte (2001) or Yatchew and Härdle (2005).



Assumptions and constraints

The conditional expected value of the option price $C(K)$ is denoted as

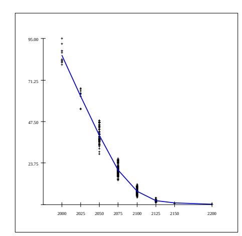
$$\mu(k_j) = E\left[C(K)|K = k_j\right], \quad j = i, \ldots, p.$$

The i-th observation (at $K = k_j$) follows the model

$$C_i(k_j) = \mu(k_j) + \varepsilon_i, \quad \varepsilon_i \sim \text{iid } N(0, \sigma^2).$$

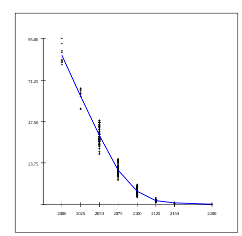
Assume that $\mu(.)$ satisfies the following constraints:

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).



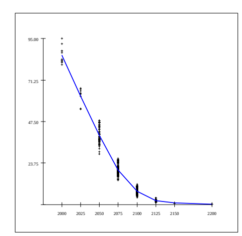
Challenges and Questions

1. Estimation is not arbitrage free
2. Can we construcs confidence intervals (C.I.)?
3. How does the SPD behave over time?
4. Is the future realization of the underlying inside a C.I.?



Outline

1. Introduction ✓
2. Linear regression
3. Constraints
4. Asymptotics
5. Put and call option prices
6. Time dependency and covariance structure
7. Dynamics of the SPD



Linear Regression

LEMMA: Suppose that f satisfies the above constraints. Then we have for its first derivative, f', that $\lim_{+\infty} f'(x) = 0$ and $\lim_{-\infty} f'(x) = -1$.

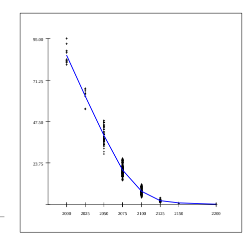
Existence and uniqueness

We restate conditions 1–4 for discrete functions, defined only on small number of points, in terms of their function values, $f(x_i)$, and estimated first derivatives, $f^{(1)}_{x_i,x_j} = \frac{f(x_i)-f(x_j)}{x_i-x_j}$,

5. $f(x_i) > 0$, $i = 1, \ldots, p$,

6. $x_i < x_j$ implies that $f_{x_i} \geq f_{x_j}$,

7. $x_i < x_j < x_k$ implies that $-1 \leq f^{(1)}_{x_i,x_j} \leq f^{(1)}_{x_j,x_k} \leq 0$.



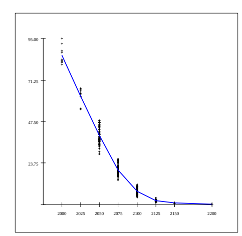
Think of the collection, $\mathcal{F}$, of functions satisfying Constraints 5–7 as a subset of a p-dimensional Euclidean space.

The constrained regression, $\hat{g}$, is in this setting the closest point of $\mathcal{F}$ to the observed g with distances measured by the usual Euclidean distance

$$d(f,g) = (f-g)^\top (f-g) = \sum_{i=1}^{n} (f_i - g_i)^2.$$

1. $\mathcal{F}$ is closed,
2. $\mathcal{F}$ is convex.

A regression, $\hat{g}$, satisfying Constraints 5–7 exists and it is unique.



Assume that $\hat{g}$ is the regression of $g(x_i)$ on $x_1 \leq \cdots \leq x_n$ under Constraints 5–7.

If a and b are constants such that $a \leq g(x_i) \leq b$, $\forall i$, then $a - (x_n - x_1) \leq \hat{g}(x_i) \leq b + (x_n - x_1)$.

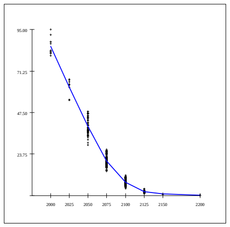
Suppose $\mathcal{F}$ is any convex set of functions on $\mathcal{X}$ and g is a given function on $\mathcal{X}$. If

$$\hat{g} = \arg\min_{f \in \mathcal{F}} d(g, f)$$

then for every $f \in \mathcal{F}$,

$$\sum_{i=1}^{n} \{g(x_i) - \hat{g}(x_i)\}^\top \{\hat{g}(x_i) - f(x_i)\} \geq 0. \qquad (3)$$

There exists at most one function $\hat{g}$ satisfying (3).
(Robertson, Wright and Dykstra 1988, Theorem 1.3.1)



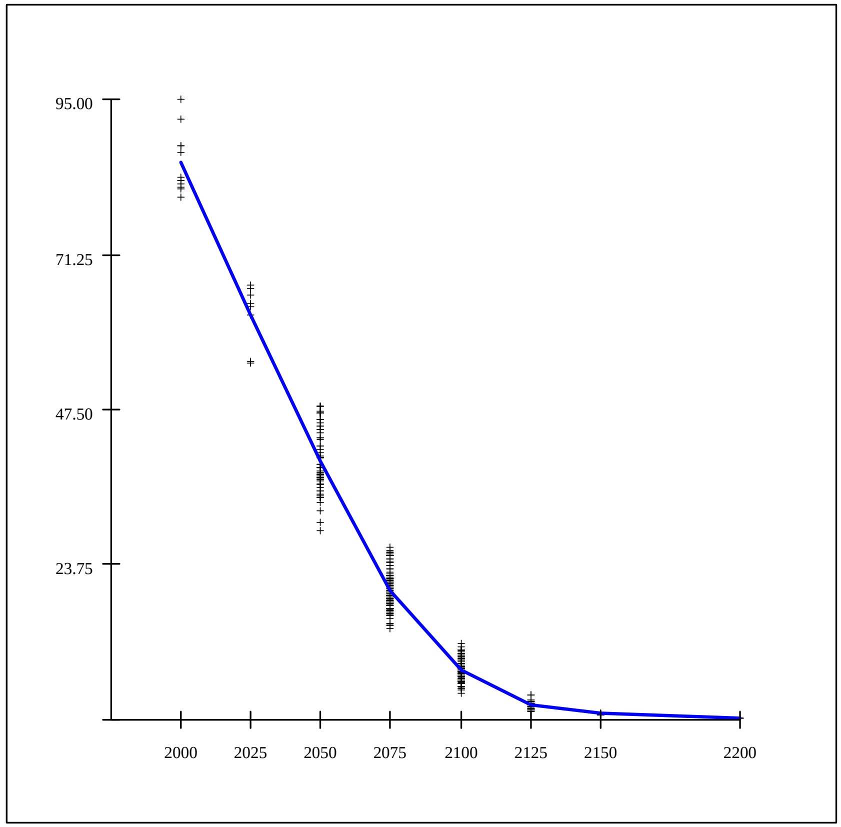


Figure 3: Option price as a function of moneyness for fixed time to maturity with linear regression model, 16th January 1995.

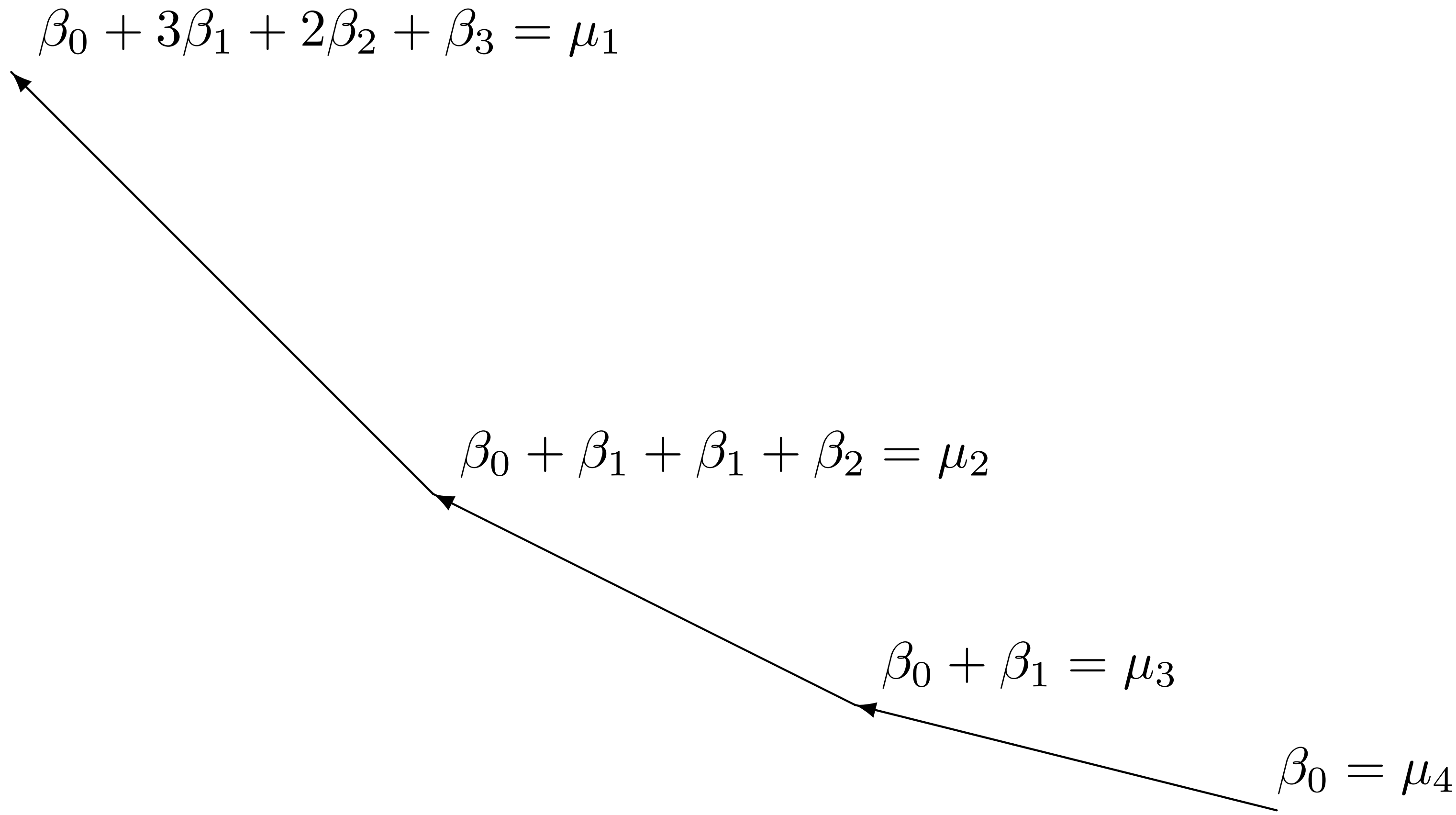
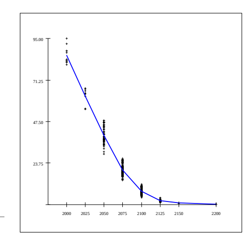


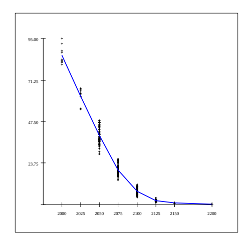
Figure 4: The four points example: illustration of the dummy variables for call options.



$$
\begin{aligned}
\mu_4 &= \beta_0, \\
\mu_3 &= \beta_0 + \beta_1, \\
\mu_2 &= \beta_0 + 2\beta_1 + \beta_2, \\
\mu_1 &= \beta_0 + 3\beta_1 + 2\beta_2 + \beta_3.
\end{aligned}
$$

The model may be written in matrix form:

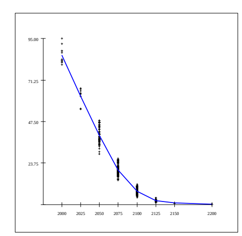
$$
\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.
$$



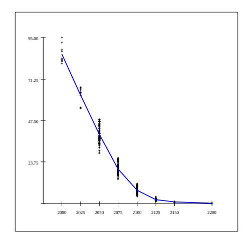
For non-equidistant data, define:

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

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



[1,]	1	275	250	225	200	175	150	125	100	75	50	25
[2,]	1	275	250	225	200	175	150	125	100	75	50	25
[3,]	1	225	200	175	150	125	100	75	50	25	0	0
[4,]	1	150	125	100	75	50	25	0	0	0	0	0
[5,]	1	150	125	100	75	50	25	0	0	0	0	0
[6,]	1	150	125	100	75	50	25	0	0	0	0	0
[7,]	1	150	125	100	75	50	25	0	0	0	0	0
[8,]	1	150	125	100	75	50	25	0	0	0	0	0
[9,]	1	150	125	100	75	50	25	0	0	0	0	0
[10,]	1	150	125	100	75	50	25	0	0	0	0	0
[11,]	1	150	125	100	75	50	25	0	0	0	0	0
[12,]	1	150	125	100	75	50	25	0	0	0	0	0
[13,]	1	150	125	100	75	50	25	0	0	0	0	0
[14,]	1	150	125	100	75	50	25	0	0	0	0	0
[15,]	1	150	125	100	75	50	25	0	0	0	0	0
[16,]	1	150	125	100	75	50	25	0	0	0	0	0
[17,]	1	150	125	100	75	50	25	0	0	0	0	0
[18,]	1	150	125	100	75	50	25	0	0	0	0	0
[19,]	1	150	125	100	75	50	25	0	0	0	0	0
[20,]	1	125	100	75	50	25	0	0	0	0	0	0
[21,]	1	125	100	75	50	25	0	0	0	0	0	0
[22,]	1	125	100	75	50	25	0	0	0	0	0	0
[23,]	1	125	100	75	50	25	0	0	0	0	0	0
[24,]	1	125	100	75	50	25	0	0	0	0	0	0
[25,]	1	125	100	75	50	25	0	0	0	0	0	0
[26,]	1	125	100	75	50	25	0	0	0	0	0	0
[27,]	1	125	100	75	50	25	0	0	0	0	0	0
[28,]	1	125	100	75	50	25	0	0	0	0	0	0
[29,]	1	125	100	75	50	25	0	0	0	0	0	0



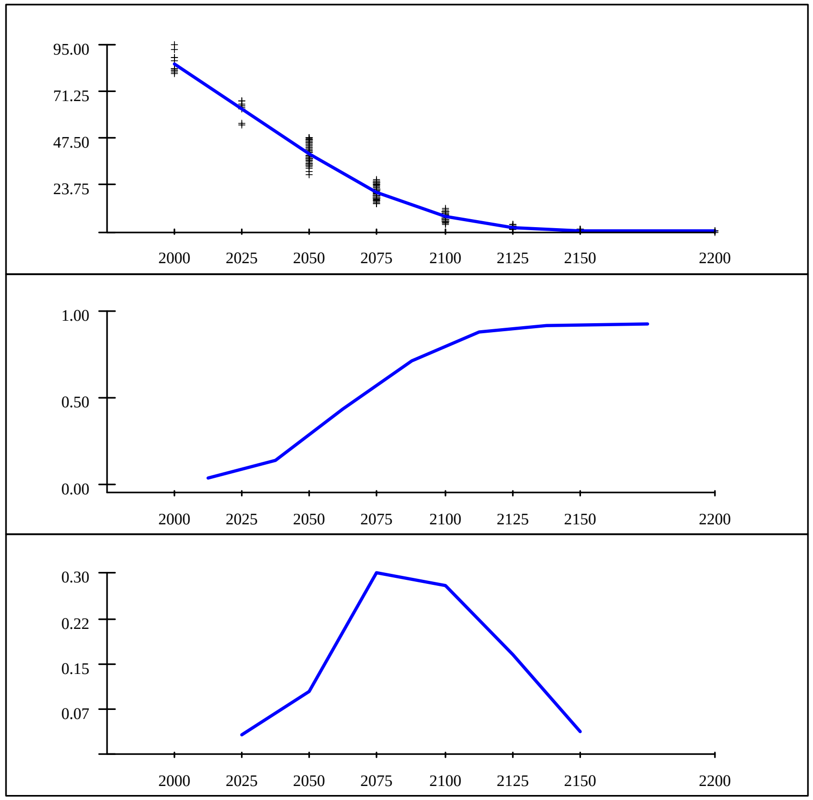
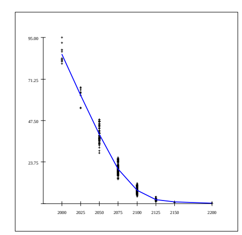


Figure 5: Option price as a function of moneyness for fixed time to maturity with linear regression model and estimates of first and second derivatives, 16th January 1995.



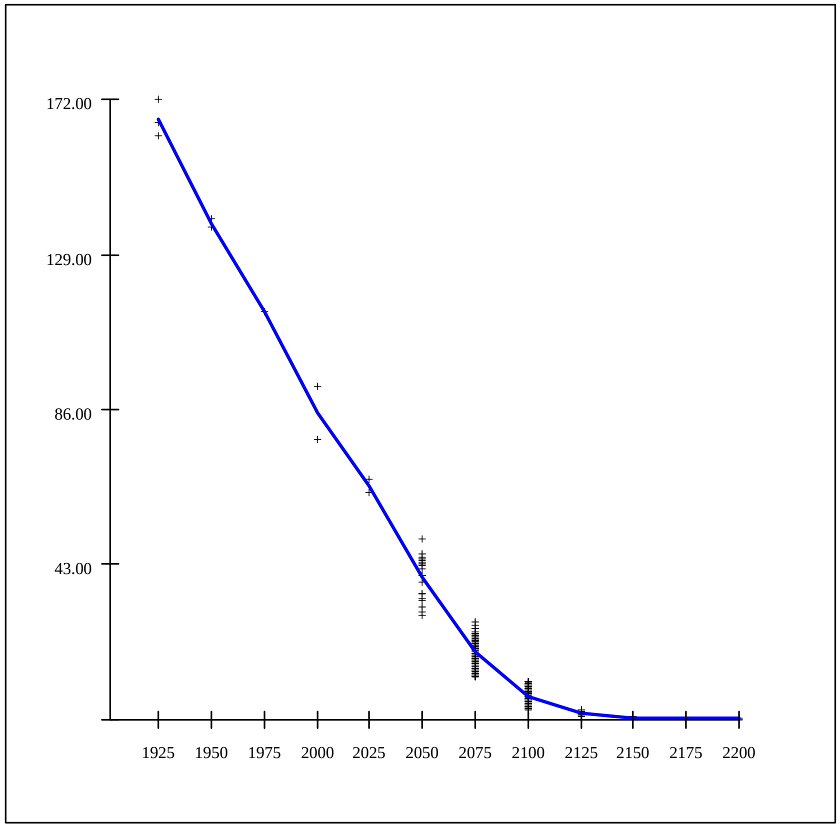
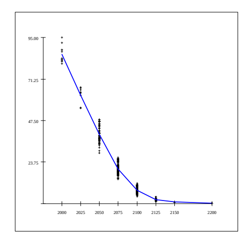


Figure 6: Option price as a function of moneyness for fixed time to maturity with linear regression model, 17th January 1995.



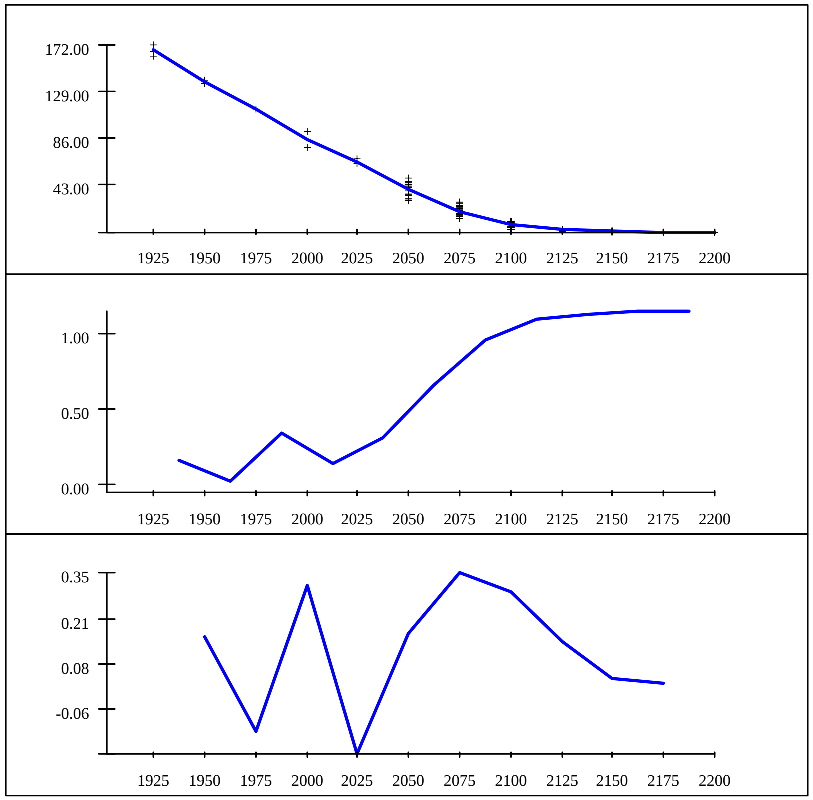
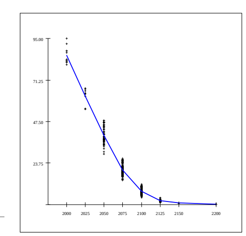


Figure 7: Option price as a function of moneyness for fixed time to maturity with linear regression model and estimates of first and second derivatives, 17th January 1995.

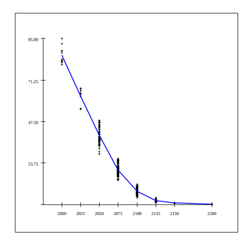


The vector of conditional means μ can then be expressed in terms of the parameters β as follows

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \mu = \Delta\beta = \Delta \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix}. \tag{4}$$

The constraints on the conditional means μ_j can now be expressed as conditions on the parameters of the model.

It suffices to request that $\beta_i > 0$, $i = 0, \dots, p-1$ and that $\sum_{j=1}^{p-1} \beta_j \leq 1$.

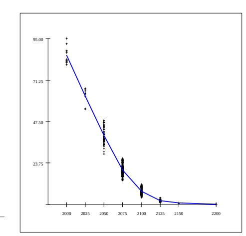


The unconstrained linear regression model for the observed option prices $\mathcal{C}$ and for the observed strike prices K can now be written as

$$\mathcal{C} = \mathcal{X}_{\Delta}\beta + \varepsilon,$$

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

Running linear regression in any statistical software and using the standard errors of the parameter estimates, we obtain confidence intervals for the SPD.



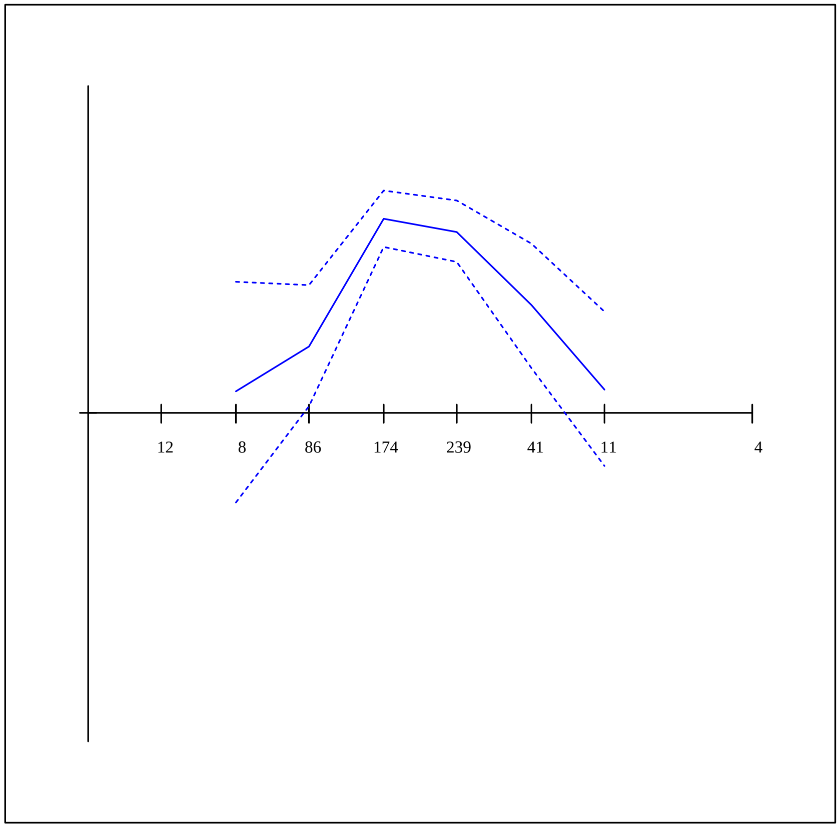
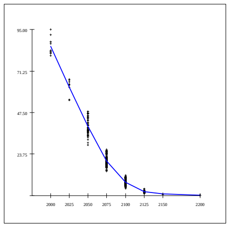


Figure 8: Estimated second derivative with confidence intervals, 16th January 1995.



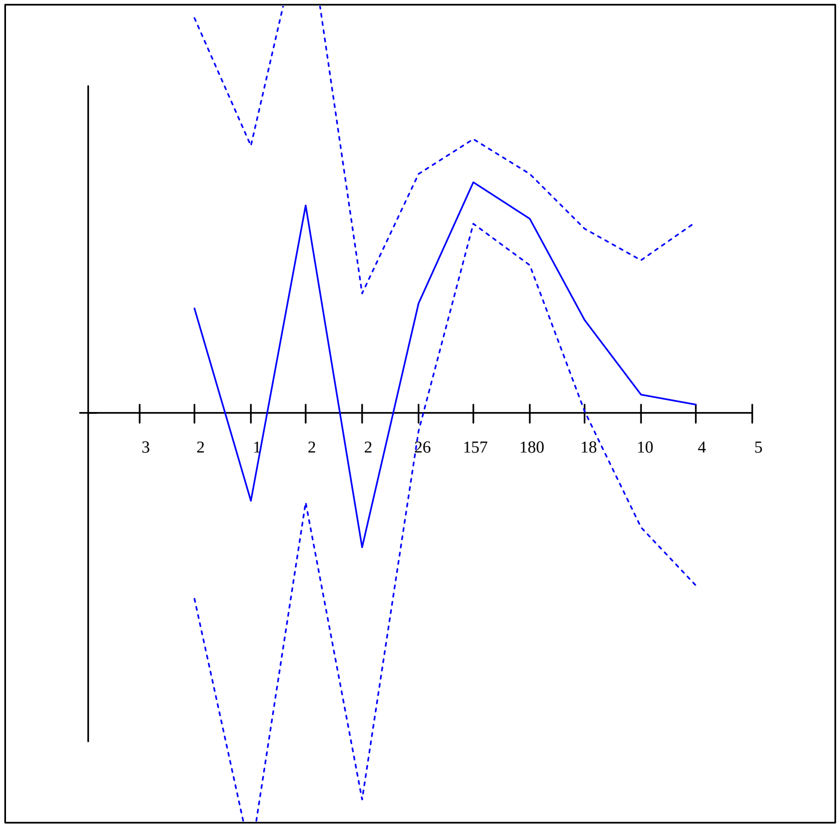
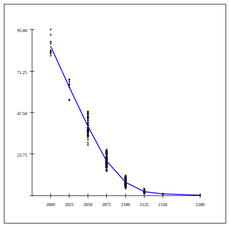


Figure 9: Estimated second derivative with confidence intervals, 17th January 1995.



Constraints

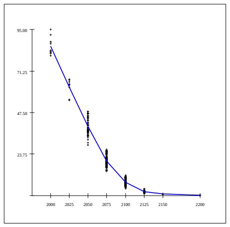
Calculation of the constrained estimate

In order to impose the constraints, we reparametrize the model:

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

This parametrization leads to a model which is not identified.

$(\sum_1^p \exp(\theta_j) = 1 \Leftrightarrow \sum_1^{p-1} \exp(\theta_j) < 1)$



Starting values for the algorithm

For the numerical algorithm, it is useful to know how to calculate θs from given βs.

Given $\beta = (\beta_1, \dots, \beta_p)^\top$, where $\beta_p = 1 - \sum_{i=1}^{p-1} \beta_i$, the parameters $\theta = (\theta_1, \dots, \theta_p)^\top$ satisfy the system of equations

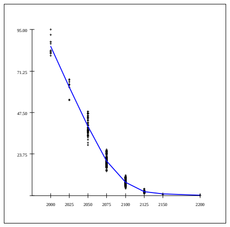
$$\left(\beta \mathbf{1}_p^\top - \mathbf{I}_p\right) \exp \theta^\top = 0 \qquad (5)$$

Furthermore, $\operatorname{rank}\left(\beta \mathbf{1}_p^\top - \mathbf{I}_p\right) = p - 1$.

Thus, the system of equations (5) has infinitely many equations which can be expressed as

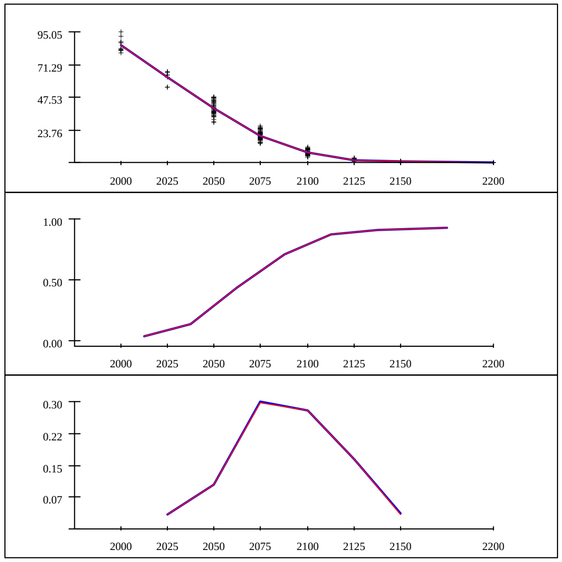
$$\exp(\theta) = \left[\left(\beta \mathbf{1}_p^\top - \mathbf{I}_p\right)^{-} \left(\beta \mathbf{1}_p^\top - \mathbf{I}_p\right) - \mathbf{I}_p\right] z,$$

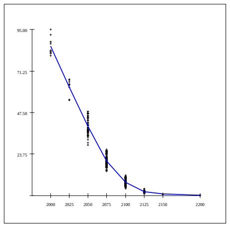
where z is an arbitrary vector in $\mathbb{R}^p$.



The algorithm

1. unconstrained estimates
2. PAV on first differences
3. transform
4. numerical minimization





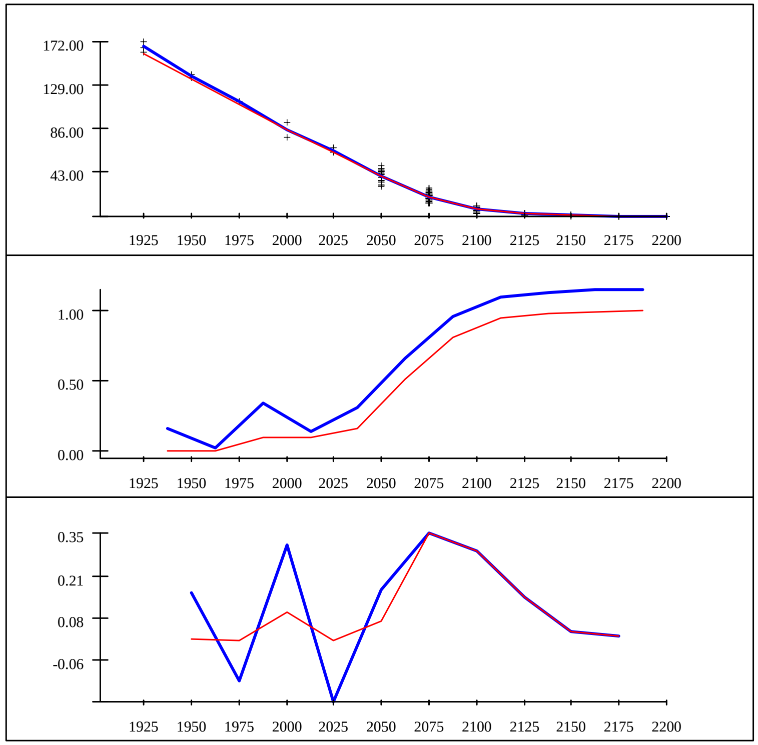
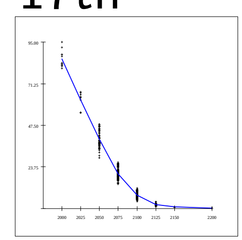
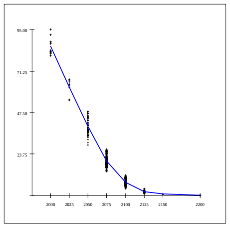


Figure 10: Option price as a function of moneyness for fixed time to maturity with linear regression model (blue) and the constrained model (red) and corresponding estimates of first and second derivatives, 17th January 1995.



Confidence Intervals

1. based on β
   - can be negative
   - leads to results which do not make sense
2. based on θ
   - difficult to calculate
   - numerical problems
   - the parameters aren't unique

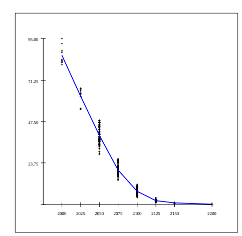


Confidence intervals

In order to calculate the confidence intervals, we propose the following parametrization of the model:

$$
\begin{aligned}
\beta_0(\theta) &= \exp(\theta_0), \\
\beta_1(\theta) &= \exp(\theta_1), \\
\beta_2(\theta) &= \exp(\theta_2), \\
&\vdots \\
\beta_{p-1}(\theta) &= \exp(\theta_{p-1}),
\end{aligned}
$$

under the constraint that $\sum_{i=1}^{p-1} \beta_i(\theta) < 1$. This is equivalent to the non-identified model defined above (by setting $\sum_{i=1}^{p} \exp\theta_i = 1$).



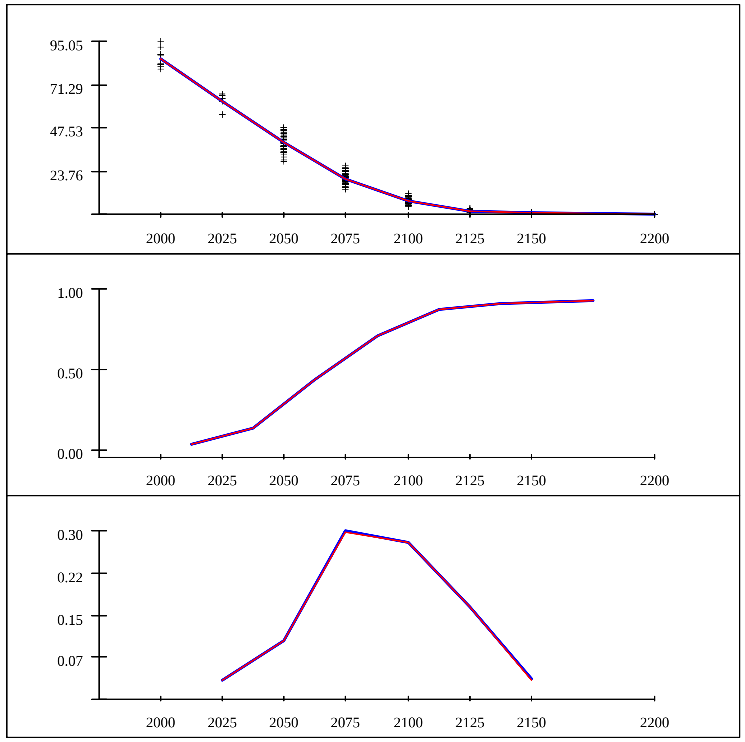
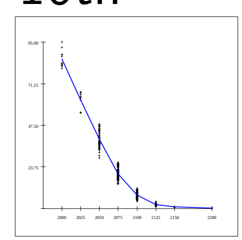


Figure 11: Option price as a function of moneyness for fixed time to maturity with linear regression model (blue) and the constrained model (red) and corresponding estimates of first and second derivatives, 16th January 1995.



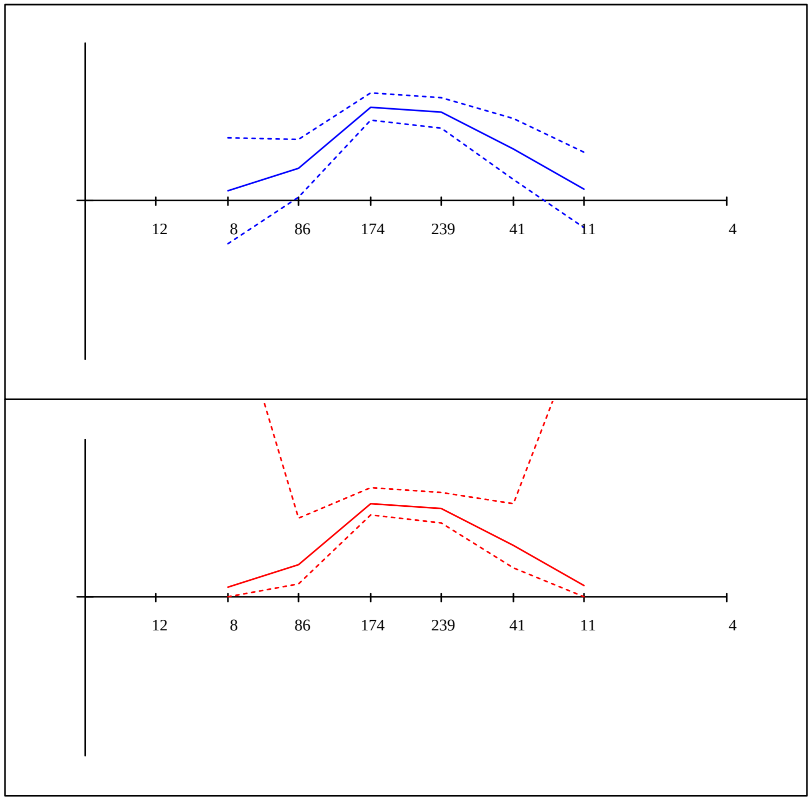
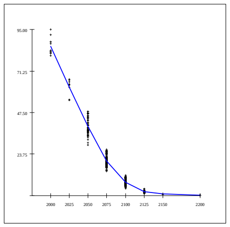


Figure 12: Comparison of the confidence intervals for the estimated SPD for the linear regression model (blue) and the constrained model (red), 16th January 1995.



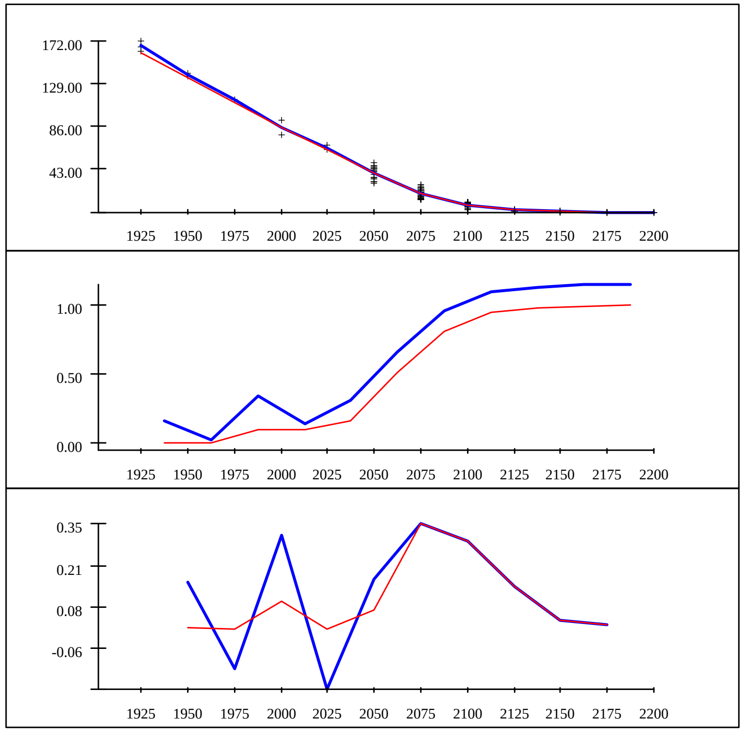
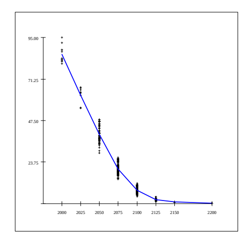


Figure 13: Option price as a function of moneyness for fixed time to maturity with linear regression model (blue) and the constrained model (red) and corresponding estimates of first and second derivatives, 17th January 1995.



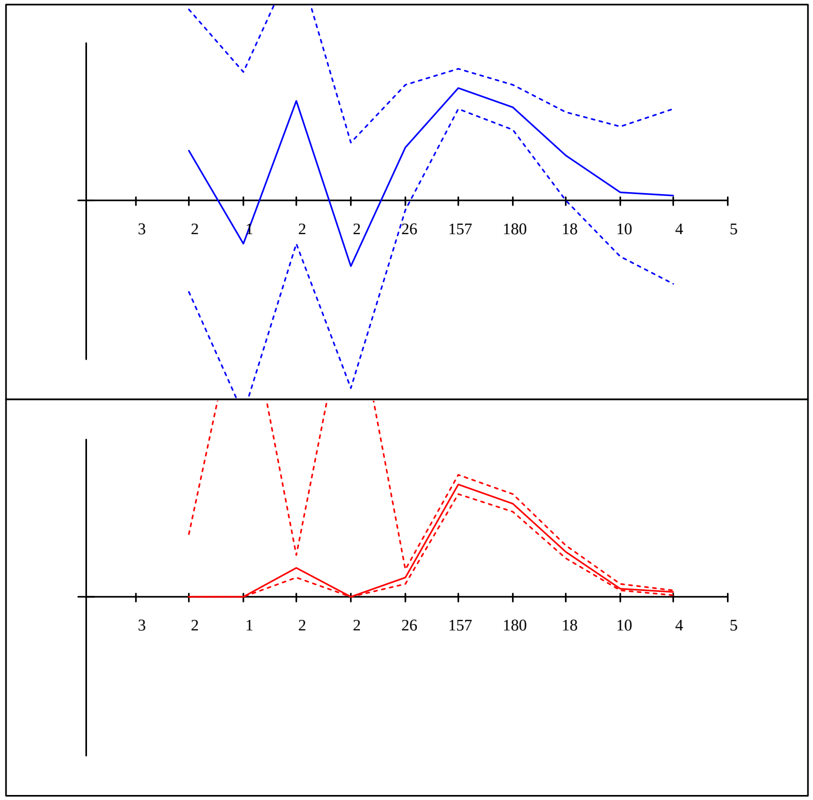
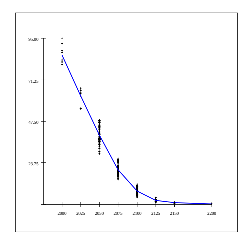
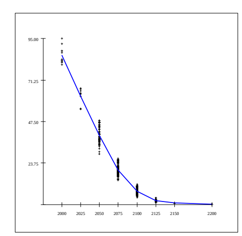


Figure 14: Comparison of the confidence intervals for the estimated SPD for the linear regression model (blue) and the constrained model (red), 17th January 1995.



	SPD estimate	C.I. (lower)	C.I. (upper)
1950	0.0048	0.0001	0.2000
1975	1.41e-16	0.0000	+INF
2000	0.0920	0.0624	0.1358
2025	4.34e-17	0.0000	+INF
2050	0.0636	0.0447	0.0904
2075	0.3543	0.3256	0.3855
2100	0.2967	0.2716	0.3241
2125	0.1428	0.1232	0.1654
2150	0.0294	0.0207	0.0419
2175	0.0152	0.0096	0.0240

Confidence intervals for the nonlinear regression model.



Put and Call Option Prices

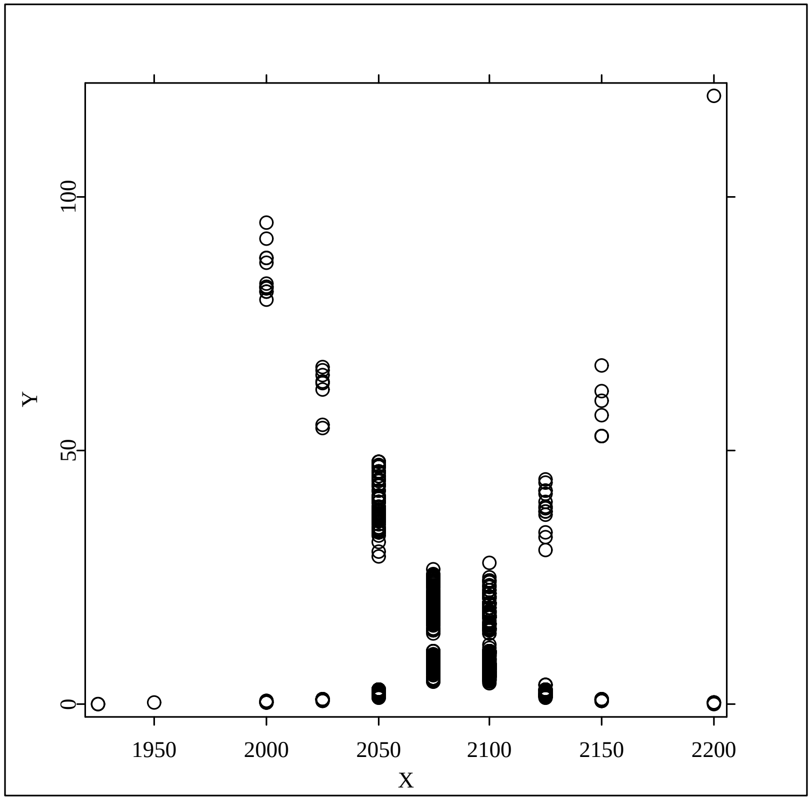
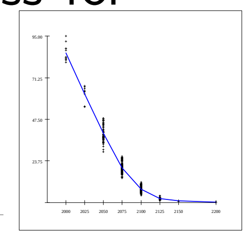


Figure 15: Option price (call and put) as a function of moneyness for fixed time to maturity



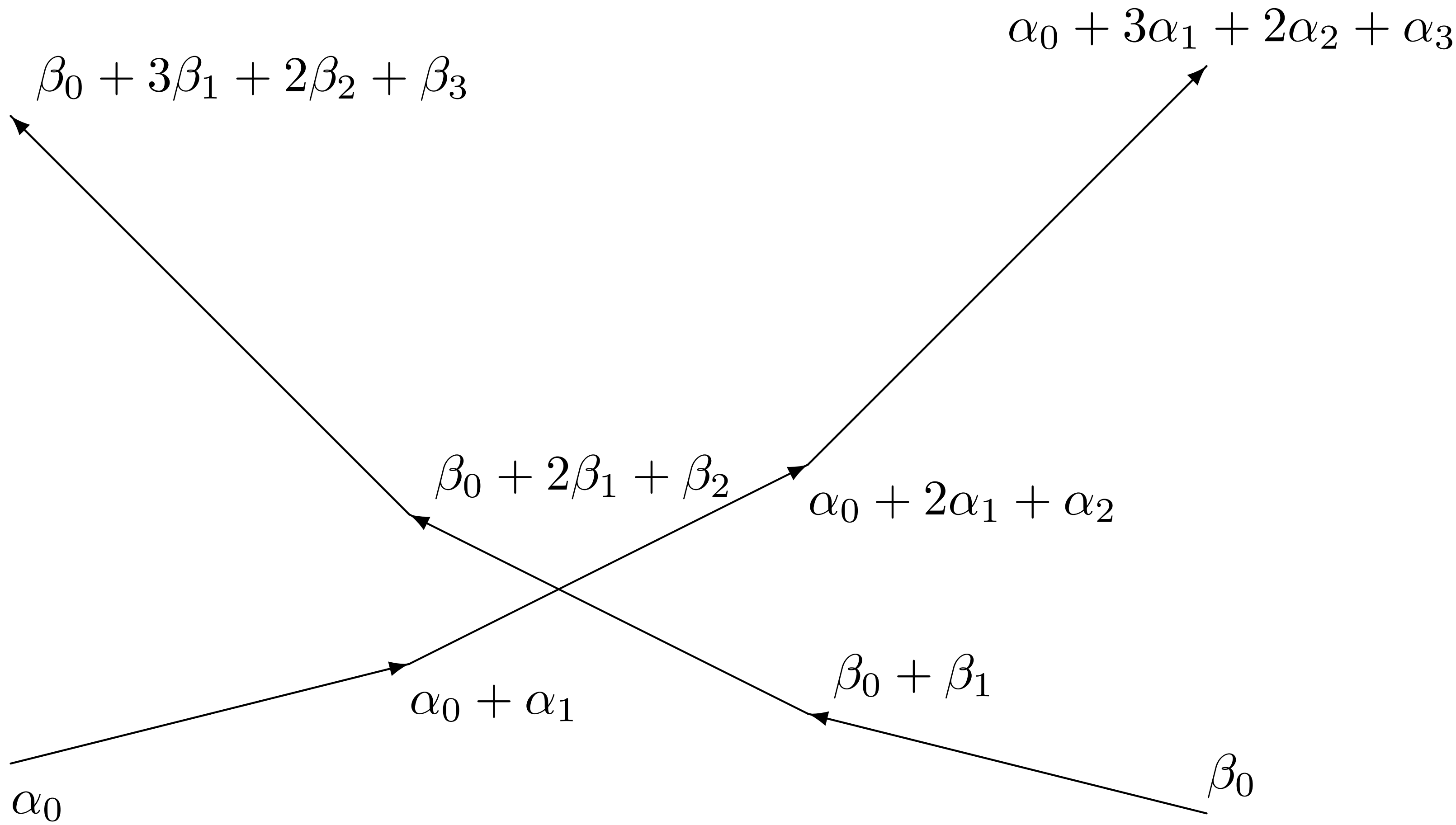
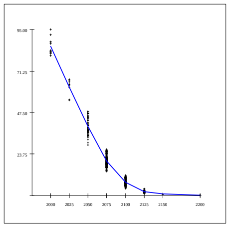


Figure 16: Illustration of the dummy variables for both call (β) and put (α) options.

0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1	275	250	225	200	175	150	125	100	75	50	25	0	0	0	0	0	0	0	0	0	0	0	0
1	275	250	225	200	175	150	125	100	75	50	25	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	25	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	50	25	0	0	0	0	0	0	0	0	0
1	225	200	175	150	125	100	75	50	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	75	50	25	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	75	50	25	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	100	75	50	25	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	100	75	50	25	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	100	75	50	25	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	100	75	50	25	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	125	100	75	50	25	0	0	0	0	0	0
1	150	125	100	75	50	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	150	125	100	75	50	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

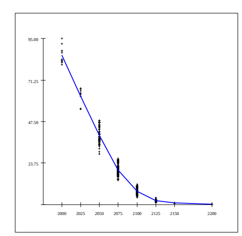


Put and call option prices

The coefficients α_i and β_i have to satisfy

$$\alpha_1 = 1 - \sum_{i=1}^{p-1} \beta_i$$

$$\alpha_i = \beta_{p-i+1}, \quad \text{for } i = 2, \ldots, p-1$$



Time dependency and covariance structure

Up to now, we worked with the model

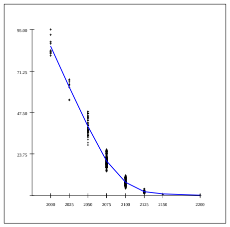
$$C_t(k_j) = \Delta_j \widetilde{\beta} + \varepsilon_t$$

or

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

where t is the time, $\widetilde{\beta}$ denotes the column vector of the unknown parameters, and Δ_j denotes the j-th row of the design matrix, i.e.,

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



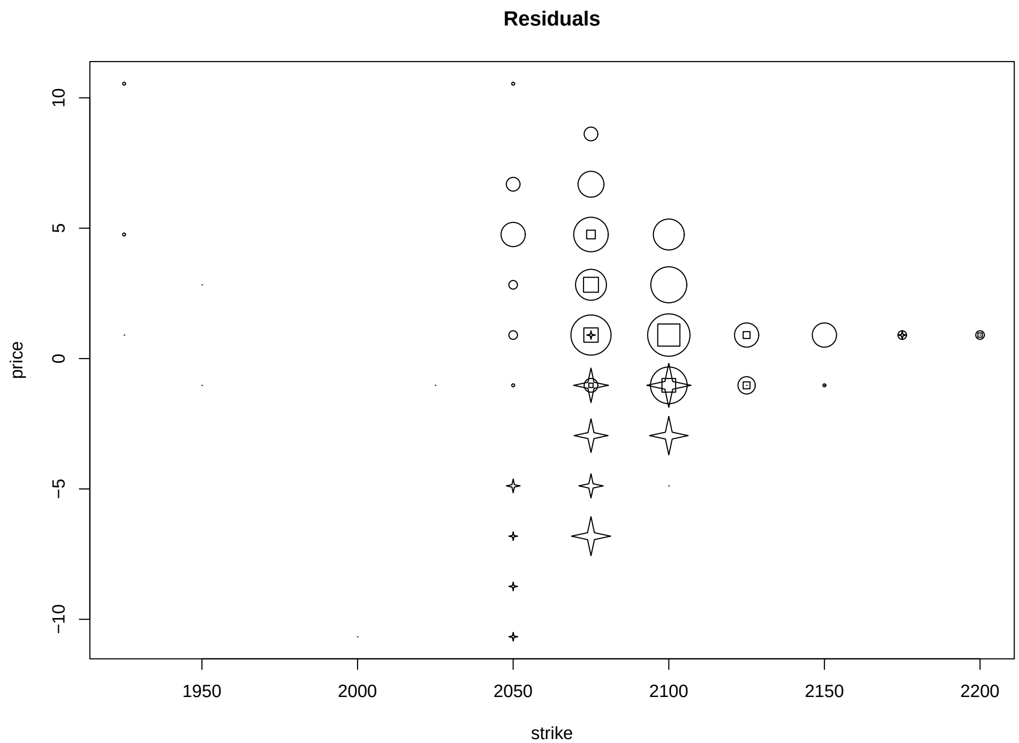
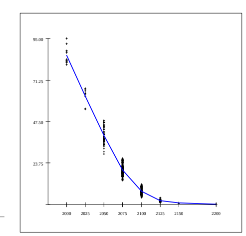


Figure 17: The time dependency and the heteroscedasticity of the residuals during one day (17th January 1995). The circle, square and star denote the trades carried out in the morning, midday and in the afternoon, respectively. Size of the symbols denotes number of residuals.

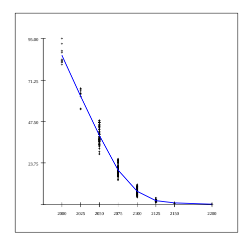


Let δ_t denote the time between the t-th and $(t-1)$-st observation. The model is

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

and it leads to the covariance matrix with elements

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



In this way, we obtain a joint estimation strategy for both the call and put option prices:

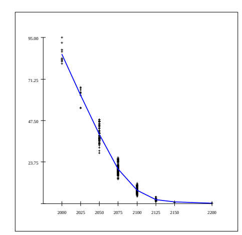
$$
\begin{aligned}
C_t(k_j) &= \Delta_j \widetilde{\beta}_t, \\
P_t(k_j) &= \Delta_j^P \widetilde{\alpha}_t, \\
\begin{pmatrix} \widetilde{\beta}_t \\ \widetilde{\alpha}_t \end{pmatrix} &= \begin{pmatrix} \widetilde{\beta}_{t-1} \\ \widetilde{\alpha}_{t-1} \end{pmatrix} + \varepsilon_t,
\end{aligned}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

which directly leads to the covariances

$$
\mathrm{Cov}\{P_{t-u}(k_j), P_{t-v}(k_i)\} = \sigma^2 \Delta_j^P (\Delta_i^P)^\top \sum_{l=1}^{\min(u,v)} \delta_{t+1-l}^2
$$

and

$$
\mathrm{Cov}\{C_{t-u}(k_j), P_{t-v}(k_i)\} = \sigma^2 \sum_{l=1}^{\min(u,v)} \delta_{t+1-l}^2 \sum_{k=2}^{p-1} \Delta_{p+1-k}^{j} \Delta_i^{p+1-k}.
$$



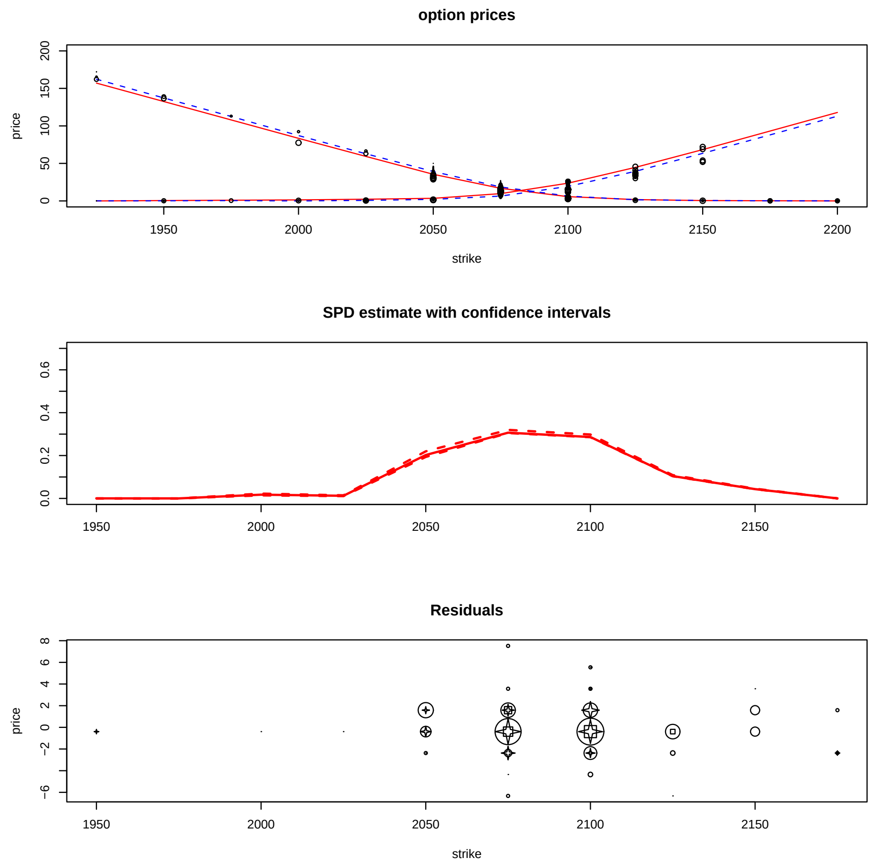
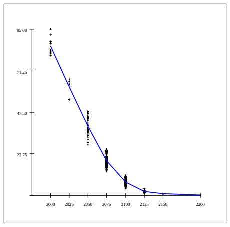


Figure 18: Option price as a function of moneyness for fixed time to maturity with the fitted lines (upper panel); the estimated SPD (middle panel), and the residuals (lower panel), 17th January 1995.



Dynamics of SPD

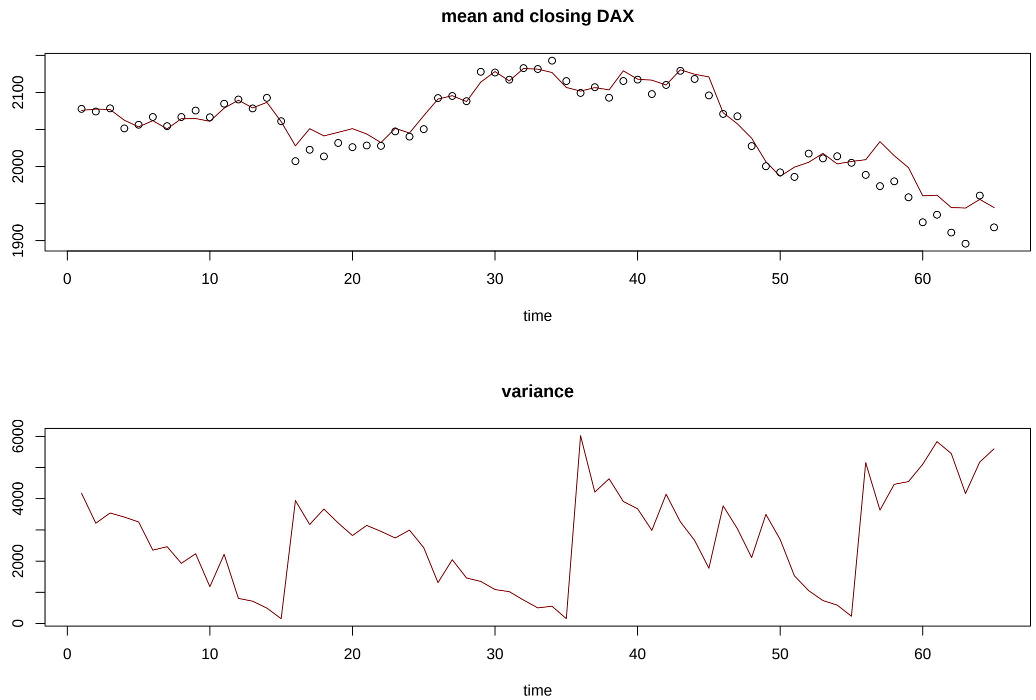
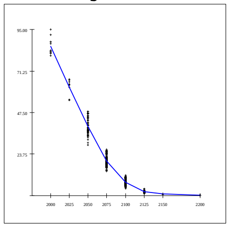


Figure 19: The shape of the SPD changes as the option moves closer to its expiry: the mean of the estimated SPD (red line) follows closely the closing value of DAX(circles) (upper panel) and the variance of SPD decreases as time to maturity decreases (lower panel).

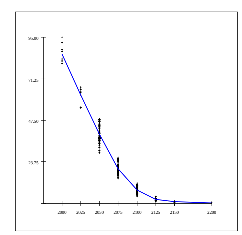


The estimates $f_{\tau_1}(.)$ and $f_{\tau_2}(.)$ of SPDs corresponding to the times of expiry τ_1 and τ_2 lead to an estimate $f_\tau(.)$ for the time of expiry $\tau \in (\tau_1, \tau_2)$

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

In this way, we construct the SPD corresponding to the constant time of expiry $\tau = 45$ days.

The estimated SPD can be used to predict the behavior of the DAX in 45 days.



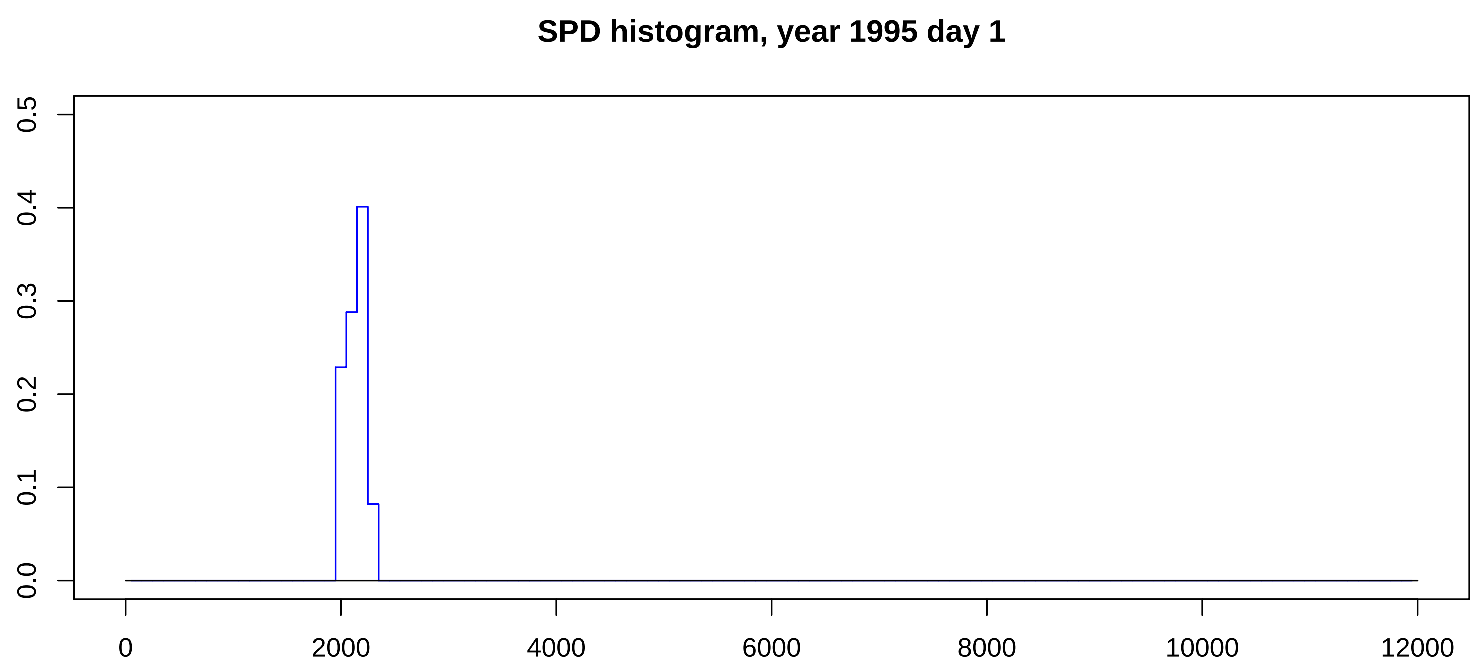
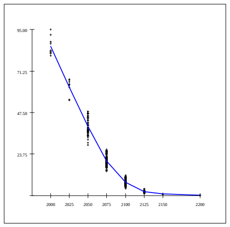


Figure 20: Histogram of the SPD estimates on the first trading day in year 1995.



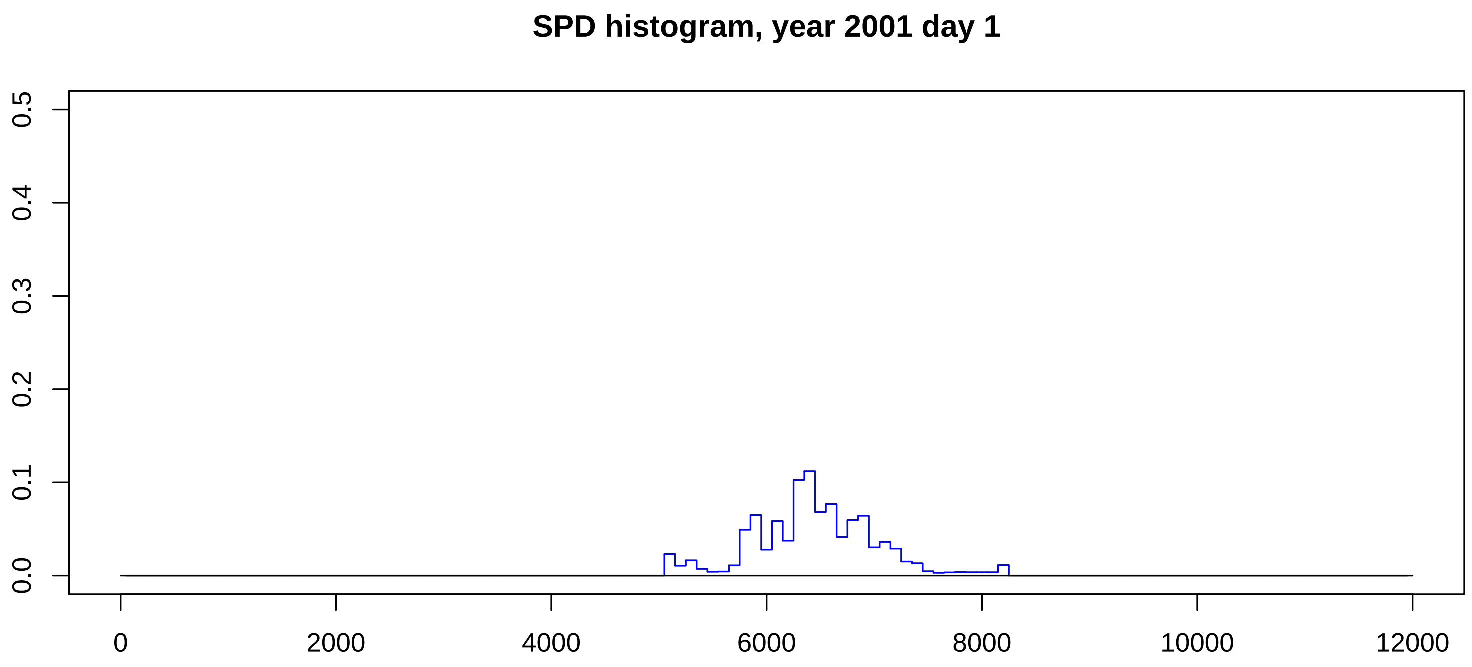
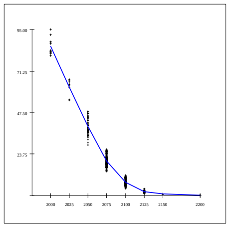


Figure 21: Histogram of the SPD estimates on the first trading day in year 2001.



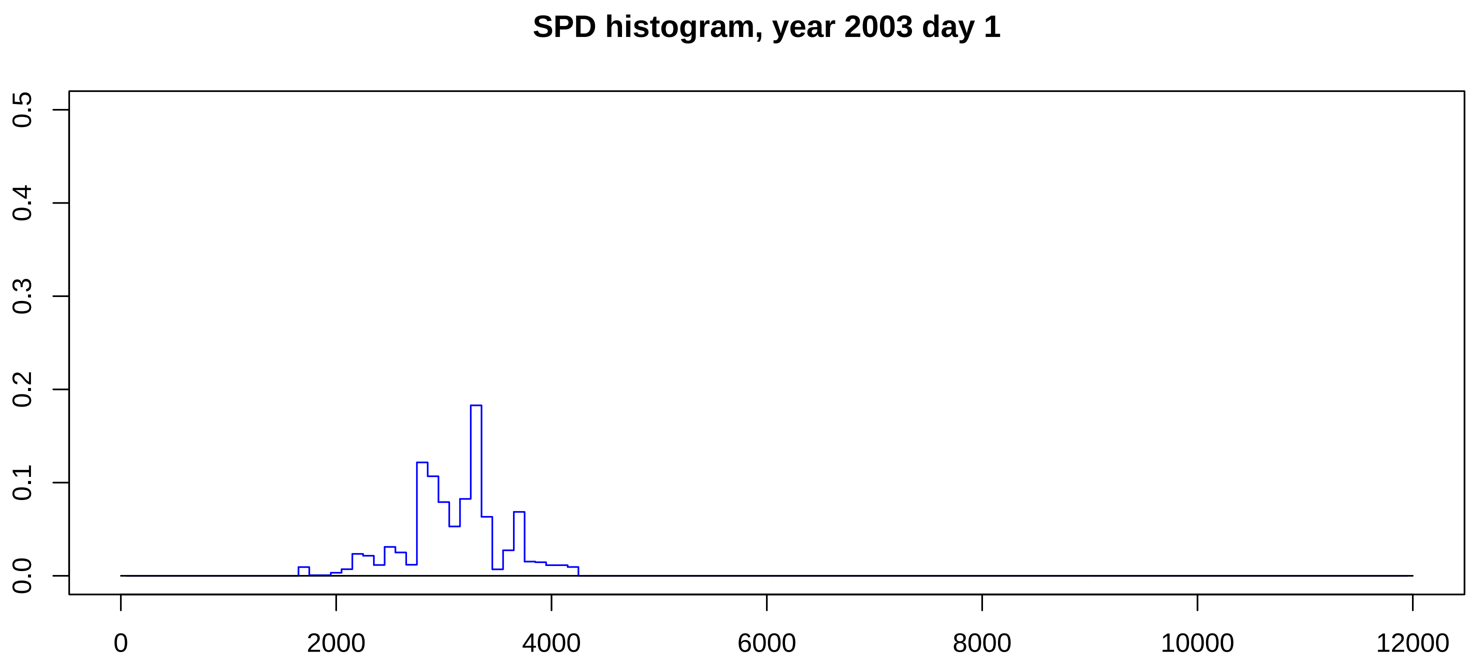


Figure 22: Histogram of the SPD estimates on the first trading day in year 2003.

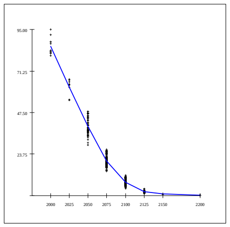
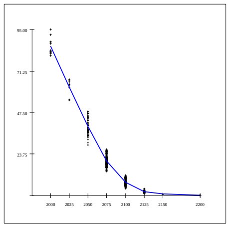


Figure 23: Predictions in year 1995 (upper panel) and in year 2001 (lower panel). The true future values are denoted by black lines, the green and red predictions are based on a quantiles of the estimated SPD.



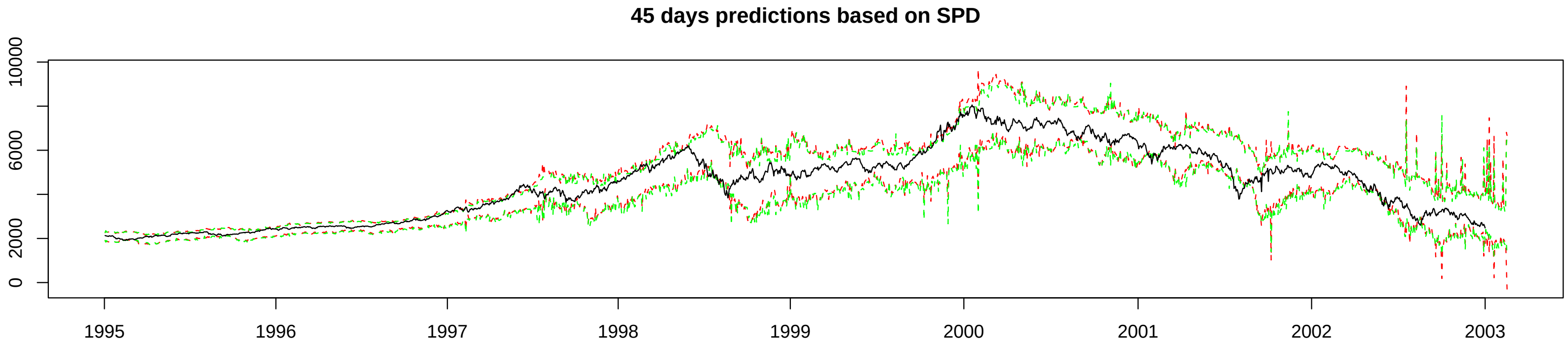
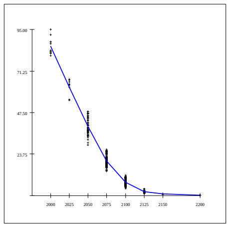


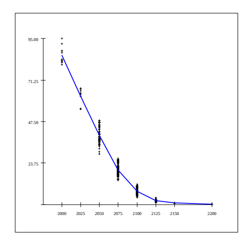
Figure 24: The changes in the market from January 1995 until April 2003 (45 days predictions based on SPD). Red intervals are calculated as the mean of the SPD $\pm$ 2 times standard deviation. The green intervals are quantiles of the SPD. The black line is the true future value of the DAX.

The market is much more volatile in 2003 than in 1995.



Conclusion

1. New model capable of dealing with the constraints
2. Model accomodates both put and call option prices
3. Residual analysis suggests covariance structure of the observations improving the model
4. The dynamics of the prices can be studied using our estimates
5. We obtain good prediction of the future development of the market



References

Aït-Sahalia, Y. & Duarte, J. (2001). Nonparametric Option Pricing under Shape Restrictions, *manuscript, Princeton University*

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

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