Arbitrage-free smoothing of the implied volatility surface

Fengler, Matthias

Arbitrage-free smoothing of the implied volatility surface

Journal:	*Quantitative Finance*
Manuscript ID:	RQUF-2005-0004.R4
Manuscript Category:	Research Paper
Date Submitted by the Author:	24-Aug-2008
Complete List of Authors:	fengler, matthias; sal oppenheim
Keywords:	Local Volatility Theory, Statistical Methods, Exotic Options, Implied Volatilities
JEL Code:	C14 - Semiparametric and Nonparametric Methods < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, G13 - Contingent Pricing|Futures Pricing < G1 - General Financial Markets < G - Financial Economics

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.
FinalSubmission.RQUF-2005-0004.R4.zip

Arbitrage-free smoothing of the implied volatility surface

Matthias R. Fengler*

Trading & Derivatives

Sal. Oppenheim jr. & Cie.

Untermainanlage 1, 60329 Frankfurt am Main, Germany

August 24, 2008

*Corresponding author: matthias.fengler@oppenheim.de, TEL ++49 69 7134 5512, FAX ++49 69 7134 9 5512. The paper represents the author's personal opinion and does not reflect the views of Sal. Oppenheim. I thank Matthias Bode, Tom Christiansen, Enno Mammen, Christian Menn, Daniel Oeltz, Kay Pilz, Peter Schwendner, and the anonymous referees for their helpful suggestions. I am indebted to Eric Reiner for making his material available to me. Support by the Deutsche Forschungsgemeinschaft and by the SfB 649 is gratefully acknowledged.

Arbitrage-free smoothing of the implied volatility surface

Abstract

The pricing accuracy and pricing performance of local volatility models depends on the absence of arbitrage in the implied volatility surface. An input implied volatility surface that is not arbitrage-free can result in negative transition probabilities and consequently into mispricings and false greeks. We propose an approach for smoothing the implied volatility smile in an arbitrage-free way. The method is simple to implement, computationally cheap and builds on the well-founded theory of natural smoothing splines under suitable shape constraints.

Key words: implied volatility surface, local volatility, cubic spline smoothing, no-arbitrage constraints

1 Introduction

The implied volatility surface (IVS) obtained by inverting the Black Scholes (BS) formula serves as a key parameter for pricing and hedging exotic derivatives. For this purpose, other models, more sophisticated than the BS valuation approach, are calibrated to the IVS. A classical candidate is the local volatility model proposed by Dupire (1994), Derman and Kani (1994), and Rubinstein (1994). Local volatility models posit the (risk-neutral) stock price evolution given by

$$\frac{dS_t}{S_t} = (r_t - \delta_t)\,dt + \sigma(S_t, t)\,dW_t\ , \tag{1}$$

where W_t denotes a standard Brownian motion, and r_t and δ_t the continuously compounded interest rate and a dividend yield respectively (both assumed to be deterministic here). Local volatility $\sigma(S_t, t)$ is a nonparametric, deterministic function depending on the asset price S_t and time t. A priori unknown, it must be computed numerically from option prices, or equivalently, from the IVS. Techniques for calibration and pricing are proposed, among others, by Andersen and Brotherton-Ratcliffe (1997); Avellaneda et al. (1997); Dempster and Richards (2000); Jiang and Tao (2001); Jiang et al. (2003).

A crucial property of the calibration data, given as an ensemble of market prices quoted for different strikes and expiries, is the absence of arbitrage. In this context, we refer to arbitrage as to any violation of the theoretical properties of option prices, such as negative butterfly and calendar spreads (see Section 2 for details). If the market data admit arbitrage, the calibration of the local volatility model can fail since negative local volatilities or negative transition probabilities ensue, which obstructs the convergence of the finite difference schemes solving the underlying generalized Black-Scholes partial differential equation. Occasional arbitrage violations may be overridden by an ad hoc approach, but the algorithm fails, when the violations become excessive. While the robustness of the calibration process can be improved by regularizing techniques (Lagnado and Osher; 1997; Bodurtha and Jermakyan; 1999; Crépey; 2003a,b), the specific numerical implementation does not solve the underlying economic problem of data contaminated with arbitrage. One may therefore obtain mispricings and noisy greeks. For illustration, we present in

Figure 1 the delta of a down-and-out put which is computed from a local volatility pricer using a volatility surface contaminated by arbitrage (the data are given in Appendix B). Comparing with a delta calculated from cleaned data (Figure 10) it is apparent that the delta displays local discontinuities which are – aside from the one at the barrier – not economically meaningful. The delta position will hence undergo sudden and unforeseeable jumps as the spot moves, which are inexplicable by changing market conditions or higher order greeks. The hedging performance can therefore deteriorate dramatically.

Unfortunately, an arbitrage-free IVS is not a natural situation in practice, since it is often computed from bid and ask prices, or derived from settlement data of poor quality, see Hentschel (2003) for an exhaustive exposition of this topic. As a strategy to overcome this deficiency, one employs algorithms to remove arbitrage violations from the raw data. Kahalé (2004) proposes an interpolation procedure based on piecewise convex polynomials mimicking the BS pricing formula. The resulting estimate of the call price function is globally arbitrage-free and so is the volatility smile computed from it. In a second step, the total (implied) variance is interpolated linearly along strikes. Crucially, for the interpolation algorithm to work, the data must be arbitrage-free from the outset. Instead of smoothing prices, Benko et al. (2007) suggest to estimate the IVS using local quadratic polynomials. Their strategy requires to solve a smoothing problem under nonlinear constraints.

Here, we propose an approach that unlike Kahalé (2004) is based on cubic spline smoothing of option prices rather than on interpolation. Therefore, the input data do not have to be arbitrage-free. More specifically, for a sample of strikes and call prices, $\{(u_i, y_i)\}$, $u_i \in [a, b]$ for $i = 1, \ldots, n$, we consider the curve estimate defined as minimizer $\widehat{g}$ of the penalized sum of squares

$$\sum_{i=1}^{n} w_i \Big\{ y_i - g(u_i) \Big\}^2 + \lambda \int_a^b \{g''(v)\}^2 \, dv \, , \qquad (2)$$

given strictly positive weights $w_1, \ldots, w_n$. The minimization is carried out with respect to appropriately chosen, linear constraints. The minimizer $\widehat{g}$ is a twice differentiable function and represents a globally arbitrage-free call price function, the smoothness of which is determined by the parameter $\lambda > 0$. To cope with calendar arbitrage across different expiries we apply (2) iteratively to each expiry in adding further constraints. More precisely, we take advantage of a monotonicity property for European options along forward-moneyness corrected strike prices. These additional inequality constraints are straightforward to add to the minimization. Finally, via the BS formula, one obtains an IVS well-suited for pricing and hedging.

In employing cubic spline smoothing, we benefit from a number of nice properties. First, it is possible to cast problem (2) into a convex quadratic program that is known to be solvable within polynomial time (Floudas and Viswewaran; 1995). Second, by virtue of convexity, we have uniqueness of the minimizer. Third, from a statistical point of view, spline smoothers under shape constraints achieve optimal rates of convergence in shape-restricted Sobolev classes (Mammen and Thomas-Agnan; 1999). Finally, since the natural cubic spline is entirely determined by its set of function values and second-order derivatives at the knots, it can be stored and evaluated at the desired grid points in an efficient way and interpolation between grid points is unnecessary. In this way, the method complements existing local volatility pricing engines. The approach is close to the literature on estimating risk neutral transition densities nonparametrically, such as Aït-Sahalia and Duarte (2003) and Härdle and Yatchew (2006), but is less complicated and also applicable when data are scarce (typically there are 20-25 observations, one for each strike, only).

The paper is organized as follows. The next section outlines the principles of no-arbitrage in the option pricing function. Section 3 presents spline smoothing under no-arbitrage constraints. In Section 4, we explore some examples and simulations, and Section 5 concludes.

2 No-arbitrage constraints on call prices and the IVS

In a dynamically complete market, the absence of arbitrage opportunities implies the existence of an equivalent martingale measure, Harrison and Kreps (1979) and Harrison and Pliska (1981), that is uniquely characterized by a risk-neutral transition probability function. We assume that its density exists, which we denote by $\phi(t, T, S_T) = \phi(t, T, S_T, \{r_s, \delta_s\}_{t \le s \le T})$, where S_t is the time-t asset price, $T = t + \tau$ the expiry date of the option, τ time-to-expiration, r_t the deterministic risk-free interest rate and δ_t a deterministic dividend yield of the asset. The valuation function of a European call with strike K is then given by

$$C_t(K, T) = e^{-\int_t^T r_s ds} \int_0^\infty \max(S_T - K, 0)\, \phi(t, T, S_T)\, dS_T \,. \tag{3}$$

From (3) the well-known fact that the call price function is a decreasing and convex function in K is immediately obtained[1]. Taking the derivative with respect to K, and together with the positivity of ϕ and its integrability to one, one gets:

$$-e^{-\int_t^T r_s ds} \le \frac{\partial C}{\partial K} \le 0 \,, \tag{4}$$

which implies monotonicity. Convexity follows from differentiating a second time with respect to K (Breeden and Litzenberger; 1978):

$$\frac{\partial^2 C}{\partial K^2} = e^{-\int_t^T r_s ds} \phi(t, T, S_T) \ge 0 \,. \tag{5}$$

Finally, by general no-arbitrage considerations, the call price function is bounded by

$$\max\left(e^{-\int_t^T \delta_s ds} S_t - e^{-\int_t^T r_s ds} K, 0\right) \le C_t(K, T) \le e^{-\int_t^T \delta_s ds} S_t \,. \tag{6}$$

These constraints are clear-cut for the option price function, but translate into nonlinear conditions for the implied volatility smile. This can be seen by computing (5) explicitly, using the BS formula and assuming a strike-dependent implied volatility function, see Brunner and Hafner (2003) and Benko et al. (2007) for details.

In the time-to-maturity direction only weak constraints on the option price function are known. The prices of American calls for the same strikes must be nondecreasing, which translates to European calls in the absence of dividends. With non-zero dividends, it can

[1] We stress that these properties do not depend on the existence of a density. In continuous time models, they hold when the discounted stock price process is a martingale, but may fail for strict local martingales (Cox and Hobson; 2005).

be shown that there exists a monotonicity property for European call prices along forward-moneyness corrected strikes. This result implies that total (implied) variance must be nondecreasing in forward-moneyness to preclude arbitrage. We define total variance as $\nu^2(\kappa,\tau) = \widehat{\sigma}^2(\kappa,\tau)\tau$, where $\kappa = K/F_t^T$ is forward-moneyness and $F_t^T = S_t e^{\int_t^T (r_s-\delta_s)ds}$ the forward price. The BS implied volatility $\widehat{\sigma}$ is derived by equating market prices with the BS formula

$$C_t^{BS}(K,T) = e^{-\int_t^T \delta_s ds} S_t \Phi(\bar{d}_1) - e^{-\int_t^T r_s ds} K \Phi(\bar{d}_2) ,$$

where Φ is the CDF of the standard normal distribution, and $\bar{d}_1 = \{\ln(S_t/K) + \int_t^T (r_s-\delta_s)ds + \frac{1}{2}\widehat{\sigma}^2\tau\}/\{\sigma\sqrt{\tau}\}$ and $\bar{d}_2 = \bar{d}_1 - \widehat{\sigma}\sqrt{\tau}$. The monotonicity property, which appears to have been found by a number of practitioners independently (Gatheral; 2004; Kahalé; 2004; Reiner; 2004), must to our knowledge be credited to Reiner (2000).

Proposition 2.1. *(Reiner; 2000): Let r_t be an interest rate and and δ_t dividend yield, both depending on time only. For $\tau_1 = T_1 - t < \tau_2 = T_2 - t$ and two strikes K_1 and K_2 related by the forward-moneyness, there is no calendar arbitrage if $C_t(K_2,T_2) \geq e^{-\int_{T_1}^{T_2} \delta_s ds} C_t(K_1,T_1)$. Furthermore, $\nu^2(\kappa,\tau_i)$ is an increasing function in τ_i.*

Proof: Given two expiry dates $t < T_1 < T_2$, construct in t the following calendar spread in two calls with same the forward-moneyness: a long position in the call $C_t(K_2,T_2)$ and a short position in $e^{-\int_{T_1}^{T_2} \delta_s ds}$ calls $C_t(K_1,T_1)$. The forward-moneyness requirement implies $K_1 = e^{\int_{T_1}^{T_2} (\delta_s - r_s)ds} K_2$. In T_1, if $S_{T_1} \leq K_1$, the short position expires worthless, while $C_{T_1}(K_2,T_2) \geq 0$. Otherwise, the entire portfolio consists of $C_{T_1}(K_2,T_2) - e^{-\int_{T_1}^{T_2} \delta_s ds}\left(S_{T_1} - e^{\int_{T_1}^{T_2}(\delta_s - r_s)ds} K_2\right) = P_{T_1}(K_2,T_2) \geq 0$ by put-call-parity. Thus, the payoff of this portfolio is always non-negative. To preclude arbitrage we must have:

$$C_t(K_2,T_2) \geq e^{-\int_{T_1}^{T_2} \delta_s ds} C_t(K_1,T_1) , \tag{7}$$

which proves the first statement. Multiplying with $e^{\int_t^{T_2} r_s ds}$ and dividing by K_2 yields:

$$\frac{e^{\int_t^{T_2} r_s ds} C_t(K_2,T_2)}{K_2} \geq \frac{e^{\int_t^{T_1} r_s ds} C_t(K_1,T_1)}{K_1} . \tag{8}$$

Replacing C_t by C_t^{BS}, define the function

$$\begin{aligned} f(\kappa, \nu^2) &= \frac{e^{\int_t^T r_s ds} C_t^{BS}(K,T)}{K} \\ &= \kappa^{-1}\Phi(\bar{d}_1) - \Phi(\bar{d}_2) \,. \end{aligned} \quad (9)$$

As can be observed, $f(\kappa, \nu^2)$ is a function in κ and ν^2 only and, for a fixed κ, is strictly monotonically increasing in ν^2, since $\partial f / \partial \nu^2 = \frac{1}{2}\varphi(\bar{d}_2)/\sqrt{\nu^2} > 0$ for $\nu^2 \in (0, \infty)$. Thus, Eq. (8) implies $\nu^2(\kappa, T_2) \geq \nu^2(\kappa, T_1)$, ruling out calendar arbitrage. $\square$

To more precisely characterize the concept of no-arbitrage in a set of option data we rely on recent work by Carr et al. (2003) who introduced the concept of 'static arbitrage'. Static arbitrage refers to a costless trading strategy which yields a positive profit with non-zero probability, but has zero probability to incur a loss. The term 'static' means that positions can only depend on time and the concurrent underlying stock price. In particular, they are not allowed to depend on past prices or on path properties. For a discrete ensemble of strikes $K_i,\ i = 1, \ldots, \infty$ and expiries $T_j,\ j = 1, \ldots, M$, static no arbitrage can be established along the line of arguments outlined in Carr and Madan (2005): Given that the data set does not admit strike arbitrage and calendar arbitrage, one constructs a convex order of risk neutral probability measures at different expiries. The convex order implies the existence of a Markov martingale by the results of Kellerer (1972). It follows that there exists a martingale measure consistent with all call price quotes and defined on a filtration that contains at least the underlying asset price and time. Hence the option call price quotes are free of static arbitrage (Carr et al.; 2003; Carr and Madan; 2005).

As a consequence of Proposition 2.1, a plot of the total variance against the forward moneyness shows calendar arbitrage when the graphs intersect. In Figure 2 we provide such a total variance plot of our IVS data. Evidently, there are a significant number of implied volatility observations with three days to expiry which violate the no-arbitrage restriction. It is typical that only the front month violates calendar arbitrage. This occurs when the short run smile is very pronounced or when the term structure of the IVS is strongly downward sloping or humped.

3 Spline smoothing

3.1 Generic set-up

Spline smoothing is a classical statistical technique that is covered in almost every monograph on smoothing, see e.g. Härdle (1990) and Green and Silverman (1994). A particularly nice resource is Turlach (2005) whose exposition we follow closely.

Assume that we observe call prices y_i at strikes $a = u_0, \ldots, u_{n+1} = b$. A function g defined on $[a, b]$ is called a cubic spline, if g, on each subinterval (a, u_1), $(u_2, u_3), \ldots, (u_n, b)$, is a cubic polynomial and if g belongs to the class of twice differentiable functions $\mathcal{C}^2([a, b])$. The points u_i are called knots. The spline g has the representation

$$\begin{aligned} g(u) &= \sum_{i=0}^{n} \mathbf{1}\{[u_i, u_{i+1})\}\, s_i(u) \qquad (10) \\ \text{where} \qquad s_i(u) &= a_i + b_i(u - u_i) + c_i(u - u_i)^2 + d_i(u - u_i)^3 \ , \end{aligned}$$

for $i = 0, \ldots, n$ and given constants a_i, b_i, c_i, d_i. There are $4(n + 1)$ coefficients to be determined. The continuity conditions on g and its first and second order derivatives in each interior segment imply $4n$ restrictions on the coefficients. The indeterminacy can be resolved by requiring that g has zero second order derivatives in the very first and the very last segment of the spline. This assumption implies that $c_0 = d_0 = c_n = d_n = 0$, in which case g is called a *natural cubic spline*. This choice is justified by the fact that the gamma of the call converges fast to zero for high and low strikes. As will be seen presently, the natural cubic spline allows for a convenient formulation of the no-arbitrage conditions to be imposed on the call price function[2].

[2]It should be noted that the literature on the numerical treatment of splines also discusses other end conditions (Wahba; 1990). A popular choice is to fix the first-order derivatives at the end points of the spline. We experimented with this solution. In this case, the smoothness penalty does not have the convenient quadratic form anymore, see Proposition 3.1, but could be approximated by the smoothness penalty given by the natural spline. Further, since in our application the two first-order derivatives are unknown, they must be estimated. As proxy we used the first-order BS derivative w.r.t. the strike evaluated at the strike implied volatility. In our simulations it turned out that the spline functions are

As is discussed in Green and Silverman (1994), a more convenient representation of (10) is given by the so called *value-second derivative representation* of the natural cubic spline. In particular it allows to formulate a quadratic program to solve (2). For the value-second derivative representation, put $g_i = g(u_i)$ and $\gamma_i = g''(u_i)$, for $i = 1, \ldots, n$. Furthermore define $\mathbf{g} = (g_1, \ldots, g_n)^\top$ and $\boldsymbol{\gamma} = (\gamma_2, \ldots, \gamma_{n-1})^\top$. By definition, $\gamma_1 = \gamma_n = 0$. The nonstandard notation of the entries in $\boldsymbol{\gamma}$ is proposed by Green and Silverman (1994). The natural spline is completely specified by the vectors $\mathbf{g}$ and $\boldsymbol{\gamma}$. In Appendix A, we give the formulae to switch between the two representations.

Not all possible vectors $\mathbf{g}$ and $\boldsymbol{\gamma}$ result in a valid cubic spline. Sufficient and necessary conditions are formulated via the following two matrices $\mathbf{Q}$ and $\mathbf{R}$. Let $h_i = u_{i+1} - u_i$ for $i = 1, \ldots, n-1$, and define the $n \times (n-2)$ matrix $\mathbf{Q}$ by its elements $q_{i,j}$, for $i = 1, \ldots, n$ and $j = 2, \ldots, n-1$, given by

$$q_{j-1,j} = h_{j-1}^{-1}, \qquad q_{j,j} = -h_{j-1}^{-1} - h_j^{-1}, \qquad \text{and} \quad q_{j+1,j} = h_j^{-1},$$

for $j = 2, \ldots, n-1$, and $q_{i,j} = 0$ for $|i-j| \geq 2$. The columns of $\mathbf{Q}$ are numbered in the same non-standard way as the vector $\boldsymbol{\gamma}$.

The $(n-2) \times (n-2)$ matrix $\mathbf{R}$ is symmetric and is defined by its elements $r_{i,j}$ for $i, j = 2, \ldots, n-1$, given by

$$\begin{aligned} r_{i,i} &= \tfrac{1}{3}(h_{i-1} + h_i) \quad \text{for} \quad i = 2, \ldots, n-1 \\ r_{i,i+1} = r_{i+1,i} &= \tfrac{1}{6}h_i \quad \text{for} \quad i = 2, \ldots, n-2, \end{aligned} \tag{11}$$

and $r_{i,j} = 0$ for $|i-j| \geq 2$. The matrix $\mathbf{R}$ is strictly diagonal dominant, so by standard arguments in linear algebra, $\mathbf{R}$ is strictly positive-definite.

Proposition 3.1. *The vectors* $\mathbf{g}$ *and* $\boldsymbol{\gamma}$ *specify a natural cubic spline if and only if*

$$\mathbf{Q}^\top \mathbf{g} = \mathbf{R}\boldsymbol{\gamma}. \tag{12}$$

If (12) holds, the roughness penalty satisfies

$$\int_a^b g''(u)^2 du = \boldsymbol{\gamma}^\top \mathbf{R} \boldsymbol{\gamma}. \tag{13}$$

very sensitive to a misspecification of the first-order derivatives and less robust than the natural spline solution.

Proof: Green and Silverman (1994, Section 2.5). $\square$

This result allows us to state the spline smoothing task as a quadratic minimization problem. Define the $(2n-2)$-vector $\mathbf{y} = (w_1 y_1, \dots, w_n y_n, 0, \dots, 0)^\top$, where the w_i are strictly positive weights, and the $(2n-2)$-vector $\mathbf{x} = (\mathbf{g}^\top, \boldsymbol{\gamma}^\top)^\top$. Further, define the matrices, $\mathbf{A} = \left(\mathbf{Q}, -\mathbf{R}^\top\right)$ and

$$\mathbf{B} = \begin{pmatrix} \mathbf{W}_n & 0 \\ 0 & \lambda \mathbf{R} \end{pmatrix} , \tag{14}$$

where $\mathbf{W}_n = \mathrm{diag}(w_1, \dots, w_n)$. The solution to (2) can then be written as the solution of the quadratic program:

$$\begin{aligned} \min_{\mathbf{x}} \quad & -\mathbf{y}^\top \mathbf{x} + \frac{1}{2}\mathbf{x}^\top \mathbf{B}\mathbf{x} , \\ \text{subject to} \quad & \mathbf{A}^\top \mathbf{x} = 0 . \end{aligned} \tag{15}$$

The quadratic program (15) serves as the basis for our arbitrage-free smoothing of the call price function. To this end we will add further restrictions on $\mathbf{x}$ that ensure the properties outlined in Section 2. Since $\mathbf{B}$ is strictly positive-definite by construction, program (15) benefits from two decisive properties, irrespective of the additional no-arbitrage constraints to be imposed. First, by positive-definiteness of $\mathbf{B}$, it belongs to the class of convex programs which are known to be solvable within polynomial time (Floudas and Viswewaran; 1995). Algorithms for solving convex quadratic programs are nowadays available in almost every statistical software package. An excellent resource is Boyd and Vandenberghe (2004). Second, and most importantly, convex programs are known to have a unique minimizer. Hence the smoothing spline, for given data and λ, is unique (Green and Silverman; 1994, Theorem 2.4).

3.2 Cubic spline smoothing under no-arbitrage constraints

It is straightforward to translate the no-arbitrage conditions for the call price function into conditions on the smoothing spline. Convexity of the spline is imposed by noting that the second derivative of the spline is linear. Hence it is sufficient to require that the second derivatives at the knot points be positive, i.e.,

$$\gamma_i \geq 0 \,, \tag{16}$$

for $i = 2, \ldots, n-1$. By definition, we have $\gamma_1 = \gamma_n = 0$.

Next, the price function must be nonincreasing in strikes. Since the convexity constraints insure that the slope is increasing, it is sufficient to constrain the initial derivatives at both end points of the spline. For a cubic spline on the segment $[u_L, u_R]$ the left boundary derivative from the right is given by $g'(u_L^+) = (g_R - g_L)/h - h(2\gamma_L + \gamma_R)/6$, and the right boundary derivative from the left by $g'(u_R^-) = (g_R - g_L)/h + h(\gamma_L + 2\gamma_R)/6$. Thus, since $\gamma_1 = \gamma_n = 0$, the necessary and sufficient constraints are given by

$$\frac{g_2 - g_1}{h_1} - \frac{h_1}{6}\gamma_2 \geq -e^{-\int_t^T r_s ds} \quad \text{and} \quad \frac{g_n - g_{n-1}}{h_{n-1}} + \frac{h_{n-1}}{6}\gamma_{n-1} \leq 0 \,. \tag{17}$$

Finally, we add the constraints:

$$e^{-\int_t^T \delta_s ds} S_t - e^{-\int_t^T r_s ds} u_1 \leq g_1 \leq e^{-\int_t^T \delta_s ds} S_t \qquad \text{and} \qquad g_n \geq 0 \,. \tag{18}$$

Including the conditions (16) to (18) into the quadratic program (15) yields an arbitrage-free call price function and, in consequence, an arbitrage-free volatility smile.

3.3 Estimating an arbitrage-free IVS

The preceding sections lead to a natural procedure to generate an arbitrage-free IVS:

1. Estimate the IVS via an initial estimate on a regular forward-moneyness grid $\mathcal{J} = [\kappa_1, \kappa_n] \times [t_1, t_m]$.

2. Iterate through the price surface from the last to the first maturity, and solve the following quadratic program.

   For t_m, solve

$$
\begin{array}{lrcl}
\min_{\mathbf{x}} & & & -\mathbf{y}^\top \mathbf{x} + \frac{1}{2}\mathbf{x}^\top \mathbf{B}\mathbf{x} \,, \\
\text{subject to} & \mathbf{A}^\top \mathbf{x} & = & 0 \\
& \gamma_i & \geq & 0 \,, \\
& \frac{g_2 - g_1}{h_1} - \frac{h_1}{6}\gamma_2 & \geq & -e^{-\int_{t_m}^T r_s ds} \\
& -\frac{g_n - g_{n-1}}{h_{n-1}} - \frac{h_{n-1}}{6}\gamma_{n-1} & \geq & 0 \\
& g_1 & \leq & e^{-\int_{t_m}^T \delta_s ds} S_t \qquad (*) \\
& g_1 & \geq & e^{-\int_{t_m}^T \delta_s ds} S_t - e^{-\int_{t_m}^T r_s ds} u_1 \\
& g_n & \geq & 0 \,.
\end{array}
\tag{19}
$$

   where $\mathbf{x} = (\mathbf{g}^\top, \boldsymbol{\gamma}^\top)^\top$;

   For t_j, $j = m-1, \ldots, 1$, solve (19) replacing condition $(*)$ by:

$$
g_i^{(j)} < e^{\int_{t_j}^{t_{j+1}} \delta_s ds} g_i^{(j+1)} \,, \quad \text{for} \quad i = 1, \ldots, n \,,
$$

   where $g_i^{(j)}$ denotes the ith spline value of maturity j.

In order to respect condition (7), Step 1 can be circumvented by evaluating each spline of the previous time-to-maturity at the desired strikes. But it might be faster to employ the initial estimate, because the IVS observations can easily be spaced on the forward-moneyness grid. As an initial estimator any two-dimensional nonparametric smoother, such as a local polynomial estimator or a thin plate spline, is a natural candidate (Wahba;

1990; Green and Silverman; 1994). The absence of strike arbitrage along the price function and the absence of calendar arbitrage at the knots is insured by Step 2. In general, it cannot be excluded that there is calendar arbitrage between the knots, but this is very unlikely given the convex, monotonic shape of the call price function.

The smoothing parameter λ can either be determined according to a subjective view or an automatic, data-driven choice of the smoothing parameter can be used. In the latter case, asymptotically optimal bandwidths can be found by 'leave-one-out' cross-validation techniques (Green and Silverman; 1994, Section 3.2). Unfortunately, due to the no-arbitrage constraints present in the program, the common and efficient calculation techniques are not applicable. For each cross validation score it is necessary to solve n separate smoothing problems which is cumbersome. However, the shape constraints we impose – monotonicity and convexity – act already as a strong smoothing device. As pointed out by Dole (1999, p. 446), bounds on second-order derivatives can be seen as smoothing parameters in their own right. Therefore, the choice of the smoothing parameter is of secondary importance. It can be fixed at some small number without large impact on the estimate (see Turlach (2005) for a related discussion). Choosing a very small number has the additional benefit that initially good data will hardly be smoothed at all.

From the perspective of financial theory, one might worry that the sum of squared differences of $y_i - g(u_i)$ in (2) may not be the right measure of loss, since an investor is only interested in relative prices. This concern can be addressed by using the underlying asset price as numéraire. By setting $w_i = S_t^{-2}$ and switching to a spot moneyness space $\widetilde{u} = u/S_t$, one can conduct the minimization on relative option prices after some obvious adjustments to the no-arbitrage constraints in (19). The resulting curve estimate $(\widetilde{\mathbf{g}}^\top, \widetilde{\boldsymbol{\gamma}}^\top)^\top$ can be inflated again via $g_i(u) = S_t \widetilde{g}_i(\widetilde{u})$ and $\gamma_i(u) = \widetilde{\gamma}_i(\widetilde{u})/S_t$, which yields a natural cubic spline as can be verified from (12). Seemingly this approach comes at the additional cost of a homogeneity assumption. However, as can be observed from (15), in choosing as smoothing parameter $\widetilde{\lambda} = \lambda S_t^{-3}$ the program in relative prices is equivalent to the former one in absolute prices (up to the aforementioned scales). This property is hidden in the value-second derivative representation of the natural cubic spline. For a discussion of the financial implications of an option pricing function that is homogeneous in spot and strikes we refer to Renault (1997), Alexander and Nogueira (2007) and Fengler et al. (2007).

4 Empirical demonstration

We demonstrate the estimator using single expiries and the entire IVS of DAX settlement data observed on June 13, 2000; see also Table 1 the Appendix B. These data represent a typical situation one faces when working with settlement data. By the conditions spelled out in Section 2, market data violating strike arbitrage conditions are found by testing in the sample of strikes and prices (K_i, C_i), $i = 1, \dots, n$, whether

$$-e^{-\int_t^T r_s ds} \leq \frac{C_i - C_{i-1}}{K_i - K_{i-1}} \leq \frac{C_{i+1} - C_i}{K_{i+1} - K_i} \leq 0 \tag{20}$$

holds. In the optimization do not use specific weights and work in absolute prices. A good initial value $\mathbf{x}_0$ for the quadratic program (19) is given by the observed market prices; the part in $\mathbf{x}_0$ containing the second-order derivatives is initialized to 1e-3. The smoothing parameter fixed at $\lambda =$1e-7. Implied volatility is computed from the smoothed call prices.

For the exposition, we pick the expiries with 68 and 398 days time-to-maturity as they have a significant vega. Figures 3 and 6 show the smoothed implied volatility data together with the original observations printed as crosses. We identify the (center) observations that allow for arbitrage according to Eq. (20) by an additional square. Since the residuals computed as differences between the raw data and the estimated spline are sometimes hardly discernible, we further present the implied volatility residuals in Figures 4 and 7 and the price residuals in Figures 5 and 8. Note that all observations marked with the square are in the positive half plane of the plot. The reason is that the simplest way to correct three observations for convexity is to pull the center observation (marked with the square) downwards and to correct the observations $i-1$ and $i+1$ into the opposite direction. This is the correction the quadratic program chooses in most cases.

The adjustments, which are necessary to achieve an arbitrage-free set of call prices, can be substantial. Measured in terms of implied volatility they amount to around 10bp in Figure 4 and to around 30bp in Figure 7. For the price residuals, the biggest deviations are observed near-the-money, where the vega sensitivity is highest.

The entire IVS is given in Figure 9. The estimate is obtained using a thin plate spline as initial estimator on the forward-moneyness grid $\mathcal{J} = [0.6, 1.25] \times [0.1, 1.6]$ with 100 grid points altogether and by applying the arbitrage-free estimation technique from the last to first time-to-maturity. For these computations the implied volatility observations with three days to expiry were deleted from the raw data sample as is regularly suggested in the literature (Andersen and Brotherton-Ratcliffe; 1997; Bodurtha and Jermakyan; 1999; Crépey; 2003b). In Figure 10 we present the delta of the down-and-out put obtained from the local volatility model based on this arbitrage-free data set. The delta is computed via a finite difference quotient, directly read from the grid of the PDE solver. As explained in the introduction, the local discontinuities vanish when using data smoothed in an arbitrage-free manner.

To give an idea of the properties of our spline smoothing approach we do a simulation comparing it with a benchmark model. As benchmark model we choose the Heston (1993) model, which is often taken as the first alternative to local volatility models. Under a risk-neutral measure, the model is given by

$$dS_t = (r_t - \delta_t)S_t dt + \sqrt{V_t} S_t dW_t^1 \tag{21}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^2 \; , \tag{22}$$

where $dW^1 dW^2 = \rho dt$. Unlike spline smoothing the Heston model is a parametric model with five parameters $\kappa, \theta, \sigma, \rho$ and the initial variance V_0. A comparison between the two models is essentially a comparison of the trade-off between variance and bias. Nevertheless it is instructive to compare both types of models.

The set-up we used is borrowed from Bliss and Panigirtzoglou (2002) developed for testing the stability of state price densities. The idea is to resample from artificially perturbed data. We consider two cases. First we fit the Heston model to the observed data using the FFT pricer by Carr and Madan (1999). From the estimated parameters we generate implied volatility smiles that are perturbed by zero mean normal errors. The standard deviation is chosen between 50bp to 10bp. Then we fit both models to the perturbed Heston data. Second, we use the market data and perturb those. Again both models are fitted. We look at single expiries only, since the Heston model displayed too much bias when fitted to the entire surface. The number of simulations is set to 100. At any time, the natural cubic spline converged, while the Heston model occasionally did not; in these cases a new set of random errors was drawn.

The results are displayed in Table 2. The trade-off between variance and bias is well obvious in the figures. In almost every case the RSME (root mean square error) of the spline model is smaller than the Heston model's. Furthermore, when comparing the RMSE* measures, which present the error w.r.t. the true smile, it is evident that the Heston model is superior to the spline in terms of identifying its own model, from which data are generated. This advantage disappears when the market data are used and perturbed. Of course, the Heston model cannot identify the unperturbed market, and the error measures are of comparable size. This is a significant virtue of the spline smoother, since as a matter of fact the market model is unknown; it underpins that the spline smoother is a the natural complement to local volatility pricers which aim at best fitting all market prices.

5 Conclusion

Local volatility pricers require as input an arbitrage-free implied volatility surface (IVS) – otherwise they can produce mispricings. This is because arbitrage violations lead to negative transition probabilities in the underlying finite difference scheme. In this paper, we propose an algorithm for estimating the IVS in an arbitrage-free manner. For a single time-to-maturity the approach consists in applying a natural cubic spline to the call price function under suitable linear inequality constraints. For the entire IVS, we first obtain the fit on a fixed forward-moneyness grid. Second the natural spline smoothing algorithm is applied by stepping from the last expiry to the first one. This precludes calendar and strike arbitrage.

The method improves on existing algorithms in three ways. First the initial data do not have to be arbitrage-free from the beginning. Second, the solution is obtained via a convex quadratic program that has a unique minimizer. Finally, the estimate can be stored efficiently via the value-second derivative representation of the natural spline. Integration into local volatility pricers is therefore straightforward.

References

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

Alexander, C. and Nogueira, L. M. (2007). Model-free hedge ratios and scale-invariant models, *Journal of Banking and Finance* **31**(6): 1839–1861.

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

Avellaneda, M., Friedman, C., Holmes, R. and Samperi, D. (1997). Calibrating volatility surfaces via relative entropy minimization, *Applied Mathematical Finance* **4**: 37–64.

Benko, M., Fengler, M. R., Härdle, W. and Kopa, M. (2007). On extracting information implied in options, *Computational Statistics* **22**(4): 543–553.

Bliss, R. and Panigirtzoglou, N. (2002). Testing the stability of implied probability density functions, *Journal of Banking and Finance* **26**: 381–422.

Bodurtha, J. N. and Jermakyan, M. (1999). Nonparametric estimation of an implied volatility surface, *Journal of Computational Finance* **2**(4): 29–60.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*, Cambridge University Press, Cambridge.

Breeden, D. and Litzenberger, R. (1978). Price of state-contingent claims implicit in options prices, *Journal of Business* **51**: 621–651.

Brunner, B. and Hafner, R. (2003). Arbitrage-free estimation of the risk-neutral density from the implied volatility smile, *Journal of Computational Finance* **7**(1): 75–106.

Carr, P., Geman, H., Madan, D. B. and Yor, M. (2003). Stochastic volatility for Lévy processes, *Mathematical Finance* **13**(3): 345–382.

Carr, P. and Madan, D. B. (1999). Option valuation using the fast Fourier transform, *Journal of Computational Finance* **2**(4): 61–73.

Carr, P. and Madan, D. B. (2005). A note on sufficient conditions for no arbitrage, *Finance Research Letters* **2**: 125–130.

Cox, A. M. G. and Hobson, D. G. (2005). Local martingales, bubbles and option prices, *Finance and Stochastics* **9**(4): 477–492.

Crépey, S. (2003a). Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization, *SIAM Journal on Mathematical Analysis* **34**(5): 1183–1206.

Crépey, S. (2003b). Calibration of the local volatility in a trinomial tree using Tikhonov regularization, *Inverse Problems* **19**: 91–127.

Dempster, M. A. H. and Richards, D. G. (2000). Pricing American options fitting the smile, *Mathematical Finance* **10**(2): 157–177.

Derman, E. and Kani, I. (1994). Riding on a smile, *RISK* **7**(2): 32–39.

Deutsche Börse (2006). *Guide to the Equity Indices of Deutsche Börse*, 5.12 edn, Deutsche Börse AG, 60485 Frankfurt am Main, Germany.

Dole, D. (1999). CoSmo: A constrained scatterplot smoother for estimating convex, monotonic transformations, *Journal of Business and Economic Statistics* **17**(4): 444–455.

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

Fengler, M. R., Härdle, W. and Mammen, E. (2007). A semiparametric factor model for implied volatility surface dynamics, *Journal of Financial Econometrics* **5**(2): 189–218.

Floudas, C. A. and Viswewaran, V. (1995). Quadratic optimization, *in* R. Horst and P. M. Pardalos (eds), *Handbook of global optimization*, Kluwer Academic Publishers, Dordrecht, pp. 217–270.

Gatheral, J. (2004). A parsimonious arbitrage-free implied volatility parametrization with application to the valuation of volatility derivatives, http://www.math.nyu.edu/fellows_fin_math/gatheral/madrid2004.pdf.

Green, P. J. and Silverman, B. W. (1994). *Nonparametric regression and generalized linear models*, Vol. 58 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, London.

Härdle, W. (1990). *Applied Nonparametric Regression*, Cambridge University Press, Cambridge, UK.

Härdle, W. and Yatchew, A. (2006). Dynamic state price density estimation using constrained least squares and the bootstrap, *Journal of Econometrics* **133**(2): 579–599.

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

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

Hentschel, L. (2003). Errors in implied volatility estimation, *Journal of Financial and Quantitative Analysis* **38**: 779–810.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* **6**: 327–343.

Jiang, L., Chen, Q., Wang, L. and Zhang, J. E. (2003). A new well-posed algorithm to recover implied local volatility, *Quantitative Finance* **3**: 451–457.

Jiang, L. and Tao, Y. (2001). Identifying the volatility of the underlying assets from option prices, *Inverse Problems* **17**: 137–155.

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

Kellerer, H. G. (1972). Markov-Komposition und eine Anwendung auf Martingale, *Mathematische Annalen* **198**: 99–122.

Lagnado, R. and Osher, S. (1997). A technique for calibrating derivative security pricing models: Numerical solution of an inverse problem, *Journal of Computational Finance* **1**(1): 13–25.

Mammen, E. and Thomas-Agnan, C. (1999). Smoothing splines and shape restrictions, *Scandinavian Journal of Statistics* **26**: 239–252.

Reiner, E. (2000). Calendar spreads, characteristic functions, and variance interpolation. Mimeo.

Reiner, E. (2004). The characteristic curve approach to arbitrage-free time interpolation of volatility, presentation at the ICBI Global Derivatives and Risk Management, Madrid, España.

Renault, E. (1997). Econometric models of option pricing errors, *in* D. M. Kreps and K. F. Wallis (eds), *Advances in Economics and Econometrics, Seventh World Congress*, Econometric Society Monographs, Cambridge University Press, pp. 223–278.

Rubinstein, M. (1994). Implied binomial trees, *Journal of Finance* **49**: 771–818.

Turlach, B. A. (2005). Shape constrained smoothing using smoothing splines, *Computational Statistics* **20**(1): 81–104.

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

A Transformation formulae

To switch from the value-second derivative representation to the piecewise polynomial representation (10) employ:

$$\begin{aligned} a_i &= g_i \\ b_i &= \tfrac{g_{i+1}-g_i}{h_i} - \tfrac{h_i}{6}(2\gamma_i + \gamma_{i+1}) \\ c_i &= \tfrac{\gamma_i}{2} \\ d_i &= \tfrac{\gamma_{i+1}-\gamma_i}{6h_i} \end{aligned} \tag{23}$$

for $i = 1, \ldots, n-1$. Furthermore,

$$a_0 = a_1 = g_1, \quad a_n = g_n, \quad b_0 = b_1, \quad c_0 = d_0 = c_n = d_n = 0 \ ,$$

and

$$b_n = s'_{n-1}(u_n) = b_{n-1} + 2c_{n-1}h_{n-1} + 3d_{n-1}h_{n-1}^2 = \frac{g_n - g_{n-1}}{h_{n-1}} + \frac{h_i}{6}(\gamma_{n-2} + 2\gamma_n) \ ,$$

where $h_i = u_{i+1} - u_i$ for $i = 1, \ldots, n-1$ and $\gamma_1 = \gamma_n = 0$.

Changing vice versa is accomplished by:

$$\begin{aligned} g_i &= s_i(u_i) &&= a_i \quad \text{for} \quad i = 1, \ldots, n \ , \\ \gamma_i &= s''_i(u_i) &&= 2c_i \quad \text{for} \quad i = 2, \ldots, n-1 \ , \\ \gamma_1 &= \gamma_n &&= 0 \ . \end{aligned} \tag{24}$$

B Data

time-to-maturity	3	28	48	68	133	198	263	398
strikes	implied volatilities							
2600	367.09							
2800	340.92							
3000	316.57							
3200	293.81							
3400	272.43							
3600	252.28							
3800	233.23							
4000	215.16							
4200	197.98				38.39			
4400	181.60				37.10			
4600	165.96				36.19		34.51	
4800	150.99			36.86	35.25		33.79	
4900	143.73			36.15				
5000	136.63			35.65	34.16		33.28	
5100	129.66			35.38				
5200	122.83			34.57	33.55		32.32	
5300	116.13			33.94				
5400	109.56			33.57	32.35		31.82	
5500	103.11			33.02				
5600	96.78	31.86		32.30	31.70		31.11	
5700	90.56			31.80				
5800	84.45	30.18		31.53	30.65		30.16	
5900	78.44			30.73				
6000	72.54	28.46	29.12	30.13	29.92	29.38	29.47	29.49
6100	66.74			29.77				
6200	61.03	27.04	28.01	29.17	29.04	28.72	28.91	28.78
6250	58.21							
6300	55.42			28.54	28.58			
6350	52.64							
6400	55.24	25.86	27.06	28.19	28.26	28.06	28.00	27.83
6450	52.88							
6500	50.76			27.66	28.04			
6550	48.08							
6600	45.41	24.70	26.02	27.08	27.44	27.24	27.39	27.87
6650	42.77	24.29	25.74	26.88				
6700	40.15	24.15	25.38	26.74	26.99	26.97		
6750	37.55	24.05	25.12	26.59				
6800	34.97	23.59	24.96	26.23	26.67	26.79	26.79	27.15
6850	34.14	23.32	24.90	25.92				
6900	31.62	23.21	24.49	25.69	26.48	26.30		
6950	29.71	23.00	24.21	25.54				

Raw DAX implied volatility data from June 13, 2000, traded at the EUREX, Germany. Time-to-maturity measured in calendar days.

time-to-maturity	3	28	48	68	133	198	263	398
strikes	implied volatilities							
7000	28.95	22.62	24.02	25.45	25.91	25.88	26.00	26.63
7050	28.46	22.41	23.94	25.15				
7100	26.85	22.42	23.65	24.83	25.50	25.53		
7150	27.11	22.01	23.34	24.59				
7200	25.56	21.74	23.14	24.43	25.24	25.30	25.54	26.32
7250	25.30	21.69	23.07	24.34				
7300	23.98	21.21	22.65	23.91	24.87	24.86		
7350	23.80	20.94	22.33	23.60				
7400	23.59	20.86	22.16	23.37	24.39	24.40	24.47	24.94
7450	23.91	20.77	22.11	23.23				
7500	24.87	20.46	21.88	23.17	24.06	24.06		
7550	25.59	20.37	21.62	22.95				
7600	26.96	20.37	21.48	22.67	23.89	23.85	23.90	24.45
7650	28.32	20.06	21.48	22.48				
7700	29.03	20.00		22.38	23.51	23.71		
7750	30.25	20.10		22.35				
7800	32.47	19.80	20.93	22.07	23.14	23.30	23.56	24.15
7850	34.69			21.84				
7900	36.37	19.89	20.65	21.72	22.93	22.97		
7950	35.67			21.67				
8000	37.85	19.55	20.48	21.48	22.76	22.76	22.91	23.67
8050	40.02			21.32				
8100	42.17			21.15	22.35	22.66		
8150	44.31							
8200	46.44	19.53	20.04	21.08	22.13	22.33	22.59	23.19
8250	48.55							
8300				20.68	22.06	22.02		
8400	54.81	19.54		20.60	21.69	21.81	22.05	22.92
8500						21.72		
8600	62.97	19.38		20.23	21.36	21.45		
8700						21.14		
8800	70.95			19.93	20.85	20.97	21.31	22.07
8900						20.90		
9000	78.75			19.78	20.61	20.63		
9100						20.37		
9200	86.37			19.74	20.25	20.21	20.64	21.46
9400	93.84			19.91	19.93	19.86		
9600	101.14			19.93	19.86	19.52	19.90	20.85
9800	108.29			19.94	19.54	19.15		
10000	115.30			20.38	19.53	18.89	19.19	20.13
10200					19.35			
10400					19.41			

Raw DAX implied volatility data from June 13, 2000, traded at the EUREX, Germany. Time-to-maturity measured in calendar days.

Figures

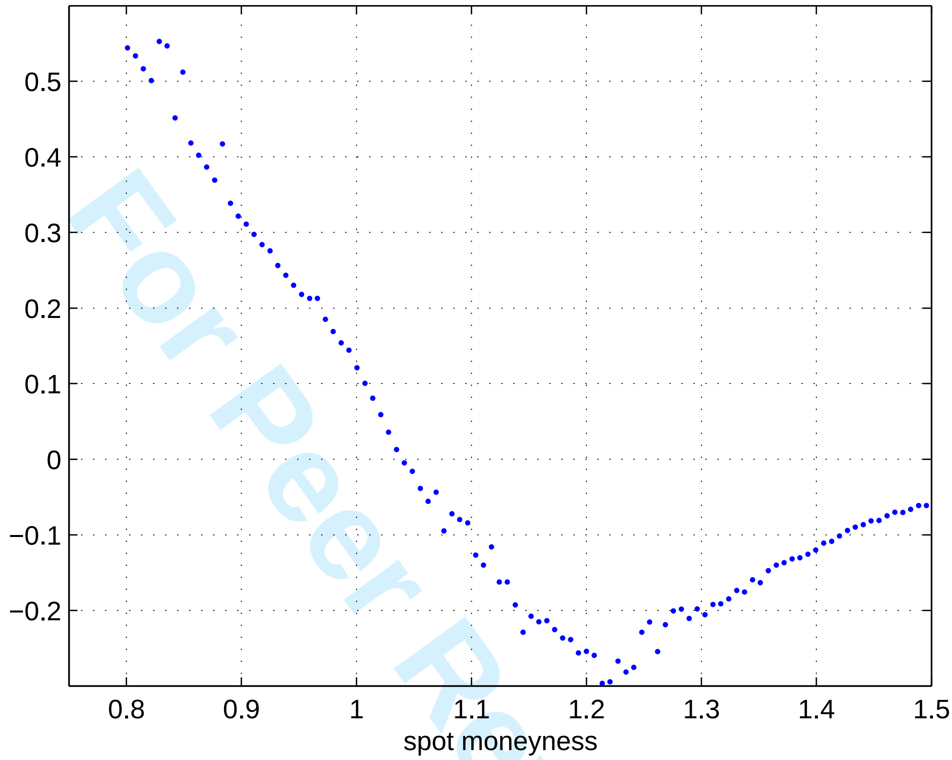

Figure 1: *Delta of a one-year down-and-out put calculated from arbitrage-contaminated IVS of DAX settlement data from June 13, 2000. Strike is at 120% and barrier at 80% of the DAX spot price at 7268.91. Pricing follows Andersen and Brotherton-Ratcliffe (1997) which is an implicit finite difference solver; delta is read from the grid.*

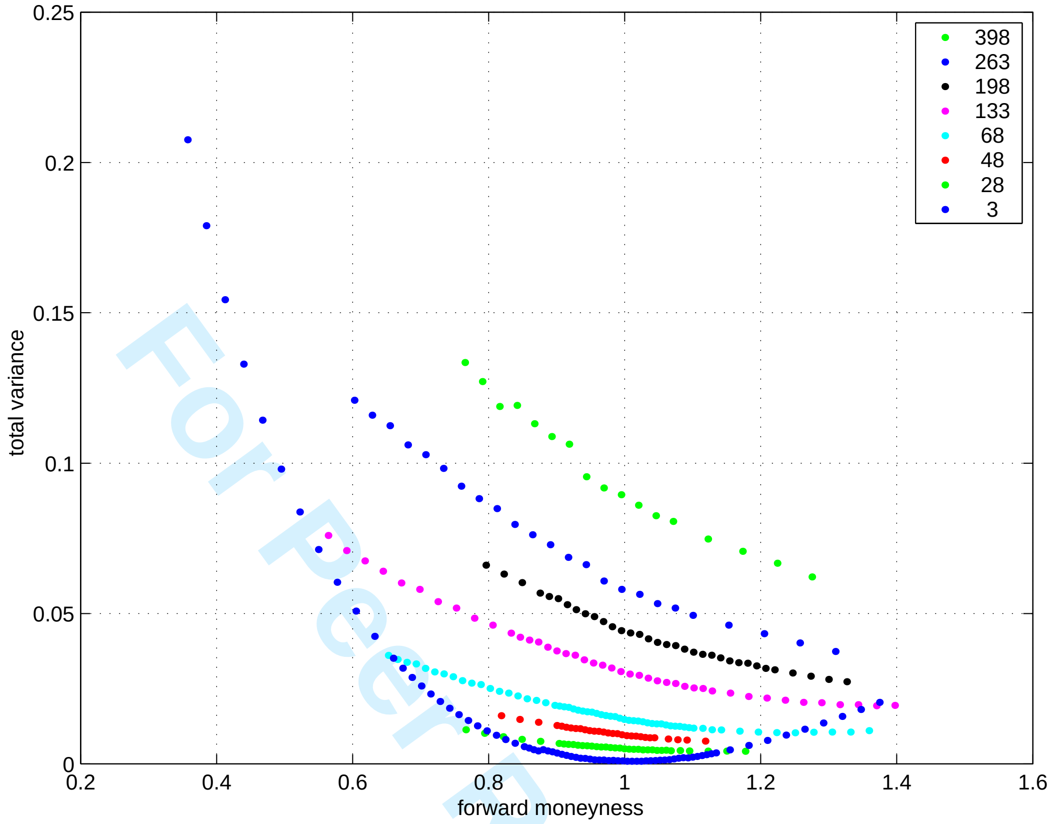

Figure 2: *Total variance plot for DAX data, June 13, 2000, see Table 1 and Appendix B for the data. Total variance is defined by* $\nu^2(\kappa,\tau) = \widehat{\sigma}^2(\kappa,\tau)\tau$. *Time-to-maturity given in calendar days; top graph corresponds to top legend entry, second graph to the second one, etc.*

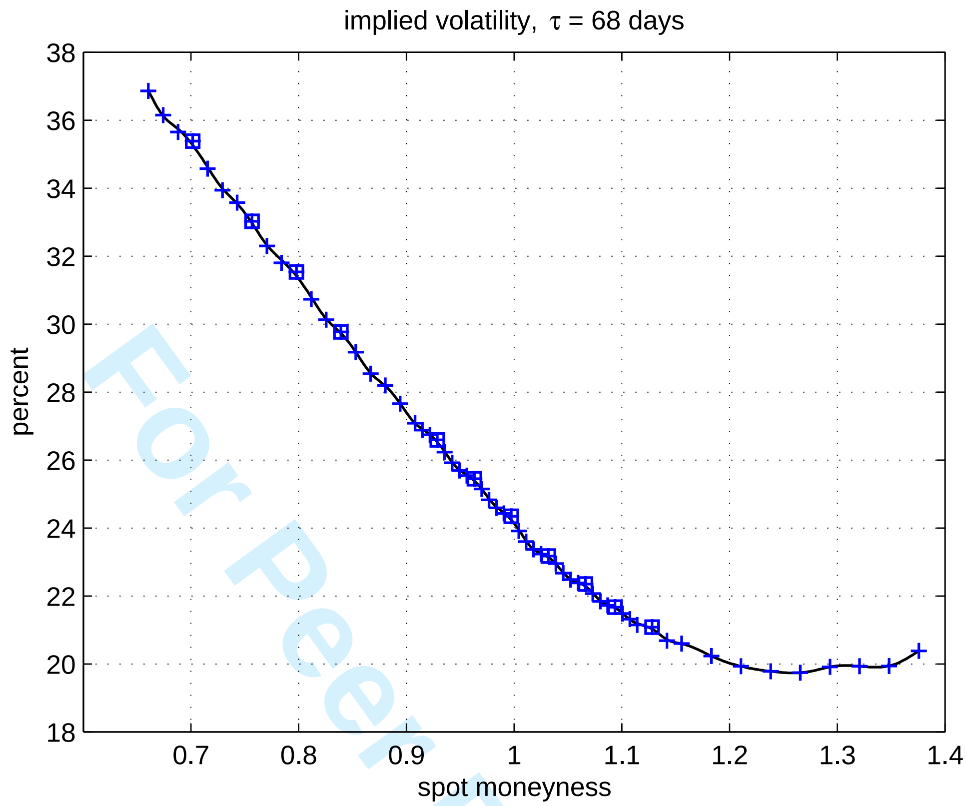

Figure 3: *Arbitrage-free implied volatility smile for a time-to-maturity of 68 days. Estimated function is shown as straight line, original observations are denoted by crosses. Observations violating strike arbitrage and belonging to the center strike price in Eq. (20) are identified by additional squares.*

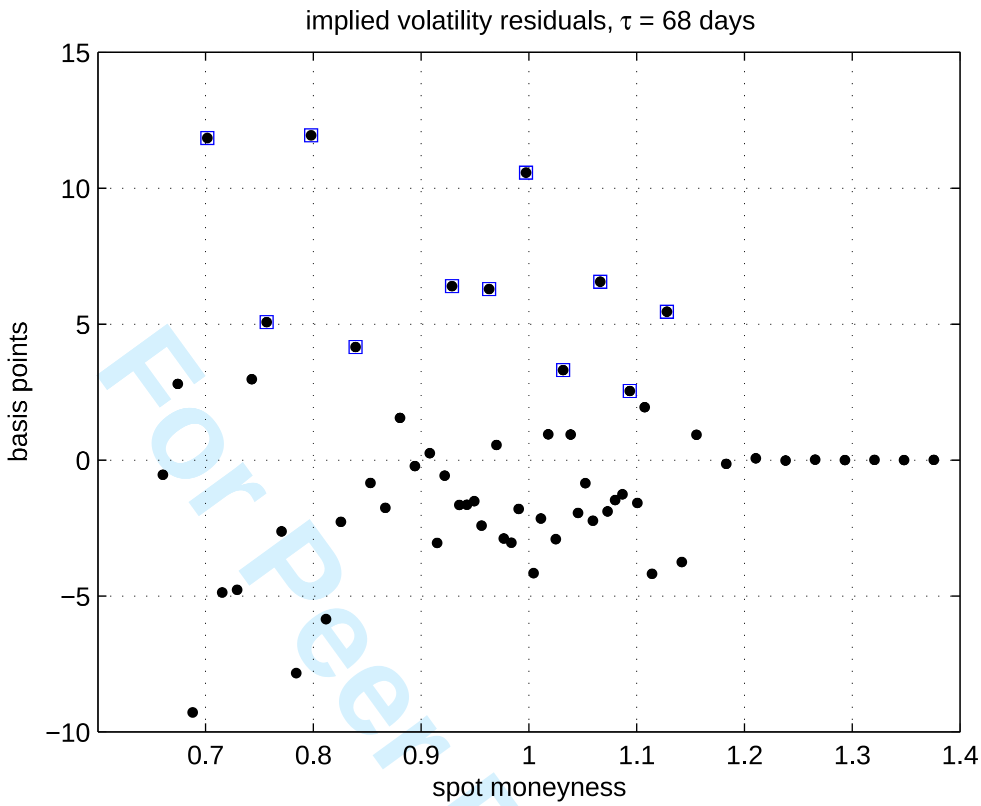

Figure 4: *Implied volatility residuals for the time-to-maturity of 68 days computed as* $\widehat{\sigma}_i - \widehat{\widehat{\sigma}}_i$*, where* $\widehat{\widehat{\sigma}}_i$ *denotes the estimator for the arbitrage-free implied volatility. Residuals belonging to observations that previously violated strike arbitrage according to Eq. (20) are identified by additional squares.*

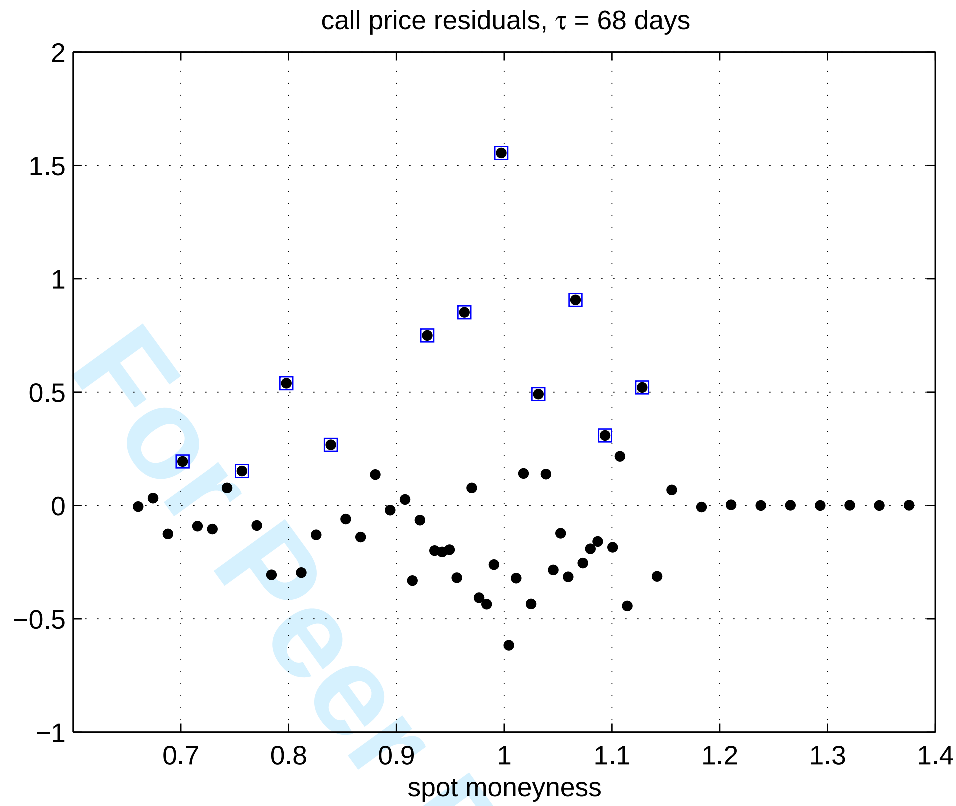

Figure 5: *Call price residuals for the time-to-maturity of 68 days computed as* $g_i - \widehat{g}_i$*, where* $\widehat{g}_i$ *denotes the value of the estimated spline. Residuals belonging to observations that previously violated strike arbitrage according to Eq. (20) are identified by additional squares.*

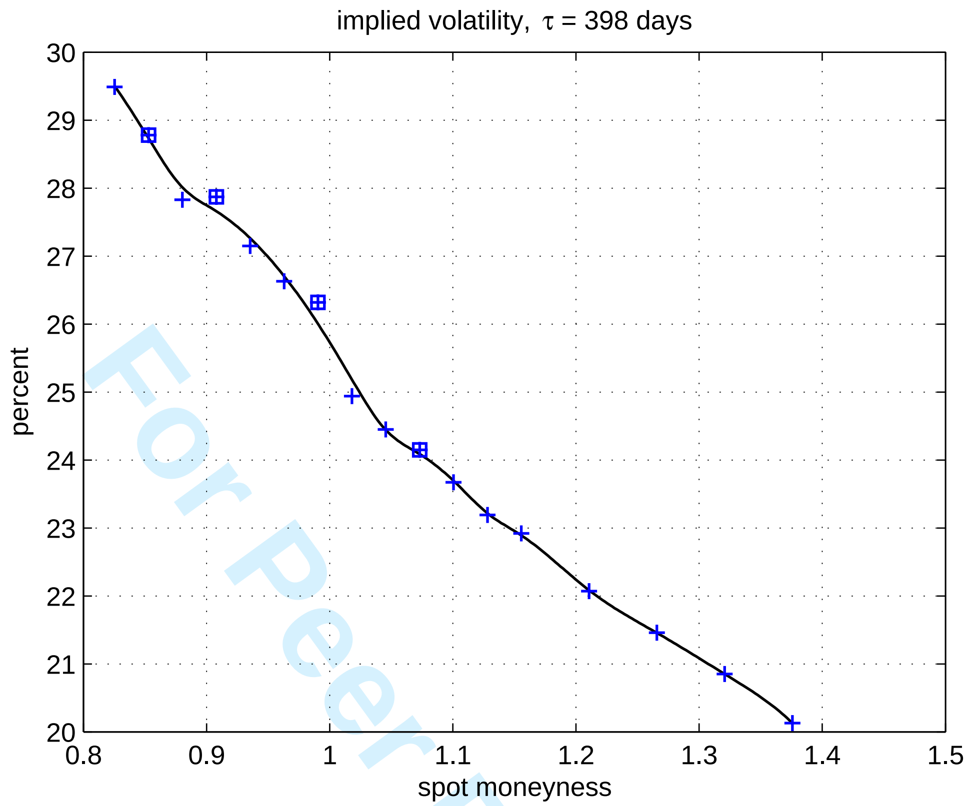

Figure 6: *Arbitrage-free implied volatility smile for a time-to-maturity of 398 days. Estimated function is shown as straight line, original observations are denoted by crosses. Observations violating strike arbitrage and belonging to the center strike price in Eq. (20) are identified by additional squares.*

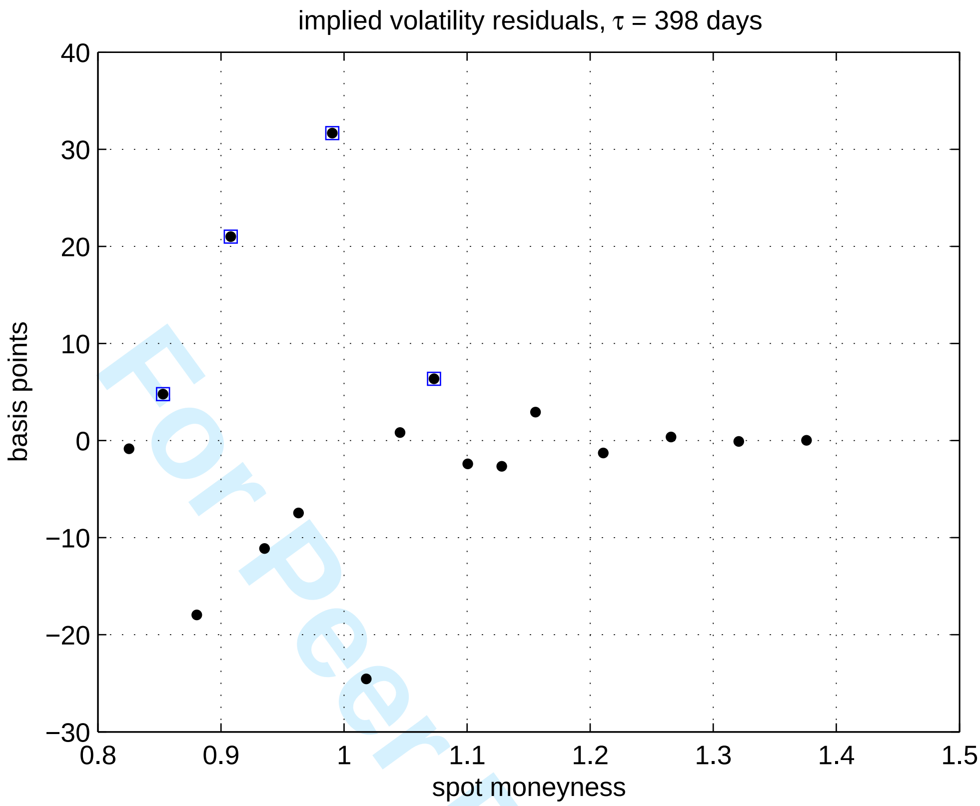

Figure 7: *Implied volatility residuals for the time-to-maturity of 398 days computed as* $\widehat{\sigma}_i - \widehat{\widehat{\sigma}}_i$, *where* $\widehat{\widehat{\sigma}}_i$ *denotes the estimator for the arbitrage-free implied volatility. Residuals belonging to observations that previously violated strike arbitrage according to Eq. (20) are identified by additional squares.*

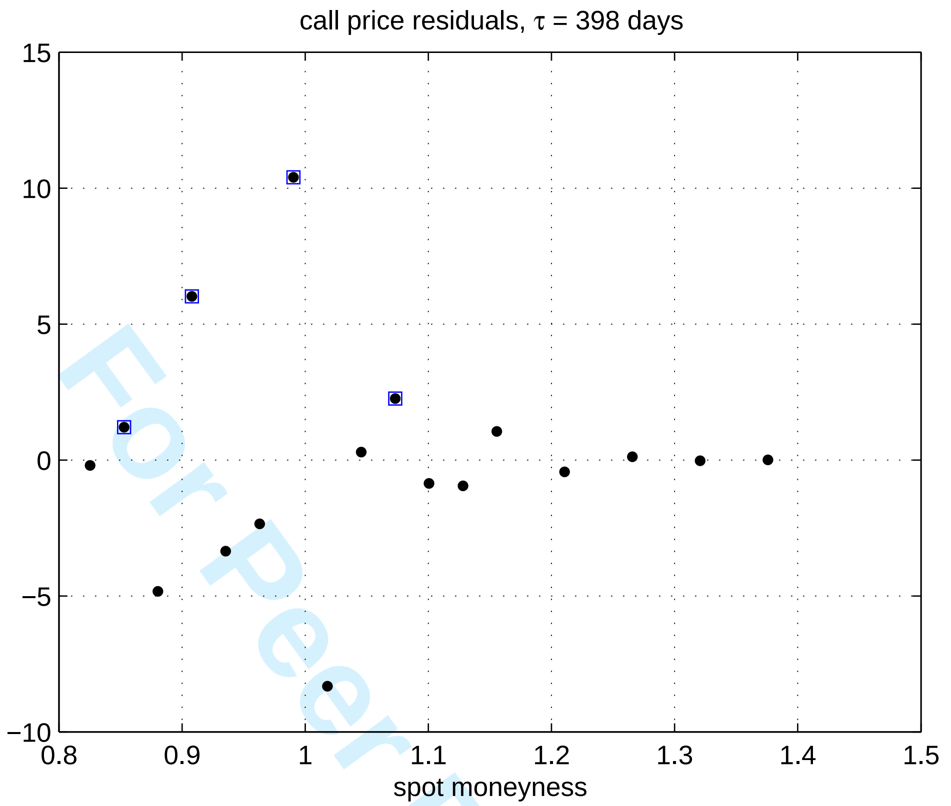

Figure 8: *Call price residuals for the time-to-maturity of 398 days computed as* $g_i - \widehat{g}_i$*, where* $\widehat{g}_i$ *denotes the value of the estimated spline. Residuals belonging to observations that previously violated strike arbitrage according to Eq. (20) are identified by additional squares.*

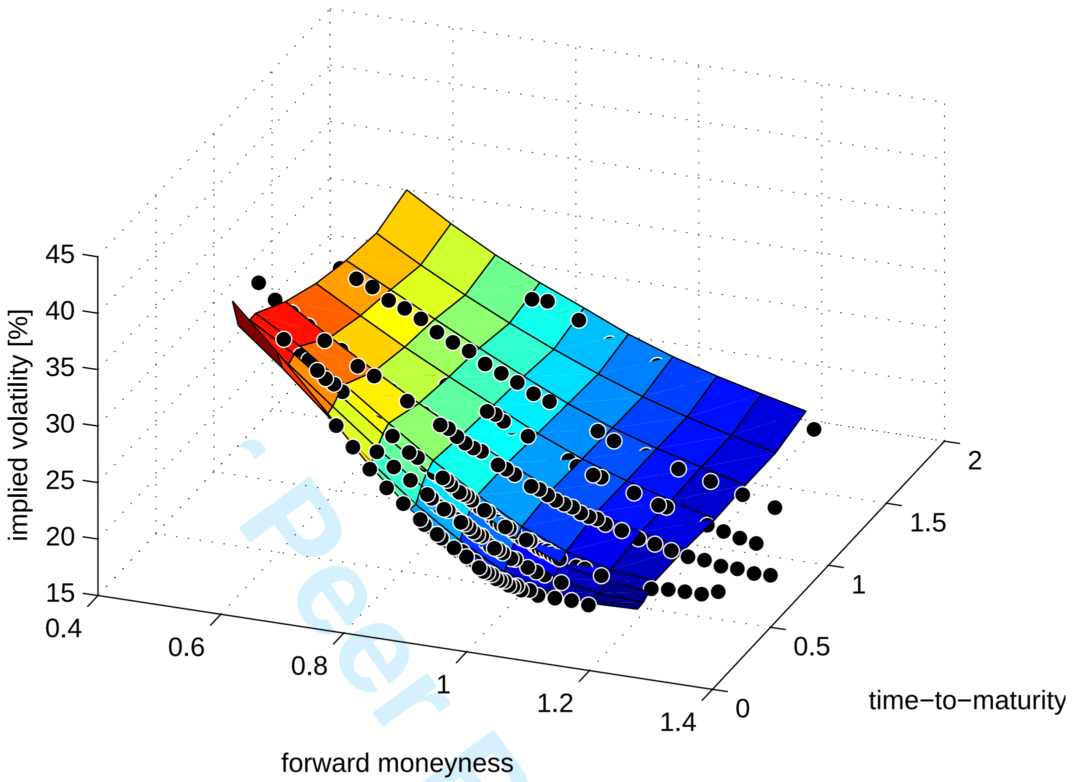

Figure 9: *Estimated arbitrage-free IVS using the constrained cubic spline applied to an initial estimate coming from a thin plate spline; DAX settlement data, June 13, 2000.*

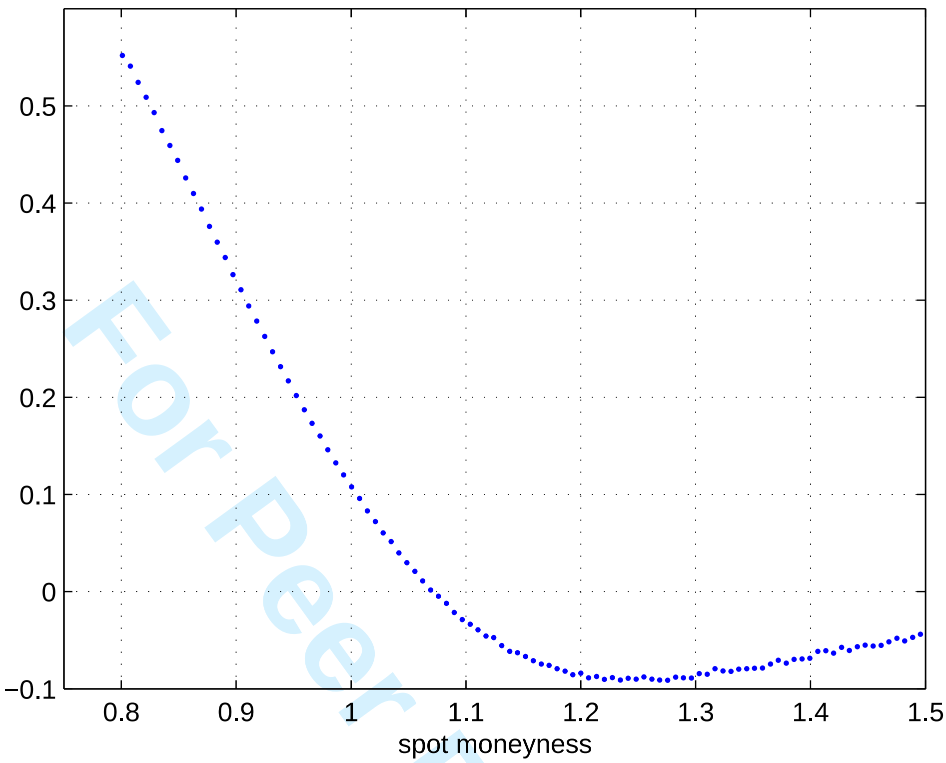

Figure 10: *Delta of a one-year down-and-out barrier put calculated from the arbitrage-free IVS of DAX settlement data from June 13, 2000. Strike is at 120% and barrier at 80% of the DAX spot at 7268.91. Pricing follows Andersen and Brotherton-Ratcliffe (1997); delta is read from the grid.*

Tables

Time-to-maturity	3	28	48	68	133	198	263	398
Interest rate	4.36%	4.47%	4.53%	4.57%	4.71%	4.85%	4.93%	5.04%

Table 1: *Data of DAX index settlement prices, June 13, 2000. Time-to-maturity is given in calendar days; the dividend yield is assumed to be zero, since the DAX index is a performance index, see Deutsche Börse (2006) for a precise description. DAX spot price is* $S_t = 7268.91$.

Time-to-mat.	3	28	48	68	133	198	263	398
NCS, RMSE (prices)	1.0749	0.3542	0.3552	0.3648	0.5623	0.6617	1.4107	3.9284
NCS, RMSE (vol)	0.0496	0.0004	0.0003	0.0004	0.0004	0.0003	0.0006	0.0012
Heston, RMSE (prices)	4.1435	0.7916	0.9091	1.0389	1.7211	2.5468	3.8228	8.1583
Heston, RMSE (vol)	0.4684	0.0040	0.0008	0.0031	0.0034	0.0017	0.0019	0.0026
κ	1.0066	9.0638	7.6691	1.2101	0.4868	1.0146	0.1302	0.1254
θ	9.2017	0.0682	0.0636	0.4323	0.5882	0.2278	1.1336	0.8087
σ	4.3041	0.8755	0.8902	0.8300	0.6445	0.6286	0.5432	0.4503
ρ	-0.3019	-0.4399	-0.5243	-0.6084	-0.6592	-0.6892	-0.7037	-0.7327
V_0	0.0001	0.0370	0.0492	0.0011	0.0010	0.0001	0.0003	0.0014
Simulation from estimated Heston								
NCS, RMSE (prices)	0.0275	0.5873	1.1637	0.6514	1.5852	2.3430	0.8430	1.4461
NCS, RMSE (vol)	0.0039	0.0031	0.0025	0.0009	0.0013	0.0014	0.0004	0.0005
NCS, RMSE* (prices)	0.5311	1.6894	2.1193	1.5639	2.2984	2.8380	1.5578	2.0008
NCS, RMSE* (vol)	0.0055	0.0045	0.0039	0.0020	0.0019	0.0017	0.0008	0.0007
Heston, RMSE (prices)	0.4668	1.6949	2.4265	1.7513	2.9378	3.9392	1.8838	2.6694
Heston, RMSE (vol)	0.0067	0.0079	0.0066	0.0026	0.0025	0.0023	0.0009	0.0010
Heston, RMSE* (prices)	0.3348	0.9430	1.1195	0.6452	0.8520	0.8915	0.4786	0.5445
Heston, RMSE* (vol)	0.0044	0.0058	0.0046	0.0014	0.0009	0.0006	0.0003	0.0002
Simulation from observed market data								
NCS, RMSE (prices)	1.3195	2.5556	3.7661	1.9850	2.4999	3.7458	1.8047	4.3230
NCS, RMSE (vol)	0.0395	0.0034	0.0034	0.0018	0.0016	0.0017	0.0008	0.0014
NCS, RMSE* (prices)	1.3510	2.0523	2.7571	1.5376	2.3321	3.2912	2.2842	4.6488
NCS, RMSE* (vol)	0.0393	0.0035	0.0028	0.0017	0.0018	0.0017	0.0010	0.0015
Heston, RMSE (prices)	4.1974	3.6398	4.7761	3.4267	5.0641	5.6090	4.4932	8.8147
Heston, RMSE (vol)	0.4794	0.0107	0.0046	0.0090	0.0081	0.0029	0.0020	0.0028
Heston, RMSE* (prices)	4.1438	1.7627	1.6052	2.2388	3.7631	2.9330	3.9613	8.7661
Heston, RMSE* (vol)	0.4793	0.0093	0.0017	0.0084	0.0076	0.0019	0.0018	0.0028
stdev. of simul. errors [bp]	50	50	50	25	25	25	10	10

Table 2: *RMSE (root mean square error) for prices and implied volatilities for the natural cubic spline (NCS) and the Heston model computed from unweighted observations. Number of simulations is 100. RMSE* is the error between the true (or the market) model and the perturbed one. Last line gives the standard deviation of the errors added to implied volatility during the simulations.*