Arbitrage-free estimation of the risk-neutral density from the implied volatility smile

Bernhard Brunner
Department of Finance and Banking, University of Augsburg, Universitätsstraße 2, D-86135 Augsburg, Germany

Reinhold Hafner
risklab germany GmbH, Nymphenburger Str. 112–116, D-80636 Munich, Germany

All methods for estimating the risk-neutral density from the volatility smile boil down to the completion of the implied volatility function by interpolating between available strike prices and extrapolating outside their range. In this paper we focus on the extrapolation and develop a new method, which is, under weak constraints, consistent with the absence of arbitrage. The method does not depend on a particular interpolation scheme and is therefore universally applicable. The implementation involves only straightforward numerical procedures. In an empirical study we apply the method to options on the German stock index DAX. The method turns out to be robust, accurate, and fast. Compared with the methods of Shimko (1993) and Bliss and Panigirtzoglou (2002), it tends to be superior.

1 Introduction

Implied risk-neutral density functions (RNDs) derived from cross-sections of observed standard option prices have gained considerable attention during the last few years. They have found various applications in finance: in the field of option pricing, the implied RND allows to price illiquid exotic derivatives in a consistent way. Central banks, among others, use the RND to assess the market participants' expectations about underlying asset prices in the future.[1] AitSahlia and Lo (2000) and Jackwerth (2000) compare the RND with the objective density to retrieve the investors' risk preferences. Bates (1996) uses the RND to estimate the parameters of the underlying stochastic process which generates this RND.

Provided that a continuum of European call options with the same time to maturity and strike prices ranging from zero to infinity exists on a single underlying asset, we can apply the fundamental result of Breeden and Litzenberger (1978) to fully recover the RND in an easy and unique way.[2] They have shown

[1] See, eg, Bahra (1997), Cooper and Talbot (1999), Levin and McManus (1998), and McManus (1999).

[2] Since Breeden and Litzenberger (1978) build their work upon the state-preference theory of Arrow (1964) and Debreu (1959), they speak of a "state-price density". We refrain from using this expression for two reasons: First, it is ambiguous. The term state-price density is also referred to the Radon–Nikodym derivative dQ/dP, where Q is the risk-neutral and P the objective probability measure (see, eg, Pliska, 2000, p. 28). Second, the term RND is more popular in the recent literature.

that the discounted RND is equal to the second derivative of the European call price function with respect to the strike price. Yet, in practice, option contracts are only available for a discrete set of strike prices within a relatively small range around the at-the-money (ATM) strike price. Therefore, all of the methods for estimating RND functions boil down to the *completion of the call price function* by *interpolating between* available strike prices and *extrapolating outside* their range.

According to the survey paper of Jackwerth (1999), estimation techniques for RNDs might be roughly classified into three approaches: the option price function approach, the volatility smile approach and the RND approximating function approach.[3] Methods belonging to the *option price function approach* fit a function of option prices across strike prices through the observed option prices. Then, the Breeden and Litzenberger (1978) result is used to recover the RND. Due to practical problems, the approach was hardly ever attempted.[4] Referring to the *RND approximating function approach*, observed option prices are fit to the theoretical option prices derived for a prespecified RND. Thereby, the RND is specified either directly or indirectly. The direct method explicitly presumes a certain functional form of the RND as, eg, a mixture of lognormal distributions[5] or a GB2 density.[6] The indirect method, on the other hand, specifies the stochastic process driving the underlying asset price. Prominent examples are stochastic volatility processes[7] and jump-diffusion processes.[8] The *volatility smile approach*, originally introduced by Shimko (1993), fits a function through observed implied volatilities. Then, this implied volatility function is translated into an option price function from which the RND is obtained as in the first approach. The volatility smile approach overcomes many of the numerical difficulties emerging in connection with the direct estimation of the call price function. The main reason is that the translation of option prices into implied volatilities eliminates a substantial amount of non-linearity.[9] Moreover, implied volatilities tend to be more smooth than option prices, as Shimko (1993) points out. Comparing the volatility smile approach and the RND approximating function approach, there is some empirical evidence that the former outperforms the latter.[10]

---

[3] The volatility smile approach is sometimes called smoothed implied volatility smile approach. For an in-depth discussion of RND estimation methods see Bahra (1997), Cont (1997) and Jackwerth (1999).
[4] Two exceptions are Bates (1991) and AitSahlia and Lo (1998). Bates (1991) interpolates the set of observed call prices by directly fitting a cubic splice to the observed data. AitSahlia and Lo (1998) apply a kernel estimation technique to the observed option prices.
[5] See, eg, Melick and Thomas (1997) and Bahra (1996).
[6] See, eg, Bookstaber and McDonald (1987).
[7] See, eg, Hull and White (1987), Heston (1993) and Stein and Stein (1991).
[8] See, eg, Bates (1996) and Trautmann and Beinert (1999).
[9] See Rosenberg (2000), pp. 51–2.
[10] Bliss and Panigirtzoglou (2002), using short sterling futures options and FTSE index options, find that smoothing volatility smile methods dominate the mixture of two lognormal densities method in terms of stability. Cooper (1999) comes to the same conclusion when running Monte-Carlo simulations.

In following the volatility smile approach, two issues have to be considered. First, the implied volatility function is unknown beyond the range of traded strike prices. In order to obtain a well-defined RND, the implied volatility function has to be extrapolated or, equivalently, the tails of the RND have to be modeled. Second, it has to be ensured that the constructed implied volatility function does not allow for arbitrage. This requires that the corresponding RND is nonnegative, integrates to one and discounted asset prices have to be martingales with respect to this RND. The last condition is known as the *martingale restriction*.

These two issues have been approached in a number of ways. Shimko (1993) presumes that implied volatility is a quadratic function of the strike price within the range of traded strike prices and constant outside this range. This is equivalent to the assumption of lognormally distributed tails. By construction, the method cannot ensure asset prices to be martingales. The same criticism applies to the method of Campa *et al* (1997), since it also emanates from a constant volatility outside the range of traded strikes. Brown and Toft (1999) estimate a high-order polynomial functional form to fit the implied volatility smile. To avoid arbitrage, the function is required to converge to some exogenously given implied volatilities for some implicitly determined lower and upper strike price. The method has the disadvantage that the shape of the implied volatility function depends on the limiting values for the implied volatilities. If they are misspecified, the fit to the observable implied volatilities becomes worse. Malz (1997) developed a method for estimating the volatility smile of currency options given only three data points containing information on the level, slope and convexity of the smile. He assumes that implied volatility is a quadratic function, but with respect to the option's delta. In contrast to the strike space, the delta space is bounded and more weight is given to observations near the ATM point. The advantage of the interpolation in delta-space is that the corresponding RND integrates to one and the martingale restriction is satisfied. While the method is considered to be the best that can be done if only three observations are available, Cincibuch (2001) shows that in more general cases the method significantly underestimates the tails of the RND. Bliss and Panigirtzoglou (2002) combined the innovations of Malz (1997) and Campa *et al* (1998). They use a smoothing natural spline method to fit the implied volatilities as a function of the options' deltas. One difficulty with this approach is the choice of the smoothing parameter. As a general conclusion, none of the methods discussed above can guarantee nonnegative probabilities.

In this paper we propose a new extrapolation method for volatility smiles, which is, under weak constraints, consistent with the absence of arbitrage. The method does not depend on a particular interpolation scheme. This makes it very flexible and universally applicable. The implementation is straightforward.

The rest of the paper is organized as follows. Section 2 reviews the relation between option prices, implied volatilities, and the RND. Additionally, we state no-arbitrage constraints and explain our estimation strategy. In Section 3 we present our extrapolation method for volatility smiles in detail. Within an empirical analysis in Section 4 we apply the method to options on the German stock index DAX. To assess the quality of our method, we compare the results with the results obtained from the Shimko (1993) and Bliss and Panigirtzoglou (2002) method. Finally, we use our method to price digital options in Section 5. A summary and concluding remarks are found in Section 6.

2 Fundamentals of the volatility smile approach

2.1 The relationship between option prices, implied volatilities and RNDs

Let S_t denote the price of a non-dividend paying stock at time t. Consider a general European-style contingent claim with payoff function g: $\mathbb{R} \to \mathbb{R}$ at maturity T. Under the assumption of no-arbitrage and frictionless markets, Ross (1976) and Cox and Ross (1976) have shown, that the time t claim price Π_t can be obtained using risk-neutral valuation. In this approach, the option price is given as the expected value of its future payoff with respect to the *risk-neutral measure* Q discounted back to the present time t. Formally,

$$\Pi_t = \mathrm{e}^{-r(T-t)} \mathbb{E}^Q\left[g(S_T) \middle| \mathcal{F}_t\right] = \mathrm{e}^{-r(T-t)} \int_0^\infty g(x) q_{S_T}(x) \mathrm{d}x \tag{1}$$

where $q_{t,T,S_T}(x)$, briefly $q_{S_T}(x)$, denotes the RND of S_T conditional on the information $\mathcal{F}_t$ at time t, and $\mathbb{E}^Q[\,\cdot \mid \mathcal{F}_t]$ is the conditional expectation operator with respect to the risk-neutral measure. The risk free interest rate r is supposed to be constant. Setting $g(S_T) = \max\{S_T - K; 0\}$ yields the price of a standard European call option with strike price K and maturity T:

$$C_t(K,T) = \mathrm{e}^{-r(T-t)} \int_0^\infty \max\{x - K; 0\} q_{S_T}(x) \mathrm{d}x \tag{2}$$

In the Black–Scholes model the RND $q_{S_T}(x)$ is assumed to be lognormal with mean $(r - \upsilon^2/2)(T - t)$ and variance $\upsilon^2(T - t)$. Evaluating (2) using this density gives the well-known Black–Scholes formula:

$$C_t^{BS}(K,T) = S_t N(d_1) - K\mathrm{e}^{-r(T-t)} N(d_2) \tag{3}$$

where

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\upsilon^2\right)(T-t)}{\upsilon\sqrt{T-t}}, \quad d_2 = d_1 - \upsilon\sqrt{T-t} \tag{4}$$

Here, $N(\cdot)$ denotes the distribution function of a standard normal random variable.

Given the market prices $C_t(K, T)$ of standard European call options with maturity T and strike prices $K \geq 0$, Breeden and Litzenberger (1978) have shown that the discounted *RND* is equal to the second derivative of the call price function with respect to the strike price K:

$$q_{S_T}(x) = \mathrm{e}^{r(T-t)} \left. \frac{\partial^2 C_t(K,T)}{\partial K^2} \right|_{K=x} \tag{5}$$

Next, we define the implied Black–Scholes volatility or simply implied volatility $\sigma_t(K, T)$ at time t of a standard European call option with strike price K and maturity T as the volatility parameter that, when put into the Black–Scholes formula (3), results in a model price equal to the market price $C_t(K, T)$. Formally,

$$C_t(K,T) = C_t^{BS}\big(K,T,\sigma_t(K,T)\big) \tag{6}$$

If the put–call parity holds, implied call and put volatilities are identical. The function $\sigma_t(K, T)$, which provides the implied volatilities for options of all possible strike prices $K \geq 0$ and maturities $T > t$ is called the v*olatility surface*.[11] For any fixed maturity, the strike price structure is called the "volatility smile" or just "smile". In this paper, it also covers a "skew", characterized by monotonically decreasing implied volatilities when the strike price rises relative to the stock price. On the other hand, for any fixed strike price, the implied volatility pattern is called the *volatility term structure*.

For any fixed maturity T, $T > t$, the relation between the RND and the volatility smile is obtained by successive application of (5) and (6)

$$\begin{aligned} q_{S_T}(x) &= \mathrm{e}^{r(T-t)} \left. \frac{\partial^2 C_t(K,T)}{\partial K^2} \right|_{K=x} \\ &= \mathrm{e}^{r(T-t)} \left. \frac{\partial^2 C_t^{BS}\big(K,T,\sigma_t(K,T)\big)}{\partial K^2} \right|_{K=x} \end{aligned} \tag{7}$$

After applying the chain rule for derivatives, we get:

$$\begin{aligned} q_{S_T}(x) = n\big(d_2(x)\big) \Bigg[& \frac{1}{x\sigma_t(x,T)\sqrt{T-t}} + \frac{2d_1(x)}{\sigma_t(x,T)} \left. \frac{\partial \sigma_t(K,T)}{\partial K} \right|_{K=x} \\ & + \frac{x d_1(x) d_2(x) \sqrt{T-t}}{\sigma_t(x,T)} \left(\left. \frac{\partial \sigma_t(K,T)}{\partial K} \right|_{K=x} \right)^2 \\ & + x\sqrt{T-t} \left. \frac{\partial^2 \sigma_t(K,T)}{\partial K^2} \right|_{K=x} \Bigg] \end{aligned} \tag{8}$$

[11] For ease of notation, we use the symbol $\sigma_t(K, T)$ for both, the implied volatility of an option with fixed strike K and maturity T and for implied volatility as a function of strike price and maturity.

for $x \geq 0$, where

$$d_1(x) = \frac{\ln\left(\frac{S_t}{x}\right) + \left(r + \frac{1}{2}\left(\sigma_t(x,T)\right)^2\right)(T-t)}{\sigma_t(x,T)\sqrt{T-t}}$$

$$d_2(x) = d_1(x) - \sigma_t(x,T)\sqrt{T-t} \tag{9}$$

and $n(x)$ is the standard normal density function. In order for equation (8) to be properly defined, the implied volatility function $\sigma_t(K, T)$ is required to be twice differentiable in K. If $\sigma_t(K, T)$ is given in closed form, so is the risk-neutral density $q_{S_T}(x)$.

From the above it should have become clear that for each maturity T, the following three functions contain essentially the same information: the RND $q_{S_T}(x)$, the call price function $C_t(K, T)$, and the volatility smile $\sigma_t(K, T)$.

2.2 No-arbitrage constraints

Let us consider a frictionless and arbitrage-free market with time horizon T^*, where we can observe at any time $t \in [0, T^*]$ a stock, a complete collection of standard European calls of strike $K \geq 0$ and maturity $T \in (t, T^*]$ and a money market account paying interest at a constant rate r. These assumptions allow us to infer at time $t \in [0, T^*]$ a unique and well-defined RND $q_{S_T}(x)$ for the state variable S_T, $T \in (t, T^*]$, with the following properties:[12]

- *Nonnegativity property:* The RND is nonnegative, ie,:

$$q_{S_T}(x) \geq 0, \quad x \geq 0 \tag{10}$$

- *Integrability property:* The RND integrates to one, ie:

$$\int_0^\infty q_{S_T}(x)\,\mathrm{d}x = 1 \tag{11}$$

- *Martingale property:* The RND reprices all calls:

$$\int_0^\infty \max\{x - K; 0\}\, q_{S_T}(x)\,\mathrm{d}x = \mathrm{e}^{r(T-t)} C_t(K,T), \quad K \geq 0 \tag{12}$$

The nonnegativity and integrability property ensure that the RND is in fact a probability density. The martingale property includes the special case $K = 0$. This implies $\mathbb{E}^{Q}(S_T \mid \mathcal{F}_t) = S_t \mathrm{e}^{r(T-t)} = F_t(T)$, where $F_t(T)$ is the time t price of a forward contract with maturity T. Therefore, the martingale property is sometimes called forward price property or forward property. On the other hand, if a continuous function $q_{S_T}(x)$ has the above properties, it is a well-defined RND and the market is free of arbitrage with respect to maturity T.

[12] See Carr (2001).

Conditions (10), (11) and (12) can also be formulated in terms of call prices and implied volatilities. Concerning call prices, a set of equivalent conditions is:[13]

- The value of a call is never greater than the stock price and never less than its intrinsic value:

$$S_t \geq C_t(K,T) \geq \max\left\{S_t - K\mathrm{e}^{-r(T-t)}; 0\right\}, \quad K \geq 0 \tag{13}$$

- The value of a call takes the value of the stock at a strike price of zero and converges to a value of zero for very large strike prices:

$$C_t(0,T) = S_t, \quad \lim_{K\to\infty} C_t(K,T) = 0 \tag{14}$$

- The value of a vertical call spread is nonpositive or the call price function is monotonically decreasing, respectively. The slope of the call price function is never less than -1:

$$-1 \leq \frac{\partial C_t(K,T)}{\partial K} \leq 0, \quad K \geq 0 \tag{15}$$

- The value of a butterfly spread is nonnegative or the call price function is convex, respectively:

$$\frac{\partial^2 C_t(K,T)}{\partial K^2} \geq 0, \quad K \geq 0 \tag{16}$$

If there exists at time t a RND for S_T with properties (10), (11) and (12) for all maturities $T \in (t, T^*]$, there can still be arbitrage opportunities between options with different maturities. In general, it is not possible to state a sufficient condition which prevents such arbitrage opportunities. However, a necessary condition is: if the market is free of arbitrage, then a calendar call spread with arbitrary maturities $T_1, T_2 \in (t, T^*]$, $T_1 \leq T_2$, has nonnegative value at time t:[14]

$$C_t(K,T_2) \geq C_t\left(\mathrm{e}^{-r(T_2-T_1)}K, T_1\right) \tag{17}$$

and the T_1 and T_2 RNDs obey

$$\int_0^\infty \max\{x-K;0\}\left(\mathrm{e}^{r(T_2-T_1)}q_{S_{T_2}}(x) - q_{S_{T_1}}(x\mathrm{e}^{-r(T_2-T_1)})\right)\mathrm{d}x \geq 0 \tag{18}$$

In the following a volatility surface at time t is called *arbitrage-free* if a RND with properties (10), (11) and (12) exists for all maturities $T \in (t, T^*]$ and condition (18) holds for all $T_1, T_2 \in (t, T^*]$, $T_1 \leq T_2$.

---

[13] See, for example, Carr (2001).

[14] For a formal proof of relations (17) and (18) see the technical appendix.

2.3 Estimation strategy

In the volatility smile approach observed option prices are first converted to implied volatilities by the use of the Black–Scholes formula (3). Then, a function is fit through these implied volatilities with respect to the strike price. Next, this continuous implied volatility function is translated back into a continuous option price function from which finally the RND is obtained by using the Breeden and Litzenberger (1978) theorem.

In general, estimating a RND from an implied volatility smile comprises two basic tasks:

- *Interpolation:* Within the range of available strike prices we have to interpolate between the observable implied volatilities such that the estimated volatility smile best possibly fits the actual smile.
- *Extrapolation:* Beyond the range of available strike prices we have to extrapolate the estimated volatility smile such that the whole implied volatility function or the corresponding RND, respectively, fulfills the no-arbitrage conditions derived in Section (2.2).

Whereas the interpolation depends on the market we are considering, the extrapolation does not. Therefore, volatility smile methods, where the extrapolation scheme impacts the interpolation scheme, as eg, Brown and Toft (1999), are often not flexible enough to accurately capture the pattern of observable implied volatilities. In contrast, we suggest to perform inter- and extrapolation as two independent consecutive steps.

3 An arbitrage-free extrapolation method for the volatility smile

3.1 General method

For a fixed maturity T, we assume that we can observe at current time t, $T > t$, the market forward price $F_t(T)$ and the market option price function $C_t^{\mathcal{M}}(K, T)$ or equivalently the market implied volatility function $\sigma_t^{\mathcal{M}}(K, T)$ within some strike price interval $\mathcal{M} = [K_L, K_U]$. The lower and upper strike price boundaries K_L and K_U thereby depend on t and T. We demand $C_t^{\mathcal{M}}(K, T)$ to be free of arbitrage, ie, $C_t^{\mathcal{M}}(K, T)$ to fulfill conditions (13)–(16) for all $K \in \mathcal{M}$. Applying equation (8) to $\sigma_t^{\mathcal{M}}(K, T)$, the middle part of the corresponding RND $q_{S_T}^{\mathcal{M}}(x)$, $x \in \mathcal{M}$, is obtained.

The basic idea of our method is to complete the RND by attaching a nonnegative function $q_{S_T}^{L}(x; \theta_L)$ with parameter vector θ_L to the lower tail and a nonnegative function $q_{S_T}^{U}(x; \theta_U)$ with parameter vector θ_U to the upper tail. The complete RND is then piecewise defined as

$$q_{S_T}(x;\theta_L,\theta_U) = \begin{cases} q_{S_T}^{L}(x;\theta_L), & x < K_L \\ q_{S_T}^{\mathcal{M}}(x), & K_L \le x \le K_U \\ q_{S_T}^{U}(x;\theta_U), & x > K_U \end{cases} \tag{19}$$

To be continuous, $q_{S_T}(x; \theta_L, \theta_U)$ must satisfy:

$$q_{S_T}^{L}(K_L;\theta_L) = q_{S_T}^{\mathcal{M}}(K_L)$$
$$q_{S_T}^{U}(K_U;\theta_U) = q_{S_T}^{\mathcal{M}}(K_U) \tag{20}$$

The absence of arbitrage requires the RND to meet the nonnegativity, integrability and martingale constraint, stated in section (2.2). While the nonnegativity constraint is fulfilled by construction, the other two conditions must be explicitly demanded. The integrability constraint implies

$$\int_0^{K_L} q_{S_T}^{L}(x;\theta_L)\mathrm{d}x + \int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\mathrm{d}x$$
$$= 1 - \int_{K_L}^{K_U} q_{S_T}^{\mathcal{M}}(x)\mathrm{d}x \tag{21}$$

The information on the probability mass that the market attributes to the lower and upper tail of the risk-neutral distribution is contained in $C_t(K_L, T)$ and $C_t(K_U, T)$, respectively:[15]

$$-\mathrm{e}^{r(T-t)} \left.\frac{\partial C_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U} = \int_{K_U}^{\infty} q_{S_T}(x;\theta_L,\theta_U)\,\mathrm{d}x$$
$$= \int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\,\mathrm{d}x \tag{22}$$

and

$$1+\mathrm{e}^{r(T-t)} \left.\frac{\partial C_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_L} = 1-\int_{K_L}^{\infty} q_{S_T}(x;\theta_L,\theta_U)\,\mathrm{d}x$$
$$= 1-\int_{K_L}^{K_U} q_{S_T}^{\mathcal{M}}(x)\mathrm{d}x - \int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\,\mathrm{d}x \tag{23}$$

Using similar arguments, we get:

[15] See Neuhaus (1995). The derivatives

$$\left.\frac{\partial C_t(K,T)}{\partial K}\right|_{K=K_L} \quad \text{and} \quad \left.\frac{\partial C_t(K,T)}{\partial K}\right|_{K=K_U}$$

have to be interpreted as right-hand and left-hand derivatives.

$$
\begin{aligned}
&\mathrm{e}^{r(T-t)}\left(C_t^{\mathcal{M}}(K_U,T)-K_U\cdot\left.\frac{\partial C_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U}\right)\\
&=\int_{K_U}^{\infty} x q_{S_T}^{U}(x;\theta_U)\,\mathrm{d}x
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
&F_t(T)-\mathrm{e}^{r(T-t)}\left(C_t^{\mathcal{M}}(K_L,T)-K_L\cdot\left.\frac{\partial C_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_L}\right)\\
&=F_t(T)-\int_{K_L}^{\infty} x q_{S_T}(x;\theta_L,\theta_U)\,\mathrm{d}x\\
&=F_t(T)-\int_{K_L}^{K_U} x q_{S_T}^{\mathcal{M}}(x)\,\mathrm{d}x-\int_{K_U}^{\infty} x q_{S_T}^{U}(x;\theta_U)\,\mathrm{d}x
\end{aligned}
\tag{25}
$$

In combination with conditions (22)–(25), the correct repricing of the forward contract

$$
F_t(T)=\int_0^{K_L} x q_{S_T}^{L}(x;\theta_L)\,\mathrm{d}x+\int_{K_L}^{K_U} x q_{S_T}^{\mathcal{M}}(x)\,\mathrm{d}x+\int_{K_U}^{\infty} x q_{S_T}^{U}(x;\theta_U)\,\mathrm{d}x \tag{26}
$$

ensures that discounted call prices of all strikes $K \geq 0$ are martingales.

Evaluating the call price derivatives in (22)–(25), leads, in conjunction with equation (20), to the following system of equations, which has to be solved for the parameter vectors θ_L and θ_U:

$$
\begin{aligned}
&q_{S_T}^{L}(K_L;\theta_L)=q_{S_T}^{\mathcal{M}}(K_L)\\
&q_{S_T}^{U}(K_U;\theta_U)=q_{S_T}^{\mathcal{M}}(K_U)\\
&\int_0^{K_L} q_{S_T}^{L}(x;\theta_L)\,\mathrm{d}x\\
&\quad=N\big(-d_2(K_L)\big)+K_L n\big(d_2(K_L)\big)\sqrt{T-t}\left.\frac{\partial\sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_L}\\
&\int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\,\mathrm{d}x\\
&\quad=N\big(d_2(K_U)\big)-K_U n\big(d_2(K_U)\big)\sqrt{T-t}\left.\frac{\partial\sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U}
\end{aligned}
\tag{27}
$$

$$\int_0^{K_L} x\, q_{S_T}^L(x;\theta_L)\,\mathrm{d}x = F_t(T)N\big(-d_1(K_L)\big) + K_L^2 n\big(d_2(K_L)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_L}$$

$$\int_{K_U}^{\infty} x\, q_{S_T}^U(x;\theta_U)\,\mathrm{d}x = F_t(T)N\big(d_1(K_U)\big) - K_U^2 n\big(d_2(K_U)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U} \tag{27}$$

Provided a solution to the system (27) exists, then the function $q_{S_T}(x;\,\theta_L,\,\theta_U)$ fulfills the no-arbitrage constraints and is indeed a well-defined RND.[16]

3.2 The mixture of two lognormals case

Shimko (1993) assumes a lognormal density for $q_{S_T}^L(x;\,\theta_L)$ and $q_{S_T}^U(x;\,\theta_U)$. This density is completely characterized by two parameters, the mean and the variance. Yet, in general, the total number of four parameters is not sufficient to solve the system (27), and therefore the method is usually not consistent with the absence of arbitrage. By construction, it satisfies the integrability condition and forces the RND to be continuous. The martingale constraint, however, is in general violated. Therefore, we propose a *mixture of two lognormal distributions* (DLN) for each tail:

$$q_{S_T}^i(x;\theta_i) = \lambda_i \ell\big(x;\eta_{i,1},\nu_{i,1}^2\big) + (1-\lambda_i)\ell\big(x;\eta_{i,2},\nu_{i,2}^2\big), \qquad \lambda_i \in [0,1] \tag{28}$$

where the lognormal density function is defined as:

$$\ell\big(x;\eta_{i,j},\nu_{i,j}^2\big) = \frac{1}{x\nu_{i,j}\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(x)-\ln(\eta_{i,j})+\frac{\nu_{i,j}^2}{2}}{\nu_{i,j}}\right)^2\right), \quad j=1,2 \tag{29}$$

and the parameter vector is given by $\theta_i = \big(\lambda_i,\,\eta_{i,1},\,\nu_{i,1}^2,\,\eta_{i,2},\,\nu_{i,2}^2\big)'$ for $i \in \{L, U\}$. The total number of ten parameters provides us with enough flexibility to always solve the system (27).[17]

Our choice of a DLN was partly motivated by the ease of computing option values. The price of an option, where the underlying follows a linear combina-

[16] For a formal proof see the technical appendix.

[17] Strictly speaking, there exists more than one solution if one exists at all. This implies that the derived RND is not unique. Yet, all possible solutions are consistent with market prices. Thus, we consider them to be equivalent, although the particular choice has impact on the higher moments of the RND.

tion of lognormal distributions is simply the weighted average of Black–Scholes prices. Therefore, the part of the option value attributable to the tails of the underlying's distribution can be calculated analytically.

The primary reason for selecting the DLN is that it combines the advantages of volatility smile methods with the advantages of mixture of lognormals methods, the latter belonging to the RND approximating function approach. On the one hand, mixture of lognormals methods may be favored if only a few data is available. This is because volatility smile methods can only reveal the structure provided by the data. In contrast, mixture of lognormals methods impose a predefined structure on the RND. However, the price to be paid is a lack of flexibility in fitting the RND when there are enough observations available. On the other hand, volatility smile methods exactly have this flexibility. Moreover, RND's produced by volatility smile methods tend to be more stable with respect to variations in the option's data than those generated by mixture of lognormal methods. In our method, we estimate the volatility smile within the range of observable data using an appropriate estimation method and apply the mixture of lognormals method where we cannot observe any data. Depending on the information available from the option's market, our method resembles more a volatility smile or a mixture of lognormals method.

To solve the system (27) for the parameter vectors θ_L and θ_U, we first evaluate the integrals on the left-hand side:

$$\begin{aligned}
&\int_0^{K_L} q_{S_T}^L(x;\theta_L)\,\mathrm{d}x = \lambda_L N(-z_{L,1}) + (1-\lambda_L)N(-z_{L,2})\\
&\int_{K_U}^{\infty} q_{S_T}^U(x;\theta_U)\,\mathrm{d}x = \lambda_U N(z_{U,1}) + (1-\lambda_U)N(z_{U,2})\\
&\int_0^{K_L} x\,q_{S_T}^L(x;\theta_L)\,\mathrm{d}x\\
&\quad = \lambda_L \eta_{L,1} N(-z_{L,1}-\nu_{L,1}) + (1-\lambda_L)\eta_{L,2}N(-z_{L,2}-\nu_{L,2})\\
&\int_{K_U}^{\infty} x\,q_{S_T}^U(x;\theta_U)\,\mathrm{d}x\\
&\quad = \lambda_U \eta_{U,1} N(z_{U,1}+\nu_{U,1}) + (1-\lambda_U)\eta_{U,2}N(z_{U,2}+\nu_{U,2})
\end{aligned} \tag{30}$$

where

$$z_{i,j} = \frac{\ln(\eta_{i,j}) - \ln(K_i) - \frac{\nu_{i.j}^2}{2}}{\nu_{i,j}}, \quad i \in \{L,U\}, \quad j = 1,2 \tag{31}$$

Next, we reduce the problem's dimension by setting:

$$z_L = z_{L,1} = z_{L,2} \quad \text{and} \quad z_U = z_{U,1} = z_{U,2} \tag{32}$$

Combining (27), (30) and (32) yields the system

$$\begin{aligned}
&\lambda_i\left(\frac{1}{K_i\nu_{i,1}\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}z_i^2}\right)+(1-\lambda_i)\left(\frac{1}{K_i\nu_{i,2}\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}z_i^2}\right)=q_{S_T}^{\mathcal{M}}(K_i),\\
&\quad i\in\{L,U\}\\
&N(z_i)=N(d_2(K_i))-K_i n(d_2(K_i))\sqrt{T-t}\left.\frac{\partial\sigma_t(K,T)}{\partial K}\right|_{K=K_i},\\
&\quad i\in\{L,U\}\\
&\lambda_L\eta_{L,1}N(-z_L-\nu_{L,1})+(1-\lambda_L)\eta_{L,2}N(-z_L-\nu_{L,2})\\
&=F_t(T)N(-d_1(K_L))+K_L^2 n(d_2(K_L))\sqrt{T-t}\left.\frac{\partial\sigma_t(K,T)}{\partial K}\right|_{K=K_L}\\
&\lambda_U\eta_{U,1}N(z_U+\nu_{U,1})+(1-\lambda_U)\eta_{U,2}N(z_U+\nu_{U,2})\\
&=F_t(T)N(d_1(K_U))-K_U^2 n(d_2(K_U))\sqrt{T-t}\left.\frac{\partial\sigma_t(K,T)}{\partial K}\right|_{K=K_U}
\end{aligned}\tag{33}$$

for $\lambda_i \in [0, 1]$, $i \in \{L, U\}$. The first two equations of (33) can be rearranged to

$$\lambda_i=\frac{q_{S_T}^{\mathcal{M}}(K_i)-\dfrac{1}{K_i\nu_{i,2}\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}z_i^2}}{\dfrac{1}{K_i\nu_{i,1}\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}z_i^2}-\dfrac{1}{K_i\nu_{i,2}\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}z_i^2}}$$

$$\eta_{i,1}=K_i\mathrm{e}^{z_i\nu_{i,1}+\frac{\nu_{i,1}^2}{2}} \qquad \eta_{i,2}=K_i\mathrm{e}^{z_i\nu_{i,2}+\frac{\nu_{i,2}^2}{2}} \tag{34}$$

for $i \in \{L, U\}$.

The remaining two free parameters can be determined in various ways. For example, in some applications we may wish the RND to be differentiable. This additional restriction can be easily incorporated by requiring that the derivatives of $q_{S_T}^L(x;\theta_L)$ and $q_{S_T}^{\mathcal{M}}(x)$ as well as $q_{S_T}^U(x;\theta_U)$ and $q_{S_T}^{\mathcal{M}}(x)$ match at the strike price boundaries K_L and K_U:

$$\left.\frac{\partial q_{S_T}^L(x;\theta_L)}{\partial x}\right|_{x=K_L}=\left.\frac{\partial q_{S_T}^{\mathcal{M}}(x)}{\partial x}\right|_{x=K_L} \quad\text{and}\quad \left.\frac{\partial q_{S_T}^U(x;\theta_U)}{\partial x}\right|_{x=K_U}=\left.\frac{\partial q_{S_T}^{\mathcal{M}}(x)}{\partial x}\right|_{x=K_U} \tag{35}$$

After some transformations, we see that:

$$\nu_{i,1} = \frac{q_{S_T}^{\mathcal{M}}(K_i) - \dfrac{1}{K_i \nu_{i,2}\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z_i^2}}{\left(q_{S_T}^{\mathcal{M}}(K_i) + K_i \left.\dfrac{\partial q_{S_T}^{\mathcal{M}}(x)}{\partial x}\right|_{x=K_i} \right) \dfrac{1}{z_i} - \dfrac{q_{S_T}^{\mathcal{M}}(K_i)}{\nu_{i,2}}}, \qquad i \in \{L, U\} \tag{36}$$

A very simple alternative is to set the two free parameters to reasonable values. From a computational perspective, this turns out to be numerically more stable than the first alternative.

In the end, the system can be reduced to a nonlinear equation, which can be solved numerically by standard one-dimensional root-finding methods.[18]

4 Empirical study

4.1 Data

Our database contains all reported transactions of options and futures on the German stock index, DAX, traded on the DTB/Eurex in the year 2000.[19] The underlying of the option, the DAX index, comprises the 30 largest and most actively traded German companies. The DAX is a capital-weighted performance index, ie, dividends are reinvested. DAX options are cash settled European-style options which expire on the third Friday of the contract month. At any point in time eight option maturities with lifetimes of up to two years are available: the three nearest calendar months, the three following months of the cycle March–June–September–December and the two following months of the cycle June–December. The futures contract on the DAX index is clearly associated with the option contract, nevertheless, one difference can be noted: the expiry months are only the three nearest months within the cycle March–June–September–December.

Apart from the option price and the strike, three parameters are required to compute the implied volatilities: the time to expiration, the risk-free rate and the level of the underlying index. Let t denote the trading day and T the option's expiration date. The time to expiration $\tau = T - t$ is measured in calendar days.[20]

Daily series of 1, 3, 6, and 12 months EURIBOR rates serve as riskless interest rates r. The τ-period interest rate is obtained by linear interpolation between

[18] For example, the Newton–Raphson method or the bisection method can be applied.

[19] We are grateful to Eurex Deutschland for providing us with these data.

[20] It is uncertain whether volatility is related to trading or calendar days. The difference between calendar and trading days, expressed as a fraction of one year, is small except for very short-term options. These are not considered in this paper. See also Hull (2000). In the calculations the time to maturity is measured as a proportion of 365 days per year.

the available rates enclosing τ. The resulting value is then converted to a continuously compounded rate.[21]

Let n ($n = 1,\ldots, N$) be the trading minute of an option's transaction.[22] The underlying index $S_{t,n}$ on day t at minute n is derived from the current price $F_{t,n}$ of the futures contract most actively traded on that day. The maturity of this contract, which is normally the nearest available, is denoted by T_F. The value $F_{t,n}(T_F)$ corresponds to the *average* transaction price observed in the T_F-futures contract in minute n on day t. To obtain the corresponding index level we solve the theoretical futures pricing model[23]

$$F_{t,n}(T_F) = S_{t,n}\mathrm{e}^{r(T_F - t)} \tag{37}$$

for $S_{t,n}$. If no future is traded at minute n, we exclude all options transactions that took place in this minute from our database. This procedure ensures simultaneous options and underlying prices, ie, their respective time stamps diverge by not more than one minute.

Empirically it is convenient to express the implied volatility of an option in terms of the option's moneyness and time to maturity rather than strike price and maturity. Therefore we define the simple moneyness $\tilde{M}$ of an option expiring in T as

$$\tilde{M} = \tilde{M}_{t,n}(K,T) = \frac{K}{F_{t,n}(T)} \tag{38}$$

where $F_{t,n}(T)$ is the theoretical T-futures price computed with (37) for a given $S_{t,n}$.[24]

At the DAX options market, liquidity is very much concentrated in short-term options and declines exponentially with increasing time to expiration. The call trades distribution across degrees of moneyness is clearly skewed to the left whereas the put trades distribution is skewed to the right. This means that out-of-the money options are traded far more frequently than in-the-money options. Since the estimation of the volatility surface requires a sufficient variety of strike prices, we include both calls and puts in our empirical study.

Put–call-parity requires that the implied call volatilities do not systematically deviate from the implied put volatilities with the same degree of moneyness. However, on a number of trading days we found systematic violations of this arbitrage relationship. After a closer inspection, the problem could be traced back to a biased index level caused by dividend payments. The DAX index calculation

---

[21] The riskless rate r is dependent on day t and the investment horizon T. For ease of exposition, we suppress these indices.

[22] Since trading hours changed through time, N is time-dependent. To keep notation simple, we suppress the time index.

[23] Using the futures-based implied index level rather than the reported index level as the underlying price has also been suggested in a study for the S&P 500 options market by Jackwerth and Rubinstein (1996).

[24] For a motivation to define moneyness in this way, see Natenberg (1994).

rests on the assumption that cash dividends are reinvested after deduction of the corporate income tax for distributed gains from the gross dividend If the marginal investor's tax rate is smaller than the corporate income tax rate, he receives an extra dividend, called difference dividend, which has the same effect as an ordinary dividend in the case of unprotected options and futures. To estimate the difference dividend for each day and each pair of options-futures expiry dates, we apply an implicit method. The method presumes that a difference dividend adjusted version of the put-call parity holds.[25] Equipped with the estimates of the difference dividends, we recalculate the underlying index level and the implied volatilities. An inspection of all scatterplots reveals that this implied estimation of the relevant underlying index level solves the problem of the difference dividend.

In a final step, we eliminate all options that violate the well-known arbitrage bounds or have implied volatilities higher than 150%.

4.2 Estimation method

In this section we propose a specific interpolation method for DAX implied volatilities. However, we want to highlight that any other suitable method can be used, too.

According to Section (2.3) we first *interpolate* between observable implied volatilities using the data of one trading day in any cross sectional analysis.[26] In their paper, Hafner and Wallmeier (2001) proposed to model the volatility smile of DAX options by the quadratic spline function[27]

$$\beta_0 + \beta_1 \tilde{M} + \beta_2 \tilde{M}^2 + \tilde{D}\beta_3\left(1 - 2\tilde{M} + \tilde{M}^2\right) \tag{39}$$

where $\tilde{D}$ is a dummy variable defined as

$$\tilde{D} = \begin{cases} 0, & \tilde{M} \le 1 \\ 1, & \tilde{M} > 1 \end{cases}$$

and β_i, $i = 0, 1, 2, 3$ are time-to-maturity dependent parameters. While the function describes the data very well, it has the disadvantage that it is not twice differentiable at the threshold $\tilde{M} = 1$. Consequently, the resulting RND is not continuous at this point. To overcome this technical problem we choose the modified form

$$\beta_0 + \beta_1 M + \beta_2 M^2 + D\beta_3 M^3 \tag{40}$$

where the dummy variable D is given by

[25] For a detailed description of the method, see Hafner and Wallmeier (2001).
[26] For a motivation to use a time window of one day, see Hafner and Wallmeier (2001), p. 14.
[27] For the ease of notation we suppress the time index t.

$$D = \begin{cases} 0, & M \le 0 \\ 1, & M > 0 \end{cases} \tag{41}$$

and the moneyness measure changes to

$$M = M_{t,n}(K,T) = \frac{\ln\left(\frac{K}{F_{t,n}(T)}\right)}{\sqrt{T-t}} \tag{42}$$

When expressed in this measure the structure of implied volatilities becomes similar for different times to maturity; this property is particularly useful if we wish to estimate the whole volatility surface rather than a single smile.[28] In practice, volatility surfaces are used to value large option portfolios or special types of exotic options.

In the following, we extend our smile function (40) and formulate a regression model for the daily volatility surface of DAX options:

$$\sigma^{\mathcal{M}} = \beta_0 + \beta_1 M + \beta_2 M^2 + D\beta_3 M^3 + \beta_4\sqrt{\tau} + \beta_5 M\sqrt{\tau} + \varepsilon \tag{43}$$

where D is given by (41), the parameter vector is specified as $\beta = (\beta_0, \beta_1, \ldots, \beta_5)'$, and ε is a random disturbance.

The implied volatility of deep in-the-money calls and puts is very sensitive to changes in the index level. Since small errors in determining the appropriate index level are unavoidable, the disturbance variance of regression model (43) is supposed to increase as options go deeper in-the-money. Residual scatterplots support this presumption. Using the White-test, the null hypothesis of homoskedasticity was rejected in about 60% of all regressions. To account for the heteroskedasticity of the disturbances we apply a weighted least squares estimation assuming that the disturbance variance is proportional to the positive ratio of the option's Black–Scholes delta and vega.[29] This ratio indicates how an increase in the index level by one (marginal) point affects the implied volatility of an option, if its price does not change.

In view of the large number of intraday transactions it is not astonishing that some extreme deviations occur representing "off-market" implied volatilities. They can, for example, be due to a faulty and unintentional input by a market participant. In this case, the trade can be annulled if certain conditions are fulfilled. To exclude such unusual events we discard all observations corresponding to large errors of more than four standard deviations of the regression residuals where the standard deviation is computed as the square root of the weighted average squared residuals. We then repeat the estimation on the basis of the reduced sample until no further observations are discarded. This procedure is

[28] See Natenberg (1994).

[29] The Black–Scholes delta and vega are computed using the implied volatility of the corresponding option. The delta of puts is multiplied by -1 to obtain a positive ratio.

known as applying the "4-sigma-rule" or "trimmed regression".[30] We examined the impact of this exclusion of outliers and found it to be negligible in all but very few cases.

A large percentage of all traded DAX options in the year 2000 features a degree of log simple moneyness $\ln(\tilde{M})$ between –0.3 and 0.2 and a time to maturity below 0.5. Therefore, we discard all observations outside this range in order to eliminate potential problems with extreme degrees of moneyness or times to maturity. The smile function for a fixed but arbitrary time to maturity $\tau^* \in (0; 0.5]$ is easily obtained by evaluating (43) for $\tau = \tau^*$. If the observable log simple moneyness range is smaller than $[-0.3, 0.2]$, the function is defined on the maximum subinterval which ensures the smile function to be consistent with the arbitrage constraints stated in Section 2.2. Outside this range, we apply the extrapolation method presented in Section 3. For the calculation of the market forward price $F_t(t + \tau^*)$ we use the average stock price on day t.

4.3 Empirical results

For each trading day $t \in \{1, \ldots, 254\}$ in 2000 we first estimate the *regression model* (43) as described in Section 4.2. Across the 254 days during the sample period the average adjusted R^2 value is 96.34%. This indicates that, using our regression model, most of the variation of implied volatilities can be attributed to a variation of the degree of moneyness and the time to maturity. An examination of the estimated implied volatility surfaces within the range of traded strikes reveals only a very few violations of the no-arbitrage constraints.

Table 1 reports the mean and the standard deviation of the daily coefficient estimates for each parameter, as well as the *t*-statistic for the mean.

TABLE 1 Mean, standard deviation and *t*-value of the daily parameter estimates over the period January 2000 to December 2000.

Parameter estimate	Mean	Standard deviation	*t*-value
$\hat{\beta}_0$	0.218	0.053	65.55
$\hat{\beta}_1$	–0.054	0.045	–19.12
$\hat{\beta}_2$	0.077	0.051	24.06
$\hat{\beta}_3$	0.206	0.113	15.93
$\hat{\beta}_4$	0.044	0.060	11.68
$\hat{\beta}_5$	–0.159	0.088	–28.79

The average general volatility level $\bar{\hat{\beta}}_0$ during the sample period amounts to 21.80%. The second parameter, β_1 may be interpreted as the overall slope of the average volatility smile. Its average estimated value of –0.054 indicates a downward-sloping smile profile. The parameters β_2 and β_3 together determine the

[30] See Kmenta (1997), p. 219 and Sachs (1972), p. 265.

curvature of the average volatility smile. On average, the parameter estimates are positive, signifying a convex smile pattern. The asymmetry of implied volatilities with moneyness $M \leq 0$ and $M > 0$ is apparent from the positive $\hat{\beta}_3$-value. The slope of the volatility term structure of at-the-money options is represented by the parameter β_4. Its mean value of 0.044 implies an upward-sloping term-structure, ie, options with longer times to expiration exhibit a higher implied volatility than shorter-term options. The last parameter β_5 describes the behavior of the smile slope with varying time to maturity. Its mean value of −0.159 suggests that the volatility smile gets flatter for longer times to maturity. This phenomenon is known as "flattening out" effect.

Given the estimated regression function for the volatility surface, we next employ our extrapolation method to the volatility smiles for 30, 90, and 180 days to maturity. The method turns out to be very fast and robust. For all smiles we could find a continuation which solves the system of equations (27). Table 2 reports the mean and the standard deviation (in parentheses) of the daily coefficients. For a time to expiration of 30 days the initial log simple moneyness interval of [− 0.3, 0.2] has to be narrowed in some cases to produce an arbitrage-free RND (see Table 2).

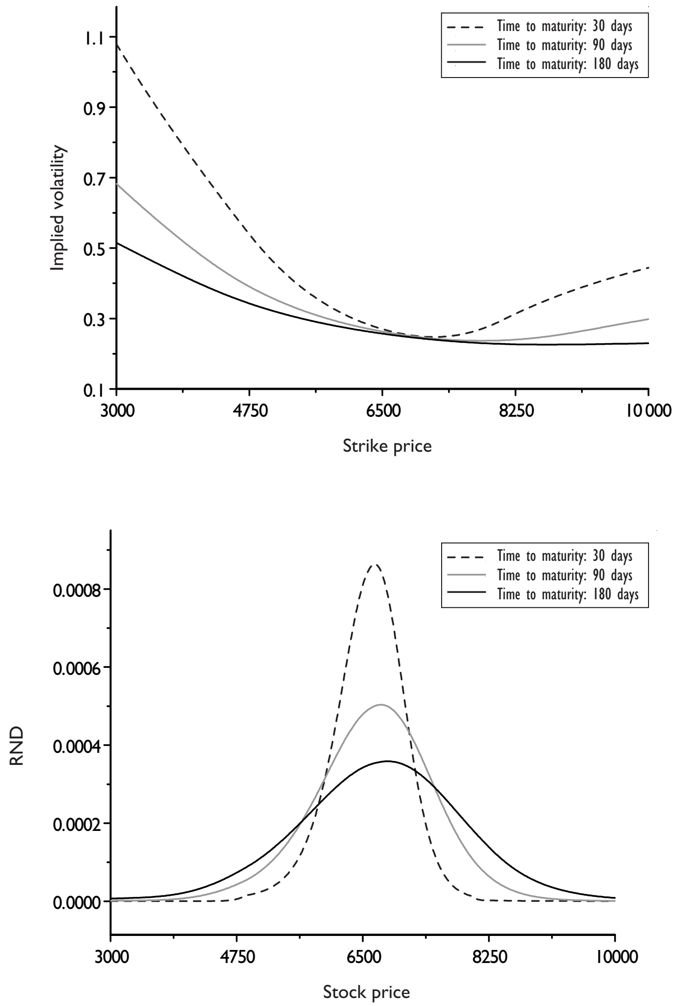
To illustrate our method, we single out one trading day from the sample, October 11, 2000, and plot the estimated volatility smile and the corresponding RND for three different times to maturity (see Figure 1). The charts were generated using the parameters of Table 3. The smooth continuation of the volatility smile and the RND, respectively, is apparent.

In a further analysis we investigate the RNDs of 30, 90 and 180 days to matu-

TABLE 2 Means and standard deviations of the extrapolation parameter values for the period January 2000 to December 2000.

Parameter	Days to maturity		
	30	90	180
λ_L	0.537 (0.299)	0.389 (0.147)	0.487 (0.135)
$\nu_{1,L}$	0.831 (0.107)	0.800 (0.000)	0.800 (0.000)
$\nu_{2,L}$	0.098 (0.066)	0.169 (0.043)	0.187 (0.027)
$\eta_{1,L}$	108 391 (179 712)	35 706 (6 526)	23 803 (2 639)
$\eta_{2,L}$	7240 (1,368)	7 515 (745)	7 205 (482)
λ_U	0.595 (0.203)	0.137 (0.127)	0.071 (0.149)
$\nu_{1,U}$	0.337 (0.098)	0.300 (0.000)	0.275 (0.077)
$\nu_{2,U}$	0.042 (0.081)	0.086 (0.011)	0.121 (0.011)
$\eta_{1,U}$	3 492 (1 146)	5 077 (356)	6 434 (911)
$\eta_{2,U}$	7 713 (1 054)	7 448 (471)	7 627 (462)
$\int_{K_L}^{K_U} q_{S_T}^{\mathcal{M}}(x)\mathrm{d}x$	0.994 (0.005)	0.947 (0.023)	0.834 (0.038)
$\ln(\tilde{M}_L)$	−0.262 (0.033)	−0.300 (0.000)	−0.300 (0.000)
$\ln(\tilde{M}_U)$	0.180 (0.022)	0.200 (0.000)	0.200 (0.000)

FIGURE 1 (*a*) Estimated volatility smiles on October 11, 2000, for 30, 90, and 180 days to maturity. (*b*) Corresponding RNDs.



(*a*)

(*b*)

TABLE 3 Parameters for Figure 2.

τ **(days)**	$F_t(T)$	K_L	K_U	$\int_{K_L}^{K_U} q_{S_T}^{\mathcal{M}}(x)dx$
30	6578.00	4873.28	8034.69	99.38%
90	6633.50	4914.31	8102.33	93.26%
180	6717.34	4976.33	8204.58	83.23%

$\ln(\tilde{M}_L) = -0.3$ and $\ln(\tilde{M}_U) = 0.2$.

TABLE 4 Summary statistics for DAX RNDs of 30, 90, and 180 days to maturity over the sample period.

Days to maturity	Mean S_T	Std dev S_T/Mean S_T	Skewness S_T	Kurtosis S_T
30 days to maturity				
Mean	7152	0.069	–0.458	4.580
Standard deviation	383	0.011	0.264	2.547
90 days to maturity				
Mean	7204	0.125	–0.510	2.350
Standard deviation	382	0.015	0.226	0.750
180 days to maturity				
Mean	7288	0.182	–0.480	1.478
Standard deviation	383	0.017	0.296	0.550

rity over the sample period. Table 4 summarizes the results. As expected, the distribution is negatively skewed and leptokurtic for all times to maturity.[31]

4.4 Comparison with other methods

According to Bliss and Panigirtzoglou (2002) an important criterion to judge the quality of a RND estimation method is its *robustness* to small errors in recorded option prices. Another one is the *goodness-of-fit* to the market data. Typically there is a trade-off between robustness and goodness-of-fit. On the one hand, if a method fits the data accurately, it is suspected not to be very stable. On the other hand, one expects a robust method to produce a worse fit to the data. If we intend to use the estimated RND to price exotic derivatives, it is essential that the estimated RND fulfills the *no-arbitrage constraints*. This property represents a third criterion to gauge the quality of an estimation method.

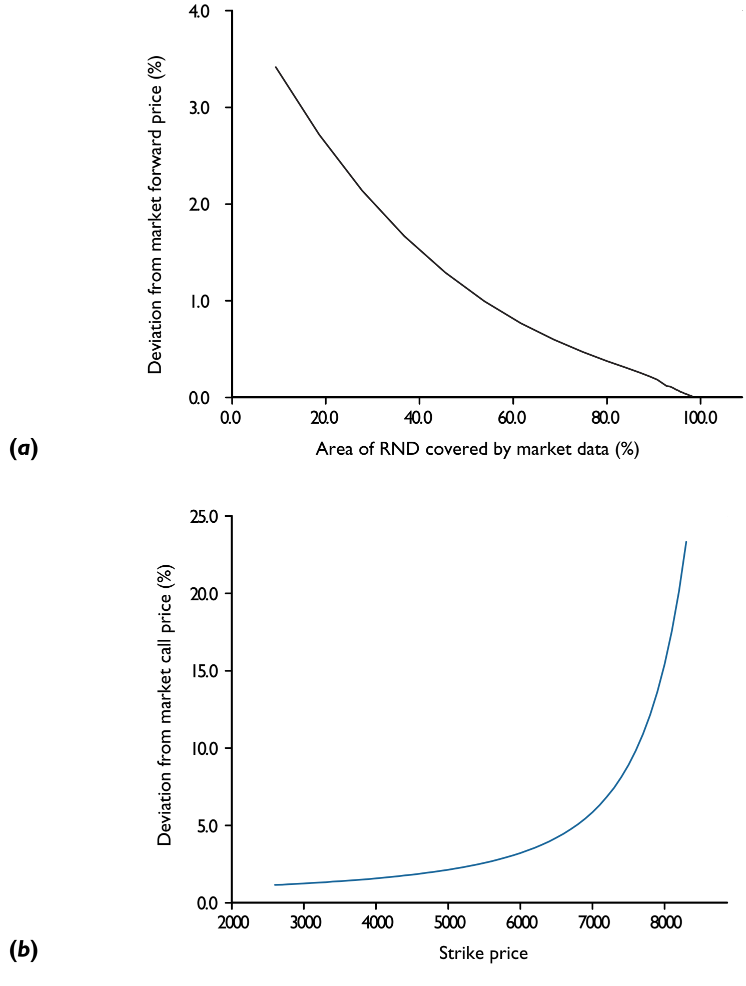
Using these criteria, we compare our method with two other very common volatility smile methods. The first method we consider was suggested by Shimko (1993) and is conceptually close to ours. Within the strike price interval $\mathcal{M} = [K_L, K_U]$ he estimates the implied volatility function for a fixed maturity T by the quadratic regression model

$$\sigma^{\mathcal{M}} = \alpha_0 + \alpha_1 K + \alpha_2 K^2 + \varepsilon \tag{44}$$

Outside this interval he assumes implied volatilities to be constant or equivalently the tails to follow a lognormal distribution. Concerning the criteria robustness and goodness-of-fit the two methods perform similarly, although the smile function (40) is in general more appropriate to describe the pattern of DAX implied volatilities than function (44). As documented in Section 3.2, RNDs generated by the Shimko (1993) method are continuous and integrate to

[31] Bliss and Panigirtzoglou (2002) report similar values for the FTSE100 index.

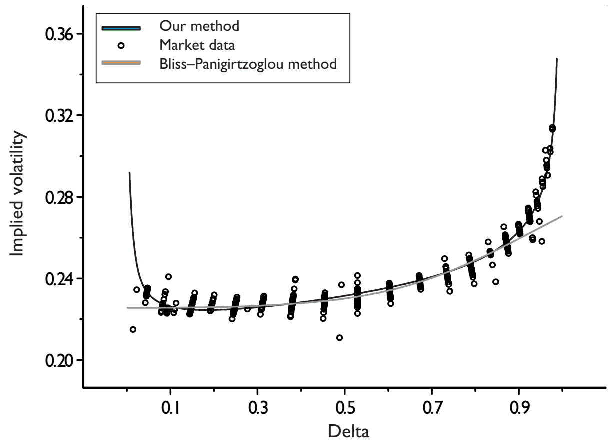
FIGURE 2 (*a*) Deviations from market forward price versus different values of $\int_{K_L}^{K_U} q_{S_T}^{\mathcal{M}}(x)dx$ on February 10, 2000, for the Shimko (1993) method; τ =127 calendar days; $F_t(T)$ = 7829.21. (*b*) Corresponding deviations from market call prices for different strike prices; K_L = 2600, K_U = 8400.



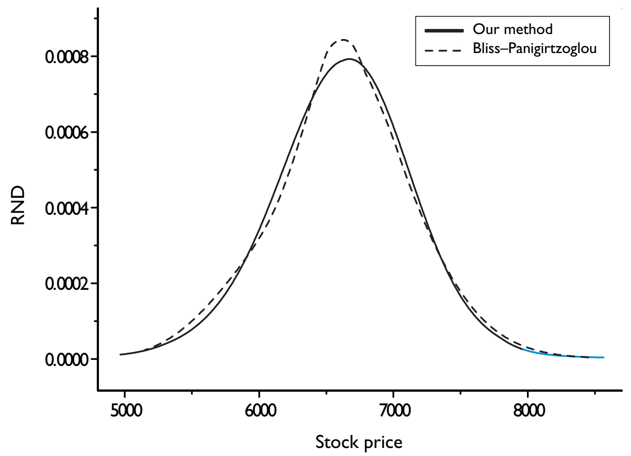
one. The martingale property, however, is not explicitly demanded and therefore often violated, indicating that arbitrage opportunities exist. In general, the pricing error increases when the strike price interval $[K_L, K_U]$ narrows. Figure 2(*a*) illustrates this effect for the DAX forward price on February 10, 2000. The pricing error of the call options is even worse as Figure 2(*b*) shows. Due to the violation of the martingale restriction, the Shimko RND is not suitable to price derivatives, not even the ones used for constructing the RND.

A second method we compare our method with is the one of Bliss and

FIGURE 3 (*a*) Estimated volatility smiles with Bliss–Panigirtzoglou (2002) and our method on December 7, 2000; τ = 43 calendar days, K_L = 5300, K_U = 7800, $F_t(T)$ = 6622.45; smoothing: high. (*b*) Corresponding RNDs.



(*a*)

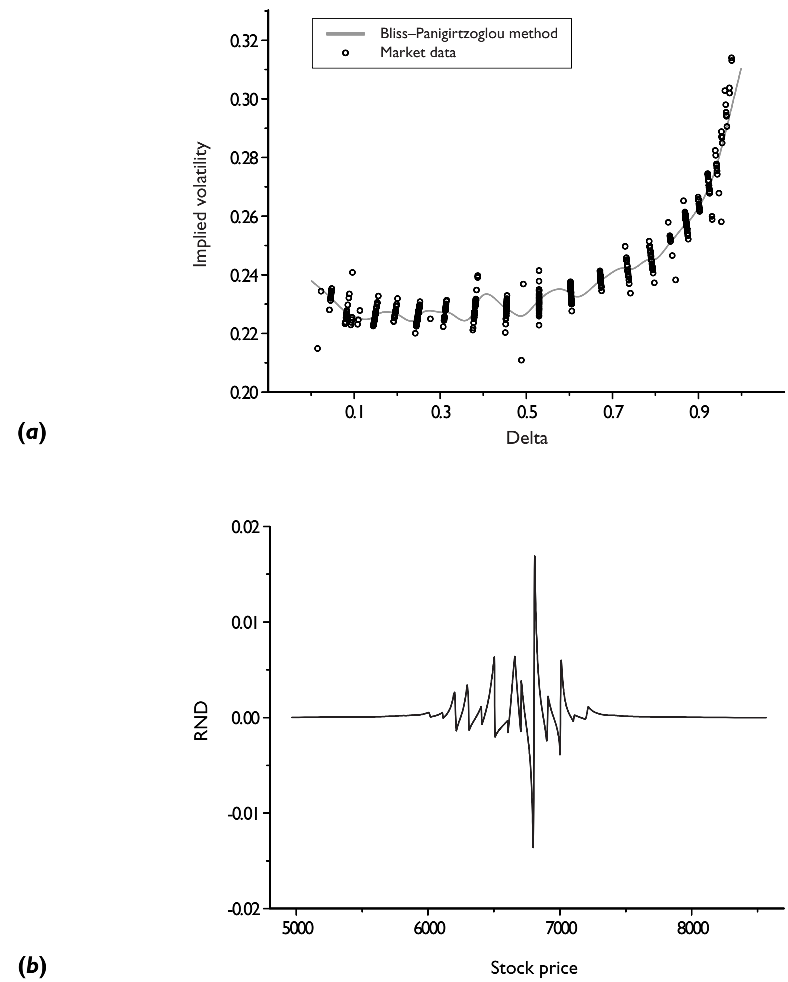


(*b*)

Panigirtzoglou (2002). Bliss and Panigirtzoglou (2002) combined the innovations of Malz (1997) and Campa *et al* (1998). They use a smoothing natural spline method to fit the implied volatilities as a function of the options' deltas. The method satisfies the integrability constraint. It also meets the martingale constraint, since the natural spline function for the implied volatility converges to constant values for very low and very high strike prices. On the other hand, the RND can become negative. In contrast to our method, such a situation is difficult to handle using the Bliss and Panigirtzoglou (2002) method. The introduction of a smoothing parameter allows to control for the trade-off between robustness and goodness-of-fit. On the one hand, forcing the volatility function to be stable has in general a negative impact on the goodness-of-fit. This is shown in Figure 3. Obviously, the method significantly underestimates the implied volatilities for options with low and high deltas. Conversely, our method matches the market data, symbolized by circles, very well.

On the other hand, when trimming the method to achieve the same goodness-of-fit as our method, the estimated RND is no longer smooth, as Figure 4 shows. Here, we can even observe negative probabilities.

FIGURE 4 (*a*) Estimated volatility smiles with Bliss/Panigirtzoglou (2002) and our method on December 7, 2000; smoothing: low. (*b*) Corresponding RND.



5 An application: pricing of digital options

In this section we apply our method to price a particular type of digital option: *a cash-or-nothing call*. This option pays off a fixed amount of money c at maturity T, when the terminal stock price S_T exceeds the strike level K, and otherwise 0. Without loss of generality, we set $c = 1$.

By the risk-neutral valuation formula (1), the time t price of a cash-or-nothing call with strike price K and maturity T is simply the discounted risk-neutral probability that the stock price is above the strike price at maturity T.

$$CoNC_t(K,T) = \mathrm{e}^{-r(T-t)} \int_K^{\infty} q_{S_T}(x)\mathrm{d}x \tag{45}$$

To compute (45), using our piecewise defined RND $q_{S_T}(x; \theta_L, \theta_U)$, let us consider three cases: the strike price K is between K_L and K_U, above K_U, or below K_L. In the first case, (45) evaluates to

$$\begin{aligned}
CoNC_t(K,T) &= \mathrm{e}^{-r(T-t)}\left(\int_K^{K_U} q_{S_T}^{\mathcal{M}}(x)\mathrm{d}x + \int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\mathrm{d}x \right) \\
&= \mathrm{e}^{-r(T-t)}\left(N(d_2(K)) - Kn(d_2(K))\sqrt{T-t}\,\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K} \right) \\
&\quad -\mathrm{e}^{-r(T-t)}\left(N(d_2(K_U)) - K_U n(d_2(K_U))\sqrt{T-t}\,\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U} \right) \\
&\quad +\mathrm{e}^{-r(T-t)} \int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\mathrm{d}x
\end{aligned} \tag{46}$$

By the specification of $q_{S_T}^{U}(x; \theta_U)$, expression (46) simplifies to

$$\begin{aligned}
CoNC_t(K,T) &= CoNC_t^{BS}(K,T,\sigma_t^{\mathcal{M}}(K,T)) \\
&\quad -\Lambda_t^{BS}(K,T,\sigma_t^{\mathcal{M}}(K,T))\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}
\end{aligned} \tag{47}$$

where $CoNC_t^{BS}(K, T, \sigma_t^{\mathcal{M}}(K, T))$ denotes the Black–Scholes price of a cash-or-nothing call, and $\Lambda_t^{BS}(K, T, \sigma_t^{\mathcal{M}}(K, T))$ is the Black–Scholes vega of a standard call option.[32] If we look at formula (47), the risk-neutral price of a cash-or-nothing call is the Black–Scholes price minus a "correction term", which is in

[32] The Black–Scholes formula for a cash-or-nothing call option can be found in Hull (2000), pp. 464–5. The Black–Scholes vega of a standard option is given by

$$\Lambda_t^{BS}(K,T,\sigma_t^{\mathcal{M}}(K,T)) = n(d_2(K))K\sqrt{T-t}\,\mathrm{e}^{-r(T-t)}$$

turn the Black–Scholes vega times the slope of the volatility smile at the strike price K. The higher the vega and the higher the slope of the volatility smile, the larger the correction.

Defining

$$z_{i,j}(K) = \frac{\ln(\eta_{i,j}) - \ln(K) - \frac{\nu_{i,j}^2}{2}}{\nu_{i,j}}, \quad i \in \{L, U\}, \quad j = 1, 2 \tag{48}$$

and

$$\begin{aligned} z_L(K_L) &= z_{L,1}(K_L) = z_{L,2}(K_L) \\ \text{and} \quad z_U(K_U) &= z_{U,1}(K_U) = z_{U,2}(K_U) \end{aligned} \tag{49}$$

equation (45) becomes

$$\begin{aligned} &CoNC_t(K,T) \\ &= e^{-r(T-t)} \int_K^\infty \left(\lambda_U \ell(x; \eta_{U,1}, \nu_{U,1}^2) + (1-\lambda_U) \ell(x; \eta_{U,2}, \nu_{U,2}^2) \right) dx \\ &= e^{-r(T-t)} \lambda_U N\left(z_{U,1}(K)\right) + e^{-r(T-t)} (1-\lambda_U) N\left(z_{U,2}(K)\right) \end{aligned} \tag{50}$$

in the second case, ie, for $K > K_U$. In particular, if $K = K_U$, we get with the help of (32):

$$CoNC_t(K_U, T) = e^{-r(T-t)} N\left(z_U(K_U)\right)$$

The pricing formula for the last case, $K < K_L$, is developed analogously. We obtain:

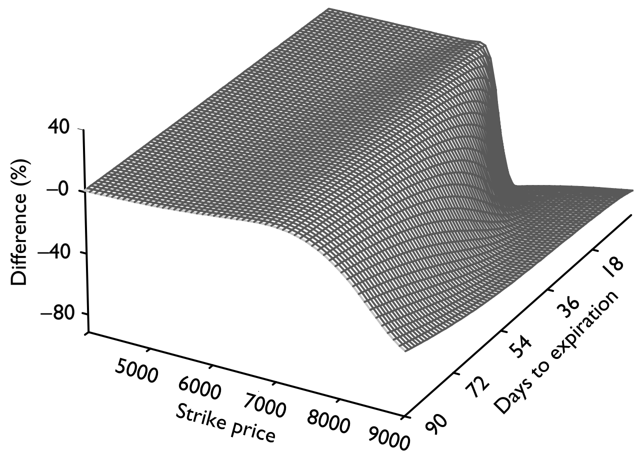
$$CoNC_t(K,T) = e^{-r(T-t)} \lambda_L N\left(z_{L,1}(K)\right) + e^{-r(T-t)} (1-\lambda_L) N\left(z_{L,2}(K)\right) \tag{51}$$

and for $K = K_L$:

$$CoNC_t(K_L, T) = e^{-r(T-t)} N\left(z_L(K_L)\right)$$

To illustrate the correction to the Black–Scholes price, we consider a cash-or-nothing call on the DAX as of December 7, 2000. The index level is 6583.55. The volatility surface and the RND are estimated as described in Section 4.2. The risk-neutral price $CoNC_t(K, T)$ is computed according to formulas (47), (50), and (51). The Black–Scholes price $CoNC_t^{BS}(K, T, \hat{\sigma}_t(K, T))$ is calculated by plugging in the estimated implied volatility into the Black–Scholes formula for the cash-or-nothing call. As is apparent from Figure 5, the price difference $(CoNC - CoNC^{BS})/CoNC^{BS}$ is quite substantial. It changes sign, depending on the slope of the implied volatility function in the respective strike price area.

FIGURE 5 Price difference $(CoNC - CoNC^{BS})/CoNC^{BS}$ (%) by strike and time to maturity on December 7, 2000.



6 Summary and conclusions

All volatility smile methods for estimating the risk-neutral density boil down to the completion of the implied volatility function by interpolating between available strike prices and extrapolating outside their range. In this paper we focus on the extrapolation and develop a new method, which is, under weak constraints, consistent with the absence of arbitrage.

In the first part of the paper we show that for any option's maturity the option price function, the function of implied volatilities, and the RND contain essentially the same information. Given the market implied volatility function within the range of observable strike prices, we derive a set of conditions any extrapolation function has to fulfill in order to be consistent with the absence of arbitrage. In terms of the RND, we propose a mixture of two lognormal distributions as a specific extrapolation function. This choice was mainly driven by the flexibility this distribution offers. It covers a lot of different shapes and has enough degrees of freedom to meet the no-arbitrage constraints. In the limit, when the observable strike price interval collapses to one point, our method migrates to the mixture of lognormals method, one of the most common representatives of the *RND function approximating approach*.

Using all call and put prices of options on the German stock index DAX, traded in the year 2000, we estimate daily the volatility surface within the range of trade strikes and maturities by a spline regression model. Note that there is nothing special with this model. Any other could be used instead. The empirical results show a very accurate fit to the data. On average, the variation of moneyness and time to maturity, explains about 96% of the cross-sectional variation of implied volatilities. Focusing on three different times to maturity, we then apply our extrapolation method. In all cases, the method generates an arbitrage-free and smooth RND.

In contrast to our method, the method of Shimko (1993) cannot ensure discounted asset prices to be martingales with respect to the estimated RND. As was empirically shown, the mispricing can become quite substantial. Thus, the generated RND must be used with care when pricing other derivatives. On the other hand, the smoothing spline method of Bliss and Panigirtzoglou (2002) usually satisfies the no-arbitrage conditions. However, in our empirical study, the method turns out to be either unstable or inaccurate. It is important to note, that these findings may be different for other data sets.

In an application, we use our constructed RND to price a digital option. For this option, as for many other path-independent exotic options, a closed-form valuation formula can be derived. When valuing such an option using the simple Black–Scholes implied volatility, the pricing error was shown to be quite large.

All in all, the proposed extrapolation method does not depend on a particular interpolation scheme and is therefore universally applicable. It does not negatively impact the goodness-of-fit of the interpolation. The implementation involves only straightforward numerical procedures and is highly computer-efficient.

Technical appendix

Proof of equations (17) and (18)

THEOREM 1 *Let us consider a frictionless market with time horizon T^*, where we can observe at any time $t \in [0, T^*]$ a stock, a complete collection of standard European calls of strike $K \geq 0$ and maturity $T \in (t, T^*]$ and a money market account paying interest at a constant rate r. If the market is free of arbitrage, then a calendar call spread with arbitrary maturities $T_1, T_2 \in (t, T^*]$, $T_1 \leq T_2$, has nonnegative value at time t:*

$$C_t(K, T_2) \geq C_t\left(\mathrm{e}^{-r(T_2 - T_1)} K, T_1\right) \tag{52}$$

and the T_1 and T_2 RNDs obey

$$\int_0^\infty \max\{x - K; 0\}\left(\mathrm{e}^{r(T_2 - T_1)} q_{S_{T_2}}(x) - q_{S_{T_1}}\left(x\mathrm{e}^{-r(T_2 - T_1)}\right)\right)\mathrm{d}x \geq 0 \tag{53}$$

PROOF By the absence of arbitrage, there exists a risk-neutral measure Q such that discounted asset prices are martingales with respect to this measure. In particular, the RNDs $q_{ST_1}(x)$ and $q_{ST_2}(x)$ represent this measure at times T_1 and T_2.

Integrating both sides of

$$\max\left\{S_{T_2} - K; 0\right\} \geq S_{T_2} - K$$

at time T_1 against $q_{S_{T_2}}(x)$ and noting that $C_{T1}(K, T_2) \geq 0$, it follows

$$C_{T_1}(K,T_2) \geq \max\left\{S_{T_1} - \mathrm{e}^{-r(T_2-T_1)}K;0\right\} \tag{54}$$

Integrating (54) with respect to $q_{S_{T_1}}(x)$ proves proposition (52)

$$\begin{aligned}
&\mathrm{e}^{-r(T_1-t)}\int_0^\infty C_{T_1}(K,T_2)q_{S_{T_1}}(x)\,\mathrm{d}x\\
&\geq \mathrm{e}^{-r(T_1-t)}\int_0^\infty \max\left\{S_{T_1} - \mathrm{e}^{-r(T_2-T_1)}K;0\right\}q_{S_{T_1}}(x)\,\mathrm{d}x\\
&\Rightarrow C_t(K,T_2) \geq C_t\left(\mathrm{e}^{-r(T_2-T_1)}K,T_1\right)
\end{aligned} \tag{55}$$

Re-expressing (55) yields:

$$\int_0^\infty \max\{x-K;0\}q_{S_{T_2}}(x)\,\mathrm{d}x \geq \int_0^\infty \max\left\{\mathrm{e}^{r(T_2-T_1)}x - K;0\right\}q_{S_{T_1}}(x)\,\mathrm{d}x \tag{56}$$

and after substituting $y = \mathrm{e}^{r(T2-T1)}x$ on the right-hand side of (56) the proposition (53) follows. □

Proof that $q_{S_T}(x; \theta_L, \theta_U)$ is a RND

THEOREM 2 *Let us assume that we can observe at time $t \in [0, T^*]$ the forward price $F_t(T)$ for maturity $T \in (t, T^*]$ and the market option price function $C_t^{\mathcal{M}}(K, T)$ or equivalently the market implied volatility function $\sigma_t^{\mathcal{M}}(K, T)$ for $K \in [K_L, K_U]$. If the call price function $C_t^{\mathcal{M}}(K, T)$ fulfills the no-arbitrage constraints (13)–(16) for all $K \in [K_L, K_U]$, the tail functions $q_{S_T}^L(x; \theta_L)$, $q_{S_T}^U(x; \theta_U)$ are nonnegative for $x \in (0, K_L)$ and $x \in (K_U, \infty)$, respectively, and $q_{S_T}^L(x; \theta_L)$, $q_{S_T}^U(x; \theta_U)$ solve the system (27), then the piecewise defined function*

$$q_{S_T}(x;\theta_L,\theta_U) = \begin{cases} q_{S_T}^L(x;\theta_L), & x < K_L\\ q_{S_T}^{\mathcal{M}}(x), & K_L \leq x \leq K_U\\ q_{S_T}^U(x;\theta_U), & x > K_U\end{cases} \tag{57}$$

is indeed a well-defined RND.

PROOF In the following we show that $q_{S_T}(x; \theta_L, \theta_U)$ has the nonnegative, integrability and martingale property. The first property is fulfilled by construction. Concerning the second one, integration over $q_{S_T}(x; \theta_L, \theta_U)$ yields

$$\int_0^\infty q_{S_T}(x;\theta_L,\theta_U)\,\mathrm{d}x = \int_0^{K_L} q_{S_T}^L(x;\theta_L)\,\mathrm{d}x + \int_{K_L}^{K_U} q_{S_T}^{\mathcal{M}}(x)\,\mathrm{d}x + \int_{K_U}^\infty q_{S_T}^U(x;\theta_U)\,\mathrm{d}x \tag{58}$$

With the continuity of $q_{S_T}^{\mathcal{M}}(x)$ ensured by the first two conditions in (27),

$$\int_{K_L}^{K_U} q_{S_T}^{\mathcal{M}}(x)\mathrm{d}x = -N\big(d_2(K_U)\big) + K_U n\big(d_2(K_U)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U}$$
$$+N\big(d_2(K_L)\big) - K_L n\big(d_2(K_L)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_L} \quad (59)$$

and

$$\int_{0}^{K_L} q_{S_T}^{L}(x;\theta_L)\mathrm{d}x = N\big(-d_2(K_L)\big) + K_L n\big(d_2(K_L)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_L}$$
$$\int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\mathrm{d}x = N\big(d_2(K_U)\big) - K_U n\big(d_2(K_U)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U} \quad (60)$$

from the third and fourth conditions in (27) we get the integrability property. In the same manner, we show $\int_0^\infty x q_{S_T}^{\mathcal{M}}(x)\mathrm{d}x = F_t(T)$. In the last step we have to prove that all calls with strike $K > 0$ have the martingale property. Obviously, all calls with $K \notin [K_L, K_U]$ have this property, since we cannot observe a market price for these options. For all other options, we use the risk-neutral valuation formula (1) to obtain

$$\mathrm{e}^{-r(T-t)}\int_0^\infty \max\{x-K;0\} q_{S_T}(x;\theta_L,\theta_U)\mathrm{d}x$$
$$= \mathrm{e}^{-r(T-t)}\int_K^{K_U}(x-K)\,q_{S_T}^{\mathcal{M}}(x)\mathrm{d}x + \mathrm{e}^{-r(T-t)}\int_{K_U}^{\infty}(x-K)\,q_{S_T}^{U}(x;\theta_U)\mathrm{d}x \quad (61)$$

Figuring out the first integral on the right-hand side, we find:

$$\mathrm{e}^{-r(T-t)}\int_0^\infty \max\{x-K;0\} q_{S_T}(x;\theta_L,\theta_U)\mathrm{d}x$$
$$= C_t(K,T)$$
$$- \mathrm{e}^{-r(T-t)}\left(F_t(T)N\big(d_1(K_U)\big) - K_U^2 n\big(d_2(K_U)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U}\right)$$
$$+ K\mathrm{e}^{-r(T-t)}\left(N\big(d_2(K_U)\big) - K_U n\big(d_2(K_U)\big)\sqrt{T-t}\left.\frac{\partial \sigma_t^{\mathcal{M}}(K,T)}{\partial K}\right|_{K=K_U}\right)$$
$$+ \mathrm{e}^{-r(T-t)}\int_{K_U}^{\infty} x q_{S_T}^{U}(x;\theta_U)\mathrm{d}x - K\mathrm{e}^{-r(T-t)}\int_{K_U}^{\infty} q_{S_T}^{U}(x;\theta_U)\mathrm{d}x \quad (62)$$

Finally, by using conditions (5) and (6) of system (27), we can eliminate the last four terms in (62) and get $C_t(K, T)$. This proves the claim. □

REFERENCES

AitSahlia, Y., and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance* **53**(2), 499–547.

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

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies* **31**(2), 91–6.

Bahra, B. (1996). Probability distributions of future asset prices implied by option prices. Working paper, Bank of England.

Bahra, B. (1997). Implied risk-neutral probability density functions from option prices: Theory and application. Technical report.

Bates, D. (1991). The crash of '87: Was it expected? The evidence from options markets. *Journal of Finance* **46**(3), 1009–44.

Bates, D. (1996). Jumps and stochastic volatility: Exchange rate process implicit in deutsch mark options. *Review of Financial Studies* **9**, 69–107.

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

Bookstaber, R., and McDonald, J. (1987). A general distribution for describing security price returns, *Journal of Business* **60**, 401–24.

Breeden, D. T., and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business* **51**(4), 621–51.

Brown, G., and Toft, K. B. (1999). Constructing binomial trees from multiple implied probability distributions. *Journal of Derivatives* **7**(2), 83–100.

Campa, J. M., Chang, K. P., and Reider, R. L. (1997). ERM bandwidth for EMU and after: Evidence from foreign exchange options. *Economic Policy* **24**, 55–89.

Campa, J. M., Chang, K. P., and Reider, R. L. (1998). Implied exchange rate distributions: Evidence from OTC option markets. *Journal of International Money and Finance* **17**, 117–60.

Carr, P. (2001). Constraints on implied volatility. Working paper, Bank of America Securities.

Cincibuch, M. (2001). Distributions implied by exchange traded options: A ghost's smile. Working paper, CERGE-EI.

Cont, R. (1997). Beyond implied volatility. Lecture notes, Ecole Polytechnique Fédérale de Lausanne.

Cooper, N. (1999). Testing techniques for estimating implied RNDs from the prices of European-style options. *Proceedings of the BIS workshop on implied PDFs*. Bank of International Settlements, Basle.

Cooper, N., and Talbot, J. (1999). The Yen, and Dollar exchange rate in 1998: Views from options markets. *Bank of England Quarterly Bulletin* pp. 68–77.

Cox, J. C., and Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics* **3**, 145–66.

Debreu, G. (1959). *Theory of Value*, Wiley, New York.

Hafner, R., and Wallmeier, M. (2001). The dynamics of DAX implied volatilities. *International Quarterly Journal of Finance* **1**(1), 1–27.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* **6**(2), 327–43.

Hull, J. C. (2000). *Options, Futures, and Other Derivatives.* 4th edn, Prentice Hall, New Jersey.

Hull, J., and White, A. (1987). The pricing of options on assets with stochastic volatilities, *Journal of Finance* **42**, 281–300.

Jackwerth, J. C. (1999). Option implied risk-neutral distributions and implied binomial trees: A literature review. Working paper, University of Wisconsin.

Jackwerth, J. C. (2000). Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* **13**(2), 433–51.

Jackwerth, J. C., and Rubinstein, M. (1996). Recovering probability distributions from option prices. *Journal of Finance* **51**(5), 1611–31.

Kmenta, J. (1997). *Elements of Econometrics.* 2nd edn, University of Michigan Press, Ann Arbor.

Levin, A., McManus, D., and Watt, D. (1998). The information content of Canadian dollar futures options. In *Information in Financial Asset Prices.* Bank of Canada, pp. 229–75.

Malz, A. M. (1997). Estimating the probability distribution of the future exchange rate from option prices. *Journal of Derivatives* **5**(2), 18–36.

McManus, D. (1999). The information content of interest rate futures options. Technical report.

Melick, W., and Thomas, C. (1997). Recovering an asset's implied PDF from option prices: An application to crude oil during the gulf crisis. *Journal of Financial and Quantitative Analysis* **32**(1), 91–115.

Natenberg, S. (1994). *Option Volatility and Pricing*. Probus, Chicago, Cambridge.

Neuhaus, H. (1995). Der Informationsgehalt von Derivaten für die Geldpolitik: Implizite Volatilitäten und Wahrscheinlichkeiten. *Diskussionspapier 3/95*, Deutsche Bundesbank.

Pliska, S. R. (2000). *Introduction to Mathematical Finance*, Blackwell Publishers, Malden, Oxford.

Rosenberg, J. V. (2000). Implied volatility functions: A reprise. *Journal of Derivatives* **7**(3), 51–64.

Ross, S. (1976). Options and efficiency. *Quarterly Journal of Economics* **90**, 75–89.

Sachs, L. (1972). *Statistische Auswertungsmethoden*, Vol. 3, Berlin, Heidelberg, New York.

Shimko, D. (1993). Bounds of probability, *Risk* **6**(4), 33–7.

Stein, E., and Stein, J. (1991). Stock price distributions with stochastic volatility: An analytical approach. *Review of Financial Studies* **4**, 727–52.

Trautmann, S., and Beinert, M. (1999). Impact of stock price jumps on option values. In W. Bühler, H. Hax, and R. Schmidt (eds), *Empirical Research on German Capital Markets*, Springer, Berlin *et al*.