Journal of Financial Econometrics, 2007, Vol. 5, No. 2, 189–218

Special Issue on Econometrics of Financial and Insurance Risks

A Semiparametric Factor Model for Implied Volatility Surface Dynamics

MATTHIAS R. FENGLER
Sal. Oppenheim jr. & Cie.

WOLFGANG K. HÄRDLE
Center for Applied Statistics and Economics

ENNO MAMMEN
University of Mannheim

ABSTRACT
We propose a semiparametric factor model, which approximates the implied volatility surface (IVS) in a finite dimensional function space. Unlike standard principal component approaches typically used to reduce complexity, our approach is tailored to the degenerated design of IVS data. In particular, we only fit in the local neighborhood of the design points by exploiting the expiry effect present in option data. Using DAX index option data, we estimate the nonparametric components and a low-dimensional time series of latent factors. The modeling approach is completed by studying vector autoregressive models fitted to the latent factors.

KEYWORDS: functional principal component analysis, implied volatility surface, semiparametric factor models

The implied volatility surface (IVS), which is derived by applying the Black-Scholes (BS) formula to a set of traded plain vanilla options across different strikes and expiries, is a key financial variable for trading, hedging, and the risk management of option portfolios. In equity markets, the IVS is usually observed as a nonflat surface with a strong skew for out-of-the-money puts, as shown in the left panel of Figure 1. Despite this observation, which invalidates the constant volatility

We gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft and the Sonderforschungsbereich 649 "Ökonomisches Risiko". We thank two anonymous referees, the editors, and Matthias Bode for their valuable comments and suggestions. The article represents the authors' personal opinion and does not reflect the views of Sal. Oppenheim. Address correspondence to Matthias R. Fengler, Trading & Derivatives, Sal. Oppenheim jr. & Cie., Untermainanlage 1, 60329 Frankfurt am Main, Germany, or e-mail: matthias.fengler@oppenheim.de.

doi:10.1093/jjfinec/nbm005
Advance Access publication March 12, 2007

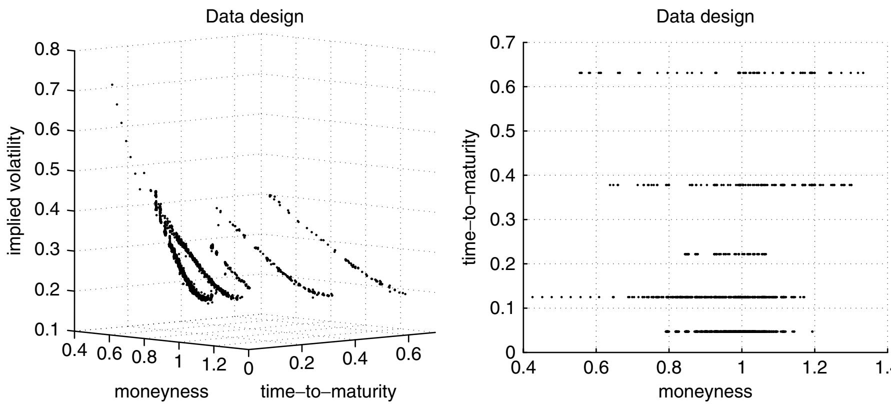

Figure 1 Left panel: DAX index call and put implied volatilities observed on May 2, 2000, plotted across futures moneyness and time-to-maturity. Right panel: same data as seen from top.

assumption in the underlying model framework by Black and Scholes (1973), the IVS is a widely accepted state variable, which—unlike estimates based on historical data—is believed to reflect the current perception of market risk [Bakshi et al. (2000); Britten-Jones and Neuberger (2000)]. Consequently, market makers at plain vanilla desks continuously monitor and update the IVS they trade on, and risk managers study the impact of large market moves on entire portfolios in terms of joint spot and implied volatility shifts or other deformations of the IVS. To be implemented in a realistic way, all these applications require a suitable statistical model of the IVS.

Modeling the IVS poses two main challenges. First and foremost, the implied volatility data have a degenerated design: by ‘‘degenerated design’’ we refer to the institutional convention that implied volatility data exist only for a small number of maturities such as 1, 2, 3, 6, 9, 12, 18, and 24 months to expiry on the date of issue. Hence, implied volatility observations appear in ‘‘strings’’; see the right panel of Figure 1, which shows the observations in the moneyness and time-to-maturity space only. Options belonging to the same string have the same time-to-maturity. As time passes, these strings move along the maturity axis toward expiry while changing their levels and shapes in a random fashion, see Figure 2. A second challenge is that the observations do not always cover the desired estimation grid completely and observations can be missing in certain sub-regions in the moneyness dimension. For instance, the options belonging to the third and fifth time-to-maturity (from top) are observed only near-the-money, while options belonging to the other time-to-maturities appear shifted either to the out-of-the-money put region or to the out-of-the-money call region. However, despite that appearance, implied volatilities are thought as being the observed structure of a smooth surface. This is because in practice one needs to price and hedge over-the-counter options whose expiry dates do not coincide with the expiry dates of options that are traded at the futures exchange.

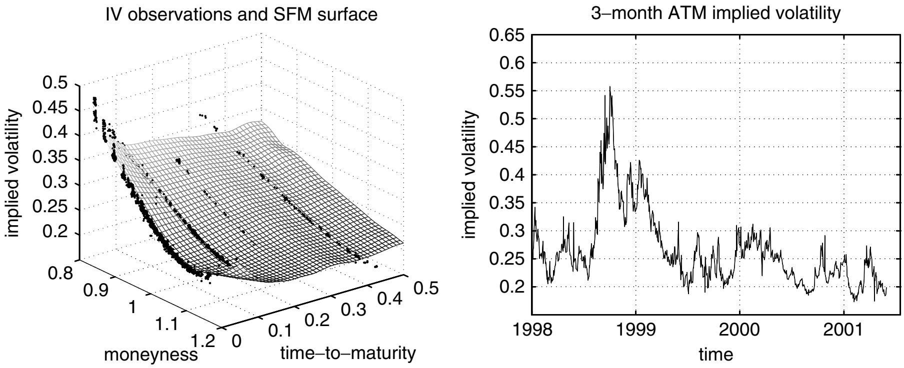

Figure 2 Left panel: DAX index IV observations and SFM fit for May 2, 2000. Bandwidths $h_1 = 0.04$ for the moneyness and $h_2 = 0.06$ for the time-to-maturity dimension. Right panel: Time series of three-months at-the-money implied volatility data obtained from linear interpolation of the raw data.

For the semi- or nonparametric approximations to the IVS used in Aït-Sahalia and Lo (1998), Aït-Sahalia et al. (2001), Cont and da Fonseca (2002), Fengler et al. (2003), and Benko et al. (2007) this design can pose difficulties. The reason is that for standard nonparametric estimators to achieve a reasonable fit to the IVS, the bandwidth in the time-to-maturity dimension must be chosen very large in order to bridge the gaps between the different time-to-maturity strings. This induces an estimation bias that becomes the more severe the more irregular the observations are distributed across moneyness.

In this article we propose a semiparametric factor model (SFM) for modeling the IVS. The dynamic structure of the IVS is approximated by unknown basis functions moving in a finite dimensional function space. To introduce our model, denote by $Y_{i,j}$ the log-implied volatility, where $i = 1, \dots, I$ is an index of time, in our case the number of the day, and $j = 1, \dots, J_i$ is an intra-day numbering of the option traded on day i. The observations $Y_{i,j}$ are regressed on two-dimensional covariables $X_{i,j}$ that contain moneyness $\kappa_{i,j}$ and time-to-maturity $\tau_{i,j}$. Moneyness is defined as $\kappa_{i,j} \stackrel{\text{def}}{=} K_{i,j}/F_{i,j}$, that is, strike $K_{i,j}$ divided by the futures price $F_{i,j}$ belonging to the option trade (i, j). The SFM approximates the IVS by

$$Y_{i,j} \approx m_0(X_{i,j}) + \sum_{l=1}^{L} \beta_{i,l}\, m_l(X_{i,j}), \tag{1}$$

where m_l are smooth basis functions $(l = 0, \dots, L)$ and $\beta_{i,l}$ are weights depending on time i. The factor loading $\beta_i \stackrel{\text{def}}{=} (\beta_{i,1}, \dots \beta_{i,L})^\top$ forms an unobserved multivariate time series. By fitting model (1) to the implied volatility observations, we obtain approximations $\widehat{\beta}_i$. Their dynamic structure can be understood by applying classical vector autoregression (VAR) techniques. We argue that the VAR estimation based on $\widehat{\beta}_i$ is asymptotically equivalent to an estimation based on the unobserved β_i. A justification for this is given in Borak et al. (2006), where the functions m_l are estimated by orthogonal series estimators. Our conjecture is that their arguments carry over to kernel smoothing. In the left panel of Figure 2, we display a fit coming from the SFM.

Factor models that represent processes whose dynamics are driven by latent variables have a long tradition in the asset pricing literature. For equity derivatives, the main approach is to model stock returns with stochastic volatility as latent factor, such as in Hull and White (1987), Wiggins (1987), Heston (1993), and Schöbel and Zhu (1999) among others. Empirical research suggests to add a jump component, [see Bates (1996, 2000) and Duffie et al. (2000)]. A further refinement includes stochastic volatility with Lévy innovations, [see inter alia Barndorff-Nielsen (1997), Barndorff-Nielsen and Shepard (2001), and Carr et al. (2003)]. For parameter identification in the presence of latent variables a number of different estimation techniques were proposed[1]: analytically tractable filtration approaches as in Harvey et al. (1994) and Fridman and Harris (1998), or the approximate maximum likelihood methodology due to Bates (2006); generalized method of moments (GMM) approaches, either based on analytical moment restrictions, such as Melino and Turnbull (1990), Andersen and Sørensen (1996), Chacko and Viceira (2003), or based on simulated moments, such as Duffie and Singleton (1993); efficient method of moments techniques as in Gallant and Tauchen (1998), Andersen et al. (1999), Chernov and Ghysels (2000), and the indirect inference method due to Gouriéroux et al. (1993); implied-state methods, for example, as an extension of the expectation-maximization methodology as in Renault and Touzi (1996) and Pastorello et al. (2003), or in a GMM sense as in Pan (2002); and Bayesian Monte Carlo Markov Chain approaches as suggested by Jacquier et al. (1994) and Eraker et al. (2003).

To better capture unknown nonlinearities, a more recent strand of literature employs non- and semiparametric techniques, for example, to estimate the option pricing function [see e.g. Broadie et al. (2000a,b) and Aït-Sahalia and Duarte (2003)], or to recover the yield curve [Linton et al. (2001)]. Semiparametric factor models were proposed by Ghysels and Ng (1989) and Jeffrey et al. (2004) for modeling the term structure of interest rates. Connor and Linton (2006) use a semiparametric model for stock returns based on observable characteristics, and in Connor et al. (2006) a heteroskedastic latent factor model is designed which has nonparametric components in volatility. We complement this literature, as low-dimensional factor approximations to the IVS are traditionally based on principal components analysis (PCA), such as in Zhu and Avellaneda (1997), Skiadopoulos et al. (1999), Cont and da Fonseca (2002), and Fengler et al. (2003). The SFM differs in the following respect: for example, in Cont and da Fonseca (2002) the IVS is fitted on a grid for each day and a functional norm is applied to the surfaces as described in Ramsay and Silverman (1997). In our approach, we use only finite dimensional fits which are obtained in the local neighborhood of strikes and maturities for

[1] This overview is by no means complete; for a comprehensive review see Garcia et al. (2007) and Gallant and Tauchen (2007).

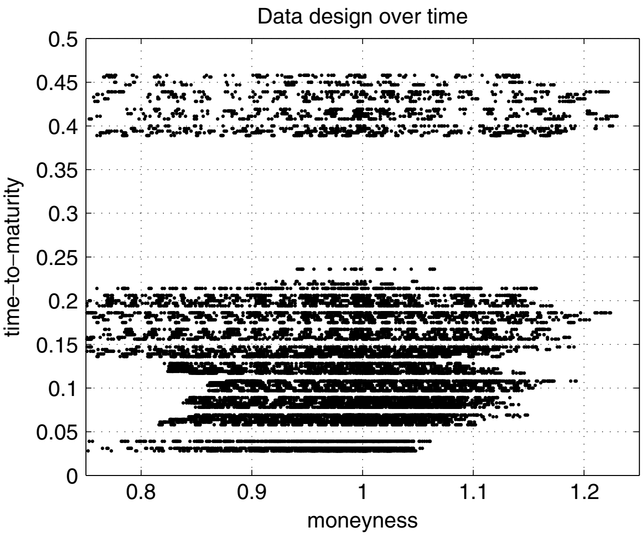

Figure 3 Evolution of data design as time passes. Data taken from January 01 to 31,1998; weekends appear as the white lines, since the scale is calendar days. Each dot corresponds to an actual trade.

which implied volatilities are recorded at the specific observation days. This leads to a minimization with respect to functional norms that depend on time i. The method can be understood as smoothing through time, which is made possible by the expiry effect present in IVS data, but which also exists in other economic variables, for example, in fixed income securities. This is illustrated in Figure 3 which shows the evolution of the data design for the period of one month. To our knowledge there do not yet exist estimation techniques that explicitly take advantage of the expiry effect.

The SFM is close to Connor and Linton (2006), but differs in that the explanatory variables $X_{i,j}$ do not depend on l. Under this assumption, the "factor coefficients" $m_l(X_{i,j,l})$ are identified and model (1) reduces to a nonparametric additive model where the additive components m_l can be estimated by smoothing techniques.[2] In our case $X_{i,j}$ does not depend on l and therefore, as in classical factor analysis, factor loadings and factor functions are not uniquely defined. Furthermore, our research has obvious relations to the literature on functional or varying-coefficient models, such as Hastie and Tibshirani (1993), Cai et al. (2000), Fan et al. (2003), or Kauermann et al. (2005).

Using transactions-based DAX index implied volatility data from January 1998 to May 2001, we recover three basis functions and three latent factors which we model with a VAR(2) model. The factor functions have natural interpretations as level, moneyness, and term structure slope effects. In comparison to a Nadaraya–Watson estimator, we demonstrate that the SFM mitigates bias

[2] For simplicity, Connor and Linton (2006) assume that $X_{i,j,l}$ does not depend on time i.

effects caused by the global daily fits, which are usually used in functional PCA.

The article is organized as follows. In the following section, implied volatilities are described and the SFM is presented. In Section 2 we estimate the factor model on DAX index option data. Section 3 concludes. The appendix gives details on the issues of model choice and on the preparation of the data.

1 IVS MODELING USING A SEMIPARAMETRIC FACTOR MODEL

1.1 Model Characterization

Implied volatilities are derived from the BS option pricing formula for European calls and puts [Black and Scholes (1973)]. European style calls and puts are contingent claims on a stock S_t (paying no dividends for simplicity, here), which yield the pay-off $\max(S_T - K, 0)$ and $\max(K - S_T, 0)$, respectively, for a strike price K at a given expiry day T. The asset price process S_t in the BS model is assumed to be a geometric Brownian motion. The BS option pricing formula for calls is given by

$$C_t^{BS}(S_t, K, \tau, r, \sigma) = S_t \Phi(d_1) - e^{-r\tau} K \Phi(d_2), \tag{2}$$

where $d_1 = \frac{\log(S_t/K)+(r+0.5\sigma^2)\tau}{\sigma\sqrt{\tau}}$ and $d_2 = d_1 - \sigma\sqrt{\tau}$. Here, $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, $\tau = T - t$ time-to-maturity of the option, r the riskless interest rate over the option's life time, and σ the diffusion coefficient of the Brownian motion. Put prices P_t are obtained via the put–call parity $C_t - P_t = S_t - e^{-\tau r}K$. Given observed market prices $\tilde{C}_t$, one defines implied volatility $\hat{\sigma}$ by:

$$C_t^{BS}(S_t, K, \tau, r, \hat{\sigma}) - \tilde{C}_t = 0\ . \tag{3}$$

Because of monotonicity of the BS price in σ, there exists a unique solution $\hat{\sigma} > 0$ which can be found using numerical methods [see the discussion by Manaster and Koehler (1982)]. Practically the computation of Equation (3) can fail, in particular, when the data violate standard arbitrage arguments or when the options are very far in or very far out-of-the-money.

In the SFM, we regress log-implied volatility $Y_{i,j} = \log\{\hat{\sigma}_{i,j}(\kappa, \tau)\}$ on the covariates $X_{i,j} = (\kappa_{i,j}, \tau_{i,j})$, where $\kappa_{i,j} = K_{i,j}/F_{i,j}$ is a moneyness metric with $F_{i,j}$ as the futures price at time (i, j) and where $\tau_{i,j}$ denotes time-to-maturity. In borrowing ideas from fitting additive models, such as as in Stone (1986), Hastie and Tibshirani (1990) and Mammen et al. (1999), we will estimate the nonparametric components m_l and the state variables $\beta_{i,l}$ with least squares methods. This is motivated by rewriting Equation (1) as

$$Y_{i,j} = m_0(X_{i,j}) + \sum_{l=1}^{L} \beta_{i,l}\, m_l(X_{i,j}) + \varepsilon_{i,j}\ , \tag{4}$$

where $\varepsilon_{i,j}$ is mean zero, conditional on $X_{i,j}$ and independent of the stationary process β_i. From an empirical point of view, one might expect that the errors exhibit heteroskedasticity and correlation patterns that depend on time i and the spatial index j.[3] This suggests the use of weighted least squares as proposed in Renault (1997, Section 2.3). In our setting, weights that correct for spatial heteroskedasticity lead to second-order corrections of the estimate only, because of the localizing smoothing approach. Furthermore, correcting for heteroskedasticity in time would require to specify an additional model for these time effects. Therefore, we will use unweighted least squares.

In an option pricing framework such as (4) a statistical singularity emerges, when the number of observed asset prices exceeds the number of state variables. This is because the state variables can be identified exactly, when the number of observed option prices matches the number of state variables [see e.g., Pan (2002) for approaches of this kind]. Therefore, rather than interpreting the error as mispricings which could be exploited by arbitrage strategies, we follow the notion of Renault (1997) and perceive the error term as an option pricing error, that is, we interpret the observed option price as the price given by a pricing formula plus an error term which might be due to a neglected heterogeneity factor. Hence, the econometric specification is with respect to another martingale measure equivalent to the actual pricing measure. This notion allows us to stay within the Harrison and Kreps (1979) framework.

In specifying the factor functions m_l as functions of moneyness, the model is homogeneous of degree zero in strikes and spot. This allows us—although it is not explicitly stated—to more precisely characterize our underlying asset pricing model [Renault (1997)]. First, an implied volatility function homogeneous of degree zero in strikes and spot is equivalent to a (European) option pricing function that is homogeneous of degree one in strikes and spot.[4] Second, by this homogeneity property, the delta hedge ratio of the option pricing function can be decomposed into the BS delta minus a vega correction, which depends on the strike slope of the implied volatility function. This relationship allows splitting of the hedging error into a component due to the misspecification of the BS model and into another component due to the smile effect.[5] Finally, a necessary and sufficient condition for this homogeneity property is that future returns and the concurrent underlying asset price be conditionally independent under the risk-neutral pricing measure, given the currently available information other than the current underlying asset price. Therefore, it is possible to characterize our setting as a multifactor, stochastic volatility framework,

[3] The index j was introduced as an intra-day numbering of the observed options; but comprising the data for a given day i, as we do, one can interpret it in a spatial sense across the surface.

[4] Strictly speaking, this homogeneity property holds only in the absence of cash dividends; since the DAX index is a performance index, this assumption can be justified, see Section 2.1 for details.

[5] Detailed discussions of the implications for delta hedging in homogeneous models are given by Garcia and Renault (1998a) and Alexander and Nogueira (2007).

where the underlying asset price process does not cause the latent variables β_i in a Granger sense; for proofs, an illustration, and the precise formulation of the notion of conditional independence we refer to Garcia and Renault (1998b).

1.2 Estimation Algorithm

We define the estimates $\widehat{m}_l, (l = 0, \ldots, L)$ and $\widehat{\beta}_{i,l}\,(i = 1, \ldots, I;\ l = 1, \ldots, L)$ of model (4) as minimizers of the following least squares criterion ($\widehat{\beta}_{i,0} \stackrel{\text{def}}{=} 1$):

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i}\int \left\{Y_{i,j} - \sum_{l=0}^{L}\widehat{\beta}_{i,l}\widehat{m}_l(u)\right\}^2 K_h(u - X_{i,j})\, du\ . \tag{5}$$

Here, K_h denotes a two-dimensional product kernel, $K_h(u) = k_{h_1}(u_1) \times k_{h_2}(u_2)$, $h = (h_1, h_2)^\top$, based on the one-dimensional kernel $k_h(v) = h^{-1}k(h^{-1}v)$. In our applications we use the quartic kernel which is defined as $k(v) = \frac{15}{16}(1 - v^2)^2\ \mathbf{1}(|v| \leq 1)$, where $\mathbf{1}(\cdot)$ denotes the indicator function.

In (5) the minimization runs over all functions $\widehat{m}_l : \mathbb{R}^2 \to \mathbb{R}$ and all values $\widehat{\beta}_{i,l} \in \mathbb{R}$. Consider the case $L = 0$: the implied volatilities $Y_{i,j}$ are approximated by a surface $\widehat{m}_0$ that does not depend on time i. In this degenerated case, $\widehat{m}_0(u) = \sum_{i,j} K_h(u - X_{i,j})Y_{i,j} / \sum_{i,j} K_h(u - X_{i,j})$, which is the Nadaraya–Watson estimate based on the pooled sample of all days. In the case $L > 0$, the implied volatility surfaces are approximated by surfaces moving in an L-dimensional affine function space $\{\widehat{m}_0 + \sum_{l=1}^{L} \beta_l \widehat{m}_l :\ \beta_1, \ldots, \beta_L \in \mathbb{R}\}$. The estimates $\widehat{m}_l$ are not uniquely defined. They can be replaced by functions that span the same affine space. In order to respond to this problem, we select $\widehat{m}_l$ such that they are orthonormal. As shall be seen in Section 2, the orthonormalization also facilitates the interpretation of the functions.

In order to derive an estimator of the SFM we proceed in the following way. For an l' with $1 \leq l' \leq L$, for $\delta > 0$ and a function g, we replace in (5) $\widehat{m}_{l'}$ by $\widehat{m}_{l'} + \delta g$ and take the first derivative with respect to δ at the point $\delta = 0$. This yields

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i}\int \left\{Y_{i,j} - \sum_{l=0}^{L}\widehat{\beta}_{i,l}\widehat{m}_l(u)\right\} \widehat{\beta}_{i,l'} K_h(u - X_{i,j}) g(u) du = 0\ . \tag{6}$$

Since this equation holds for all functions g, it implies that for $1 \leq l' \leq L$ and for all u

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left\{Y_{i,j} - \sum_{l=0}^{L}\widehat{\beta}_{i,l}\widehat{m}_l(u)\right\} \widehat{\beta}_{i,l'} K_h(u - X_{i,j}) = 0\ . \tag{7}$$

Second, by replacing $\widehat{\beta}_{i',l'}$ by $\widehat{\beta}_{i',l'} + \delta$ in (4) and by taking derivatives with respect to δ we get, for $1 \leq l' \leq L$ and $1 \leq i' \leq I$:

$$\sum_{j=1}^{J_{i'}} \int \left\{ Y_{i',j} - \sum_{l=0}^{L} \widehat{\beta}_{i',l} \widehat{m}_l(u) \right\} \widehat{m}_{l'}(u) K_h(u - X_{i',j}) du = 0. \tag{8}$$

For convenience, we introduce the notation

$$\widehat{p}_i(u) = \frac{1}{J_i} \sum_{j=1}^{J_i} K_h(u - X_{i,j}) \,, \tag{9}$$

$$\widehat{q}_i(u) = \frac{1}{J_i} \sum_{j=1}^{J_i} K_h(u - X_{i,j}) Y_{i,j} \,, \tag{10}$$

for $1 \leq i \leq I$. The quantity $\widehat{p}_i(u)$ is recognized as the kernel density estimate of the data of day i, while $\widehat{q}_i(u)$ is a weighted sum of the dependent variable known from the numerator of a Nadaraya–Watson estimator. Then, we obtain from Equations (7)–(8), for $1 \leq l' \leq L, 1 \leq i \leq I$:

$$\sum_{i=1}^{I} J_i \widehat{\beta}_{i,l'} \widehat{q}_i(u) = \sum_{i=1}^{I} J_i \sum_{l=0}^{L} \widehat{\beta}_{i,l'} \widehat{\beta}_{i,l} \widehat{p}_i(u) \widehat{m}_l(u) \,, \tag{11}$$

$$\int \widehat{q}_i(u) \widehat{m}_{l'}(u) \, du = \sum_{l=0}^{L} \widehat{\beta}_{i,l} \int \widehat{p}_i(u) \widehat{m}_{l'}(u) \widehat{m}_l(u) \, du \,. \tag{12}$$

We calculate the estimates by iterative use of Equations (11) and (12). We start with initial values $\widehat{\beta}_{i,l}^{(0)}$ for $\widehat{\beta}_{i,l}$. A possible choice of the initial $\widehat{\beta}$ could correspond to fits of surfaces that are piecewise constant on time intervals $I_1, \ldots, I_L$. This means, for $l = 1, \ldots, L$, put $\widehat{\beta}_{i,l}^{(0)} = 1$ (for $i \in I_l$), and $\widehat{\beta}_{i,l}^{(0)} = 0$ (for $i \notin I_l$). Here $I_1, \ldots, I_L$ are pairwise disjoint subsets of $\{1, \ldots, I\}$ and $\bigcup_{l=1}^{L} I_l$ is a strict subset of $\{1, \ldots, I\}$. For $r \geq 0$, we put $\widehat{\beta}_{i,0}^{(r)} = 1$. Now, we define the matrix $B^{(r)}(u)$ by its elements

$$\left(B^{(r)}(u)\right)_{l,l'} \stackrel{\text{def}}{=} \sum_{i=1}^{I} J_i \widehat{\beta}_{i,l'}^{(r-1)} \widehat{\beta}_{i,l}^{(r-1)} \widehat{p}_i(u) \,, \qquad 0 \leq l,\ l' \leq L \,, \tag{13}$$

and introduce a vector $Q^{(r)}(u)$ with elements

$$Q^{(r)}(u)_l \stackrel{\text{def}}{=} \sum_{i=1}^{I} J_i \widehat{\beta}_{i,l}^{(r-1)} \widehat{q}_i(u) \,, \qquad 0 \leq l \leq L. \tag{14}$$

In the r-th iteration, the estimate $\widehat{m} = (\widehat{m}_0, \ldots, \widehat{m}_L)^\top$ is given by:

$$\widehat{m}^{(r)}(u) = B^{(r)}(u)^{-1} Q^{(r)}(u) \,. \tag{15}$$

This update step is motivated by Equation (11). The values of $\widehat{\beta}$ are updated in the r-th cycle as follows: for the matrix $M^{(r)}(i)$ with elements

$$\left(M^{(r)}(i)\right)_{l,l'} \stackrel{\text{def}}{=} \int \widehat{p}_i(u)\widehat{m}_{l'}^{(r)}(u)\widehat{m}_l^{(r)}(u)\,du\,, \qquad 1 \le l, l' \le L, \tag{16}$$

and for a vector $S^{(r)}(i)$ with elements

$$S^{(r)}(i)_l \stackrel{\text{def}}{=} \int \widehat{q}_i(u)\widehat{m}_l(u)\,du - \int \widehat{p}_i(u)\widehat{m}_0^{(r)}(u)\widehat{m}_l^{(r)}(u)\,du, \qquad 1 \le l \le L, \tag{17}$$

we compute

$$\left(\widehat{\beta}_{i,1}^{(r)}, \ldots, \widehat{\beta}_{i,L}^{(r)}\right)^\top = M^{(r)}(i)^{-1}S^{(r)}(i)\,, \tag{18}$$

which is motivated by Equation (12). The algorithm is iterated, until only minor changes in an L^2-fitting criterion occur (see Appendix A for the details). In the implementation, we choose a regularly spaced grid in the moneyness and time-to-maturity space and calculate $\widehat{m}_l$ for these points. In the calculation of $M^{(r)}(i)$ and $S^{(r)}(i)$, we replace the integrals by Riemann sums.

As discussed above, $\widehat{m}_l$ and $\widehat{\beta}_{i,l}$ are not uniquely defined. Therefore, we orthonormalize $\widehat{m}_0, \ldots, \widehat{m}_L$ in $L^2(\hat{p})$, where $\hat{p}(u) = I^{-1}\sum_{i=1}^{I}\hat{p}_i(u)$ is the average design density in our sample, and order them according to the variance of the corresponding factor loadings $\widehat{\beta}_l$. This can be achieved by the following two steps. First, replace

$$\begin{aligned}
\widehat{m}_0 \quad &\text{by} \quad & \widehat{m}_0^* &= \widehat{m}_0 - \gamma^\top\Gamma^{-1}\widehat{m}\,,\\
\widehat{m} \quad &\text{by} \quad & \widehat{m}^* &= \Gamma^{-1/2}\widehat{m}\,,\\
\begin{pmatrix} \widehat{\beta}_{i,1} \\ \vdots \\ \widehat{\beta}_{i,L} \end{pmatrix} \quad &\text{by} \quad & \begin{pmatrix} \widehat{\beta}_{i,1}^* \\ \vdots \\ \widehat{\beta}_{i,L}^* \end{pmatrix} &= \Gamma^{1/2}\left\{\begin{pmatrix} \widehat{\beta}_{i,1} \\ \vdots \\ \widehat{\beta}_{i,L} \end{pmatrix} + \Gamma^{-1}\gamma\right\},
\end{aligned} \tag{19}$$

where $\widehat{m} = (\widehat{m}_1, \ldots, \widehat{m}_L)^\top$ and the $(L \times L)$ matrix $\Gamma = \int \widehat{m}(u)\widehat{m}(u)^\top \hat{p}(u)\,du$. The vector γ has elements $\gamma_l = \int \widehat{m}_0(u)\widehat{m}_l(u)\hat{p}(u)du$. By applying steps (19), $\widehat{m}_0$ is replaced by a function that minimizes $\int \widehat{m}_0^2(u)\hat{p}(u)du$. This is evident because $\widehat{m}_0^*$ is orthogonal to the linear space spanned by $\widehat{m}_1, \ldots \widehat{m}_L$. By the second equation of steps (19), $\widehat{m}_1, \ldots, \widehat{m}_L$ are replaced by orthonormal functions in $L^2(\hat{p})$.

In a second step, we proceed as in PCA and define a matrix B^* with $B_{l,l'}^* = \sum_{i=1}^{I}\widehat{\beta}_{i,l}^*\widehat{\beta}_{i,l'}^*$. We calculate the eigenvalues of B^*, $\lambda_1 > \cdots > \lambda_L$, and the corresponding eigenvectors $z_1, \ldots z_L$. Put $Z = (z_1, \ldots, z_L)$. Replace

$$\widehat{m}^* \quad \text{by} \quad \widehat{m}^{**} = Z^\top\widehat{m}^*\,, \tag{20}$$

and

$$\begin{pmatrix} \widehat{\beta}^{*}_{i,1} \\ \vdots \\ \widehat{\beta}^{*}_{i,L} \end{pmatrix} \quad \text{by} \quad \begin{pmatrix} \widehat{\beta}^{**}_{i,1} \\ \vdots \\ \widehat{\beta}^{**}_{i,L} \end{pmatrix} = Z^{\top} \begin{pmatrix} \widehat{\beta}^{*}_{i,1} \\ \vdots \\ \widehat{\beta}^{*}_{i,L} \end{pmatrix}. \tag{21}$$

After application of (20) and (21), the orthonormal basis $\widehat{m}^{**}_1, \ldots, \widehat{m}^{**}_L$ is chosen so that $\widehat{\beta}^{**}_{i,1}$ has maximum variance, that is, $\widehat{m}^{**}_1$ is chosen so that as much as possible is explained by $\widehat{\beta}^{**}_{i,1}\,\widehat{m}^{**}_1$. Next $\widehat{m}^{**}_2$ is chosen to achieve maximum explanation by $\widehat{\beta}^{**}_{i,1}\,\widehat{m}^{**}_1 + \widehat{\beta}^{**}_{i,2}\,\widehat{m}^{**}_2$, and so forth, up to L. We point out that in all figures and tables and in the time series analysis to be presented in Section 2 we shall work with the quantities obtained after applying these orthonormalization steps.

The functions $\widehat{m}_l$ are not eigenfunctions of an operator as in the usual functional PCA [Ramsay and Silverman (1997)]. This is because we use a different norm, namely $\int f^2(u)\hat{p}_i(u)du$, for each day. Through the norming procedure, the functions are chosen as eigenfunctions in an L-dimensional approximating linear space. The L-dimensional approximating spaces are not necessarily nested for increasing L. For this reason the estimates cannot be calculated by an iterative procedure that starts by fitting a model with L components, and that uses these components in the iteration step from L to $L+1$ components. The calculation of $\widehat{m}_0, \ldots, \widehat{m}_L$ has to be redone for different choices of L.

2 THE FACTORS OF THE DAX INDEX IVS

2.1 Data Description and Preparation

The data set contains tick statistics on DAX futures contracts and DAX index options traded at the German-Swiss futures exchange EUREX in Frankfurt (Main) over the period from January 1998 to May 2001. The DAX option (ODAX) is of European type, is settled in cash, expires on the third Friday of the contract month, and has a minimum price movement of 0.1 index points. Expiries are the three nearest calendar months, the three following months of the cycle March, June, September, and December and the two following months of the cycle from June to December. For our purposes, we disregard data beyond half a year to expiry, as the liquidity is thin, in particular in the early years of our data set. The average daily trading volume over all strikes and expiries was about 140 000 contracts as of May, 2001. The futures contract is settled in cash as well, has a minimum price movement of half of an index point and the expiry months available are the three nearest months within the cycle March, June, September, December. The contract with the shortest time-to-maturity is by far the most liquid one, with an average daily trading volume of around 50 000 traded contracts as of May, 2001. In the data set, futures price and option price data are recorded per transaction, that is, each single trade is registered together with its price, contract size, and time of settlement (time is measured in seconds). The interest rate data we employ are one-, three-, and six-months FIBOR rates for the years 1998–1999 and EURIBOR rates[6] for the period 2000–2001 in daily frequency. The source is Thomson Financial Datastream. The interest rate data are linearly interpolated to approximate the riskless interest rate for the option's expiry. This is common practice [Dumas et al. (1998)].

Given the price data, implied volatilities have to be computed. To this end we cannot apply the Black (1976) formula, as it requires a liquid futures contract for each option expiry. This condition is not met by our data. Therefore, we work with the standard BS formula and derive implied DAX index values from the recorded futures prices. Since a futures price differs from the spot by the basis [Hull (2003, p. 75)], it may seem as a limitation that the analysis is based on futures prices rather than on spot prices. The DAX, however, is not a traded security: it is an index comprising 30 German blue chips. Thus, for option traders and market makers, trading the index at low transaction costs is possible only through futures contracts.[7] Backing out index implied volatilities from futures price data therefore is a natural process and corresponds to daily practice.

The implied DAX index values are recovered following Hafner and Wallmeier (2001), who use the same data: we pair each option price observation with the futures price F_t of the nearest available futures contract that was traded within a one-minute interval. The futures price observation is taken from the most heavily traded futures contract on that particular day. The no-arbitrage price of the underlying index in a frictionless market without dividends can be computed from the fair forward price formula via $S_t = F_t e^{-r_{T_F,t}(T_F - t)}$, where S_t and F_t denote the index and the futures price, respectively, T_F the expiry date of the futures contract, and $r_{T,t}$ the interest rate for the period $T - t$. It needs to be remarked that owing to the daily settlement of futures contracts, futures prices and forward prices are not equal, when interest rates are stochastic. However, for the time-to-maturities we consider in this work (up to half a year), we believe this difference to be negligible [see Hull (2003, p. 51--52) for a more detailed discussion and further references to this topic].

The DAX index is a capital-weighted performance index, that is, dividends less corporate tax are reinvested into the index [Deutsche Börse (2006)]. Therefore, dividend payments should have no impact on index options. However, when one uses only the discounted futures price to recover implied volatilities, implied volatilities of calls and puts can differ significantly. To accommodate for this fact, we apply a correction algorithm due to Hafner and Wallmeier (2001), which is described in Appendix B. The data are stored in the financial data base MD*base at the Center for Applied Statistics and Economics (CASE), Berlin.

Since trade-by-trade data may contain misprints or outliers, a filter is applied before estimating the model: observations with an implied volatility outside the

[6] FIBOR is the Frankfurt Interbank Offered Rate, which was a daily reference rate based on the interest rates at which banks offer to lend unsecured funds to other banks in the Deutsche Mark money market. With introducing the Euro in 1999, it was replaced by the EURIBOR, the European Interbank Offered Rate.

[7] To be very exact, there are now exchange-traded funds on the DAX that track the index remarkably close; they serve, however, as investment vehicles and are not used for hedging.

Table 1 Summary statistics of implied volatility data.

		Min.	Max.	Mean	Median	Stdd.	Skewn.	Kurt.
All	Time-to-maturity	0.028	2.014	0.131	0.083	0.148	3.723	23.373
	Moneyness	0.325	1.856	0.985	0.993	0.098	−0.256	5.884
	Implied volatility	0.041	0.799	0.279	0.256	0.090	1.542	6.000
1998	Time-to-maturity	0.028	2.014	0.134	0.081	0.148	3.548	22.957
	Moneyness	0.386	1.856	0.984	0.992	0.108	−0.030	5.344
	Implied volatility	0.041	0.799	0.335	0.306	0.114	0.970	3.471
1999	Time-to-maturity	0.028	1.994	0.126	0.083	0.139	4.331	32.578
	Moneyness	0.371	1.516	0.979	0.992	0.099	−0.595	5.563
	Implied volatility	0.047	0.798	0.273	0.259	0.076	0.942	4.075
2000	Time-to-maturity	0.028	1.994	0.130	0.083	0.151	3.858	23.393
	Moneyness	0.325	1.611	0.985	0.992	0.092	−0.337	6.197
	Implied volatility	0.041	0.798	0.254	0.242	0.060	1.463	7.313
2001	Time-to-maturity	0.028	0.978	0.142	0.083	0.159	2.699	10.443
	Moneyness	0.583	1.811	1.001	1.001	0.085	0.519	6.762
	Implied volatility	0.043	0.789	0.230	0.221	0.049	1.558	7.733

Summary statistics of the implied volatility data computed from DAX index option data for the period from January 1998 to May 2001, entire data set and annual subsamples. The figures are computed after filtering and data preparation.

range [4%, 80%] are dropped. Furthermore, we disregard all observations with a maturity of less than ten days [for similar approaches see Dumas et al. (1998) and Skiadopoulos et al. (1999)]. The filtering removed 635 000 observations from file, leaving us with a total number of about 2.24 million contracts, that is, around 2 600 observations per daily IVS. Clearly, this filter does not remove all potential outliers, but since our methodology relies on data of different observation days for the estimation of the factor functions, it is robust against spurious data points.

Table 1 gives a short summary of our IVS data. Most heavy trading occurs in short-term contracts, as is seen from the skewness of the term structure distribution of the observations. Median time-to-maturity is 30 days (0.083 years). Across moneyness, the distribution is slightly negatively skewed. This is because out-of-the-money puts are more heavily traded than out-of-the-money calls. Mean implied volatility over the sample period is 27.9%.

2.2 Model Setup

The implementation of the SFM requires decisions on a number of issues, most importantly the model size and the bandwidths. In this section, we outline the general procedure. Technicalities are transferred to Appendix A.

Because of skewness and positivity constraints, we regress log-implied volatility $Y_{i,j}$ on $X_{i,j} = (\kappa_{i,j}, \tau_{i,j})^\top$. The grid is regularly spaced and covers in moneyness $\kappa \in [0.80, 1.20]$ and in time-to-maturity $\tau \in [0.05, 0.5]$ measured in years. To determine the model size L, we compute the portion of variance explained $1 - RV(L)$, where $RV(L)$ is the residual sum of squares per total variance [Equation (A1) in Appendix A]. We find that up to 89% of the variation are explained by $L = 1$ basis function, 94% by $L = 2$, and 97% by $L = 3$ basis functions. These findings are in line with the textbook literature of the standard PCA on the IVS [see e.g., Alexander (2001)]. Therefore, we decide for $L = 3$ basis functions as a sufficiently good approximation to the IVS. For measuring the goodness of fit and the progress of convergence we compute the L^2-measure Equation (A5) in the Appendix. Convergence is typically achieved after 25 cycles.

The choice of the smoothing parameter is essential in the practice of non- and semiparametric modeling. The literature offers a number of approaches, such as cross-validation procedures, penalization and bootstrap techniques, or plug-in methods for determining the bandwidths [see Härdle (1990, Section 5) for detailed discussions]. Since in this study we are interested in fitting the IVS, a natural approach for bandwidth choice can be based on the Akaike (1970) Information Criterion (AIC) and on the Schwarz (1978) Criterion (SC). More precisely, we work with the two computationally slightly differing versions of a weighted AIC and SC given in Equations (A3) and (A4). The reason for using weighted criteria is the uneven distribution of the data, which can lead to nonconvexities and which might result in unacceptably small bandwidths [see Fengler et al. (2003) for a first description of this problem]. We choose as weight function $p^{-1}(u)$ where $p(u)$ is the average design density. It can be shown that this choice ensures an equal weight everywhere.

Table 2 presents the criteria for a grid of bandwidths. It is seen—from bottom to top—that for a given bandwidth in moneyness (h_1) all criteria drop as the bandwidth in time-to-maturity (h_2) decreases, but that the criteria increase as the latter approaches the length of two weeks, i.e. when $h_2 \approx 0.03$ years. Criterion Ξ_{AIC_1} assumes its global minimum at $h = (h_1, h_2) = (0.04, 0.05)^\top$, Ξ_{SC_1} at $h = (0.04, 0.06)^\top$. The second type of criteria proposes much smaller bandwidths: Ξ_{AIC_2} suggests $h = (0.02, 0.03)^\top$ and Ξ_{SC_2} proposes $h = (0.03, 0.04)^\top$, respectively. Since for the second criteria the functions still look quite volatile, we prefer the suggestion of the SC among the first group of criteria, that is, we take Ξ_{SC_1} at $h^* = (0.04, 0.06)^\top$ as our optimal bandwidth vector. For further analysis, we check how the factor loadings and the basis functions change relative to this optimal bandwidth h^*, columns seven and eight in Table 2. As is seen, the bandwidth suggestions of Ξ_{AIC_1} and Ξ_{SC_1} are hardly distinguishable and the estimates are stable.

The choices for h and L are not independent. From this point of view, one may think about minimizing the information criteria over both parameters. Practical experience shows that for a given L, changes in the criteria from variations in h are small, compared to variations in L for a given h. Therefore, separating the two selection procedures can be justified. As an alternative to our constant bandwidth, one could consider bandwidths that are functions of the grid $h = h(u)$. This is discussed in Borak et al. (2005).

Table 2 Bandwidth selection for the factor model.

h_1	h_2	$\Xi_{AIC_1} \times 10^4$	$\Xi_{AIC_2} \times 10^3$	$\Xi_{SC_1} \times 10^4$	$\Xi_{SC_2} \times 10^3$	$V_{\widehat{\beta}}$	$V_{\widehat{m}}$
0.01	0.02	8.127	1.7562	8.3540	2.0461	0.042	0.096
0.01	0.03	5.386	1.3717	5.4825	1.5142	0.039	0.069
0.01	0.04	5.381	1.3732	5.4518	1.4762	0.038	0.043
0.01	0.05	5.373	1.3816	5.4282	1.4625	0.038	0.028
0.01	0.06	5.371	1.3952	5.4162	1.4620	0.038	0.021
0.01	0.07	5.376	1.4134	5.4144	1.4707	0.038	0.019
0.01	0.08	5.390	1.4369	5.4235	1.4874	0.038	0.019
0.02	0.02	8.037	1.7338	8.1459	1.8689	0.042	0.090
0.02	0.03	5.347	1.3560*	5.3946	1.4239	0.039	0.035
0.02	0.04	5.346	1.3636	5.3803	1.4134	0.038	0.020
0.02	0.05	5.345	1.3746	5.3718	1.4140	0.002	0.014
0.02	0.06	5.346	1.3893	5.3687	1.4220	0.001	0.010
0.02	0.07	5.353	1.4081	5.3718	1.4362	0.001	0.011
0.02	0.08	5.367	1.4319	5.3841	1.4568	0.002	0.014
0.03	0.02	8.009	1.7331	8.0811	1.8213	0.042	0.090
0.03	0.03	5.329	1.3585	5.3598	1.4031	0.038	0.025
0.03	0.04	5.327	1.3667	5.3494	1.3996*	0.002	0.015
0.03	0.05	5.326	1.3777	5.3435	1.4038	0.001	0.008
0.03	0.06	5.327	1.3921	5.3417	1.4138	0.000	0.005
0.03	0.07	5.333	1.4107	5.3457	1.4293	0.001	0.007
0.03	0.08	5.348	1.4342	5.3586	1.4507	0.002	0.012
0.04	0.02	8.018	1.7476	8.0720	1.8135	0.042	0.090
0.04	0.03	5.321	1.3743	5.3446	1.4079	0.038	0.022
0.04	0.04	5.318	1.3821	5.3354	1.4069	0.002	0.012
0.04	0.05	5.316*	1.3925	5.3294	1.4121	0.001	0.006
0.04	**0.06**	5.316	1.4062	**5.3273***	1.4226	n.a.	n.a.
0.04	0.07	5.322	1.4242	5.3312	1.4382	0.001	0.006
0.04	0.08	5.336	1.4472	5.3440	1.4596	0.002	0.011
0.05	0.02	8.086	1.7833	8.1292	1.8366	0.041	0.092
0.05	0.03	5.333	1.4083	5.3515	1.4357	0.038	0.020
0.05	0.04	5.328	1.4154	5.3420	1.4355	0.026	0.012
0.05	0.05	5.324	1.4249	5.3350	1.4409	0.026	0.007
0.05	0.06	5.323	1.4378	5.3320	1.4511	0.001	0.005
0.05	0.07	5.328	1.4550	5.3355	1.4664	0.001	0.008
0.05	0.08	5.341	1.4772	5.3479	1.4873	0.002	0.013
0.06	0.02	8.183	1.8424	8.2190	1.8880	0.041	0.094
0.06	0.03	5.373	1.4667	5.3887	1.4903	0.038	0.020
0.06	0.04	5.367	1.4728	5.3781	1.4902	0.026	0.014
0.06	0.05	5.361	1.4815	5.3697	1.4952	0.026	0.010
0.06	0.06	5.358	1.4934	5.3656	1.5048	0.026	0.009
0.06	0.07	5.362	1.5096	5.3685	1.5195	0.002	0.011
0.06	0.08	5.375	1.5309	5.3806	1.5396	0.003	0.015

Bandwidth selection for different choices of h_1 (moneyness) and h_2 (time-to-maturity in years); model size $L = 3$. For each selection criterion the optimal value is labeled with an asterisk, the optimal bandwidth eventually chosen according to SC_1 (column 5) is printed in bold. $V_{\widehat{\beta}}$ and $V_{\widehat{m}}$ measure the standard deviation of $\widehat{\beta}_h$ and $\widehat{m}_h$ relative to $\widehat{\beta}_{h^*}$ and $\widehat{m}_{h^*}$, i.e. $V^2_{\widehat{\beta}}(h_k) = (LI)^{-1}\sum_{l,i}(\widehat{\beta}_{i,l}(h_k) - \widehat{\beta}_{i,l}(h^*))^2$ and $V^2_{\widehat{m}}(h_k) = (LN)^{-1}\sum_{l,n}(\widehat{m}_l(u_n, h_k) - \widehat{m}_l(u_n, h^*)du)^2$, where $n = 1, \ldots, N$ runs over all points in the estimation grid and h_k over the values displayed in column one and two.

2.3 Estimation Results and VAR Modeling

Unlike the traditional literature on smoothing the IVS, our approach exploits the expiry effect of the implied volatility data and thus avoids global daily fits. Indeed, our bandwidth in the time-to-maturity dimension is so small that in a fit of a particular day, data belonging to contracts with two adjacent expiries do not enter together $\widehat{p}_i(u)$ in Equation (9) and $\widehat{q}_i(u)$ in Equation (10). To visualize the mechanics behind the algorithm consider Figure 4. Here, we illustrate the situation for a small number of trading days and for four highlighted grid points, which are depicted by the big bullets. The neighborhood that is realized using the kernel functions is sketched by the ellipses. It is seen for a given u', that is, for a given bullet point, that the quantities $\widehat{p}_i(u')$ and $\widehat{q}_i(u')$ must be zero most of the time, and only assume positive values for dates i, when observations are in the local neighborhood of u'. This happens because of the simple fact that the contracts move towards expiry.

Figures 5 and 6 display the estimated factor functions $\widehat{m}_1$ to $\widehat{m}_3$ after orthonormalization in $L^2(\hat{p})$. We do not display the invariant function $\widehat{m}_0$, as it turns out to be the zero function of the affine space fitted by the data. The remaining three functions exhibit interesting patterns: $\widehat{m}_1$ in Figure 5 is negative throughout, U-shaped and roughly constant across the term structure. Since this function belongs to the weights with highest variance, it can be interpreted as the time-dependent mean of the (log)-IVS, that is, as a shift function. Function $\widehat{m}_2$, depicted in Figure 6, changes sign around the at-the-money region, which implies that it drives the smile deformations of the IVS. Hence, it is the moneyness

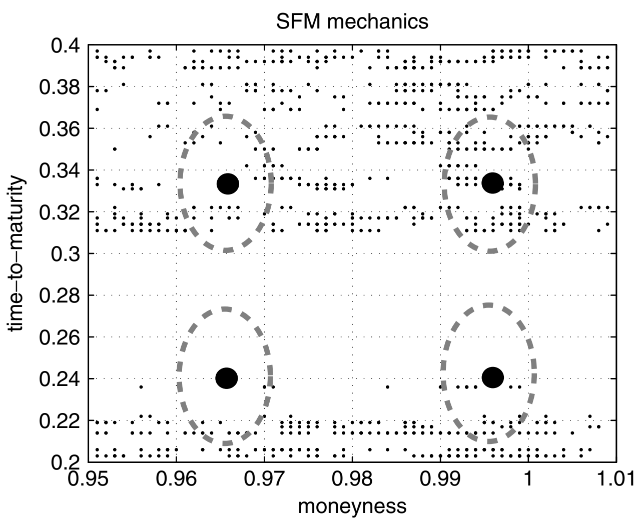

Figure 4 Sketch of the computational mechanics behind the SFM algorithm. The big bullets denote points from the estimation grid. The neighborhood is realized via the kernel functions and illustrated by the dotted ellipses. The small dots denote implied volatility data as observed for a sequence days, February to March 1998.

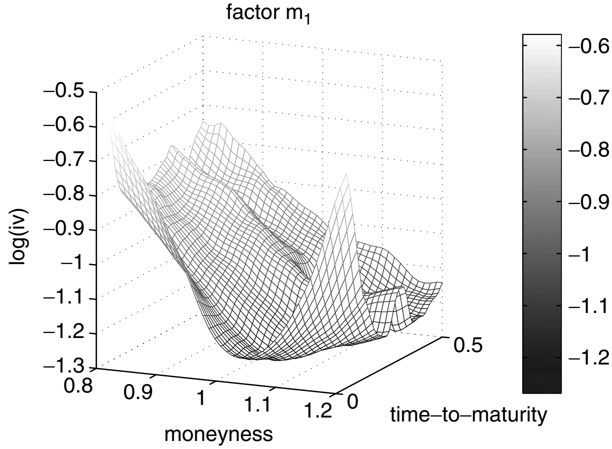

Figure 5 Factor $\widehat{m}_1$ of the SFM after orthonormalization in $L^2(\hat{p})$, where $\hat{p}(u) = I^{-1}\sum_{i=1}^{I}\hat{p}_i(u)$ is the average design density, that is, after applying steps (19) to (21). Bandwidths $h = (0.04, 0.06)$ chosen optimally according to SC_1, cf. Table 2. Estimated from ODAX data from January 1998 to May 2001.

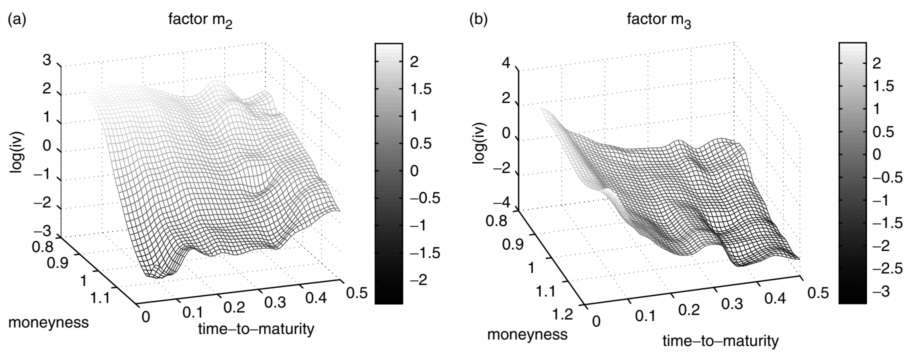

Figure 6 Factor $\widehat{m}_2$ (left) and $\widehat{m}_3$ (right) of the SFM after orthonormalization in $L^2(\hat{p})$, i.e. after applying steps (19) to (21). Bandwidths $h = (0.04, 0.06)$ chosen optimally according to SC_1, cf. Table 2. Estimated from ODAX data from January 1998 to May 2001.

slope function of the IVS. Finally, $\widehat{m}_3$ (also Figure 6) is positive for the very short term contracts and negative for contracts with a time-to-maturity of more than 0.1 years. Consequently, $\widehat{m}_3$ generates the term structure dynamics of the IVS [for similar interpretations see Skiadopoulos et al. (1999), Cont and da Fonseca (2002), and Fengler et al. (2003)].

To give an impression of the properties of the SFM, we compare the fit of the SFM with that of a Nadaraya–Watson estimator for a randomly chosen date from the sample, June 18, 1998. For the bandwidth choice in case of the Nadaraya–Watson estimator we compute a classical AIC and SC which

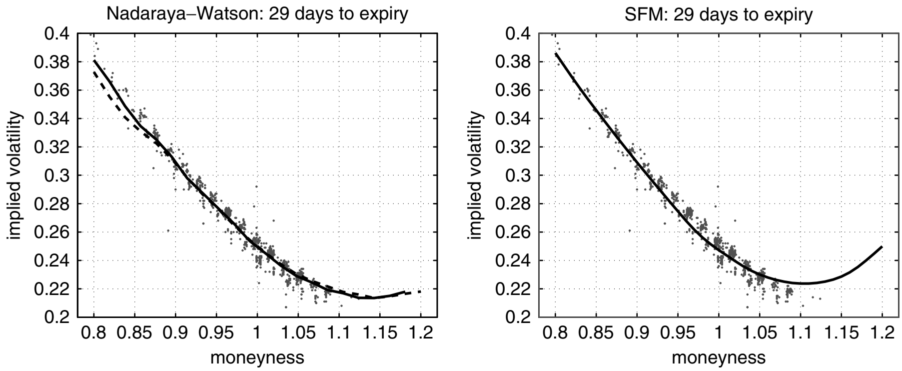

Figure 7 Fit comparisons. Left panel: Nadaraya–Watson estimator with bandwidths $h^*_{NW:AIC} = (h_1, h_2)^\top = (0.04, 0.12)^\top$ (solid line) and $h^*_{NW:SC} = (0.05, 0.19)^\top$ (dotted line), where h_1 is the bandwidth in moneyness and h_2 for time-to-maturity. Right panel: SFM using bandwidths $h^* = (0.04, 0.06)^\top$. Observed implied volatility data shown as gray dots, time-to-maturity is 29 days, data from June 18, 1998.

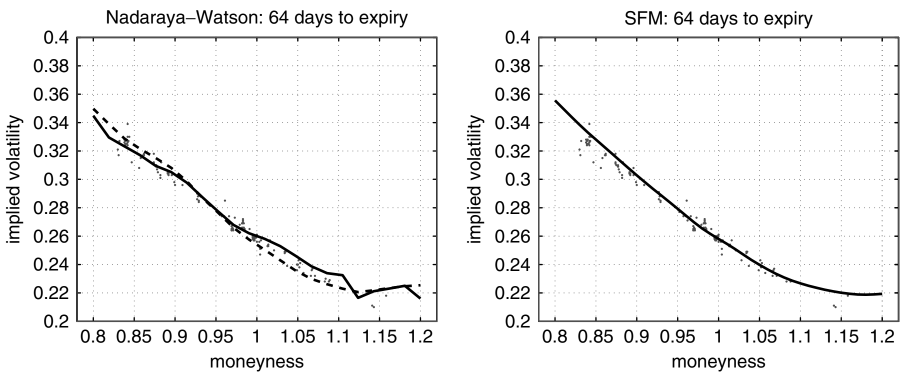

Figure 8 Fit comparisons. Left panel: Nadaraya–Watson estimator with bandwidths $h^*_{NW:AIC} = (h_1, h_2)^\top = (0.04, 0.12)^\top$ (solid line) and $h^*_{NW:SC} = (0.05, 0.19)^\top$ (dotted line), where h_1 is the bandwidth in moneyness and h_2 for time-to-maturity. Right panel: SFM using bandwidths $h^* = (0.04, 0.06)^\top$. Observed implied volatility data shown as gray dots, time-to-maturity is 64 days, data from June 18, 1998.

yield $h^*_{NW:AIC} = (h_1, h_2)^\top = (0.04, 0.12)^\top$ and $h^*_{NW:SC} = (0.05, 0.19)^\top$, respectively, as optimal bandwidths for moneyness h_1 and time-to-maturity h_2. Figures 7 to 10 show the slices through the surfaces at the observed expiries. Before studying the differences, we recall that the Nadaraya–Watson estimator is based on data of one single day, whereas the SFM takes advantage of the entire sample. Therefore, any reasonable comparison is necessarily unbalanced. Nevertheless, we believe that such a comparison can reveal the principal advantages the SFM offers over using a standard nonparametric estimator for fitting the IVS.

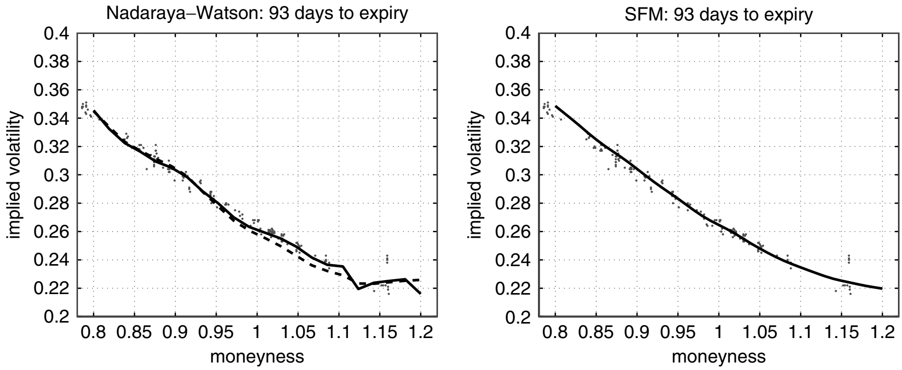

Figure 9 Fit comparisons. Left panel: Nadaraya–Watson estimator with bandwidths $h^*_{NW:AIC} = (h_1, h_2)^\top = (0.04, 0.12)^\top$ (solid line) and $h^*_{NW:SC} = (0.05, 0.19)^\top$ (dotted line), where h_1 is the bandwidth in moneyness and h_2 for time-to-maturity. Right panel: SFM using bandwidths $h^* = (0.04, 0.06)^\top$. Observed implied volatility data shown as gray dots, time-to-maturity is 93 days, data from June 18, 1998.

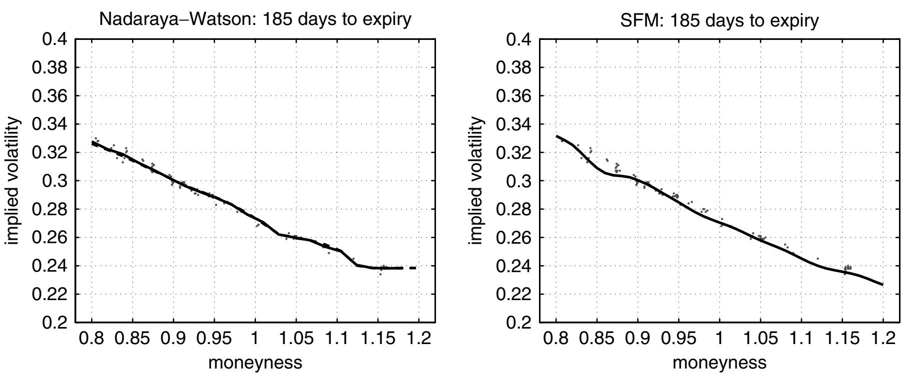

Figure 10 Fit comparisons. Left panel: Nadaraya–Watson estimator with bandwidths $h^*_{NW:AIC} = (h_1, h_2)^\top = (0.04, 0.12)^\top$ (solid line) and $h^*_{NW:SC} = (0.05, 0.19)^\top$ (dotted line), where h_1 is the bandwidth in moneyness and h_2 for time-to-maturity. Right panel: SFM using bandwidths $h^* = (0.04, 0.06)^\top$. Observed implied volatility data shown as gray dots, time-to-maturity is 185 days, data from June 18, 1998.

Figures 7 to 10 show that despite the smaller bandwidths, the SFM fit is smoother than the Nadaraya–Watson fit, in particular in comparison with the Nadaraya–Watson estimator based on the bandwidths as suggested by the AIC (solid lines). Necessarily, the Nadaraya–Watson estimate must be rough when the data are sparse, such as for the moneyness region greater than 1.1 in Figures 8 and 9. However, increasing the bandwidths to the size as proposed by the SC results in directional biases: for moneyness greater than 1.1 in Figures 8 and 9 the Nadaraya–Watson estimator (dotted line) tends to be below the actual data cloud. The SFM is not entirely unbiased either, as is seen in a small bump around

Table 3 Summary statistics of the factor series.

	Min.	Max.	Mean	Median	Stdd.	Skewn.	Kurt.
$\widehat{\beta}_1$	−1.504	−0.454	−1.191	−1.229	0.200	1.099	4.079
$\widehat{\beta}_2$	−0.109	0.074	−0.002	−0.003	0.033	0.009	2.710
$\widehat{\beta}_3$	−0.085	0.073	−0.002	0.002	0.023	−0.494	3.319

Descriptive summary statistics on the time series of the factor loadings $(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3)^\top$ after application the orthonormalization of the factor functions, i.e. after applying steps (19) to (21). Data from January 1998 to May 2001.

moneyness 0.85 in Figure 10, but overall, the fit is superior. Another noteworthy feature can be observed in Figure 7 for moneyness bigger than 1.1. Note that on this day and for this time-to-maturity almost no observations are found beyond that point and the estimates of the Nadaraya–Watson estimator and the SFM are an extrapolation. The Nadaraya–Watson yields a flat extrapolation, which is due to the flat structure in the second time-to-maturity (presented in Figure 8), while the SFM produces the typical smile pattern. Of course, one cannot tell which extrapolation is the better one. But observing that the extrapolation of the SFM is based on a large sample and on data observations from the local neighborhood (though from other dates), the smiley SFM extrapolation may be considered as the more likely functional structure. This interpretation corresponds to the general knowledge about smile patterns in equity option data [see e.g. Rubinstein (1985) and Tompkins (2001)].

Figure 11 shows the time series of $\widehat{\beta}_1$ to $\widehat{\beta}_3$ and their correlograms.[8] By comparing the plot of $\widehat{\beta}_1$ with the time series of the mean ATM-implied volatility in Figure 2, the first factor loading is recognized as a shifted and reflected version of mean implied volatility levels. This observation neatly matches the aforementioned interpretation of the first factor function as a mean function of the IVS. Summary statistics of the loading time series are given in Table 3. An augmented Dickey–Fuller (ADF) test indicates a unit root for $\widehat{\beta}_1$ and $\widehat{\beta}_2$ at the 5% level (left part of Table 4). Thus, one can model first differences of the first two loading series. Alternatively, since the test statistics are not strongly significant, a model in levels is an equally good choice. Moreover, one avoids possible over-differencing. Therefore we prefer a VAR model of order p in the levels of the series[9], that is,

$$\beta_t = c + A_1 \beta_{t-1} + \cdots + A_p \beta_{t-p} + u_t \ , \tag{22}$$

[8] For all figures and in all subsequent steps in the time series modeling the factor loadings obtained after the orthonormalization of the factor functions are used, see Section 1.

[9] In slight abuse of notation, we remove the ''hat'' notation of the β from now on. We point out that the VAR estimation is carried out on the estimated series $\widehat{\beta}_t$.

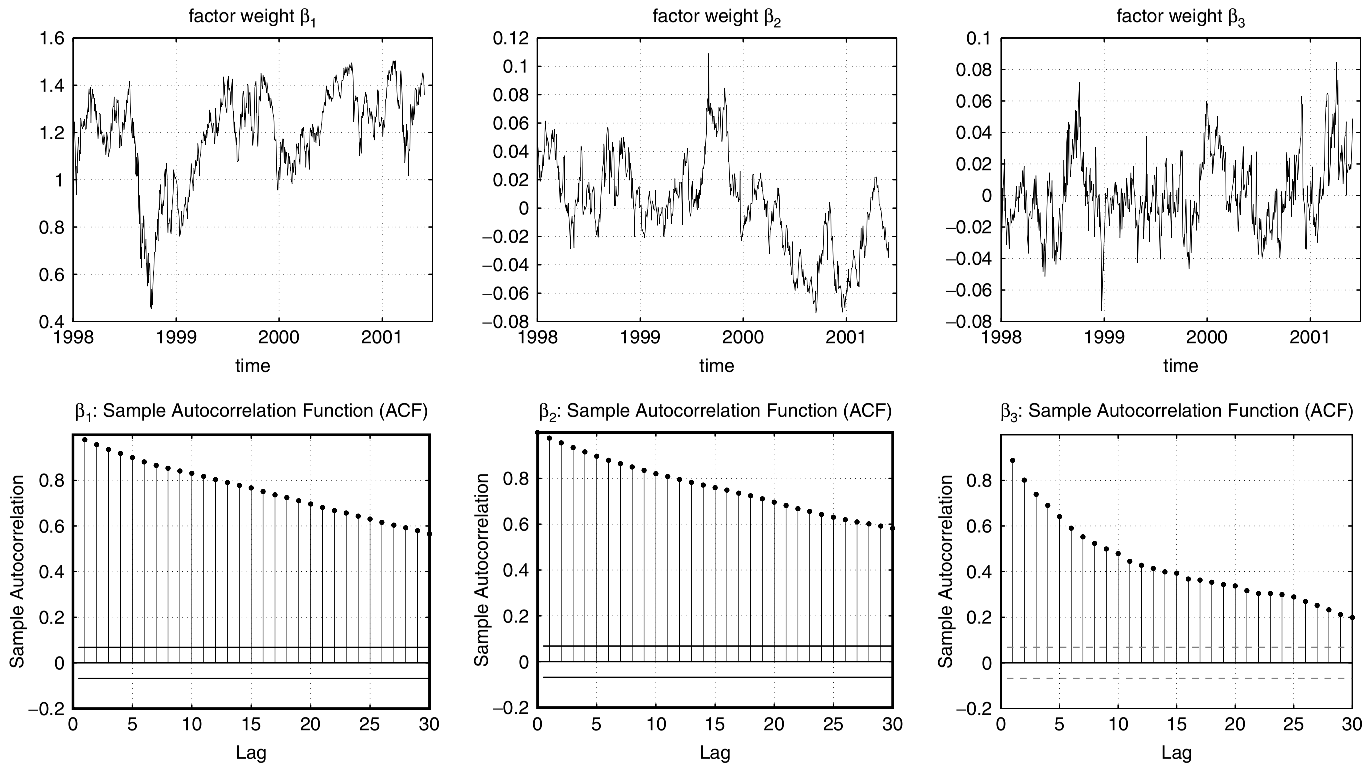

Figure 11 Upper panel: Time series of weights $(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3)^\top$ after application of the steps (19) to (21), i.e. after orthonormalization of the factor functions. Bandwidths $h = (0.04, 0.06)$ chosen according to SC_1, cf. Table 2. Lower panel: autocorrelation functions; dotted horizontal lines indicate 95% confidence interval under the null hypothesis that the data are white noise.

Table 4 ADF diagnostics and correlation.

	ADF test		Contemp. correlation		
Coefficient	Test statistic	# of lags	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$
$\widehat{\beta}_1$	-2.67	3	1.00	0.27^{***}	0.39^{***}
$\widehat{\beta}_2$	-3.01^{**}	1		1.00	0.00
$\widehat{\beta}_3$	-5.63^{***}	2			1.00

Left part: Augmented Dickey−Fuller (ADF) tests on $\widehat{\beta}_1$ to $\widehat{\beta}_3$ for the full IVS model, intercept included in each case. Third column gives the number of lags included in the ADF regression. For the choice of lag length, we start with four lags, and subsequently delete lag terms, until the last lag term is significant at least at the 5% level. MacKinnon critical values for rejecting the hypothesis of a unit root are −2.87 at the 5% significance level, and −3.44 at the 1% significance level. Right part: Contemporaneous correlation matrix; three asterisks denote an estimate significantly different from zero at the 1% level.

Table 5 Size of VAR model.

Suggested lag order	LR	AIC	HQ	SC
p	2	2	2	1

Choice of lag order p in VAR as suggested by the sequential modified likelihood ratio (LR) test, the Akaike (AIC), Hannan-Quinn (HQ), and Schwarz (SC) information criterion with maximum lag order $p_{\max} = 12$; for details on these criteria in the context of VAR see Lütkepohl (1991).

where c is an $L \times 1$ vector of intercepts, $A_i, i = 1, ..., p$ are $L \times L$ parameter matrices and $u_t = (u_{1t}, u_{3t}, u_{3t})^\top$ an unobservable error term with mean zero and nonsingular covariance matrix Σ_u.

For the choice of the lag length p in VAR models, we employ standard information criteria as described, for example, in Lütkepohl (1991, Chap. 4.3). The sequential modified likelihood ratio test, the Akaike and Hannan-Quinn information criteria suggest a lag length of order $p = 2$, except for the SC, which prefers $p = 1$, see Table 5. Since it is known that the SC tends to propose lag orders which are too low to capture the salient dynamics in VAR models, we choose a VAR(2) model.

The residual diagnostics are given in Table 6. The adjusted Box−Pierce−Ljung Q-statistic and the LM test on serial correlation do not indicate the presence of serial correlation up to order 4 and 8, respectively. White's test for heteroskedasticity [White (1980)] rejects the null hypothesis of homoskedasticity in the residuals. Therefore we report in Table 7, together with the estimated coefficient matrix, heteroskedasticity-consistent standard errors which are obtained by applying a correction in the spirit of White (1980) [see Hafner and Herwartz (2002) for the consistent estimation of covariance matrices in heteroskedastic VAR models]. Besides the expected autoregressive patterns in own lagged observations, there are a number of remarkable cross dynamics. For instance, a positive shock in

Table 6 Portmanteau statistics.

	Q(4)	Q(8)	LM(4)	LM (8)
Test statistic	26.45	57.90	6.15	3.46
p-value	0.090	0.333	0.725	0.943
df of χ^2	18	54	9	9

We report the adjusted multivariate Box–Pierce–Ljung Q-statistics for residual serial correlation and the multivariate LM test statistic described in Greene (2003, Chap. 12.7). The number in parentheses gives the order up to which there is no serial correlation under the null hypothesis; all tests are chi-squared distributed with degrees of freedom (df) as given.

Table 7 VAR model.

	Equation		
Dependent variable	$\beta_{1,i}$	$\beta_{2,i}$	$\beta_{3,i}$
$\beta_{1,i-1}$	0.958***	0.002	−0.023*
	(0.049)	(0.010)	(0.014)
$\beta_{1,i-2}$	0.024	−0.005	0.024*
	(0.049)	(0.011)	(0.013)
$\beta_{2,i-1}$	−0.150	0.902***	0.114*
	(0.235)	(0.066)	(0.067)
$\beta_{2,i-2}$	0.104	0.071*	−0.109*
	(0.235)	(0.064)	(0.066)
$\beta_{3,i-1}$	−0.223	−0.055*	0.795***
	(0.157)	(0.046)	(0.062)
$\beta_{3,i-2}$	0.329*	0.053*	0.107*
	(0.155)	(0.043)	(0.058)
c	−0.022***	−0.004**	0.001
	(0.011)	(0.002)	(0.003)
$\overline{R}^2$	0.958	0.956	0.797

Estimation results of an VAR(2) of the factor loadings β_i and an intercept denoted by c. Heteroskedasticity-consistent standard errors are given in parentheses. $\overline{R}^2$ denotes the adjusted coefficient of determination. Three asterisks denote significance at a 1% level, two asterisks significance at a 5% level, and one significance at a 10% level.

β_1, that is, in the level component of the IVS, has a negative impact on the term structure component β_3. Given the shape of $\widehat{m}_3$, this shock implies a flattening term structure. This interpretation is well understood by noting that the IVS is typically upward sloping and that an increase in the overall risk perception of the market causes short-term implied volatilities to rise much stronger than long-term implied volatilities. This pattern necessarily implies a flatter term structure; furthermore, this case has to be distinguished from a positive shock in downside risk, that is, β_2, which implies a more pronounced term structure (positive impact on β_3).

There are also significant relationships from lagged β_3 on β_1 and β_2. However, an impulse–response analysis (not shown for want of space) suggests that only the shocks on β_3 have a permanent impact, while shocks from β_3 on β_1 and β_2 have none. Contrary to shocks on the term structure, it is to be observed that the interactions between the lags of level and moneyness shocks, that is, β_1 and β_2, are not significant, while their contemporaneous observations are significantly positively correlated (see the right part of Table 4). Thus, positive shocks tend to increase both the levels, β_1, and the downside risk, β_2, always instantly. This is common market experience. Overall, the VAR(2) model appears to capture the dynamics of the factor loadings, which generate the propagation of the IVS through time, in a suitable and well-interpretable way.

3 SUMMARY AND OUTLOOK

In this study we present a modeling approach to the IVS called SFM. Unlike other methods, our approach is tailored to the degenerated design of implied volatility data by fitting basis functions in the local neighborhood of the design points only. In comparison to a Nadaraya–Watson estimator, we demonstrate that the SFM can mitigate bias effects. Using transactions-based DAX index implied volatility data from January 1998 to May 2001, we recover three basis functions and three time series of factor loadings that generate the dynamics of the IVS. The factor functions allow for an interpretation as level, moneyness, and term structure slope effects. We study the time series properties of the parameter weights and complete the modeling approach by fitting vector autoregressive models. Standard lag-length criteria suggest that the VAR(2) model captures best the dynamics of the three-dimensional factor series.

There are important topics for further research. First, given the VAR(2) time series model of the factor loadings, the SFM can be integrated into scenario and worst case simulations of portfolios, for example, along the lines of Jamshidian and Zhu (1997). The SFM can be used to generate plausible deformations of the IVS, which should allow for better vega hedges of option portfolios. Second, from a theoretical point of view, the SFM has a natural relation with Kalman filtering. Writing our model more compactly as

$$\Theta_p(B)\beta_i = u_i \tag{23}$$

$$Y_{i,j} = m_0(X_{i,j}) + \sum_{l=1}^{L} \beta_{i,l} m_l(X_{i,j}) + \varepsilon_{i,j}\ , \tag{24}$$

where u_i and $\varepsilon_{i,j}$ are noise, and where $\Theta_p(B) \stackrel{\text{def}}{=} 1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_p B^p$ denotes a polynomial of order p in the backshift operator B, we get the state-space representation. The state equation (23) depending on a parameter vector θ relates the (unobservable) state i of the system to the previous step $i-1$, while the measurement equation (24) relates the state to the measurement, the IVS in our case. In difference to our work the time series modeling of β_i could be achieved jointly with the estimation of the factor functions.

We think that this modeling approach has potential in other fields of financial research, most naturally in interest rate modeling. Since interest rates are inferred from bond prices, there too exists an expiry effect. In this sense, the SFM might complement the modeling strategies proposed by Ghysels and Ng (1989), Linton et al. (2001), and Diebold and Li (2006).

APPENDIX A: MODEL SELECTION

Here, we outline the selection procedures for the SFM. For the model size we use the residual sum of squares for different L:

$$RV(L) \stackrel{\text{def}}{=} \frac{\sum_i^I \sum_j^{J_i} \left\{ Y_{i,j} - \sum_{l=0}^{L} \widehat{\beta}_{i,l}\, \widehat{m}_l(X_{i,j}) \right\}^2}{\sum_i^I \sum_j^{J_i} (Y_{i,j} - \overline{Y})^2}\,, \tag{A1}$$

where $\overline{Y}$ denotes the overall mean of the observations.

For a data-driven choice of bandwidths, we propose a weighted AIC, since the distribution of the observations is very uneven. For a given weight function w, consider:

$$\triangle(m_0, \ldots, m_L) \stackrel{\text{def}}{=} \mathsf{E}\frac{1}{N} \sum_{i,j} \{Y_{i,j} - \sum_{l=0}^{L} \beta_{i,l} m_l(X_{i,j})\}^2\, w(X_{i,j})\,, \tag{A2}$$

for functions $m_0, \ldots, m_L$. The expectation operator is denoted by E. We choose bandwidths such that $\triangle(\widehat{m}_0, \ldots, \widehat{m}_L)$ is minimal. According to the AIC this is asymptotically equivalent to minimizing:

$$\Xi_{AIC_1} \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i,j} \{Y_{i,j} - \sum_{l=0}^{L} \widehat{\beta}_{i,l} \widehat{m}_l(X_{i,j})\}^2\, w(X_{i,j}) \exp\left\{ 2\frac{L}{N} K_h(0) \int w(u) du \right\}\,. \tag{A3}$$

Alternatively, one might consider the computationally more convenient criterion:

$$\Xi_{AIC_2} \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i,j} \{Y_{i,j} - \sum_{l=0}^{L} \widehat{\beta}_{i,l} \widehat{m}_l(X_{i,j})\}^2\, \exp\left\{ 2\frac{L}{N} K_h(0) \frac{\int w(u)\, du}{\int w(u) p(u)\, du} \right\}\,. \tag{A4}$$

Putting $w(u) \stackrel{\text{def}}{=} 1$, delivers the common AIC. This, however, does not take into account the quality of the estimation at the boundary regions or in regions where data are sparse, since in these regions $p(u)$ is small. We choose $w(u) \stackrel{\text{def}}{=} p^{-1}(u)$, which can be shown to give equal weight everywhere. We do not punish for the number parameters $\widehat{\beta}_{i,l}$ that are employed to model the time series. This can be neglected because we use a finite dimensional model for the dynamics of $\beta_{i,l}$. The corresponding penalty term is negligible compared to the smoothing penalty term. The SC are obtained by replacing the ''2'' in Equations (A3) and (A4) by log(N).

Convergence of the iterations is measured by

$$Q^2(r) \stackrel{\text{def}}{=} \sum_{i=1}^{I} \int \left(\sum_{l=0}^{L} \widehat{\beta}_i^{(r)}\, \widehat{m}_l^{(r)}(u) - \widehat{\beta}_i^{(r-1)}\, \widehat{m}_l^{(r-1)}(u) \right)^2 du\ . \tag{A5}$$

The rth cycle of the estimation is denoted by (r). Iterations are stopped when $Q^2(r) \leq 10^{-5}$.

B DIVIDEND CORRECTION SCHEME

When the DAX spot price is approximated by simply discounting the DAX futures price for calculating implied volatilities, it can turn out that implied volatilities of calls and puts at same strikes diverge. As an explanation, Hafner and Wallmeier (2001) argue that due to German tax laws which discriminated foreign investors against domestic investors and which were effective till end of 2002, the marginal investor can be different from the one actually assumed to compute the DAX index. To foreign investors this tax scheme has the same impact as a dividend payment on an unprotected option, that is, it drives a wedge into the option prices and hence into implied volatilities. In order to obtain this unknown residual dividend $\Delta\hat{D}$, Hafner and Wallmeier (2001) solve the dividend-adjusted formula of a fair forward price for the spot and insert into the dividend-adjusted put−call parity.

For an estimate of $\Delta\hat{D}$, one identifies pairs of puts and calls of the same strikes and the same expiry date traded within a five minutes interval. The residual dividend for each pair is computed. $\Delta\hat{D}$ is estimated by the median of these residual dividends of all pairs for a given maturity at day t. Implied volatilities are recovered using as index value $\tilde{S}_t = F_t e^{-r_{T_F,t}(T_F - t)} + \Delta\hat{D}$. This procedure is applied on a daily basis throughout the entire data set from January 1998 to May 2001. Observations, for which the put−call pairs could not be identified, are eliminated.

Received February 28, 2006; revised September 21, 2006; accepted January 22, 2007.

REFERENCES

Aït-Sahalia, Yacine, and Jefferson Duarte. (2003). ''Nonparametric option pricing under shape restrictions.'' *Journal of Econometrics 116*, 9−47.

Aït-Sahalia, Yacine, and Andrew Lo. (1998). ''Nonparametric estimation of state-price densities implicit in financial asset prices.'' *Journal of Finance 53*, 499−548.

Aït-Sahalia, Yacine, Yubo Wang, and Francis Yared. (2001). ''Do options markets correctly price the probabilities of movement of the underlying asset?'' *Journal of Econometrics 102*, 67−110.

Akaike, Hirotugo. (1970). ''Statistical predictor identification.'' *Annals of the Institute for Statistical Mathematics 22*, 203−217.

Alexander, Carol. (2001). *Market Models*, John Wiley & Sons, New York.

Alexander, Carol, and Leonardo M. Nogueira. (2007). ''Model-free hedge ratios and scale-invariant models.'' *Journal of Banking and Finance*. Forthcoming.

Andersen, Torben G., and Bent E. Sørensen. (1996). ''GMM estimation of a stochastic volatility model: A Monte Carlo study.'' *Journal of Business and Economic Statistics 14*, 328–352.

Andersen, Torben G., Hyung-Jin Chung, and Bent E. Sørensen. (1999). ''Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study.'' *Journal of Econometrics 91*, 61–87.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen. (2000). ''Do call and underlying prices always move in the same direction?'' *Review of Financial Studies 13*(3), 549–584.

Barndorff-Nielsen, Ole E. (1997). ''Normal inverse Gaussian distributions and stochastic volatility modelling.'' *Scandinavian Journal of Statistics 24*, 1–13.

Barndorff-Nielsen, Ole E. and Neil Shepard. (2001). ''Non-Gaussian Ornstein-Uhlenbeck-models and some of their uses in financial economics.'' *Journal of the Royal Statistical Society B 63*, 167–241.

Bates, David. S. (1996). ''Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options.'' *Review of Financial Studies 9*, 69–107.

Bates, David. S. (2000). ''Post-'87 crash fears in the S&P 500 futures option market.'' *Journal of Econometrics 94*, 181–238.

Bates, David. S. (2006). ''Maximum likelihood estimation of latent affine processes.'' *Review of Financial Studies 19*(3), 909–965.

Benko, Michal, Matthias R. Fengler, Wolfgang Härdle, and Milos Kopa. (2007). ''On extracting information implied in options.'' *Computational Statistics*. Forthcoming.

Black, Fischer. (1976). ''Studies of stock price volatility changes.'' *Proceedings of the 1976 Meetings of the American Statistical Association* 177–181.

Black, Fischer, and Myron Scholes. (1973). ''The pricing of options and corporate liabilities.'' *Journal of Political Economy 81*, 637–654.

Borak, Szymon, Matthias R. Fengler, and Wolfgang Härdle. (2005). ''DSFM fitting of implied volatility surfaces.'' *Conference proceedings of the Fifth International Conference on Intelligent Systems Design and Applications*.

Borak, Szymon, Wolfgang Härdle, Enno Mammen, and Beyong Park. (2006). ''Time series modelling with semiparametric factor dynamics.'' *CASE Discussion Paper*, Humboldt-Universität zu Berlin.

Britten-Jones, Mark, and Anthony J. Neuberger. (2000). ''Option prices, implied price processes, and stochastic volatility, *Journal of Finance 55*(2), 839–866.

Broadie, Mark, Jérôme Detemple, Eric Ghysels, and Olivier Torrès. (2000a). ''American options with stochastics dividends and volatility: A nonparametric investigation.'' *Journal of Econometrics 94*, 53–92.

Broadie, Mark, Jérôme Detemple, Eric Ghysels, and Olivier Torrès. (2000b). ''Nonparametric estimation of American options exercise boundaries and call prices.'' *Journal of Economic Dynamics and Control 24*, 1829–1857.

Cai, Zongwu, Jianqing Fan, and Qiwei Yao. (2000). ''Functional-coefficient regression models for nonlinear time series, *Journal of the American Statistical Association 95*, 941–956.

Carr, Peter, Héliette Geman, Dilip B. Madan, and Marc Yor. (2003). ''Stochastic volatility for Lévy processes.'' *Mathematical Finance 13*(3), 345–382.

Chacko, George, and Luis M. Viceira. (2003). ''Spectral GMM estimation of continuous-time processes.'' *Journal of Econometrics 116*, 259–292.

Chernov, Mikhail, and Eric Ghysels. (2000). ''A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation, *Journal of Financial Economics 56*(3), 407–458.

Connor, Gregory, and Oliver Linton. (2006). ''Semiparametric estimation of a characteristic-based factor model of stock returns.'' *Empirical Finance*. Forthcoming.

Connor, Gregory, Robert A. Korajczyk, and Oliver Linton. (2006). ''The common and specific components of dynamic volatility.'' *Journal of Econometrics 132*, 231–255.

Cont, Rama, and José da Fonseca. (2002). ''The dynamics of implied volatility surfaces.'' *Quantitative Finance 2*(1), 45–60.

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

Diebold, Francis X., and Canlin Li. (2006). ''Modeling and forecasting the term structure of government bond yields.'' *Journal of Econometrics 130*, 337–364.

Duffie, Darrell, and Kenneth Singleton. (1993). ''Simulated moments estimation of markov models of asset prices.'' *Econometrica 61*, 929–952.

Duffie, Darrell, Jun Pan, and Kenneth Singleton. (2000). ''Transform analysis and asset pricing for affine jump-diffusions.'' *Econometrica 68*, 1343–1376.

Dumas, Bernard, Jeff Fleming, and Robert E. Whaley. (1998). ''Implied volatility functions: Empirical tests.'' *Journal of Finance 53*(6), 2059–2106.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson. (2003). ''The impact of jumps in volatility and returns.'' *Journal of Finance 58*, 1269–1300.

Fan, Jianqin, Qiwei Yao, and Zongwu Cai. (2003). ''Adaptive varying-coefficient linear models.'' *Journal of the Royal Statistical Society B 65*, 57–80.

Fengler, Matthias R., Wolfgang Härdle, and Christophe Villa. (2003). ''The dynamics of implied volatilities: A common principle components approach.'' *Review of Derivatives Research 6*, 179–202.

Fridman, Moshe, and Lawrence Harris. (1998). ''A maximum likelihood approach for non-Gaussian stochastic volatility models.'' *Journal of Business and Economic Statistics 16*, 284–291.

Gallant, A. Ronald, and George Tauchen. (1998). ''Reprojecting partially observed systems with application to interest rate diffusions.'' *Journal of the American Statistical Association 93*, 10–24.

Gallant, A. Ronald, and George Tauchen. (2007). ''Simulated score methods and indirect inference for continuous-time models.'' *in* Yacine Aït-Sahalia and Lars Hansen (eds).'' *Handbook of Financial Econometrics*, North Holland, Amsterdam. Forthcoming.

Garcia, René, and Eric Renault. (1998a). ''A note on hedging in ARCH and stochastic volatility option pricing models.'' *Mathematical Finance 8*(2), 153–161.

Garcia, René, and Eric Renault. (1998b). ''Risk aversion, intertemporal substitution and option pricing.'' *Working Paper 98s-02*, CIRANO.

Garcia, René, Eric Ghysels, and Eric Renault. (2007). ''The econometrics of option pricing models.'' *in* Yacine Aït-Sahalia and Lars Hansen (eds).'' *Handbook of Financial Econometrics*, North Holland, Amsterdam. Forthcoming.

Ghysels, Eric, and Serena Ng. (1989). ''A semiparametric factor model of interest rates and tests of the affine term structure.'' *Review of Economics and Statistics 80*, 535–548.

Gouriéroux, Christian, Alain Monfort, and Eric Renault. (1993). ''Indirect inference.'' *Journal of Applied Econometrics 8*, S85–S118.

Greene, William. (2003). *Econometric Analysis*, 5th edn, Prentice Hall, New Jersey.

Hafner, Christian M., and Helmut Herwartz. (2002). ''Testing for vector autoregressive dynamics under heteroskedasticity, Econometric Institute Report EI 2002-36, Erasmus University Rotterdam, Netherlands.

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

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

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

Harvey, Andrew, Esther Ruiz, and Neil Shephard. (1994). ''Multivariate stochastic variance models.'' *Review of Economic Studies 61*, 247–264.

Hastie, Trevor, and Robert Tibshirani. (1990). *Generalized additive models*, Chapman and Hall, London.

Hastie, Trevor, and Robert Tibshirani. (1993). ''Varying-coeffcient models.'' *Journal of the Royal Statistical Society B 55*(4), 757–796.

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

Hull, John. (2003). *''Options, Futures, and Other Derivatives*, 5th edn, Prentice Hall, New Jersey, USA.

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

Jacquier, Eric, Nicholas Polson, and Peter Rossi. (1994). ''Bayesian analysis of stochastic volatility models.'' *Journal of Business and Economic Statistics 12*, 1–19.

Jamshidian, Farshid, and Yu Zhu. (1997). ''Scenario simulation: Theory and methodology.'' *Finance and Stochastics 1*, 43–67.

Jeffrey, Andrew, Dennis Kristensen, Oliver Linton, Thong Nguyen, and Peter Phillips. (2004). ''Nonparametric estimation of a multifactor Heath-Jarrow-Morton model: An integrated approach.'' *Journal of Financial Econometrics 2*(2), 251–289.

Kauermann, Göran, Gerhard Tutz, and Josel Brüderl. (2005). ''The survival of newly founded firms: a case-study into varying-coefficient models.'' *Journal of the Royal Statistical Society A 168*(1), 145–158.

Linton, Oliver, Enno Mammen, Jens Perch Nielsen, Carsten Tanggaard. (2001). ''Yield curve estimation by kernel smoothing.'' *Journal of Econometrics 105*(1), 185–223.

Lütkepohl, Helmut. (1991). *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin, Heidelberg.

Mammen, Enno, Oliver Linton, and Jens Perch Nielsen. (1999). ''The existence and asymptotic properties of a backfitting projection algorithm under weak conditions.'' *Annals of Statistics 27*(5), 1443–1490.

Manaster, Steven, and Gary Koehler. (1982). ''The calculation of implied variances from the Black-and-Scholes model: A note.'' *Journal of Finance 37*, 227–230.

Melino, Angelo, and Stuart Turnbull. (1990). ''Pricing foreign currency options with stochastic volatility.'' *Journal of Econometrics 45*, 239–265.

Pan, Jun. (2002). ''The jump-risk premia implicit in options: evidence from an integrated time-series study.'' *Journal of Financial Economics 63*, 3–50.

Pastorello, Sergio, Valentin Patilea, and Eric Renault. (2003). ''Iterative and recursive estimation of structural non-adaptive models.'' *Journal of Business and Economic Statistics 21*(4), 449–509.

Ramsay, Jim O. and Bernard Silverman. (1997). *Functional Data Analysis*, Springer-Verlag, Berlin, Heidelberg.
Renault, Eric. (1997). ''Econometric models of option pricing errors.'' *in* David M. Kreps and Kenneth F. Wallis (eds).'' *Advances in Economics and Econometrics, Seventh World Congress*, *Econometric Society Monographs*, Cambridge University Press 223–278.
Renault, Eric, and Nizar Touzi. (1996). ''Option hedging and implied volatilities in a stochastic volatility model.'' *Mathematical Finance 6*(3), 279–302.
Rubinstein, Mark (1985). ''Nonparametric tests of alternative option-pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978.'' *Journal of Finance 40*, 455–480.
Schöbel, Rainer, and Jianwei Zhu. (1999). ''Stochastic volatility with an Ornstein-Uhlenbeck process: An extension.'' *European Finance Review 3*, 23–46.
Schwarz, Gideon. (1978). ''Estimating the dimension of a model.'' *Annals of Statistics 6*, 461–464.
Skiadopoulos, George, Stuart Hodges, and Les Clewlow. (1999). ''The dynamics of the S&P 500 implied volatility surface.'' *Review of Derivatives Research 3*, 263–282.
Stone, Charles. (1986). ''The dimensionality reduction principle for generalized additive models.'' *The Annals of Statistics 14*, 592–606.
Tompkins, Robert. (2001). ''Implied volatility surfaces: Uncovering regularities for options on financial futures.'' *European Journal of Finance 7*(3), 198–230.
White, Halbert. (1980). ''A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity.'' *Econometrica 48*, 817–838.
Wiggins, James. (1987). ''Option values under stochastic volatility.'' *Journal of Financial Economics 19*, 351–372.
Zhu, Yingzi, and Marco Avellaneda. (1997). ''An E-ARCH model for the term-structure of implied volatility of FX options.'' *Applied Mathematical Finance 4*, 81–100.