Implied volatility and state price density estimation - arbitrage analysis

Miloš Kopa · **Sebastiano Vitali** · **Tomáš Tichý** · **Radek Hendrych**

Abstract This paper deals with implied volatility (IV) estimation using no-arbitrage techniques. The current market practice is to obtain implied volatility of liquid options as based on Black-Scholes type (BS hereafter) models. Such volatility is subsequently used to price illiquid or even exotic options. Therefore, it follows that the BS model can be related simultaneously to the whole set of IVs as given by maturity/moneyness relation of tradable options. Then, it is possible to get IV curve or surface (a so called smirk or smile). Since the moneyness and maturity of IV often do not match the data of valuated options, some sort of estimating and local smoothing is necessary. However, it can lead to arbitrage opportunity, if no-arbitrage conditions on state price density (SPD) are ignored. In this paper, using option data on DAX index, we aim to analyse the behavior of IV and SPD with respect to different choices of bandwidth parameter h, time to maturity and kernel function. A set of bandwidths which violates no-arbitrage conditions is identified. We document that the change of h implies interesting changes in the violation interval of moneyness. We also perform the analysis after removing outliers, in order to show that not only outliers cause the violation of no-arbitrage conditions. Moreover, we propose a new measure of arbitrage which can be considered either for the SPD curve (arbitrage area measure) or for the SPD surface (arbitrage volume measure). We highlight the impact of h on the proposed measures considering the options on a German stock index. Finally, we propose an extension of the IV and SPD estimation for the case of options on a dividend-paying stock.

Keywords Option pricing · Implied volatility · State price density · No-arbitrage conditions · Local polynomial smoothing

1 Introduction

The classic option pricing model of Black and Scholes (Black & Scholes, 1973) and Merton (Merton, 1973) (BS-type model henceforth) has been introduced more than 40 years ago and despite its drawbacks (especially the assumption of the Gaussianity for the underlying log-returns, their constant and deterministic volatility as well as deterministic riskless rate) it has become a standard of the market practice to valuate various options using the model itself or its extensions.

Notwithstanding the answer of the practice to empirical observations of the inconsistency of BS type models with the behavior of market prices of underlying assets was to invert the formula – take the

Miloš Kopa
Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Prague, Czech Republic
E-mail: kopa@karlin.mff.cuni.cz

Sebastiano Vitali
Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Prague, Czech Republic E-mail: sebastiano.vitali@unibg.it

Tomáš Tichý
Department of Finance, Faculty of Economics, VŠB-Technical University Ostrava, Sokolská 33, 701 21 Ostrava, Czech Republic. E-mail: tomas.tichy@vsb.cz

Radek Hendrych
Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Prague, Czech Republic E-mail: hendrych@karlin.mff.cuni.cz

market price of several liquid options (i.e. plain vanilla calls and puts), insert them into inverted BS type model and obtain a so called implied volatility (IV). That is, many traders currently do not quote the option prices as such but state rather IVs, which can be, in turn, subsequently used to valuate options that are not traded in the market.

At any time, for a single underlying asset, several dozens of call and put options with different strike prices and maturities are traded. Similarly, at any time, we can obtain a whole set of IVs. A so called moneyness (a ratio of the strike price and current or future underlying asset price) is used instead of strike prices in order to simplify a comparison among different options and underlying assets. If IVs are plot just against the moneyness we get a curve whose shape often resemble a smile (especially for forex) or smirk (a non-symmetric smile common for equities). By contrast, when both moneyness and maturity are used, we get an IV surface. Subsequently, the IV curve or surface can be used to construct a local volatility model which can be further adopted for option pricing (see eg. Dupire (1994)).

Various parametric and semi-parametric representations of the volatility surface have been recently introduced, see Fengler (2012) and Homescu (2011) **2012??** for a review. Such methodologies usually include some kind of parametrization. For example, Dumas et al. (1998) suggested to model IV surface as a quadratic function of the moneyness and a linear function of time to maturity. A similar semi-parametric representation with quadratic expressions was considered for the oil market in Borovkova & Permana (2009). Alternatively, Avellaneda et al. (1997) proposed an algorithm that yields an arbitrage-free diffusion process by minimizing the relative entropy distance to a prior diffusion. In practitioner papers of ?, Bloch et al. (2011), a weighted sum of interpolation functions employing a parametric family was considered in order to generate a surface without arbitrage in time and in space, while corresponding as closely as possible to market data. Each function from the family is required to satisfy arbitrage-free constraints. Several families can satisfy these constraints, for instance a sum of lognormal distributions.

Anyway, there is no continuum of data (with respect to moneyness and maturity) and thus some kind of interpolation and smoothing is necessary before the evaluation of non-traded options. The deficiency is that in some regions the smoothed IV curve or surface can lead to a violation of arbitrage-free conditions. Benko et al. (2007) combine the IV smoothing with a state price density (SPD) estimation in order to avoid these problems. They proposed to use a local polynomial smoothing technique with Epanechnikov (bivariate) kernels. The advantage of this approach is that it provides all quantities needed to calculate the corresponding SPD. Moreover, the approach operates only on the IVs - a major improvement compared to the earlier multistep approaches moving through the BS formula from the prices to IVs and vice-versa. For the implied volatility surface arbitrage-free estimation Benko et al. (2007) followed Kahalé (2004) and Fengler (2005) in applying an implied total variance condition to preclude the calendar arbitrage; however, including constraints on both SPD and the implied total variance makes the kernel local quadratic estimation quite computationally demanding. Moreover, a strict application of total variance constraints leads to a semi-infinite programming problem. Therefore, Benko et al. (2007) suggested a discrete approximation that simplifies the algorithm solving a non-linear programming problem.

Kim & Lee (2013) extended and applied the notion of Benko et al. (2007) to KOSPI 200 index options no-arbitrage implied volatility modelling. The other no-arbitrage IV and SPD estimation procedures were recently presented in e.g. Fengler (2012), Glaser & Heider (2012), Fengler & Hin (2015), Ludwig (2015).

In this paper we follow Benko et al. (2007) focusing on relatively classic approach of local polynomial smoothing techniques and we study the bandwidth selection process in more details for more recent data of DAX option prices (December 2011). In particular, we change h and the kernel function and we examine their impact on the presence of arbitrage (violation of arbitrage-free conditions). We introduce new measures of arbitrage for IV curve or surface **IV or SPD??**. Using these measures, we empirically analyze the values of these measures for different choices of bandwidth, time-to-maturity and kernel function. Moreover, we exclude the outliers to show that the arbitrage occurrence is not caused only by the outliers. We also introduce another way of a discrete approximation (grid) and compare its results with the results of the original approximation (Benko et al. (2007)). Finally, we extend the notion to the case of dividend-paying underlying stock.

The rest of the paper is structured as follows. Section 2 recalls the theory of implied volatility and state price density estimation. Moreover, it introduces the new measures of arbitrage. Section 3 presents the empirical analysis and the last section concludes the paper.

2 Implied volatility estimation

Implied volatility can be derived utilizing the BS-type model and observed price data of European call (or put) options. Assuming that the stock price dynamic evolves as a geometric Brownian motion with a

constant drift and volatility, the value C_t of a call option on a non-dividend-paying stock is expressed as (Black & Scholes, 1973):

$$C_t(S_t, K, \tau, r, \sigma) = S_t \Phi(d_1) - K e^{-r\tau} \Phi(d_2), \quad t < T, \tag{1}$$

$$d_1 = \frac{\log(S_t/K) + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \tag{2}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}, \tag{3}$$

where S_t is the spot price of the underlying asset, K stands for the strike price, $\tau = T - t$ represents the time-to-maturity, r is the risk free interest rate, σ is the volatility of the log-returns of the underlying asset (an unknown and constant parameter), and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The price of an associated put option P_t can be calculated following the put-call parity: $P_t = C_t - S_t + e^{-r\tau}K$, $t < T$. The implied volatility of the underlying asset log-returns $\tilde{\sigma}$ is simply defined as the solution of the system (1)-(3) where the price of call option C_t is substituted by its known counterpart $\tilde{C}_t$ observed at the option market.

In contrast to the declared assumptions, the implied volatility derived by the BS model is commonly not constant as was demonstrated in various practically or academically oriented publications, see Brockhaus et al. (2000) and Fengler (2006). The implied volatility is obviously a u-shaped function of the strike fixing the maturity, and the volatility surface (i.e. the implied volatility illustrated as a function of maturities and strikes) is indeed not flat.

According to the previous discussion, we assume the following model framework:

$$\tilde{\sigma}(K_i, \tau_i) = \sigma(K_i, \tau_i) + \varepsilon_i, \quad i = 1, \dots, n, \tag{4}$$

where $\tilde{\sigma}_i = \tilde{\sigma}(K_i, \tau_i)$ stands for the implied volatility observed on K_i and τ_i, $\sigma_i = \sigma(K_i, \tau_i)$ is its theoretical counterpart, ε_i denotes the disturbance of the model, and n is a number of realized observations. Following Benko et al. (2007), any explicit model expression for σ is not introduced. Instead, a familiar instrument of non-parametric regression analysis is considered to describe the dynamics in a given model, i.e. the concept of local polynomial fitting presented by Fan & Gijbels (1996), Hardle (1990) or Fan & Yao (2003). Particularly, local polynomial estimators have been successfully applied in calibration of volatility surface model (4), see Shimko (1993), Fengler et al. (2003) or Cont et al. (2002). Finally, Benko et al. (2007) extend this modelling scheme imposing constraints which eliminate arbitrage opportunities, a key assumption in financial theories.

These constraints are grounded on the state price density, i.e. a density function defined as:

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

As can be seen from (1)-(3) and (5), the state price density is linked to the implied volatility surface. More precisely, the relation of these two quantities can be explicitly written as:

$$q_{t,S_T}(x, \tau) = e^{r\tau} S_t \sqrt{\tau} \varphi(d_1(x, \tau)) \left\{ \frac{1}{x^2 \sigma(x, \tau)\tau} + \frac{2d_1(x, \tau)}{x\sigma(x, \tau)\sqrt{\tau}} \left. \frac{\partial \sigma(K, \tau)}{\partial K} \right|_{K=x} \right.$$

$$\left. + \frac{d_1(x, \tau)d_2(x, \tau)}{\sigma(x, \tau)} \left(\left. \frac{\partial \sigma}{\partial K} \right|_{K=x} \right)^2 + \left. \frac{\partial^2 \sigma}{\partial K^2} \right|_{K=x} \right\}, \tag{6}$$

$$d_1(x, \tau) = \frac{\log(S_t/x) + (r + \sigma^2(x, \tau)/2)\tau}{\sigma(x, \tau)\sqrt{\tau}}, \tag{7}$$

$$d_2(x, \tau) = d_1(x, \tau) - \sigma(x, \tau)\sqrt{\tau}, \tag{8}$$

where $\varphi(\cdot)$ denotes the probability density function of the standard normal distribution (Brunner & Hafner, 2003). Clearly, the state price density is truly expressed as a function of the implied volatility surface and its derivatives. A substantial body of literature concerning the option-implied state price densities is reviewed in Jackwerth (2004).

As shown in Benko et al. (2007), a negative value of the state price density function coincides with an arbitrage opportunity and the non-negativity of the state space price density naturally delivers the required constraint of nonexistence of arbitrages. Combining local polynomial estimators of (4) and the state price density concept, the effective estimation technique of implied volatility surfaces can be introduced. In the following sections, we present particular implementations of this modelling framework.

Furthermore, we illustrate the performance of this approach by means of empirical applications. Before that, we remark that the implied volatility should be standardized with respect to the underlying stock.

We consider the *futures moneyness* $\kappa_t := K/F_t$ with $F_t = S_t e^{r\tau}$, i.e. the relative position of the future price of the underlying asset with respect to the strike price.

2.1 Implied volatility for fixed maturities

An estimate of the implied volatility function is constructed using a dataset of observed option prices with a fixed time-to-maturity τ. In greater detail, the general framework of the model (4) can be further simplified due to the fixed τ and the adopted futures moneyness:

$$\tilde{\sigma}(\kappa_i) = \sigma(\kappa_i) + \varepsilon_i, \quad i = 1, \ldots, n_\tau, \tag{9}$$

where $\tilde{\sigma}_i := \tilde{\sigma}(\kappa_i)$ and $\sigma_i := \sigma(\kappa_i)$ may be assumed as functions of κ_i only, ε_i represents the error term, and n_τ stands for the number of observed implied volatilities with the fixed time-to-maturity τ. If $\tau = T - t$ is given, the definition of the moneyness delivers: $K_i = \kappa_i S_t e^{r\tau}$, where $F_t = S_t e^{r\tau}$ is naturally supposed to be constant observing daily data, i.e. $S_t := S$, $F_t := F$ and $K_i = \kappa_i F$. In the case of intra-day data, τ still stays fixed, but S_t can vary over time; therefore S is simply given as the daily average of S_t.

As was previously stated, the introduced model (9) can be calibrated by a local polynomial fitting procedure. The local quadratic estimator seems to be suitable in this context, see Benko et al. (2007). The estimator $\hat{\sigma}(\kappa)$ of the function $\sigma(\kappa)$ in the grid point κ is then delivered as the solution of the following minimization task:

$$\min_{\alpha_0, \alpha_1, \alpha_2} \sum_{i=1}^{n_\tau} \left[\tilde{\sigma}_i - \alpha_0 - \alpha_1(\kappa_i - \kappa) - \alpha_2(\kappa_i - \kappa)^2 \right]^2 \mathcal{K}_h(\kappa - \kappa_i), \tag{10}$$

where $\mathcal{K}$ denotes a kernel function and the weighting term $\mathcal{K}_h(\kappa - \kappa_i)$ is defined as $\frac{1}{h}\mathcal{K}\left(\frac{\kappa - \kappa_i}{h}\right)$. The kernel is a weighting function generally used in non-parametric estimation techniques. In particular, it is a non-negative real-valued integrable function $\mathcal{K}$ satisfying: (i) $\int_{-\infty}^{\infty} \mathcal{K}(x)dx = 1$, and (ii) $\mathcal{K}(x) = \mathcal{K}(-x)$, $\forall x \in \mathbb{R}$. Various kernel functions exist, see Fan & Gijbels (1996) for their review. The quantity $h > 0$ is the so-called bandwidth; it controls the size of the local neighborhood included into estimation. The choice of the bandwidth h plays an important role in the local polynomial fitting procedure. On the one hand, a large bandwidth causes over-smoothing, i.e. induces modelling bias. On the other hand, a small bandwidth results in under-smoothing, i.e. produces noisy estimates. If $h = \infty$, the local polynomial fitting becomes a global quadratic fitting (Fan & Yao, 2003). Additionally, comparing the Taylor expansion of σ with the problem (10), one can explicitly derive these formulas: $\alpha_0 = \hat{\sigma}(\kappa)$, $\alpha_1 = \hat{\sigma}'(\kappa)$, and $\alpha_2 = \hat{\sigma}''(\kappa)/2$.

To eliminate eventual arbitrage opportunities, the minimization task (10) should be resolved subject to $q_{t,S_T}(\kappa, \tau) \geq 0$ (it is possible to simplify the notation since τ is fixed, i.e. $q_{t,S_T}(\kappa, \tau) =: q(\kappa)$). The equation (6) and the following formulas can be rewritten more concisely as:

$$q(\kappa) = \sqrt{\tau}\varphi(d_1(\kappa))\left\{ \frac{1}{\kappa^2\sigma(\kappa)\tau} + \frac{2d_1(\kappa)}{\kappa\sigma(\kappa)\sqrt{\tau}}\frac{\partial\sigma(\kappa)}{\partial\kappa} + \frac{d_1(\kappa)d_2(\kappa)}{\sigma(\kappa)}\left(\frac{\partial\sigma(\kappa)}{\partial\kappa}\right)^2 + \frac{\partial^2\sigma(\kappa)}{\partial\kappa^2} \right\}, \tag{11}$$

$$d_1(\kappa) = \frac{\sigma^2(\kappa)\tau/2 - \log(\kappa)}{\sigma(\kappa)\sqrt{\tau}}, \tag{12}$$

$$d_2(\kappa) = d_1(\kappa) - \sigma(\kappa)\sqrt{\tau}, \tag{13}$$

regarding to the easily derived relations: $K = F\kappa$, $\frac{\partial K}{\partial \kappa} = F$, $\frac{\partial \sigma}{\partial K} = \frac{\partial \sigma}{\partial \kappa}\frac{1}{F}$, $\frac{\partial^2 \sigma}{\partial K^2} = \frac{\partial^2 \sigma}{\partial K^2}\frac{1}{F^2}$. After analytical computations, the state price density should be standardized to fulfil $\int_{-\infty}^{\infty} q(\kappa)d\kappa = 1$.

Obviously, the true implied volatility function $\sigma(\kappa)$ and its derivatives are not generally known and must be substituted by suitable estimates. Naturally, one can apply α_0, α_1 and α_2 as approximations according to their characterization above. Consequently, the constrained optimization problem solving

model (9) with respect to the discussed no-arbitrage condition is specified as follows:

$$\min_{\alpha_0, \alpha_1, \alpha_2} \sum_{i=1}^{n_\tau} \left[\tilde{\sigma}_i - \alpha_0 - \alpha_1(\kappa_i - \kappa) - \alpha_2(\kappa_i - \kappa)^2 \right]^2 \mathcal{K}_h \left(\kappa - \kappa_i \right), \tag{14}$$

$$\text{s.t.} \quad \hat{q}(\kappa) = \sqrt{\tau}\varphi(\hat{d}_1(\kappa)) \left\{ \frac{1}{\kappa^2 \alpha_0 \tau} + \frac{2\hat{d}_1(\kappa)}{\kappa \alpha_0 \sqrt{\tau}}\alpha_1 + \frac{\hat{d}_1(\kappa)\hat{d}_2(\kappa)}{\alpha_0}\alpha_1^2 + 2\alpha_2 \right\} \geq 0, \tag{15}$$

$$\hat{d}_1(\kappa) = \frac{\alpha_0^2 \tau/2 - \log(\kappa)}{\alpha_0 \sqrt{\tau}}, \tag{16}$$

$$\hat{d}_2(\kappa) = \hat{d}_1(\kappa) - \alpha_0 \sqrt{\tau}. \tag{17}$$

If IV is estimated without no-arbitrage constraints (15) - (17) then the estimated SPD may be negative in some points. Let $\alpha_0^*(\kappa), \alpha_1^*(\kappa), \alpha_2^*(\kappa)$ be the optimal solutions of (14) without constraints (15) - (17) for all considered κ. Let $\{\cdot\}^- = \min(\cdot, 0)$. Then the arbitrage area measure is defined as a magnitude of area of the estimated SPD below zero:

$$\text{AAM} = \int -\sqrt{\tau}\varphi(\hat{d}_1(\kappa)) \left\{ \frac{1}{\kappa^2 \alpha_0^*(\kappa)\tau} + \frac{2\hat{d}_1(\kappa)}{\kappa \alpha_0^*(\kappa)\sqrt{\tau}}\alpha_1^*(\kappa) + \frac{\hat{d}_1(\kappa)\hat{d}_2(\kappa)}{\alpha_0^*(\kappa)}\alpha_1^{*2}(\kappa) + 2\alpha_2^*(\kappa) \right\}^- d\kappa \tag{18}$$

where

$$\hat{d}_1(\kappa) = \frac{\alpha_0^{*2}(\kappa)\tau/2 - \log(\kappa)}{\alpha_0^*(\kappa)\sqrt{\tau}}, \quad \forall \kappa \tag{19}$$

$$\hat{d}_2(\kappa) = \hat{d}_1(\kappa) - \alpha_0^*(\kappa)\sqrt{\tau}, \quad \forall \kappa. \tag{20}$$

The empirical performance of the considered implied volatility estimation formulated in (14)-(17) is examined in Section 3. In particular, various kernel functions $\mathcal{K}$ and configurations of the smoothing parameter h have been tested studying real datasets. The bandwidth h has been selected regarding two goals: (i) to keep the bias small (with a reasonable variance), and (ii) to guarantee a sufficient number of observations around each κ entering into the calibration procedure. To estimate the bias and variance of our estimates, we adopt the idea of (pre)asymptotic versions of these quantities (Fan & Yao, 2003):

$$\widehat{bias}(\kappa) = \left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1} \mathbf{X}^\top \mathbf{W} \hat{\mathbf{r}}(\kappa), \quad \widehat{\mathbf{\Sigma}}(\kappa) = \hat{v}^2(\kappa) \left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1} \mathbf{X}^\top \mathbf{W}^2 \mathbf{X} \left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1}. \tag{21}$$

The introduced formulas are apparently closely connected to the general weighted ordinary least squared estimation scheme. In fact, the task (10) is a special case of a weighted linear regression. Particularly, the $(n_\tau \times 3)$ matrix $\mathbf{X}$ consists of rows $\left(1, \kappa_i - \kappa, (\kappa_i - \kappa)^2 \right)$ and the $(n_\tau \times n_\tau)$ weighting matrix $\mathbf{W}$ is diagonal with the elements $\mathcal{K}_h(\kappa_i - \kappa)$. The term $\hat{\mathbf{r}}(\kappa)$ in the bias equation is the $(n_\tau \times 1)$ vector composed of associated estimates of $r_i(\kappa) = \sigma_i - \alpha_0 - \alpha_1(\kappa_i - \kappa) - \alpha_2(\kappa_i - \kappa)^2$. These are simply derived applying the corresponding Taylor expansion of σ_i of a higher order; for more details consult Fan & Yao (2003). More precisely, we put $\hat{r}_i(\kappa) = \hat{\alpha}_3(\kappa_i - \kappa)^3 + \hat{\alpha}_4(\kappa_i - \kappa)^4$, where α_3 and α_4 are the optimal solutions of (14)-(17) with the objective function adjusted to:

$$\sum_{i=1}^{n_\tau} \left[\tilde{\sigma}_i - \alpha_0 - \alpha_1(\kappa_i - \kappa) - \alpha_2(\kappa_i - \kappa)^2 - \alpha_3(\kappa_i - \kappa)^3 - \alpha_4(\kappa_i - \kappa)^4 \right]^2 \mathcal{K}_h \left(\kappa - \kappa_i \right). \tag{22}$$

The term $\hat{v}^2(\kappa)$ appearing in the asymptotic variance estimate $\widehat{\mathbf{\Sigma}}(\kappa)$ is an estimated variance of the residuals calculated in (9), see Fan & Gijbels (1996). Explicitly, we suppose that $\hat{v}^2(\kappa) = \hat{\beta}_0$, where β_0 is the optimal solution of the following minimization task:

$$\min_{\beta_0, \beta_1, \beta_2} \sum_{i=1}^{n_\tau} \left[(\tilde{\sigma}_i - \hat{\alpha}_0)^2 - \beta_0 - \beta_1(\kappa_i - \kappa) - \beta_2(\kappa_i - \kappa)^2 \right]^2 \mathcal{K}_h \left(\kappa - \kappa_i \right), \tag{23}$$

$$\text{s.t.} \quad \beta_0 > 0, \tag{24}$$

where $\hat{\alpha}_0$ is the optimal solution of (14)-(17). Undoubtedly, more sophisticated selection criteria can be applied to choose a proper bandwidth, see Fan & Yao (2003). On the other hand, these do not obviously ensure an adequate number of observations around each κ.

2.2 Implied volatility surface

In the previous section, we described the methodology to estimate the implied volatility function for fixed maturities. In this section, we proceed to calibrate the two-dimensional implied volatility surface $\sigma(\kappa, \tau)$, i.e. the function of the futures moneyness and of the time-to-maturity. This extension allows us to include the so-called calendar arbitrage, i.e. the arbitrage in the τ direction.

To be more precise, let us assume the following model (compare with (4)):

$$\tilde{\sigma}(\kappa_i, \tau_i) = \sigma(\kappa_i, \tau_i) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{25}$$

where $\tilde{\sigma}_i = \tilde{\sigma}(\kappa_i, \tau_i)$ denotes the implied volatility observed on κ_i and τ_i, $\sigma_i = \sigma(\kappa_i, \tau_i)$ is its theoretical counterpart, ε_i corresponds to the model disturbance, and n is a number of realized observations.

Analogously as before, we adopt the instruments of non-parametric regression to estimate the model (25) given observed data $\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n$. In particular, the two-dimensional local quadratic estimator $\hat{\sigma}(\kappa, \tau)$ is supposed to be suitable in this context. It is defined as the optimal solution of the following minimization problem:

$$\min_{\boldsymbol{\alpha}} \sum_{i=1}^{n} \left[\tilde{\sigma}_i - \alpha_0 - \alpha_1(\kappa_i - \kappa) - \alpha_2(\tau_i - \tau) \right.$$
$$\left. - \alpha_{11}(\kappa_i - \kappa)^2 - \alpha_{12}(\kappa_i - \kappa)(\tau_i - \tau) - \alpha_{22}(\tau_i - \tau)^2 \right]^2 \mathcal{K}_{\mathbf{H}}\left(\kappa - \kappa_i, \tau - \tau_i\right), \tag{26}$$

where $\boldsymbol{\alpha}$ denotes the vector $(\alpha_0, \alpha_1, \alpha_2, \alpha_{11}, \alpha_{12}, \alpha_{22})^{\top}$, $\mathcal{K}_{\mathbf{H}}(\mathbf{x}) := \frac{1}{\det \mathbf{H}} \mathcal{K}\left(\mathbf{H}^{-1}\mathbf{x}\right)$ is a bivariate kernel function with a symmetric positive definite bandwidth matrix $\mathbf{H}$. Apparently, the kernel represents a weighting term in the estimation formula. Comparing (26) with the associated Taylor expansion of $\sigma(\kappa, \tau)$ yields: $\alpha_0 = \hat{\sigma}(\kappa, \tau)$, $\alpha_1 = \frac{\partial \hat{\sigma}}{\partial \kappa}(\kappa, \tau)$, $\alpha_2 = \frac{\partial \hat{\sigma}}{\partial \tau}(\kappa, \tau)$, $\alpha_{11} = \frac{\partial^2 \hat{\sigma}}{2\partial \kappa^2}(\kappa, \tau)$, $\alpha_{12} = \frac{\partial^2 \hat{\sigma}}{\partial \kappa \partial \tau}(\kappa, \tau)$, and finally $\alpha_{22} = \frac{\partial^2 \hat{\sigma}}{2\partial \tau^2}(\kappa, \tau)$.

In order to extend the available portfolio of no-arbitrage conditions, the concept of the total variance $w(\kappa, \tau) := \sigma^2(\kappa, \tau)\tau$ advocated by Kahale (2004) or Fengler (2009) is introduced. It has been shown that the total variance $w(\kappa, \tau)$ should be strictly increasing in the variable τ under the no-arbitrage settings assuming a deterministic (time-varying) interest rate, i.e. explicitly $\frac{\partial w}{\partial \tau}(\kappa, \tau) > 0$. In summary, the estimated state price density jointly with the estimated total variance should be incorporated into the original minimization problem (26) in order to eliminate eventual arbitrage opportunities.

Explicitly, an estimator of the implied volatility surface σ given κ and a set of increasing time-to-maturity $\mathcal{T} = \{\tau_1, \ldots, \tau_m\}$ is constructed as the solution of the following optimization task (Benko et al., 2007):

$$\min_{\boldsymbol{\alpha}(\ell)} \sum_{\ell=1}^{m} \sum_{i=1}^{n} \left[\sigma_i - \alpha_0(\ell) - \alpha_1(\ell)(\kappa_i - \kappa) - \alpha_2(\ell)(\tau_i - \tau_\ell) \right.$$
$$\left. - \alpha_{11}(\ell)(\kappa_i - \kappa)^2 - \alpha_{12}(\ell)(\kappa_i - \kappa)(\tau_i - \tau_\ell) - \alpha_{22}(\ell)(\tau_i - \tau_\ell)^2 \right]^2 \mathcal{K}_{\mathbf{H}}\left(\kappa - \kappa_i, \tau_\ell - \tau_i\right), \tag{27}$$
$$\text{s.t.} \quad \hat{q}(\kappa, \tau_\ell) \geq 0, \quad \ell = 1, \ldots, m, \tag{28}$$
$$2\tau_\ell \alpha_0(\ell)\alpha_2(\ell) + \alpha_0(\ell)^2 \geq 0, \quad \ell = 1, \ldots, m, \tag{29}$$
$$\alpha_0(\ell)^2 \tau_\ell \leq \alpha_0(\ell+1)^2 \tau_{\ell+1}, \quad \ell = 1, \ldots, m-1, \tag{30}$$

where the estimated counterpart of the state price density is defined as:

$$\hat{q}(\kappa, \tau_\ell) = \sqrt{\tau_\ell} \varphi(\hat{d}_1(\ell)) \left\{ \frac{1}{\kappa^2 \alpha_0(\ell)\tau_\ell} + \frac{2\hat{d}_1(\ell)}{\kappa \alpha_0(\ell)\sqrt{\tau_\ell}} \alpha_1(\ell) + \frac{\hat{d}_1(\ell)\hat{d}_2(\ell)}{\alpha_0(\ell)} \alpha_1(\ell)^2 + 2\alpha_{11}(\ell) \right\}, \tag{31}$$

$$\hat{d}_1(\ell) := \hat{d}_1(\kappa, \tau_\ell) = \frac{\alpha_0(\ell)^2 \tau_\ell/2 - \log(\kappa)}{\alpha_0(\ell)\sqrt{\tau_\ell}}, \tag{32}$$

$$\hat{d}_2(\ell) := \hat{d}_2(\kappa, \tau_\ell) = \hat{d}_1(\ell) - \alpha_0(\ell)\sqrt{\tau_\ell}, \quad \ell = 1, \ldots, m. \tag{33}$$

It should be noted that $\alpha_0(\ell) = \hat{\sigma}(\kappa, \tau_\ell)$, $\alpha_1(\ell) = \frac{\partial \hat{\sigma}}{\partial \kappa}(\kappa, \tau_\ell)$, etc. Moreover, the formulas (31)-(33) are directly derived using (6); compare also with (15)-(17). The constraint (29) follows the argument of the increasing estimated total variance (see above), i.e. $\hat{w}(\kappa, \tau_\ell) := \hat{\sigma}^2(\kappa, \tau_\ell)\tau_\ell$ satisfies $0 \leq \frac{\partial \hat{w}}{\partial \tau_\ell}(\kappa, \tau_\ell) = 2\tau_\ell \alpha_0(\ell)\alpha_2(\ell) + \alpha_0(\ell)^2$. Additionally, for a given κ and a set of times-to-maturity $\{\tau_1, \ldots, \tau_m\}$ the condition (30) guarantees $\hat{w}(\kappa, \tau_\ell) \leq \hat{w}(\kappa, \tau_{\ell+1})$ for all $\ell = 1, \ldots, m-1$.

Let $\alpha_0^*(\kappa,\tau), \alpha_1^*(\kappa,\tau), \alpha_2^*(\kappa,\tau), \alpha_{11}^*(\kappa,\tau), \alpha_{12}^*(\kappa,\tau), \alpha_{22}^*(\kappa,\tau)$ be the optimal solutions of unconstrained optimization problem (27) for all considered κ and $\tau \in \mathcal{T}$. Then, similarly to (18) - (20) we define the arbitrage volume measure as follows:

$$\text{AVM} =$$
$$\lambda_1 \int -\sqrt{\tau}\varphi(\hat{d}_1(\kappa,\tau)) \left\{ \frac{1}{\kappa^2 \alpha_0^*(\kappa,\tau)\tau} + \frac{2\hat{d}_1(\kappa,\tau)}{\kappa \alpha_0^*(\kappa,\tau)\sqrt{\tau}} \alpha_1^*(\kappa,\tau) + \frac{\hat{d}_1(\kappa,\tau)\hat{d}_2(\kappa,\tau)}{\alpha_0^*(\kappa,\tau)} \alpha_1^*(\kappa,\tau)^2 + 2\alpha_{11}^*(\kappa,\tau) \right\}^- d(\kappa,\tau)$$

$$+\lambda_2 \int - \left\{ 2\tau\alpha_0^*(\kappa,\tau)\alpha_2^*(\kappa,\tau) + \alpha_0^*(\kappa,\tau)^2 \right\}^- d(\kappa,\tau)$$

$$+\lambda_3 \int - \sum_{\ell=1}^{m-1} \left\{ -\alpha_0^*(\kappa,\tau_\ell)^2\tau_\ell + \alpha_0^*(\kappa,\tau_{\ell+1})^2\tau_{\ell+1} \right\}^- d\kappa \qquad (34)$$

where

$$\hat{d}_1(\kappa,\tau) = \frac{\alpha_0^*(\kappa,\tau)^2\tau/2 - \log(\kappa)}{\alpha_0^*(\kappa,\tau)\sqrt{\tau}}, \quad \forall \kappa, \tau \qquad (35)$$

$$\hat{d}_2(\kappa,\tau) = \hat{d}_1(\kappa,\tau) - \alpha_0^*(\kappa,\tau)\sqrt{\tau}, \quad \forall \kappa, \tau \qquad (36)$$

and $\lambda_1, \lambda_2, \lambda_3$ serve as the weights, that is, they are the parameters of the convex combination of three arbitrage volume measure components.

In Section 3, the proposed method is analysed from the practical point of view. It should be noted that in our empirical study we assume a diagonal matrix $\mathbf{H}$. Thus, the bivariate kernel function $\mathcal{K}_{\mathbf{H}}$ can be expressed as the product of two univariate kernels: $\mathcal{K}_{\mathbf{H}}(\kappa - \kappa_i, \tau - \tau_i) = \mathcal{K}_{h_{11}}(\kappa - \kappa_i)\mathcal{K}_{h_{22}}(\tau - \tau_i)$. The bandwidth $h_{11} := h_\kappa$ associated with κ is selected analogously as in Section 2.1. The bandwidth $h_{22} =: h_\tau$ corresponding to τ is chosen to be constant, see below. These choices assure enough data points entering into the calibration. Obviously, other (more sophisticated) bandwidth choices can be supposed; however, they would be likely computationally much more involving.

2.3 Note on options on dividend-paying assets

The previously introduced methodology can be further extended to the case of options on dividend-paying underlying assets. For simplicity, let us assume that dividends are paid continuously, and that the payments are proportional to the price of underlying asset. Hence, the value C_t of a European call option can be expressed as (Merton, 1973):

$$C_t(S_t, K, \tau, r, q, \sigma) = S_t e^{-q\tau}\Phi(d_1) - Ke^{-r\tau}\Phi(d_2), \quad t < T, \qquad (37)$$

$$d_1 = \frac{\log(S_t/K) + (r - q + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \qquad (38)$$

$$d_2 = d_1 - \sigma\sqrt{\tau}, \qquad (39)$$

where $q \geq 0$ denotes the dividend yield and similarly as before S_t is the spot price of the underlying asset, K stands for the strike price, $\tau = T - t$ is the time to maturity, r represents the risk free interest rate, σ is the volatility of the returns of the underlying asset (an unknown and constant parameter), and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The price of the associated put option P_t can be obviously calculated by the put-call parity, see above.

The derivation of the key state price density appropriate for the above case is relatively straightforward using (5) and (6). If we put $\tilde{S}_t = S_t e^{-q\tau}$, we obtain:

$$C_t = \tilde{S}_t\Phi(d_1) - Ke^{-r\tau}\Phi(d_2), \quad t < T, \qquad (40)$$

$$d_1 = \frac{\log(\tilde{S}_t/K) + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \qquad (41)$$

$$d_2 = d_1 - \sigma\sqrt{\tau}. \qquad (42)$$

Clearly, formulas (40)-(42) are similar to (1)-(3) if we observe $\tilde{S}_t$ instead of S_t. Accepting this convention, the corresponding state price density follows formula (6). Consequently, all optimization tasks introduced in Section 2.1 and Section 2.2 to model the implied volatility curve and surface can be applied directly in the considered context. Additionally, it should be noted that F_t appearing in the definition of the future moneyness is naturally transformed to $\tilde{S}_t e^{r\tau} = S_t e^{(r-q)\tau}$.

3 Implementation using data of German option market

In this section we show the empirical results of our analysis. The proposed methods are applied to options on DAX index. The data consists in 10 years of daily observations, for each day we observe on average the prices of call and put options on 10 different maturities. In particular, our computations regard December 8, 2008 which consists of a subset of 1991 options for 12 maturities. For each option, we consider information about strike, time to maturity (expressed in calendar days), market price, underlying asset price, implied volatility and moneyness. As risk free rate we use the implied risk free rate obtained using BS model.

In further sections, we divide our study about DAX options into two parts: the one dimension case when only one maturity is considered, and the two dimensions case when the whole set of available maturities is considered. In the first case we will produce curves of IV and SPD, in the second we obtain surfaces. Finally, we study the case of options on a dividend-paying stock.

3.1 Estimating implied volatility curve

For the implied volatility curve analysis we need to consider the options written for a single maturity. We select the options with maturity equal to 375 days and we collect 123 options. In Figure 1 (top left) we propose the estimation of the implied volatility using the Epanechnikov kernel function with bandwidth $h_\kappa = 0.12$ in line with the discussion in Section 2. For each moneyness we observe two IVs (*empty dots*), for call (above) and for put (below), and we estimate the smoothed unconstrained IV using (14) (*black line*) and the constrained IV using (14)-(17) (*red dots*). All the performed analysis are strongly influenced by the chosen bandwidth. Our specific choice is supported by the analysis shown in the left chart of Figure 2 where we show the evolution of the variance and the bias of the estimation for bandwidth increasing from 0.05 to 0.25. The choice of $h_\kappa = 0.12$ represents a good balance between the two indicators.

In the other plots of Figure 1 we compare the constrained SPD estimated using (14)-(17) (*black line*) and the unconstrained SPD estimated using (14) (*red dots*). The estimation is performed using the local quadratic smoother with the Epanechnikov kernel and with increasing bandwidth $h_\kappa = 0.08, 0.09, \ldots, 0.14$.

We can observe that in all cases there exists a discrepancy for moneyness around 1.7 where the unconstrained SPD estimation is negative and then an arbitrage opportunity occurs. This opportunity disappears for $h_\kappa > 0.145$. In order to study this behavior in more detail, we describe the intervals in which we observe the arbitrage chance according to several bandwidths, see Figure 2 (chart in the middle). Increasing the bandwidth, the length of the intervals are not strictly decreasing. This means that the size of the interval is not directly correlated with the bandwidth and than the arbitrage evidence is genuine. In order to investigate the dimension of the arbitrage for each bandwidth we use (18)-(20) to compute the arbitrage measure presented in Figure 2 (right). We notice that also the arbitrage measure is not strictly decreasing and it shows again that the presence of an arbitrage is genuine and not induced only by the bandwidth choice.

To analyze the impact of the choice of the kernel function we compute the arbitrage measure for four representative bandwidths and for a set of kernel functions. In Table 1 we report the computed measures and the mathematical definition of each kernel function. The triweight kernel function produced the highest measures for all considered bandwidths. The Epanachnikov kernel function adopted in the other analysis has a balanced behavior inducing measures which are in the middle with respect to the other kernel functions. A common behavior is that increasing the bandwidth the arbitrage measures decrease. However, as shown in Figure 2, this fact is generally true but not always.

In Figure 1 (top left), in correspondence to moneyness 1.7, we observe two values of the put option implied volatilities which can be considered as outliers. In order to analyse if the arbitrage opportunity was induced by this two observations or if it was genuine, we remove these two anomalies (compare Figure 1 (top left) and Figure 3 (top left)) and we estimate again the SPD. Figure 3 illustrates the evolution of the SPD shape estimated with Epanechnikov kernel function and for increasing bandwidth $h_\kappa = 0.08, 0.09, \ldots, 0.14$ respectively. Furthermore, we compare the constrained SPD estimation (*black line*) and the unconstrained one (*red dots*).

Figure 3 shows that the arbitrage opportunity persists even after removing the two outliers. The explanation could be that the remotion of the two put options brakes the equilibrium between put and call option data. However, this motivation could explain the first peak, but hardly it could explain the second one. Then, this implies that the whole data structure offers the arbitrage opportunity and not only the two outliers.
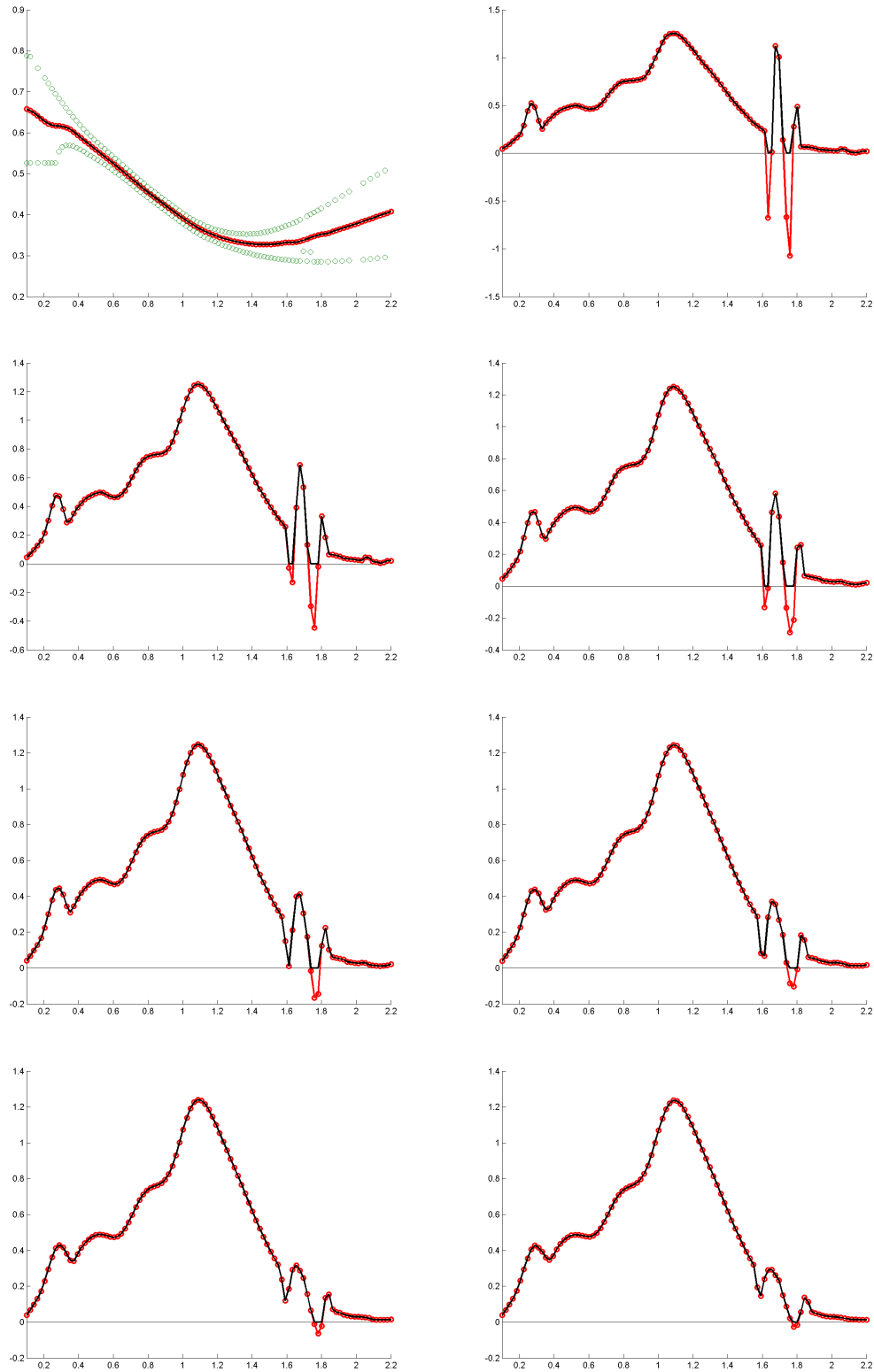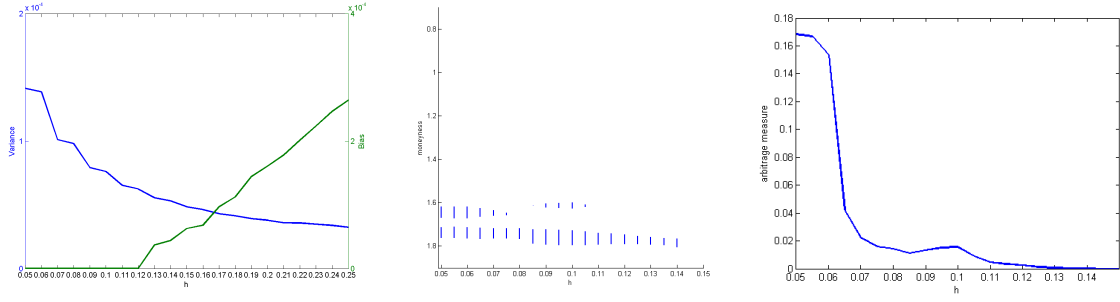
Fig. 1 Estimations of the IV curve (top left) for fixed maturity $\tau = 375$ days and of SPD curves with $h_\kappa = 0.08, 0.09, \ldots, 0.14$.

3.2 Implied volatility surface - Original calendar grid

The second part of our analysis concerns the surface case, i.e. we relax the constant maturity hypothesis. We consider all option maturities quoted in a single day and using (27) we estimate the IV surface with

Fig. 2 Variance and bias (left), arbitrage intervals on moneyness (center) and arbitrage measure (right) in dependency on selection of h_κ.

Table 1 Arbitrage area measure for various kernels and bandwiths

	0.05	0.08	0.11	0.14	$K(u)$
Uniform	0.169984	0.020431	0.000670	0.000027	$\frac{1}{2}\mathbf{1}_{\{|u|\leq 1\}}$
Triangular	0.167258	0.020295	0.005707	0.000442	$(1-|u|)\mathbf{1}_{\{|u|\leq 1\}}$
Epanechnikov	0.168257	0.014204	0.004629	0.000385	$\frac{3}{4}(1-u^2)\mathbf{1}_{\{|u|\leq 1\}}$
Quartic	0.176555	0.028206	0.009611	0.001449	$\frac{15}{16}(1-u^2)^2\mathbf{1}_{\{|u|\leq 1\}}$
Triweight	0.193120	0.051936	0.011914	0.003269	$\frac{35}{32}(1-u^2)^3\mathbf{1}_{\{|u|\leq 1\}}$
Tricube	0.185314	0.031527	0.010978	0.002164	$\frac{70}{81}(1-|u|^3)^3\mathbf{1}_{\{|u|\leq 1\}}$
Gaussian	0.168189	0.017430	0.001024	0.000020	$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}$
Cosine	0.168164	0.015570	0.004987	0.000448	$\frac{\pi}{4}\cos\left(\frac{\pi}{2}u\right)\mathbf{1}_{\{|u|\leq 1\}}$
Logistic	0.169098	0.018924	0.000833	0.000026	$\frac{1}{e^u+2+e^{-u}}$

Epanechnikov kernel function and with bandwidth $h_\kappa = 0.12$ for moneyness and $h_\tau = 1$ for maturity (Figure 4, top left). The historical data (*black dots*) are well described by the estimated surface. The non-symmetry of the IV smile is clear for small maturities and becomes less noticeable as the maturity increases.

Using (27) we produce several estimates of the SPD (Figure 4). The computation are done again with Epanechnikov kernel function adopting for the moneyness five representative bandwidths $h_\kappa = 0.08, 0.09, \ldots, 0.14$ and for maturity bandwidth $h_\tau = 1$.

The results still show some turbulence for moneyness around 1.7. Increasing the moneyness bandwidth the surface becomes more smooth but the anomalies on the right tail of the distribution persist. To better evaluate the quality and magnitude of the turbulence and to understand if it corresponds to an arbitrage opportunity, we study more in deep the case with $h_\kappa = 0.12$. In particular, in Figure 5 (left) we show the length of the arbitrage intervals in term of moneyness for each maturity, i.e. the intervals in which the surface is negative. On the one hand, it is evident that the arbitrage chance changes according to the maturity. On the other hand, the arbitrage opportunity is a common features of all the maturities. To evaluate the size of the arbitrage we use (34)-(36). With our data, we find that only the first component of (34) is positive, then we fix the convex combination as $\lambda_1 = 1, \lambda_2 = 0, \lambda_3 = 0$. Then, we compute the arbitrage measure for the surface according to different bandwidths, see Figure 5 (right). Again, the curve is not strictly decreasing and this evidence reinforces the idea that the arbitrage does not depend on the choice of the bandwidth.

In Table 2 we investigate the relation between the arbitrage volume and the type of the kernel function.

Table 2 Arbitrage volume measure for various kernels and bandwidths

	0.10	0.12	0.14
Uniform	0.117659	0.048795	0.023270
Triangular	0.057902	0.025572	0.012987
Epanechnikov	0.058283	0.026027	0.014453
Quartic	0.087371	0.035185	0.020242
Triweight	0.138065	0.053638	0.026578
Tricube	0.104245	0.042488	0.023836
Gaussian	0.088522	0.027045	0.009573
Cosine	0.058616	0.026581	0.014917
Logistic	0.104045	0.037728	0.017006
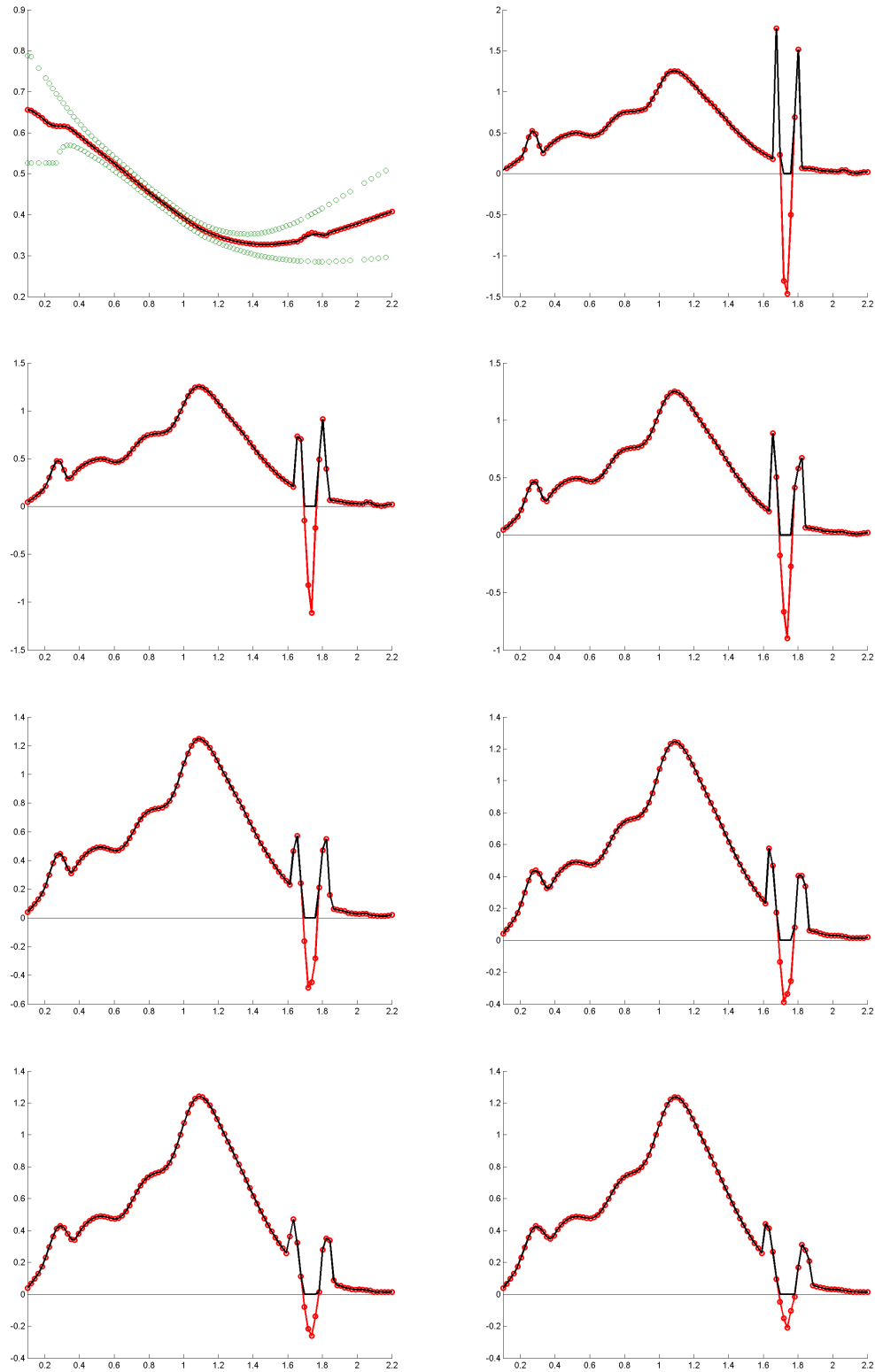
Fig. 3 Adjustment of the analysis proposed in Figure 1 after removing the observable outliers. Estimation of SPD for $h_\kappa = 0.08, 0.09, \ldots, 0.14$.

Also in the surface case, the triweight kernel function produces the highest value and the Epanechnikov kernel function represents a good balance among the considered kernel functions.
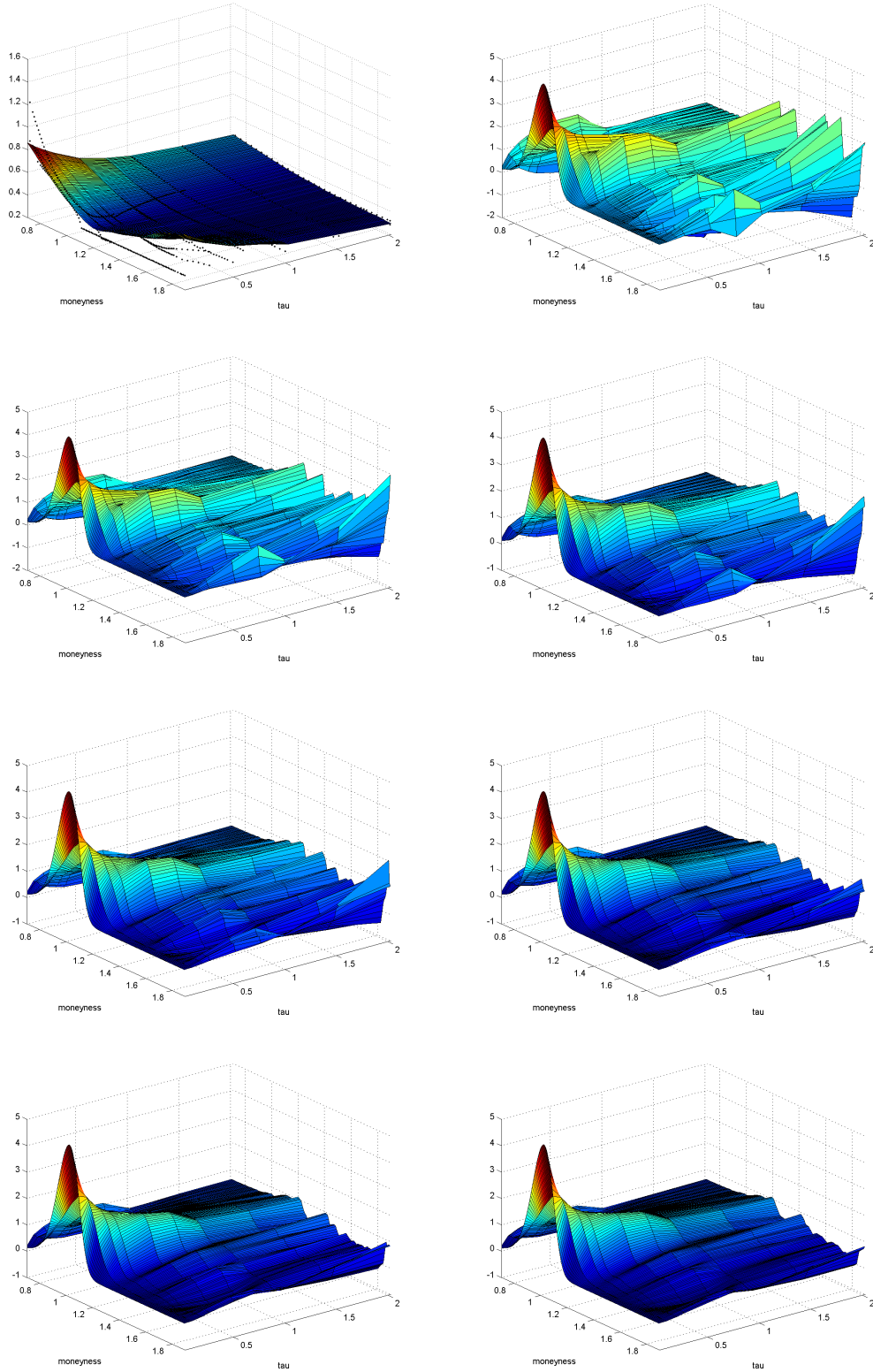
Fig. 4 Estimations using the original calendar grid - IV surface (top left) with $h_\kappa = 0.12$ and SPD surfaces with $h_\kappa = 0.08, 0.09, \ldots, 0.14$.

3.3 Implied volatility surface - Artificial calendar grid

Thus far, all the surface computations have been done using as calendar grid all the maturities available in the data. Therefore, the accuracy was strictly related with the data structure. In the following analysis
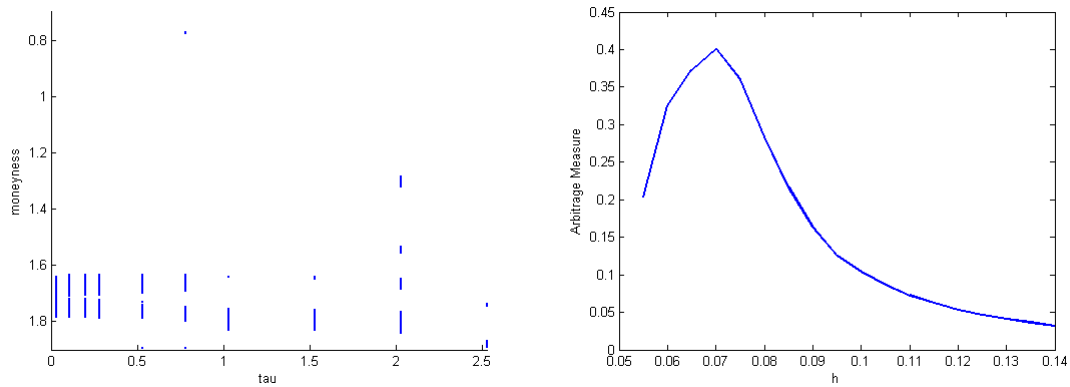
Fig. 5 Surface SPD with original calendar grid - Arbitrage interval (left) and arbitrage measure (right)

we artificially discretize the calendar grid by constructing a monthly step grid till two years. Then, we obtain a calendar grid with twenty-four points. In Figure 6 we show the estimation of the IV surface computed with (27) using as moneyness bandwidth $h_\kappa = 0.12$ (top left), the other graphs present the estimations of the SPD surfaces for a set of moneyness bandwidths $h_\kappa = 0.08, 0.09, \dots, 0.14$. The results show again some turbulence for moneyness around 1.7.

In order to compare these results with the previous ones, in Figure 7 we propose some cuts of the surfaces. In particular, we choose the three shortest maturities of the option data and the corresponding maturities in the artificial grid. These maturities are $\tau = 11$, 39 and 74 days. The estimations are performed with $h_\kappa = 0.12$ and with the Epanechnikov kernel functions. The cuts of the SPD estimated with the artificial calendar grid (*blue line*) and the cuts of the SPD estimated with the original grid (*black line with circles*) become very similar when we consider relatively long maturities. For the shortest one, the cut extracted from the SPD estimated with the original grid has fatter tails.

To analyze in more details the behaviour of the estimates in the artificial calendar case, we repeat the computation that we did for the previous case. Similarly to previous analysis, Figure 8 shows the length of the arbitrage intervals (left) and arbitrage measure (right) in term of moneyness. The arbitrage intervals highlight two sectors of arbitrage: around moneyness 1.7 till 1-year maturities and around 1.8 from 1-year to 2-year maturities. The curve which represents the volume is again not strictly decreasing for increasing bandwidth. However, both charts in Figure 8 are more regular than those in Figure 5 obtained using the original calendar grid. This means that a well defined grid in the calendar direction induces a more regular behavior of the arbitrage evidence throughout the moneyness-maturity space.

3.4 Estimating implied volatility surface for dividend-paying stocks

Finally, we propose the computation of the IV and SPD surfaces for options written on a dividend-paying stock (SAP SE) listed at the German option market on December 8, 2008. We consider all the available maturities till 376 days since after this maturity we have too few data. In Figure 9 (top left) We propose the estimation of the IV surface assuming the moneyness bandwidth $h_\kappa = 0.12$ and calendar bandwidth $h_\tau = 1.3$ since using the same bandwidth as with the DAX case ($h_\tau = 1$) we would not have enough data for the estimation.

In the same figure we present the estimation of the SPD surfaces. The computation is done in the same way as previously except the calendar bandwidth ($h_\tau = 1.3$). For medium and long maturities, we notice a remarkable peak for moneyness 1.3 and a small turbulence for moneyness around 0.9. Increasing the moneyness bandwidth the small ripple at 0.9 moneyness becomes smoothed, whilst the big one persists. A reasonable explanation for this anomaly is the huge market crash occurred few days before our observed date, therefore the market expected a rebound in the long run. We focus a further analysis on a set of shorter maturities in which the classical shape of the SPD is more evident.

In particular, we compare the estimation according to three maturities by cutting the surfaces and overlapping the cut with the corresponding one dimension estimation on the same maturities. In Figure 10, we present results for $\tau = 39$, 74 and 102 days, and adopting three representative moneyness bandwidths $h_\kappa = 0.10, 0.12, 0.14$. In each chart, the blue line with dots represents the cut of the surface estimations and the black line represents the one dimension estimation. In order to make a fair comparison, also for the one dimension estimation we consider the Epanechnikov kernel function.
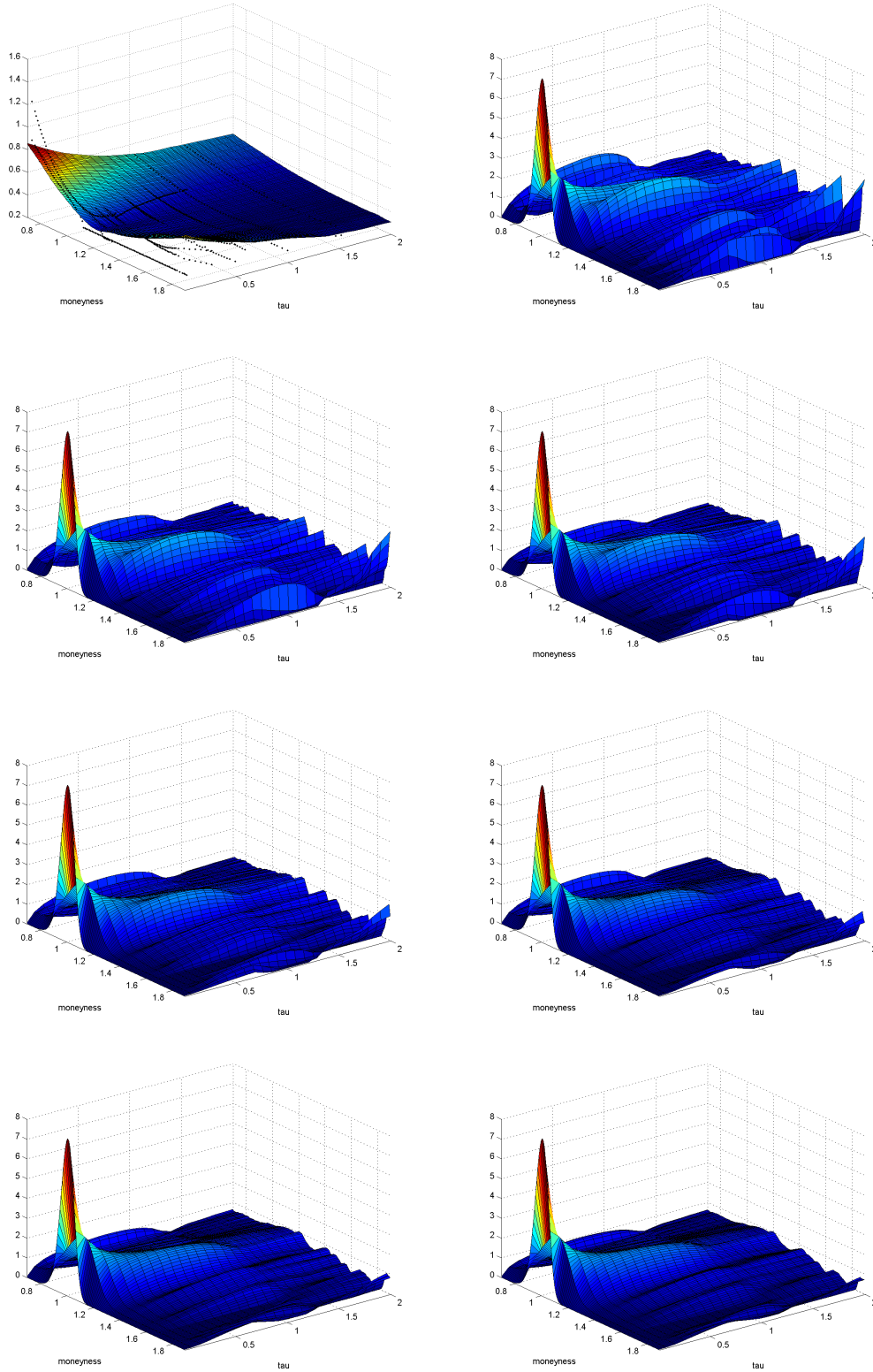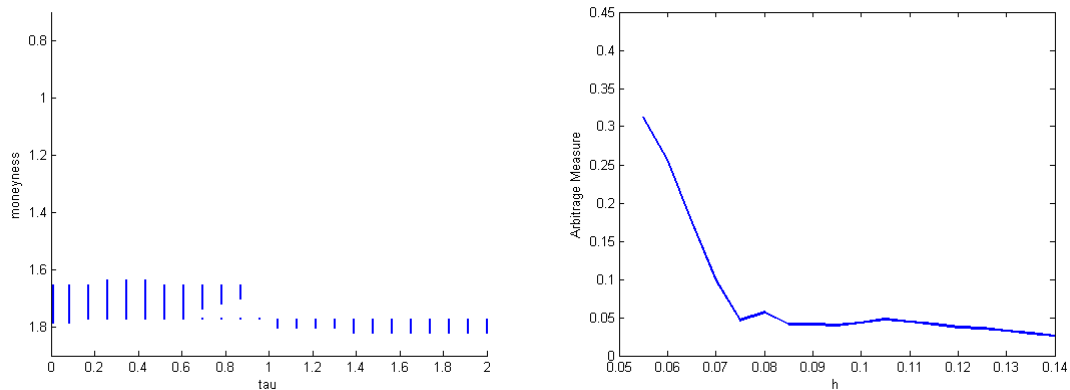
Fig. 6 Estimations using the artificial calendar grid - IV surface (top left) with $h_\kappa = 0.12$ and SPD surfaces with $h_\kappa = 0.08, 0.09, \ldots, 0.14$.

According to the results shown in Figure 10, it is clear that the surface estimation highlights an arbitrage opportunity which does not exist in the one dimension estimation. In particular, we notice that for moneyness bandwidth $h_\kappa = 0.10$ (first row of the figure) the surface cut is negative at the moneyness values of 0.8 and 1.35. Therefore, in this case, the behavior of the estimation for the same maturity
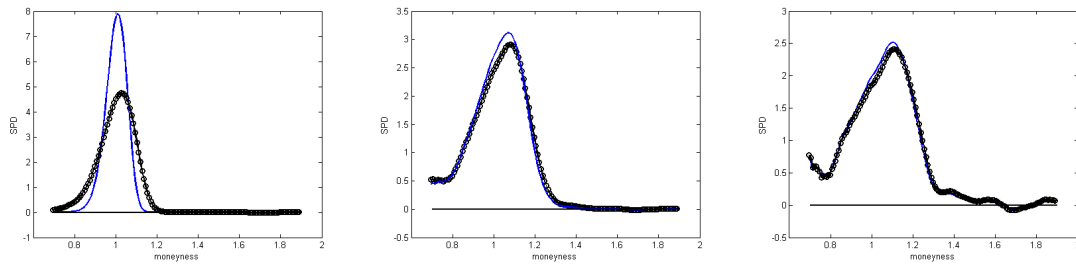
Fig. 7 Comparison between SPD estimated with the artificial calendar grid (*blue line*) and the SPD estimated with the original grid (*black line with circles*) - Cuts of the surfaces in correspondence of three maturities: $\tau = 11$, 39 and 74 days


Fig. 8 Surface SPD with artificial calendar grid - Arbitrage interval (left) and arbitrage measure (right)

differs significantly according to the type of estimation we compute: one or two dimensions. For the other moneyness bandwidths, $h_\kappa = 0.12, 0.14$, even if there in no arbitrage opportunity, the difference of the shape of the two estimations is clear. The one dimension SPD appears to be more smooth, whilst the surface cut is affected by the calendar fitting and this fact enhances its irregular behavior.

Conclusion

In this paper we propose a deep analysis of the estimation of the IV and of the SPD of options quoted in the market. The methodology proposed in Benko et al. (2007) has been applied to a set of put and call options written on DAX index in a specific day with a specific maturity. We shown that the SPD estimated within an unconstrained setting, i.e. not constrained to be positive, presents a negative part which corresponds to an arbitrage opportunity. To prove that the arbitrage is a genuine characteristic and is not induced by a particular bandwidth selection, we show that the arbitrage intervals are not strictly decreasing for increasing bandwidth; and to prove that are not induced by outliers, we perform the estimations after removing the observable outliers but the arbitrage persists. The further step of the analysis consists in the extension of the model to the surface estimation for both the IV and the SPD, i.e. considering jointly all the observable maturities in the considered day. As shown, the arbitrage is evident not only for the maturity adopted in the one dimension estimation, but also for all the others in correspondence to the same moneyness level. The study has been enriched changing the calendar grid from the one composed of the maturities of the data options to an artificial grid which splits the 2-year horizon in 24 points. Several comparisons have been proposed regarding the differences between the one dimension and the two dimensions estimations, and between the original and the artificial calendar grid. In order to compare the magnitude of the arbitrage between different choices of kernel function and bandwidth, we propose an arbitrage opportunity measure for the one dimension estimation and for the surface estimation. Finally, we further extended the methodology considering the case of options written on dividend-paying stock and we propose an estimation of IV and SPD surfaces for a German stock.
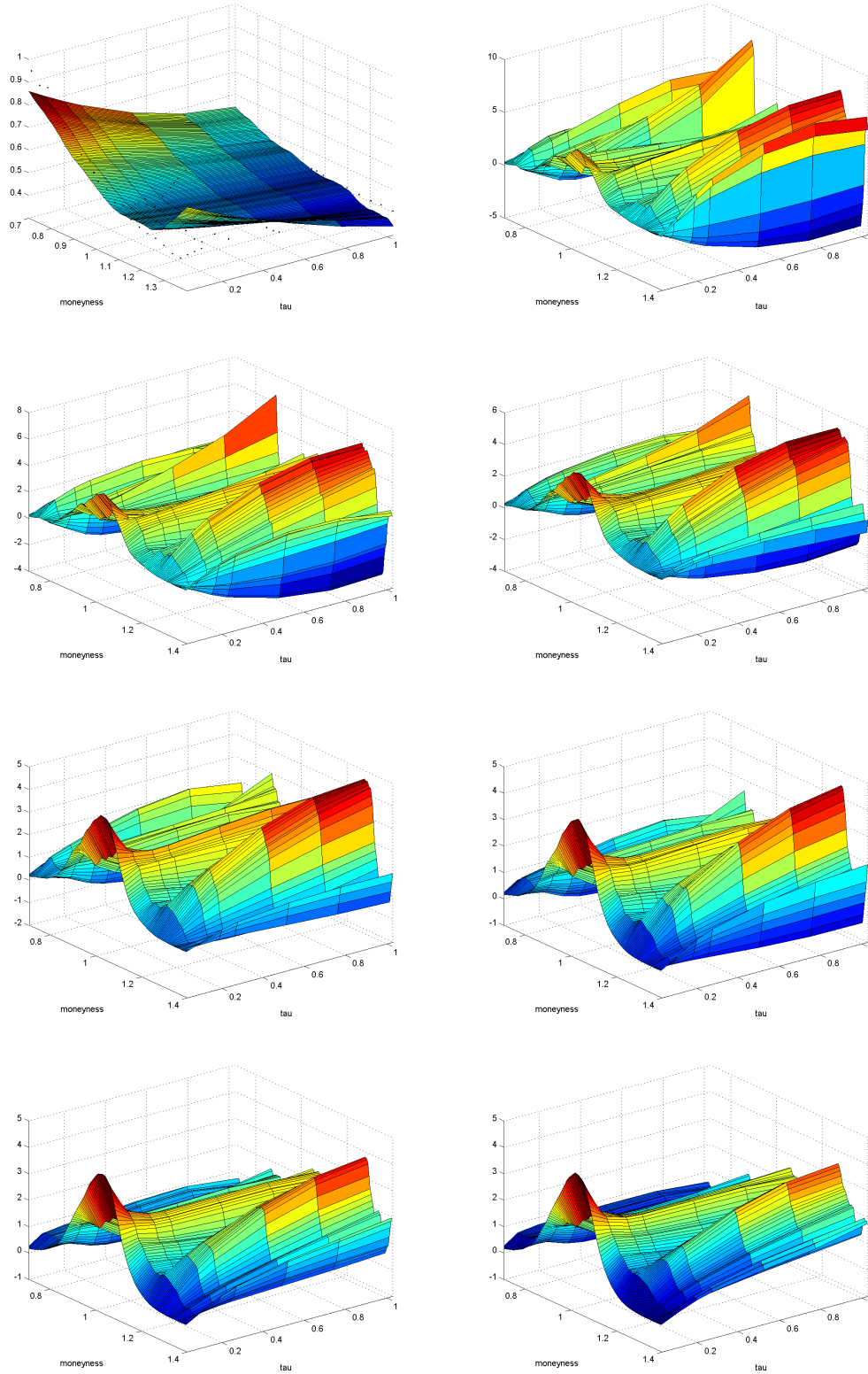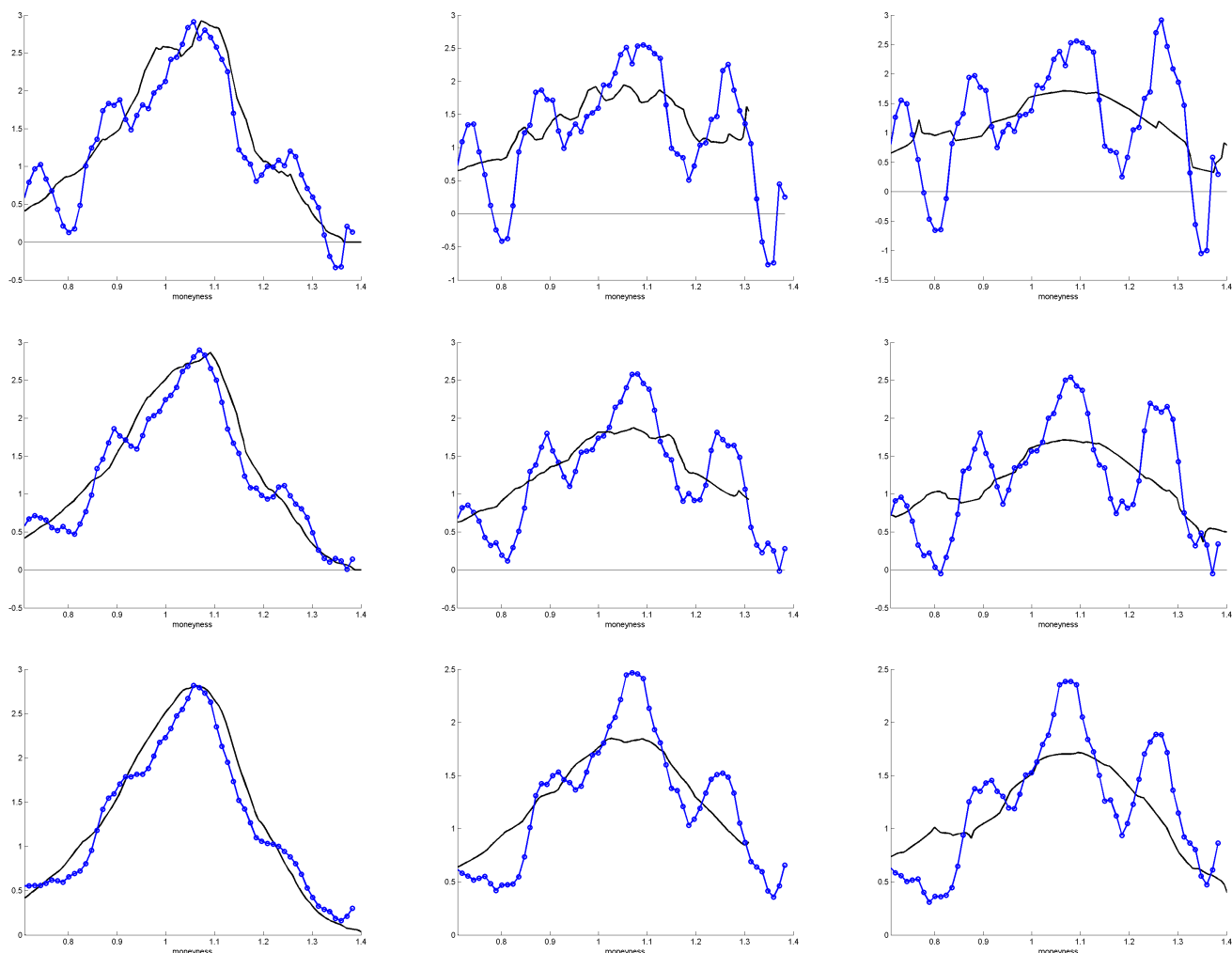
Fig. 9 Estimations for dividend-paying stock - IV surface (top left) with $h_\kappa = 0.12$ and SPD surfaces with $h_\kappa = 0.08, 0.09, \ldots, 0.14$.

Acknowledgment

Fig. 10 Dividend-paying stock analysis: comparison between SPD surface cut (blue line with dots) and one dimension SPD estimation (black line) for $\tau = 39$, 74, and 102 days (from left to right) and $h_\kappa = 0.10$, 0.12, and 0.14 (from top to bottom).

References

Avellaneda, M., Friedman, C., Holmes, R., & Samperi, D. (1997). Calibrating volatility surfaces via relative-entropy minimization. *Applied Mathematical Finance*, *4*(1), 37–64.

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

Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities. *The journal of political economy*, *81*(3), 637–654.

Bloch, D. A., Coello, C. C., Securities, M., et al. (2011). Smiling at evolution. *Applied Soft Computing*, *11*(8), 5724–5734.

Borovkova, S. & Permana, F. J. (2009). Implied volatility in oil markets. *Computational Statistics & Data Analysis*, *53*(6), 2022–2039.

Brockhaus, O., Farkas, M., Ferraris, A., Long, D., & Overhaus, M. (2000). *Equity Derivatives and Market Risk Models*. London: Risk.

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

Cont, R., Da Fonseca, J., et al. (2002). Dynamics of implied volatility surfaces. *Quantitative finance*, *2*(1), 45–60.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *The Journal of Finance*, *53*(6), 2059–2106.

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

Fan, J. & Gijbels, I. (1996). *Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability 66*. London: CRC Press.

Fan, J. & Yao, Q. (2003). *Nonlinear Time Series*. Springer.

Fengler, M. (2005). *Semiparametric Modeling of Implied Volatility*, volume 1. Ed of *Springer Finance*. Heidelberg: Springer Bln.

Fengler, M. R. (2006). *Semiparametric Modeling of Implied Volatility*. Springer.

Fengler, M. R. (2009). Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, *9*(4), 417–428.

Fengler, M. R. (2012). Option data and modeling bsm implied volatility. In *Handbook of computational finance* (pp. 117–142). Springer.

Fengler, M. R., Härdle, W. K., & Villa, C. (2003). The dynamics of implied volatilities: A common principal components approach. *Review of Derivatives Research*, *6*(3), 179–202.

Fengler, M. R. & Hin, L.-Y. (2015). Semi-nonparametric estimation of the call-option price surface under strike and time-to-expiry no-arbitrage constraints. *Journal of Econometrics*, *184*(2), 242–261.

Glaser, J. & Heider, P. (2012). Arbitrage-free approximation of call price surfaces and input data risk. *Quantitative Finance*, *12*(1), 61–73.

Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge: Cambridge Univ Press.

Homescu, C. (2011). Implied volatility surface: Construction methodologies and characteristics. *Available at SSRN 1882567*.

Jackwerth, J. C. (2004). *Option-implied Risk-neutral Distributions and Risk Aversion*. Charlotteville (USA): Research Foundation of AIMR.

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

Kahale, N. (2004). Option pricing an arbitrage-free interpolation of volatilities. *Risk*, *17*(5), 102–107.

Kim, N. & Lee, J. (2013). No-arbitrage implied volatility functions: Empirical evidence from kospi 200 index options. *Journal of Empirical Finance*, *21*, 36–53.

Ludwig, M. (2015). Robust estimation of shape-constrained state price density surfaces. *The Journal of Derivatives*, *22*(3), 56–72.

Merton, R. C. (1973). Rational theory of option pricing. *Bell Journal of Economics and Management Science*, *4*, 141–183.

Shimko, D. (1993). Bounds of probability. *Risk*, *6*(4), 33–37.