Uniform Confidence Band for Pricing Kernels

A Master Thesis Presented

by

Weining Wang

(522049)

to

Prof. Dr. Wolfgang Härdle

Prof. Dr. Ostap Okhrin

CASE - Center of Applied Statistics and Economics

Humboldt University, Berlin

in partial fulfillment of the requirements

for the degree of

Master of Statistics

Berlin, 19. Jan, 2010

Contents

1 Abstract **2**

2 Introduction **3**

3 Empirical Pricing Kernel Estimation **7**
3.1 EPK and option pricing 8

4 Confidence intervals and confidence bands **14**

5 The sheet in maturity dimension **17**

6 Confidence bands based on smoothing implied volatility **19**

7 Bootstrap confidence bands **22**

8 Monte-Carlo study **24**
8.1 How well is the BS model covered? 24
8.2 How well is the band in reality? 27

9 An illustration with DAX data **30**
9.1 Data 30
9.2 Estimation of DAX EPK and its uniform confidence band 32
9.3 Linking economic conditions to EPK dynamics 34

10 Conclusions **36**

11 Appendix **38**
11.1 Proof of Lemma 1 39
11.2 Proof of Lemma 2 40
11.3 Proof of Theorem 2 41
11.4 Expressions for the semiparametric estimator of SPD 43

List of Figures **45**

List of Tables **47**

Bibliography **49**

1 Abstract

Pricing kernels implicit in option prices play a key role in assessing the risk aversion over equity returns. We deal with nonparametric estimation of the pricing kernel (Empirical Pricing Kernel) given by the ratio of the risk-neutral density estimator and the subjective density estimator. The former density can be represented as the second derivative w.r.t. the European call option price function, which we estimate by nonparametric regression. The subjective density is estimated nonparametrically too. In this framework, we develop the asymptotic distribution theory of the EPK in the L^∞ sense. Particularly, to evaluate the overall variation of the pricing kernel, we develop a uniform confidence band of the EPK. Furthermore, as an alternative to the asymptotic approach, we propose a bootstrap confidence band. The developed theory is helpful for testing parametric specifications of pricing kernels and has a direct extension to estimating risk aversion patterns. The established results are assessed and compared in a Monte-Carlo study. As a real application, we test risk aversion over time induced by the EPK.

Keywords: Empirical Pricing Kernel; Confidence band; Bootstrap; Kernel Smoothing; Nonparametric Fitting

2 Introduction

A challenging task in financial econometrics is to understand investors' attitudes towards market risk in its evolution over time. Such a study naturally involves stochastic discount factors, empirical pricing kernels (EPK) and state price densities, see Cochrane (2001). Asset pricing kernels summarize investors' risk preferences and exhibit when estimated from data, the so called "EPK paradox", as several studies including Aït-Sahalia and Lo (2000), Brown and Jackwerth (2004), Rosenberg and Engle (2002) have shown. Although in all these studies the EPK paradox (non-monotonicity) became evident, a test for the functional form of the pricing kernel has not been devised yet. A uniform confidence band drawn around an EPK is a very simple tool for shape inspection. Confidence bands based on asymptotic theory and bootstrap is the subject of our paper. In addition, we relate critical values of our test to changing market conditions given by exogenous time series.

The common difficulty is that the investors' preference is implicit in the goods traded in the market and thus can not be directly observed from the path of returns. A profound martingale based pricing theory provides us one approach to attack the problem from a probabilistic perspective. An important concept involved is the State Price Density (SPD) or Arrow-Debreu prices reflecting fair prices of one unit gain or loss for the whole market. Under no arbitrage assumption, there exists at least one SPD, and when a market is complete, there is a unique SPD. Assuming markets are complete, pricing is done under a risk neutral measure, which is related to the pdf of the historical measure by multiplying with a stochastic discount factor, see section 2 for a detailed illustration. >From an economic perspective, the price is formulated according to utility maximization theory, which admits that the risk preference of consumers is connected to the shape of utility functions. Specifically, a concave, convex or linear utility function describes the risk averse, risk seeking or risk neutral behavior. Importantly, a stochastic discount factor can be expressed via

a utility function (Marginal Rate of Substitution), which links the shape of pricing kernel (PK) to the risk patterns of investors, see Kahneman and Tversky (1979), Jackwerth (2000), Rosenberg and Engle (2002) and others.

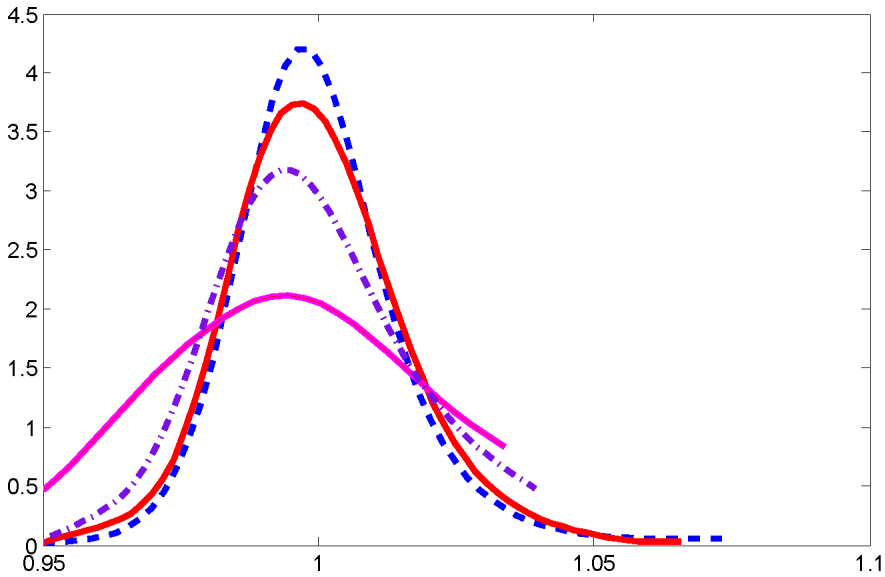


Figure 1: Examples of inter-temporal pricing kernels (as functions of moneyness) with fixed maturity 0.00833 (3 days) in years respectively on 17-Jan-2006 (blue), 18-Apr-2006 (red), 16-May-2006 (magenta), 13-June-2006 (black), see Grith, Härdle and Park (2009).

The above mentioned theory allows us to relate prices processes of assets traded in the market to risk preference of investors. This amounts to fit a flexible model to make inference on the dynamics of EPKs over time in different markets. A well-known but restrictive approach is to assume the underlying following a Geometric Brownian Motion. In this setting, risk neutral densities and historical densities are log normal distributions with different drifts, and the pricing kernel has the form of a derivative of a power utility. Thus it is decreasing in return and implies overall-risk averse behavior. However, across different markets, one observes quite often a non-decreasing pattern for EPKs, a phenomenon is called the EPK paradox, see Chabi-Yo, Garcia and Renault (2008).

Two plots of pricing kernels are shown in Figure 1 and Figure 2. Figure 1 depicts inter-temporal pricing kernels with fixed maturity, while Figure 2 depicts pricing kernels with two different maturities and their confidence bands. The figures are shown on a returns scale. The curves present a bump in the middle and a switch from convexity to concavity in all cases. Especially, this shows that very unlikely the bands contain a monotone decreasing curve.

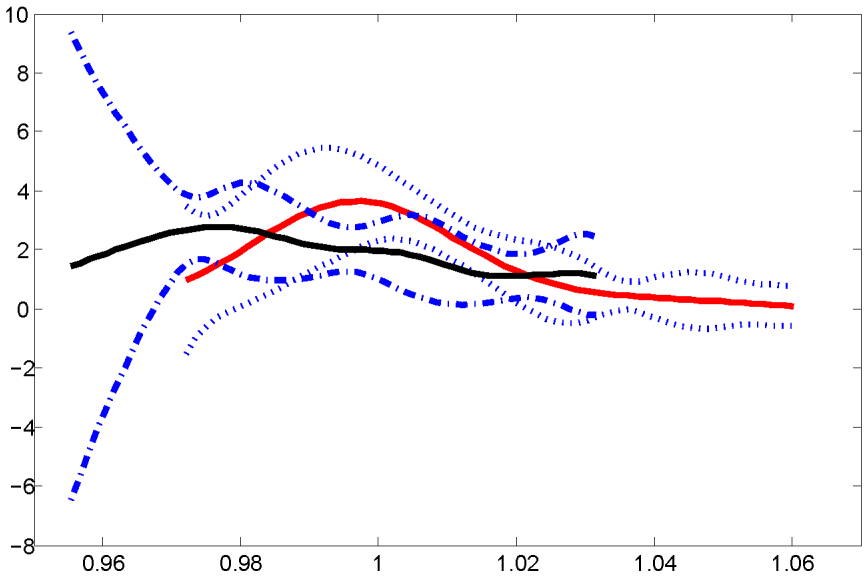


Figure 2: Examples of inter-temporal pricing kernels with various maturities in years: 0.02222 (8 days, red) 0.1 (36 days, black) on 12-Jan-2006 and their confidence bands.

In order to study the EPK paradox, the time varying coverage probability of a uniform confidence band gives us reliable information. At a fixed point in time, it helps us to test alternatives for a PK and thus yields insights into time varying risk patterns. A test on monotonicity of the PK has been proposed by Golubev, Härdle and Timofeev (2009), the extracted time varying parameter, realized either from a low dimensional model for PK or given the coverage probability, may thus be economically analyzed in connection with exogenous macroeconomic business cycle

indicator, e.g. credit spread, yield curve, etc, see also Grith, Härdle and Park (2009).

Several econometric studies are concerned with estimating PKs by estimating a risk neutral density and historical density separately. See section 2 for details. It is stressed in Aït-Sahalia and Lo (1998) that nonparametric inference from pricing kernels gives unbiased insights into the properties of asset markets. The stochastic fluctuation of EPK as measured by the maximum deviation has not been studied yet. However, the asymptotic distribution of the maximum deviation and the uniform confidence band linked to it are very useful for model check.

Uniform confidence bands for smooth curves have first been developed for kernel density estimators by Bickel and Rosenblatt (1973), Extension to regression smoothing can be found in Liero (1982) and Härdle (1989). But only recently, the results have been carried over to derivative smoothing by Claeskens and Van Keilegom (2003). Our theoretical path follows largely their results, but our results are applied to a ratio estimator instead of just a local polynomial estimator. Additionally, we extend their results into 2 dimension by including the maturity. Also we have a realistic data situation that relates coverage to economic indicators. In addition we perform the smoothing in an implied volatility space which brings by itself an interesting modification of the results of that paper.

The paper is organized as follows: In section 2, we describe the theoretical connection between utility functions and pricing kernels. In section 3, we present a nonparametric framework for the estimation of both the historical and the risk neutral density and derive the asymptotic distribution of the maximum deviation. In section 4, we simulate the asymptotic behavior of the uniform confidence band and compare it with the bootstrap method. Moreover, we also compare the result with other parametric estimations. In section 5, we conclude and discuss our results.

3 Empirical Pricing Kernel Estimation

Consider an arbitrary risky financial security with the price process $\{S_t\}_{t\in[0,T]}$. We assume that $\{S_t\}$ is a nonnegative semi-martingale with continuous marginals. In a dynamic equilibrium model the price of the security at time t is equal to the expected net present value of its future payoffs. The interest rate process r is deterministic. We assume that the market is complete, i.e. there exists one positive r.v. π such that

$$\begin{aligned} \mathsf{E}\left[\pi\right] &= 1, \\ \mathsf{E}\left[S_T \frac{\pi}{\mathsf{E}[\pi|S_t]}|S_t\right] &= e^{r\tau} S_t. \end{aligned}$$

From the risk neutral valuation principle, for the nonnegative payoff $\psi(S_T)$, it holds

$$\mathsf{E}\left[e^{-r\tau}\psi(S_T)\pi\right] = \mathsf{E}\left[e^{-r\tau}\psi(S_T)\,\mathsf{E}[\pi|S_T]\right].$$

By factorization, we obtain $\mathsf{E}[\pi|S_T] = \mathcal{K}_\pi(S_T)$, implying

$$\mathsf{E}\left[e^{-r\tau}\psi(S_T)\pi\right] = \int_0^{+\infty} e^{-r\tau}\psi(x)\mathcal{K}_\pi(x)p_{S_T}(x)dx,$$

where $p_{S_T}(x)$ is the pdf of the price S_T. The last expression justifies the notion of *pricing kernel* used for $\mathcal{K}_\pi(x)$. Following Ait-Sahalia and Lo (2000) consider the distribution $Q_{S_T}(x) = P_Q(S_T \leq x) = \int_{-\infty}^{x} \pi(z)p_{S_T}(z)dz$. Its density function q_{S_T} is commonly called the *risk neutral distribution* of S_T or *state price density* (SPD). This follows by observing that, for any ψ,

$$\mathsf{E}\left[e^{-r\tau}\psi(S_T)\pi\right] = \int_0^{+\infty} e^{-r\tau}\psi(x)q_{S_T}(x)dx. \tag{1}$$

This implies that $\mathcal{K}_\pi = \frac{q_{S_T}}{p_{S_T}}$ and both the pdf of the future payoff and the SPD are required to compute the EPK. Several approaches are available to determine

the EPK explicitly. First, we can impose strict parametric restrictions on the dynamics of the asset prices and on the distribution of the future payoff. An example are mixture normal distributions, Jackwerth (2000). In the case of more complex stochastic processes, usually no explicit solution is available. A possible technique though is to use the Brownian motion setup as prior model. Subsequently the SPD is estimated by minimizing the distance to the prior SPD subject to the constraints characterizing the underlying securities, see Rubinstein (1994) and Jackwerth and Rubinstein (1996). Second, the EPK can be determined from the utility function of the agent. Let the aim the investor be to solve the problem:

$$\max_{W_t}\{u(W_t) + \mathsf{E}_t[\beta u(W_T)]\},$$

where $u(\cdot)$ denotes the utility function, W_t the wealth and β the subjective discount factor. It can be shown that the stochastic discount factor is proportional to the EPK and is given by

$$e^{-r\tau}\frac{q_{S_T}(s)}{p_{S_T}(s)} = \beta\frac{u'(s)}{u'(S_t)}.$$

This implies that by fixing the utility of the investor we can determine the EPK. In practice, however, usually the opposite procedure is applied. The EPK is statistically estimated and used to determine the utility function or the risk aversion coefficient of the investor.

3.1 EPK and option pricing

Here we consider the SPD q_{S_T} and the pdf p_{S_T} separately. For notational convenience we drop the index and write simply q and p. The latter can easily be estimated either parametrically or nonparametrically from the time series of payoffs. On the contrary, the SPD depends on risk preferences and therefore the past observed stock time series do not contain information. Option prices do reflect preferences and,

therefore, can be used to estimate the SPD q. Let $C(S_t, X, \tau, r, \sigma^2)$ denote the European call-option price as a function of the strike price X, price S_t, maturity τ, interest rate r.

$$q(S_T) = e^{r\tau} \frac{\partial^2 C}{\partial K^2}\Big|_{K=S_T}.$$

In a Black-Scholes (BS) framework, where the underlying asset price S_t follows a geometric Brownian motion. The European options are priced via:

$$C(S_t, X, \tau, r, \sigma^2) = S_t \Phi(d_1) - X e^{r\tau} \Phi(d_2),$$

where d_1 and d_2 are known functions of σ^2, τ, X and S_t. This implies that both $q(S_T)$ and $p(S_T)$ are lognormal distributions:

$$q(S_T) = \frac{1}{S_T\sqrt{2\pi\sigma^2\tau}} \exp\left[-\frac{\{\log(S_T/S_t) - (r - \sigma^2/2)\tau\}^2}{2\sigma^2\tau}\right] \tag{2}$$

and $p(S_T)$ with μ replacing r in (2).

Beside the modeling bias that is implicit in the BS model, it is also not possible to reflect the implicit volatility smile (surface) as a function of X and τ via (2). The latter may be derived in stochastic volatility models of Heston or Bates type or even more complex parametrizations. In order to study unbiased risk patterns, we need to guarantee models for the pricing kernel that are rich enough to reflect local risk aversion in time and space. This leads naturally to a smoothing approach.

Consider call options with maturity maturity τ. The intraday call options are observed

$$Y_i = C_\tau(X_i) + \sigma(X_i)\varepsilon_i, \quad i = 1, \ldots, n_q, \tag{3}$$

where Y_i denotes the observed option price and K_i the strike price. Y_i and X_i are assumed to be i.i.d. in the cross section with $\mathsf{Var}(Y_i|X = X_i) = \sigma^2(X_i)$. It should be said that the perceived errors are due to neglected heterogeneity factors, rather

than mispricings exploited by arbitrage strategies, see Renault (1997). Figure 3 depicts the call option prices data used to calculate a SPD. The observations are distributed with different variances at discrete grid points of strikes prices. For sim-

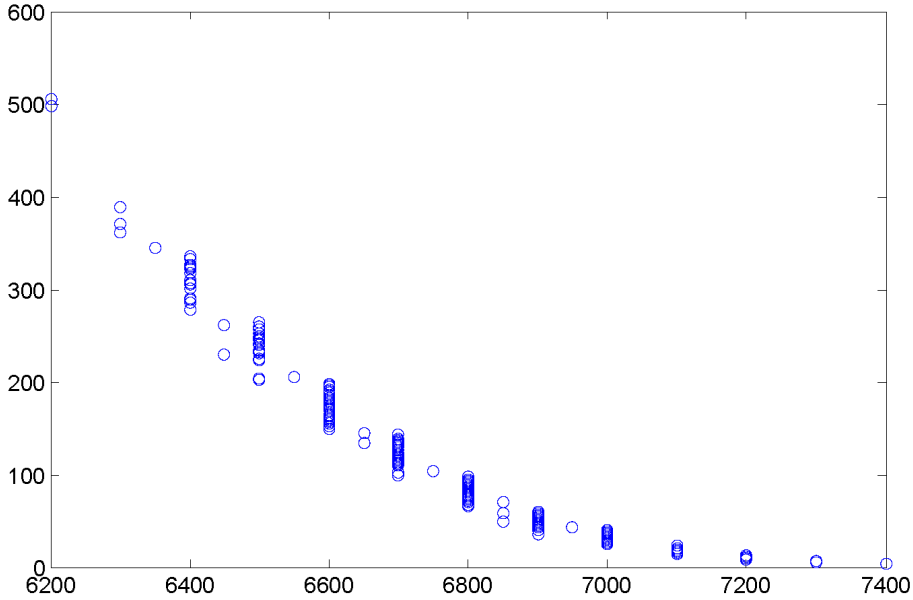


Figure 3: Plot of call option prices against strike prices 20010117

plicity of notation, we write $C(X)$ for $C_\tau(X)$. Assuming that $C(X)$ is continuously differentiable of order $p = 3$, it can be locally approximated by

$$C(X, X_0) = \sum_{j=0}^{p} C_j(X)(X_0 - X)^j, \tag{4}$$

where $C_j(X) = C^{(j)}(X)/j!$, $j = 0, \ldots, p$.

See Cleveland (1979), Fan (1992), Fan (1993), Ruppert and Wand (1994) for more details. Assuming local Gaussian quasi-likelihood model, we arrive at:

$$B_{n_q}\{C(X)\} = \frac{1}{n_q} \sum_{i=1}^{n_q} K_{h_{n_q}}(X_i - X) Q\{Y_i, C(X, X_i)\}, \tag{5}$$

where $K_{h_{n_q}}(X_i - X)$ is a (Gaussian) kernel function with a bandwidth sequence h_{n_q}.

The vector of solutions is $\boldsymbol{C}(X) = \{C_0(X), C_1(X), \ldots, C_p(X)\}^\top$ is obtained via the optimization problem

$$\hat{\boldsymbol{C}}(X) =_{X \in E} B_{n_q}\{\boldsymbol{C}(X)\}, \tag{6}$$

where E is a certain compact set. This is equivalent to solving

$$\mathbf{A}_{n_q}(X) \stackrel{\text{def}}{=} \frac{1}{n_q} \sum_{i=1}^{n_q} K_{h_q}(X_i - X) \frac{\partial Q\{Y_i, C(X, X_i)\}}{\partial C} \mathbf{X}_i \quad = \quad \mathbf{0}, \tag{7}$$

with $\mathbf{X}_i \stackrel{\text{def}}{=} (1, X_i - X, (X_i - X)^2, (X_i - X)^3)^\top$. We are concerned with $2!\hat{C}_2(x) = \frac{\partial^2 C(X)}{\partial X^2}\Big|_{X=x}$, which is shown by Breeden and Litzenberger (1978) to be proportional to $q(x)$.

In practice we assume a Gaussian quasi-likelihood function with $Q\{Y, C(X)\} = 1/2\{Y - C(X)\}^2/\sigma^2(X)$, which is equivalent to local least squares smoothing. Additionally note that we assume the parameter $C(.)$ and $\sigma(.)$ to be orthogonal to each other. Thus we can estimate them separately as in a single parameter case. Let $A_{n_q,j}$ denote the jth component of the vector equation $\mathbf{A}_{n_q}$. This component corresponds to the j-th derivative of the option price evaluated at X. The following lemma states the results on the existence of the solution and its consistency.

Lemma 1. *Under conditions* $(A1) - (A5)$*, there exists a sequence of solutions to the equations*

$$\mathbf{A}_{n_q}(x) = \mathbf{0}$$

such that

$$\sup_{x \in E} |\hat{q}(x) - q(x)| = \mathcal{O}\Big[h_{n_q}^{-2}\{\log n_q/(n_q h_{n_q})\}^{1/2} + h_{n_q}^2\Big] \quad a.s.$$

The density of the returns can be estimated separately from the SPD using historical prices $S_1, \ldots, S_{n_p}$ of the underlying asset. The nonparametric kernel estimator of

p_{S_T} is given by

$$\hat{p}(x) = n_p^{-1} \sum_{j=1}^{n_p} K_{h_{n_p}}(x - S_j),$$

where h_{n_p} is the bandwidth of the kernel L which not necessarily coincides with the kernel for q. Under assumption (A5), we know that

$$\sup_{x \in E} |\hat{p}(x) - p(x)| = \mathcal{O}\{(n_p h_{n_p} / \log n_p)^{-1/2} + h_{n_p}^2\}. \tag{8}$$

The estimator of the EPK is then given by the ratio of the estimated SPD and the risk-neutral density, i.e. $\hat{\mathcal{K}}(x) = \hat{q}(x)/\hat{p}(x)$. The next lemma provides the linearization of the ratio, which is important for further statements about the uniform confidence band of the EPK.

Lemma 2. *Under conditions (A1)-(A5) it holds*

$$\begin{aligned} & \sup_{x \in E} |\hat{\mathcal{K}}(x) - \mathcal{K}(x)| \\ = & \sup_{x \in E} \Big| \frac{\hat{q}(x) - q(x)}{p(x)} - \frac{\hat{p}(x) - p(x)}{p(x)} \cdot \frac{q(x)}{p(x)} - \frac{\{\hat{q}(x) - q(x)\}\{\hat{p}(x) - p(x)\}}{p^2(x)} \Big| \\ & + o[\max\{(n_p h_{n_p} / \log n_p)^{-1/2} + h_{n_p}^2, h_{n_q}^{-2}\{n_q h_{n_q} / \log n_q\}^{-1/2} + h_{n_q}^2\}]. \end{aligned} \tag{9}$$

This lemma implies that the stochastic deviation of $\widehat{\mathcal{K}}$ can be linearized into a stochastic part containing the estimator of the SPD and a deterministic part containing $\mathsf{E}[\hat{p}(x)]$. The uniform convergence can be proved by dealing separately with the two parts. The convergence of the deterministic part is shown by imposing mild smoothness conditions, while the convergence of the stochastic part is proved by following the approach of Claeskens and Van Keilegom (2003). Theorem 1 formalizes this uniform convergence of the EPK.

Theorem 1. *Under conditions* $(A1) - (A5)$ *and for all* $x \in E$*, it holds*

$$\sup_{x \in E} |\hat{\mathcal{K}}(x) - \mathcal{K}(x)| = \mathcal{O}[\max\{(n_p h_{n_p}/\log n_p)^{-1/2} + h_{n_p}^2, h_{n_q}^{-2}\{n_q h_{n_q}/\log n_q\}^{-1/2} + h_{n_q}^2\}] \quad a.s.$$

The proof is given in the appendix.

4 Confidence intervals and confidence bands

Confidence intervals characterize the local precision of the EPK for a given fixed value of the payoff. This allows to test EPKs at each particular return, but does not allow conclusions about the global shape. The confidence bands, however, characterize the whole EPK curve and offer therefore the possibility to test for shape characteristics. In particular, it is a way to check the persistence of the bump as observed. Give a certain shape rejection, one may verify the restriction imposed by the power utility and obtain insights about the risk aversion of the agents. In addition, the confidence bands can be used to measure the global variability of the EPK. Also, the proportion of BS fitting covered in nonparametric bands can be used as a measure of global risk aversion.

A confidence interval for the EPK at a fixed value x requires the asymptotic distribution of $\hat{p}(x)$ and $\hat{q}(x)$. Hereafter, we use $\mathcal{L}$ to denote the convergence in law. Under (A1)-(A5):

$$\sqrt{n_p h_{n_p}}\{\hat{p}(x) - p(x)\} \xrightarrow{\mathcal{L}} N\{0, p(x) \int K^2(u)du\}$$

and

$$\sqrt{n_q h_{n_q}^5}\{\hat{q}(x) - q(x)\} \xrightarrow{\mathcal{L}}, N\{0, \sigma_q^2(x)\},$$

where $\sigma_q^2 = [B(x)^{-1}N^{-1}TN^{-1}]_{(3,3)}$, with $B(x)$ equal to the product of the density $f_X(x)$ of the strike price and the local Fisher information matrix $I\{C(x)\}$. The matrices N and T are given by $N \stackrel{\text{def}}{=} [\int u^{i+j}K(u)du]_{i,j}$ and $T \stackrel{\text{def}}{=} [\int u^{i+j}K^2(u)du]_{i,j}$ with $i, j = 0, \ldots, 3$. This implies the asymptotic normality of the estimated EPK at a fixed payoff x. More precisely

$$\sqrt{n_q h_q^5}\{\hat{\mathcal{K}}(x) - \mathcal{K}(x)\} \xrightarrow{\mathcal{L}} N\{0, \sigma_q^2(x)/p^2(x)\}.$$

Let the time point t and the time to maturity τ be fixed. The respective EPK is denoted by $\hat{\mathcal{K}}(x))$. The variance of $\hat{\mathcal{K}} = \hat{\mathcal{K}}(x)$ is given by

$$\mathsf{Var}\{\hat{\mathcal{K}}(x)\} \approx \{p(x)\}^{-2} B^{-1}(x) N^{-1} T N^{-1}. \tag{10}$$

Let $\mathcal{D}_n(x)$ be the standardized process:

$$\mathcal{D}_n(x) \stackrel{\text{def}}{=} n_q^{1/2} h_{n_q}{}^{5/2} \{\hat{\mathcal{K}}(x) - \mathcal{K}(x)\} \;/\; [\mathsf{Var}\{\hat{\mathcal{K}}(x)\}]^{1/2}.$$

Relying on the linearization in Lemma 2, we derive the confidence band for $\mathcal{K}$.

Theorem 2. *Under assumptions A(1)-(A5) it follows*

$$\mathrm{P}\left[(-2\log h_{n_q})^{1/2} \left\{\sup_{x\in E} |\hat{\mathcal{D}}_n(x)| - c_{nt}\right\} < z\right] \longrightarrow \exp\{-2\exp(-z)\},$$

where $c_{nt} = (-2\log h_{n_q})^{1/2} + (-2\log h_{n_q})^{-1/2}\{x_\alpha + \log(R/2\pi)\}$.

The $(1-\alpha)100\%$ confidence band for the pricing kernel $\mathcal{K}$ is thus:

$$[f : \sup_{x\in E}\{|\hat{\mathcal{K}}(x) - f(x)| \;/\; \widehat{\mathsf{Var}}(\hat{\mathcal{K}})^{1/2}\} \le L_\alpha],$$

where $L_\alpha = 2!(n_q h_{n_q}^5)^{-1/2} c_{nt}$, $x_\alpha = -\log\{-1/2\log(1-\alpha)\}$ and $R = (N^{-1}MN^{-1})_{3,3}/(N^{-1}TN^{-1})_{3,3}$ with $M = [\int u^{i+j}\{K'_{h_{n_q}}(u)\}^2 du - \frac{1}{2}\{i(i-1) + j(j-1)\}\int u^{i+j-2} K^2_{h_{n_q}}(u)du]_{i,j=0,\ldots,3}$.

For the implementation with real data we need a consistent estimator of $\mathsf{Var}(\hat{\mathcal{K}})$. For fixed τ, we rely on the delta method and use the empirical sandwich estimator, see Carroll, Ruppert and Welsh (1998). The latter method provides the variance estimator for the parameters obtained from estimating equations given by (7).

For the data points $(X_{it}, Y_{it}),\ \ i = 1, \cdots, n;\ \ t = t+1, \cdots, t+\tau$, we have

$$\widehat{\mathsf{Var}}\{\hat{\mathcal{K}}(x)\} = \{\hat{p}(x)\}^{-2} V(x)^{-1} U(x) V(x)^{-1},$$

where

$$\begin{aligned} V(x) &= \frac{1}{n_q \tau} \sum_{i=1}^{n_q} \sum_{j=t+1}^{t+\tau} K^2_{h_{n_q}}(X_{ij} - x) \left[\frac{\partial}{\partial C} Q\{Y_{ij}; \hat{C}(x, X_{ij})\} \right]^2 (H_{n_q}^{-1} \mathbf{X}_{ij})(H_{n_q}^{-1} \mathbf{X}_{ij})^\top, \quad (11) \\ U(x) &= \frac{1}{n_q \tau} \sum_{i=1}^{n_q} \sum_{j=t+1}^{t+\tau} K^2_{h_{n_q}}(X_{ij} - x) \left[\frac{\partial^2}{\partial^2 C} Q\{Y_{ij}; \hat{C}(x, X_{ij})\} \right] (H_{n_q}^{-1} \mathbf{X}_{ij})(H_{n_q}^{-1} \mathbf{X}_{ij})^\top, \quad (12) \end{aligned}$$

where $\mathbf{X}_{ij} = (1, \cdots, (X_{ij} - x)^3)^\top$ and $h_{n_q} = \mathrm{diag}\{1, \ldots, h^3_{n_q}\}$. The estimator is consistent in our setup as motivated in Appendix A.2 of Carroll et al. (1998).

5 The sheet in maturity dimension

Note that the asymptotic behavior of $\hat{q}(x) - q(x)$ if correctly standardized does not depend on τ. If we estimate the variance function of $\hat{\mathcal{K}}_\tau(x)$ in time, we can extend the confidence band as a sheet to the τ dimension. Let $\mathfrak{x}$ be the set of maturities of interest. The joint confidence sheet over payoff and maturity is given by

$$[f : \sup_{x \in E, \tau \in \mathfrak{x}} \{|\hat{\mathcal{K}}_\tau(x) - f(x)| \ / \ \mathsf{Var}(\hat{\mathcal{K}}_\tau)^{1/2}\} \leq L_\alpha].$$

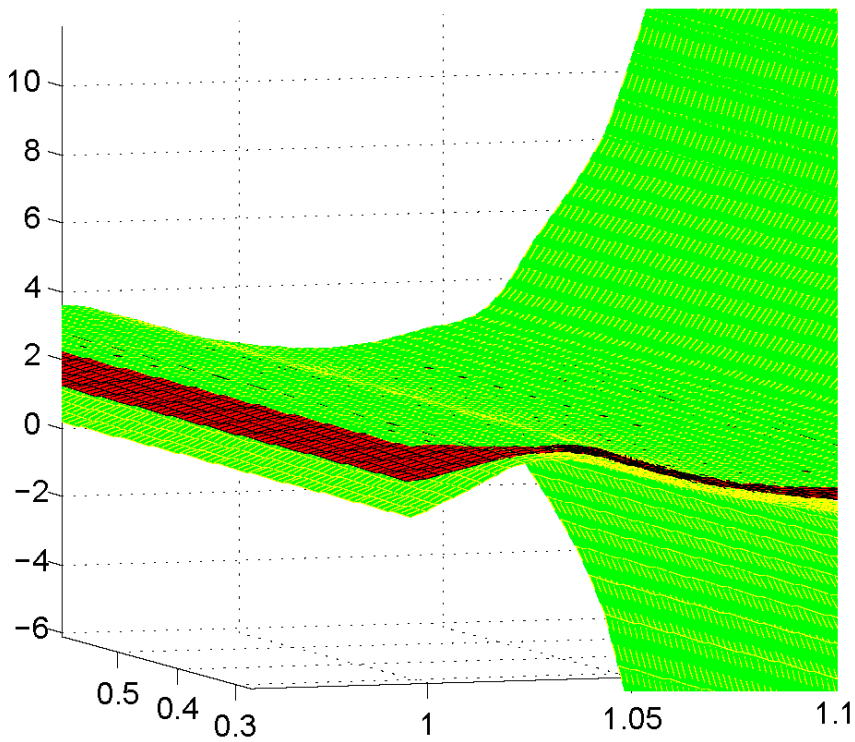


Figure 4: Examples of sheet for pricing kernels in 060228

In the BS setup we can actually provide an explicit link between the EPKs for different maturities. For fixed maturity τ, interest rate r we obtain from the normal

form of p and q:

$$\mathcal{K}_\tau(x) = \exp\{\frac{(\mu - r)(\mu + r - \sigma^2)\tau}{2\sigma^2}\}(\frac{x}{S_t})^{(\mu-r)/\sigma^2}.$$

This implies

$$\begin{aligned} \mathcal{K}_{\tau_1}(x)/\mathcal{K}_{\tau_2}(x) &= \exp\{(\mu - r)(\mu + r - \sigma^2)(\tau_1 - \tau_2)/(2\sigma^2)\} \\ &= [\exp\{\tau_1 - \tau_2\}]^{c(\mu,r,\sigma)} = g(\tau_1 - \tau_2), \end{aligned}$$

for any fixed τ_1 and τ_2, i.e. the log difference of the PKs is proportional to the difference of the maturities. In fact, this gives us some insight into the evolution of the sheet over maturities. More precisely, for known characteristics of the band for fixed τ_1, the confidence band for τ_2 is given by

$$[f : \hat{g}(\tau_1 - \tau_2)\{-L_\alpha \widehat{\mathsf{Var}}(\hat{\mathcal{K}}_{\tau_1}) + \hat{\mathcal{K}}_{\tau_1}(x)\} \leq f(x) \leq \hat{g}(\tau_1 - \tau_2)\{L_\alpha \widehat{\mathsf{Var}}(\hat{\mathcal{K}}_{\tau_1}) + \hat{\mathcal{K}}_{\tau_1}(x)\}]$$

for all $x \in E$.

6 Confidence bands based on smoothing implied volatility

The construction of the EPK estimator can be stabilized by a two-step procedure as in Rookley (1997), Fengler (2010). At the first step, we estimate the implied volatility (IV) function by a local polynomial regression. At the second step, we plug the smoothed IV into the BS formula to obtain a semiparametric estimator of the option price. Since the BS model is homogeneous with respect to the asset price and the strike price we smooth the IV using a local polynomial regression in moneyness (M_t). In the absence of dividends, it is defined at time t as $M_{it} = S_t/X_i$. The heteroscedastic model for the IV is given by:

$$\sigma_i = \sigma(M_{it}) + \sqrt{\eta(M_{it})}\upsilon_i, \quad i = 1, \ldots, n_q, \tag{13}$$

where υ_i are the i.i.d. errors with zero mean and unit variance and $\eta(\cdot)$ is the volatility function.

Defining the rescaled call option price by $c(M_{it}) = C_i/S_t$, we obtain from the BS formula

$$c(M_{it}) = c\{M_{it}; \sigma(M_{it})\} = \Phi\{d_1(M_{it})\} - \frac{e^{-r\tau}\Phi\{d_2(M_{it})\}}{M_{it}},$$

where

$$d_1(M_{it}) = \frac{\log(M_{it}) + \left\{r_t + \frac{1}{2}\sigma(M_{it})^2\right\}\tau}{\sigma(M_{it})\sqrt{\tau}}, \qquad d_2(M_{it}) = d_1(M_{it}) - \sigma(M_{it})\sqrt{\tau}.$$

Combining the result of Breeden and Litzenberger (1978) with the expression for

$c(M_{it})$ leads to the SPD

$$q(X) = e^{r\tau}\frac{\partial^2 C}{\partial X^2} = e^{r\tau}S_t\frac{\partial^2 c}{\partial X^2} \tag{14}$$

with

$$\frac{\partial^2 c}{\partial X^2} = \frac{\mathrm{d}^2 c}{\mathrm{d}M^2}\left(\frac{M}{X}\right)^2 + 2\frac{\mathrm{d}c}{\mathrm{d}M}\frac{M}{X^2}. \tag{15}$$

As it is shown in the appendix the derivatives in the last expression can be determined explicitly and are functions of $V = \sigma(M)$, $V' = \partial\sigma(M)/\partial M$ and $V'' = \partial^2\sigma(M)/\partial M^2$. We estimate the latter quantities by the nonparametric local polynomial regression for the IV of the from

$$\sigma(M_{it}) = V(M) + V'(M)(M_{it} - M) + \frac{1}{2}V''(M)(M_{it} - M)^2.$$

The respective estimators are denoted by $\hat{V}$, $\hat{V}'$ and $\hat{V}''$. Plugging the results into (14)-(15) we obtain the estimator of SPD in the smoothed IV space. Assuming that the IV process fulfills the the assumptions (A1)-(A5) in the appendix, we conclude that Theorem 2.1 of Claeskens and Van Keilegom (2003) holds also for $\hat{V}$, $\hat{V}'$ and $\hat{V}''$. Note that the convergence rate of $\hat{V}$ and $\hat{V}'$ is lower than of $\hat{V}''$. Relying on this fact, we state the asymptotic behavior of $\hat{q}(x) - q(x)$ in the next theorem.

Theorem 3. *Let $\sigma(M_{it})$ satisfy the assumptions (A1)-(A5). Then*

$$\begin{aligned}\hat{q}(x) - q(x) \;=\;& e^{r\tau}S_t\frac{M^2}{X^2}\Big[\phi\{\hat{d}_1(M)\}\Big\{\sqrt{\tau}/2 - \frac{\log(M) + r\tau}{\hat{V}(M)^2\sqrt{\tau}}\Big\} \\ & -e^{-r\tau}\phi\{\hat{d}_2(M)\}\Big\{-\sqrt{\tau}/2 - \frac{\log(M) + r\tau}{\hat{V}(M)^2\sqrt{\tau}}\Big\}\Big]\{\hat{V}''(M) - V''(M)\} \\ & +o\{\hat{V}''(M) - V''(M)\}.\end{aligned}$$

Theorem 3 allows us to construct the confidence bands of the SPD estimated semi-parametrically using the confidence bands for the IV. The variance of the estimator

is obtained by the delta method in the following way

$$\mathsf{Var}\{\hat{q}(x) - q(x)\} = \left(\frac{\partial q}{\partial V''}\right)^2 \mathsf{Var}\{\hat{V}''(M) - V''(M)\}.$$

Here it is sufficient to consider only the variance of second derivative of V. The first derivative and V itself can be neglected. The variance $\mathsf{Var}\{\hat{V}''(M) - V''(M)\}$ is estimated using sandwich estimator similarly to (10).

7 Bootstrap confidence bands

Hall (1991a) showed that for density estimators, the supremum of $\{\widehat{q}(x) - q(x)\}$ converges at the slow rate $(\log n_q)^{-1}$ to the Gumbel extreme value distribution. Therefore the confidence band may exhibit poor performance in finite samples. An alternative approach is to use the bootstrap method. Claeskens and Van Keilegom (2003) used smooth bootstrap for the numerical approximation to the critical value. Here we consider the bootstrap technique of the leading term in Lemma 2

$$\sup_{x \in E} |\frac{\widehat{q}(x) - q(x)}{p(x)}|.$$

We resample data from the smoothed bivariate distribution of (X, Y), the density estimator is:

$$\hat{f}(x, y) = \frac{\hat{\sigma}_X}{n_q h_{n_q} h_{n_q} \hat{\sigma}_Y} \sum_{i=1}^{n_q} K\Big\{\frac{X_i - x}{h_{n_q}}, \frac{(Y_i - y)\hat{\sigma}_X}{h_{n_q} \hat{\sigma}_Y}\Big\},$$

where $\hat{\sigma}_X$ and $\hat{\sigma}_Y$ are the estimated standard deviations of the distributions of X and Y. The motivation of using the smooth bootstrap procedure is that a Rosenblatt transformation requires the resampled data (X^*, Y^*) to be continuously distributed.

From the re-sampled data sets, we calculate the bootstrap analogue:

$$\sup_{x \in E} |\frac{\hat{q}^*(x) - \hat{q}(x)}{\hat{p}(x)}|.$$

One may argue that this resampling technique does not correctly reflect the bias arising in estimated q, Härdle and Marron (1991) use therefore a resampling procedure based on a larger bandwidth g. This refined bias-correcting bootstrap method is not required in our case, since the bandwidth conditions ensure a negligible bias asymptotically.

Correspondingly, we define the one-step estimator for the stochastic deviation by:

$$h_{n_q}^2\{\hat{\mathcal{K}}(x)^* - \hat{\mathcal{K}}(x)\} = -\{p(x)\}^{-2}\{U(x)^{-1}H_{n_q}^{-1}A_{n_q}^*(x)\}_{3,3}$$

with the variance estimated from the bootstrap sample as:

$$\mathsf{Var}\{\widehat{\mathcal{K}}(x)\} \approx \{p(x)\}^{-2}B(x)^{-1}N^{-1}M^*N^{-1}. \tag{16}$$

Lemma 3. *Assume conditions (A1)-(A5), a* $(1-\alpha)100\%$ *bootstrap confidence band for the EPK* $\mathcal{K}(x)$ *is:*

$$[f(x) : \sup_{x\in E}\{|\widehat{\mathcal{K}}(x) - f(x)|\widehat{Var}(\widehat{\mathcal{K}})^{-1/2}\} \leq L_\alpha^*],$$

where the bound $L_{\alpha_j}^*$ *satisfies*

$$\mathrm{P}^*[-\{U(x)^{-1}H_{n_q}^{-1}A_{n_q}^*(x)\}_{3,3} \,/\, \{B(x)^{-1}N^{-1}M^*N^{-1}\}_{3,3} \leq L_\alpha^*] = 1-\alpha.$$

8 Monte-Carlo study

The practical performance of the above theoretical considerations is investigated via two Monto-Carlo studies. The first simulation aims at evaluating the performance under the BS hypothesis, while the second simulation setup does the same under a realistically calibrated surface. The confidence bands are applied to DAX index options. We first study the confidence bands under a BS null model. Naturally, without volatility smile, both the BS estimator and nonparametric estimator are expected to be covered by the bands. While in the presence of volatility smile, we expect our tests to reject the BS hypothesis in most cases.

8.1 How well is the BS model covered?

In the first setting, we calibrate a BS model on day 20010117 with the interest rate set equal to the short rate $r = 0.0481$, $S_0 = 6500$, strike prices in the interval $[6000, 7400]$. We refer to Aït-Sahalia and Duarte (2003) on the sources of the noise and use an identical simulation setting, with the noise being uniformly distributed in the interval $[0, 6]$. Fig 5 is a scatter plot of generated observations, the data is clustered in discrete values of the strike price.

Figure 6 shows a nonparametric estimator for the SPD and a parametric BS estimator. The two estimators roughly coincide except for a small wiggle, thus the bands drawn around the nonparametric curve also fully cover the parametric one. The accuracy is evaluated by calculating the coverage probabilities and average area within the bands, see Table 1 and Table 2. The coverage probabilities is determined via 500 simulations, whenever the hypothesized curve calculated on a grid of 100. The coverage probability approaches its nominal level with the sample size. The bands get narrower with increasing sample sizes. However, the coverage probabilities never reach the nominal level, which may well be attributed to the above mentioned poor

convergence of Gaussian maxima to the Gumbel distribution. The area within the bands reflects the stability of the estimation procedure.

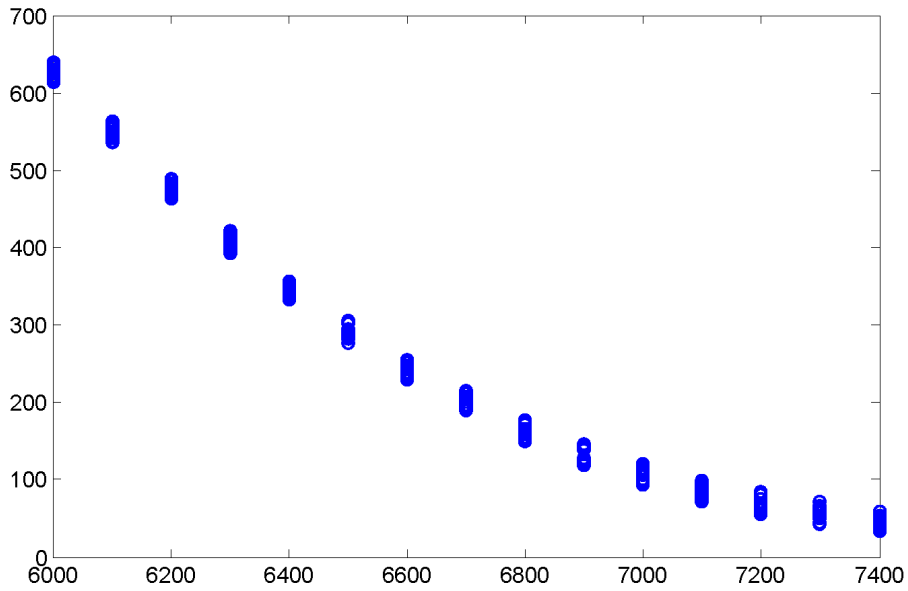


Figure 5: Generated noisy BS call option prices against strike prices

Table 1: Coverage probability (area) of the uniform confidence band at 10% with annualized volatility = 0.1878 for SPD

τ / n	300	450	600
3M	0.7945(2.17)	0.8123(2.06)	0.8533(1.88)
6M	0.8002(2.34)	0.8143(2.08)	0.8595(1.94)

Historical densities are estimated from simulated stock prices following geometric Brownian motion with $\mu = 0.23$. Therefore, a BS EPK estimator could be tested using the above procedure. Due to boundary effects, we concentrate on moneyness ($M_t = S_t/X$) in $[0.95, 1.1]$. Figure 7 displays the nonparametric EPK with confidence band and the BS EPK covered in the band. We observe that the BS EPK is strictly monotonically decreasing. The summary statistics is given in Table 3 and Table 4, due to the additional source of randomness introduced through the esti-

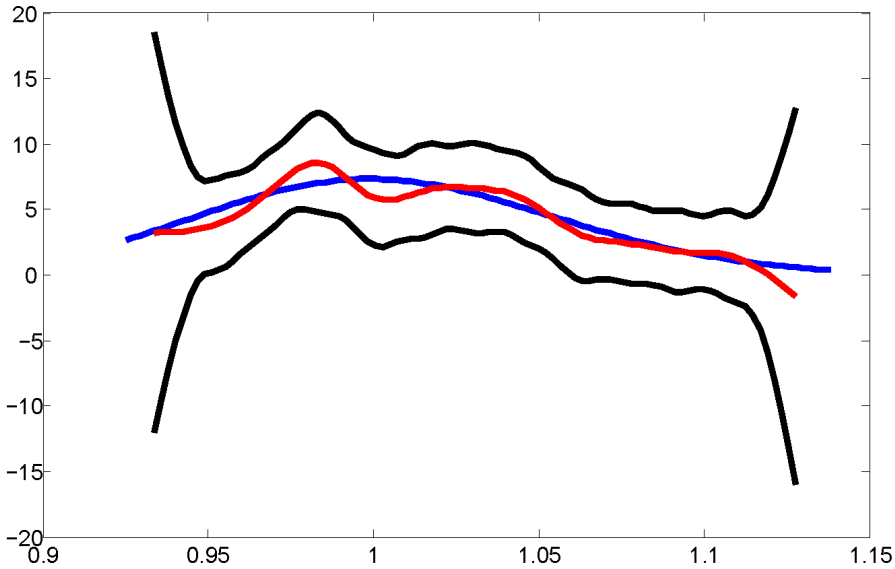


Figure 6: Estimation of SPD (red), bands (black) and the BS SPD (blue), with $h_{n_q} = 0.085$, $\alpha = 0.05$, $n_q = 300$

Table 2: Coverage probability (area) of the uniform confidence band at 5% with annualized volatility = 0.1878 for SPD

τ / n	300	450	600
3M	0.9063(2.402)	0.9144(2.204)	0.9233(1.998)
6M	0.8964(2.438)	0.9056(2.134)	0.9203(2.069)

mation of $p(x)$, the coverage probabilities are less precise than the corresponding coverage probabilities for SPD. Nevertheless, the probabilities are getting closer to their nominal values and the bands get narrower when the sample size increases.

Table 3: Coverage probability (area) of the uniform confidence band for the EPK at 5% with volatility(annualized) = 0.1878

M/n	300	450	600
3	0.7820(2.5434)	0.7980(2.4978)	0.8020(2.3876)
6	0.8602(2.4987)	0.8749(2.4307)	0.8900(2.4131)

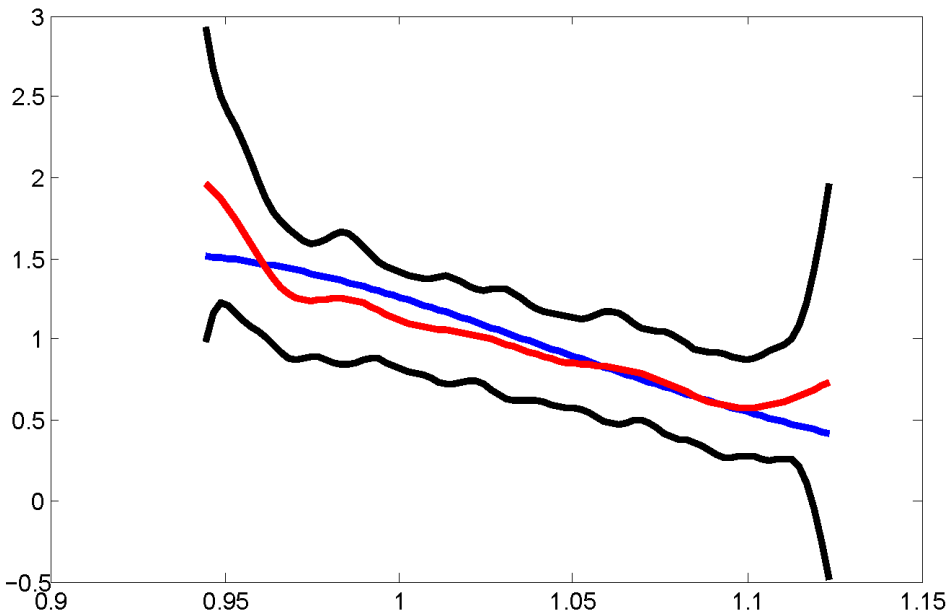


Figure 7: Estimation of EPK (red), bands (black) and the BS EPK (blue), with $h_{n_q} = 0.085$, $\alpha = 0.05$, $n = 300$

Table 4: Coverage probability (area) of the uniform confidence band for the EPK at 10% with volatility(annualized) = 0.1878

M/n	300	450	600
3	0.7062(2.4714)	0.7356(2.3410)	0.7620(2.2310)
6	0.7289(2.5020)	0.7740(2.2304)	0.8290(2.3131)

8.2 How well is the band in reality?

Section 4.1 studied the performance of the bands under the BS null, while this section is designed to investigate the performance of the bands when the null hypothesis is violated by a realistic volatility smile observed in the market. For the features of the implied volatility smile we refer to e.g. Fengler (2005). Keeping the parameters identical to the setup of the first study, we generated the data with a smoothed volatility function based on 20010117 with $\tau = 3M, 6M$ to maturity.

Figure 8 and 9 report the estimators for SPD and EPK. The bands do not cover the

BS estimator. Correspondingly, Table 5 and Table 6 show the coverage probabilities, which rapidly decrease when sample sizes are increasing. However, the area within the bands does not change significantly when compared with the results of Section 4.1. We conclude that the confidence bands are useful for detecting the deviation from the BS model.

Table 5: Coverage probability (area) of the uniform confidence band for the EPK at 5%

τ / n	300	450	600
3M	0.5120(2.4320)	0.1784(2.2340)	0.0500(2.0176)
6M	0.5920(2.5304)	0.4100(2.1678)	0.1780(2.0234)

Table 6: Coverage probability (area) of the uniform confidence band for the EPK at 10%

τ / n	300	450	600
3M	0.2580(2.1196)	0.0500(2.0483)	0.0300(2.0132)
6M	0.3746(2.2234)	0.1700(2.1324)	0.0910(2.0030)

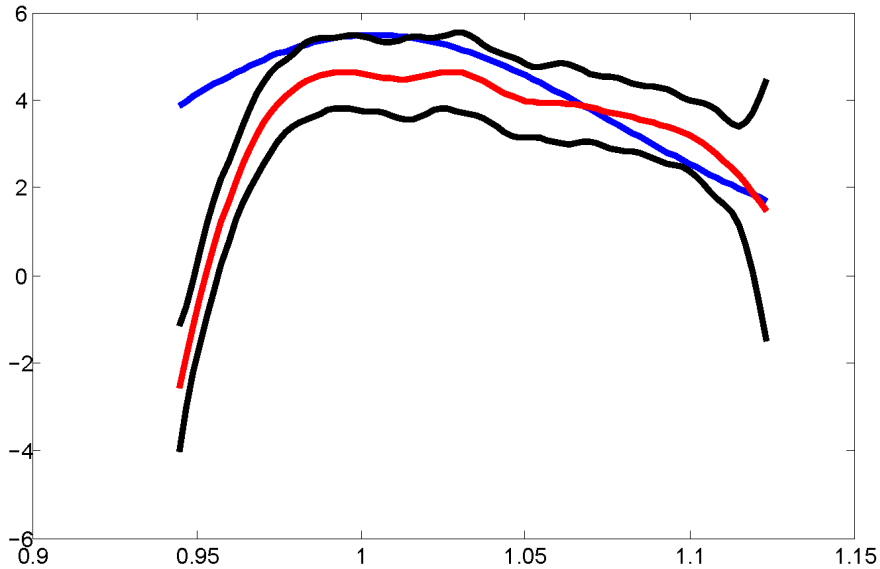


Figure 8: Plot of confidence bands (black), estimated value (red), the BS (blue) SPD with simulated volatility smile, $n_p = 2000$, $n_q = 300$, $h_{n_q} = 0.06$, $h_{n_p} = 0.0106$, $\alpha = 0.05$.

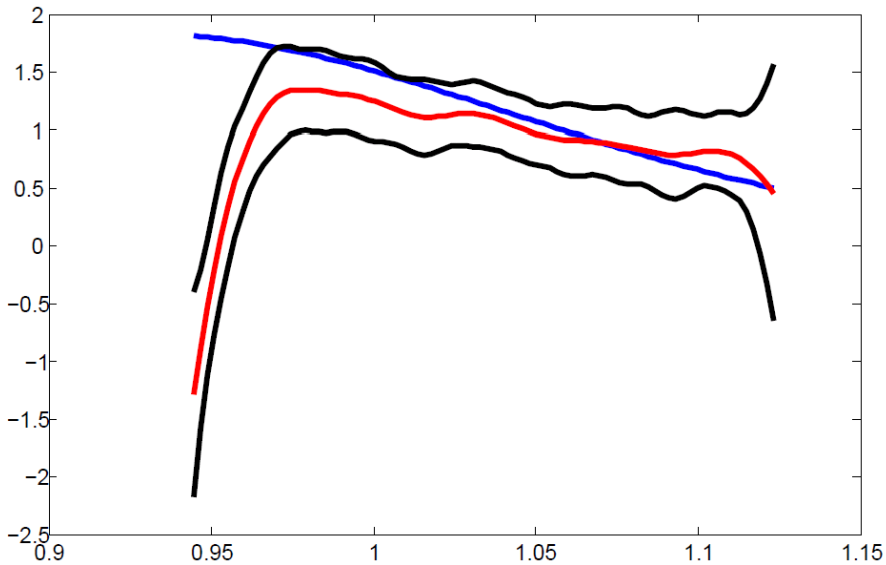


Figure 9: Plot of confidence bands (black), estimated value (red), the BS (blue) EPK with simulated volatility smile $n_p = 2000$, $n_q = 300$, $h_{n_q} = 0.06$, $h_{n_p} = 0.0106$, $\alpha = 0.05$.

9 An illustration with DAX data

9.1 Data

In contrast to previous studies that are mainly based on S&P500 data, we focus on intraday European options based on the DAX. The source is the European Exchange EUREX and data available by C.A.S.E., RDC SFB 649 (http://sfb649.wiwi.hu-berlin.de) in Berlin. The extracted observations for our analysis cover the period between 1998 and 2008. Figure 10 shows the DAX index. The semiparametric SPD estimates of Rookley (1997) are applied to estimate the EPK. We fix maturity and concentrate on the moneyness dimension. As we cannot find traded options with the same maturity on each day, we consider options with maturity 15 days (10 trading days) across several years. Specifically, we extract a time series of options for every month from Jan 2001 to Dec 2006; this adds up to 63 days.

To make sure that the data correctly represents the market conditions, we use several cleaning criteria. In our sample, we eliminate the observations with $\tau < 1D$ and $IV > 0.7$. Also, we skip the option quotes violating general no-arbitrage condition i.e. $S > C > \max\{0, S - Ke^{-r\tau}\}$. Due to the put-call parity, out-of-the-money call options and in-the-money puts are used to compute the smoothed volatility surface. The median is used to compute the risk neutral density. We use a window of 500 returns for nonparametric kernel density estimators of HD.

Figure 11 describes the relative position of the HD and SPD on a specific day, the EPK peak is apparently created through the different probability mass contributions at different moneyness states.

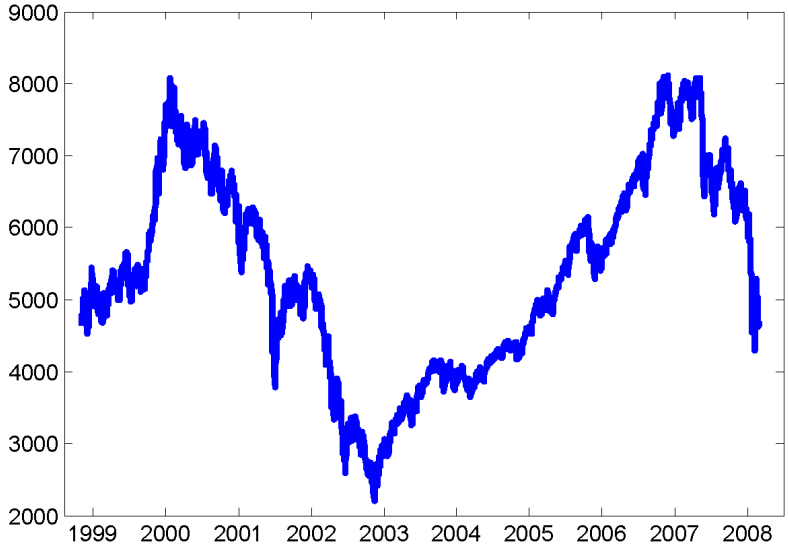


Figure 10: Plot of DAX Index

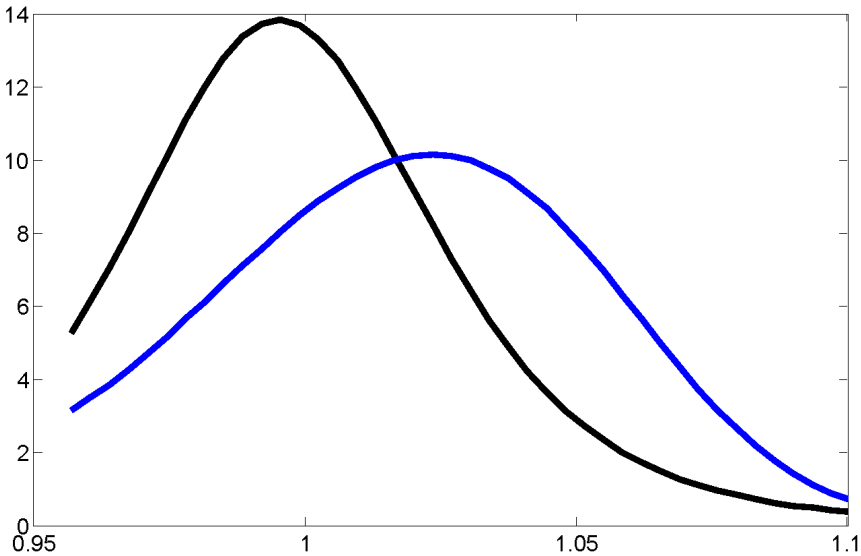
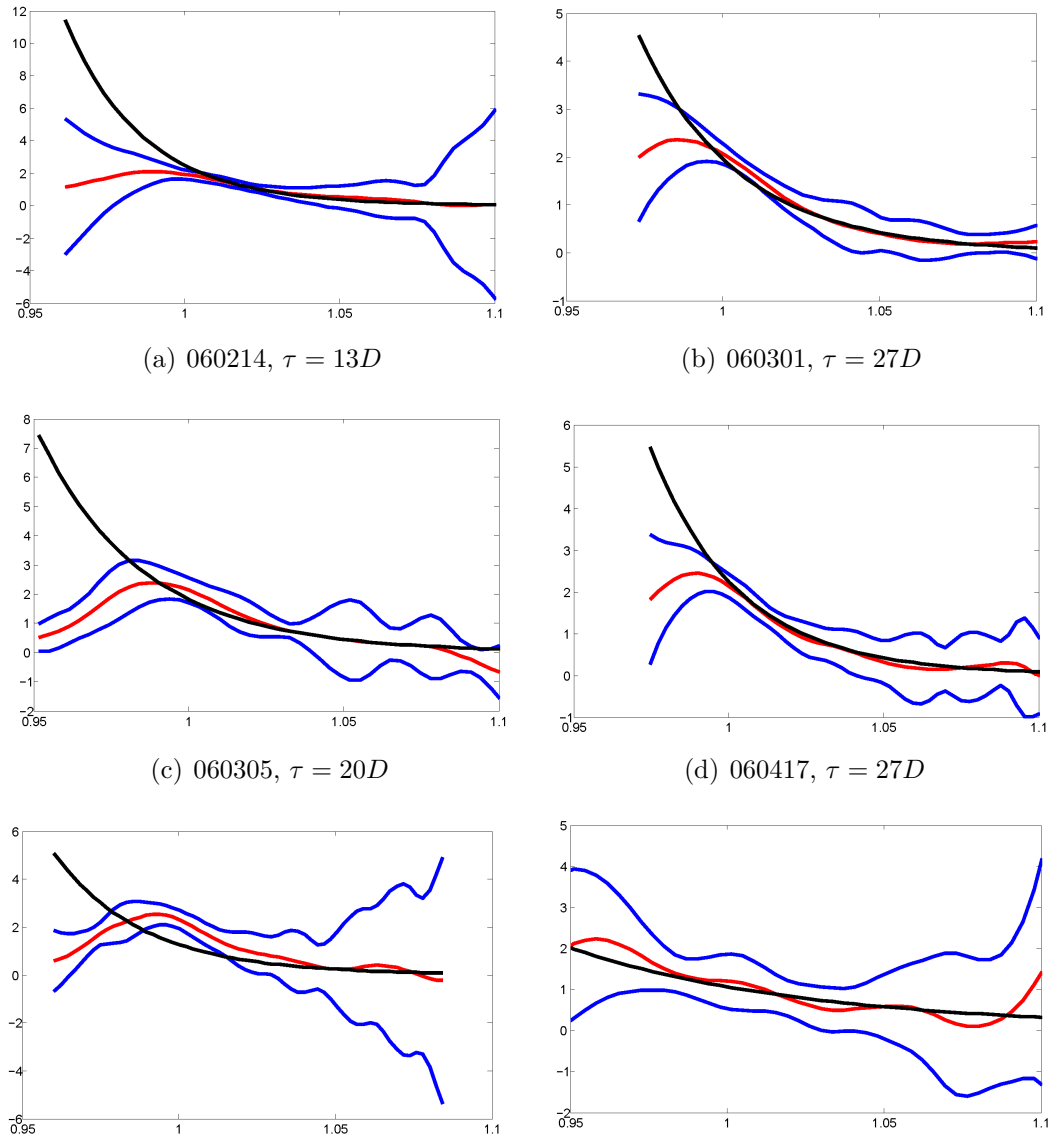


Figure 11: Plot of estimated state price density (Rookley's method, $h_{n_q} = 0.0600$) (black) and historical density ($h_{n_p} = 0.0106$)(blue) for 20060228

9.2 Estimation of DAX EPK and its uniform confidence band



(a) 060214, $\tau = 13D$

(b) 060301, $\tau = 27D$

(c) 060305, $\tau = 20D$

(d) 060417, $\tau = 27D$

(e) 060419, $\tau = 13D$

(f) 060512, $\tau = 20D$

Figure 12: BS EPK (black), Rookley EPK (red), Uniform Confidence band (blue) for $\alpha = 0.05$

We consider two specifications for the pricing kernels. In the first specification, the BS pricing kernels have a marginal rate of substitution with power utility function:

$$\mathcal{K}(M) = \beta_0 M^{-\beta_1}, \tag{17}$$

where β_0 is a scaling factor and β_1 determines the slope of pricing kernel. Thus the BS calibration is realized by linearly regressing the (ordered) log-EPK on log-moneyness. In the second specification, we construct the nonparametric confidence bands as described in Section **??**,. A sequence of EPKs and corresponding bands are shown in Figure 12. In most of the cases, the BS EPKs are rejected via the confidence bands. The amount of deviation from the hypothesized BS specification though provides us valuable information about how risk hungry investors are. Besides, the area of the bands varies over time, which gives us insights into the variabilities of the prevailing risk patterns. In sum, the bands not only provide a simple test for hypothesizes EPKs but also help us to study the dynamics risk patterns over time.

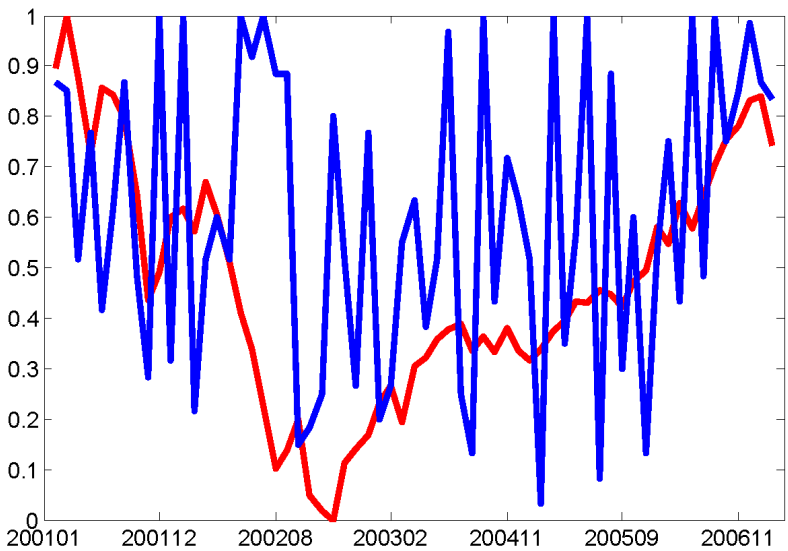


Figure 13: Coverage probability and the DAX index (rescale to $[0, 1]$)

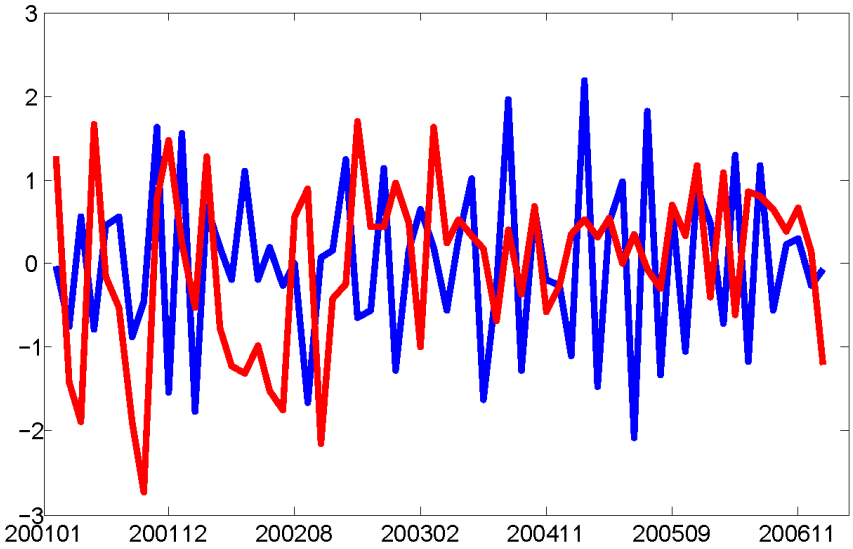


Figure 14: First difference of coverage probability and the DAX index return (standardized)

9.3 Linking economic conditions to EPK dynamics

We use two different indicators for the deviation from the simple BS model. As an approximation to the coverage probability, we calculate the proportion of grid points of the band which covers the BS EPK. As a second measure, we introduce the average width of the confidence band over the moneyness interval $[0.95, 1.1]$ as a proxy for the area between the confidence bands. This provides us with a measure of variability, see also Theorem 2.

The first risk pattern time series are given in Figure 13, where we display the DAX index (scaled to $[0, 1]$) together with the coverage probability. We discover that the coverage probability becomes less volatile when the DAX index level is high. Figure 14 shows the differenced time series. From a simple correlation analysis, we argue that the change in coverage probability and DAX return (with a lag of 3M) are highly negatively correlated (correlation -0.3543) when the DAX index goes down (200101-200302). On the contrary, in the period when the DAX goes up, one

observes a large positive correlation (0.3151). What does this mean economically? This implies in a period of worsening economic condition, a positive part of the monthly DAX returns induces a greater hunger for risk in a delay horizon of 3 months. Positive returns have just the opposite effects. With boosting and bullish markets, the positive correlation indicates a 3-month horizon of decreasing risk aversion.

As far as the average width of the bands is concerned, we may conclude from Figure 15 and Figure 16 that in periods of clearly bullish or bearish momentum, the volatility of the width of the confidence band is higher. This may be caused by the uncertainty of the market participants about the long-term persistence of the trend. The lag effect on risk hunger is also detectable for this constructed indicator. Over the whole observation interval, the correlation between the monthly DAX return and the change in the average width is -0.3230 for a 1M lag and -0.2717 for 3M.

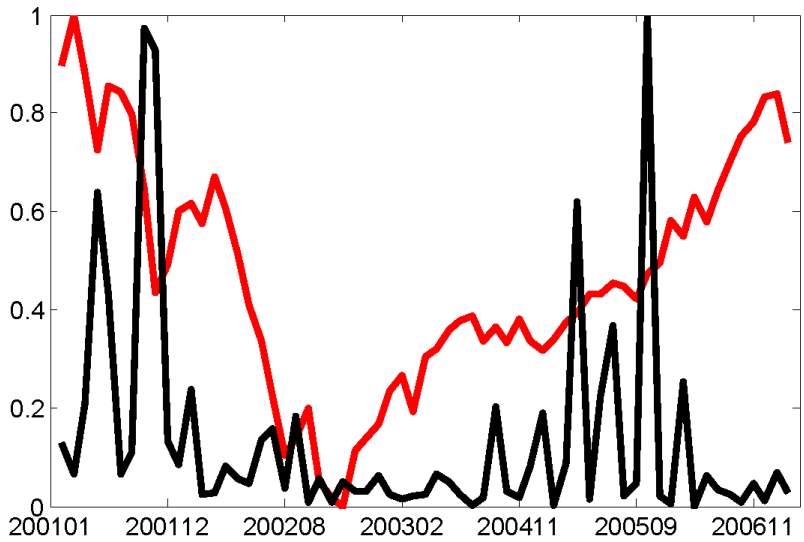


Figure 15: Area between the bands and the DAX index (scaled to $[0, 1]$)

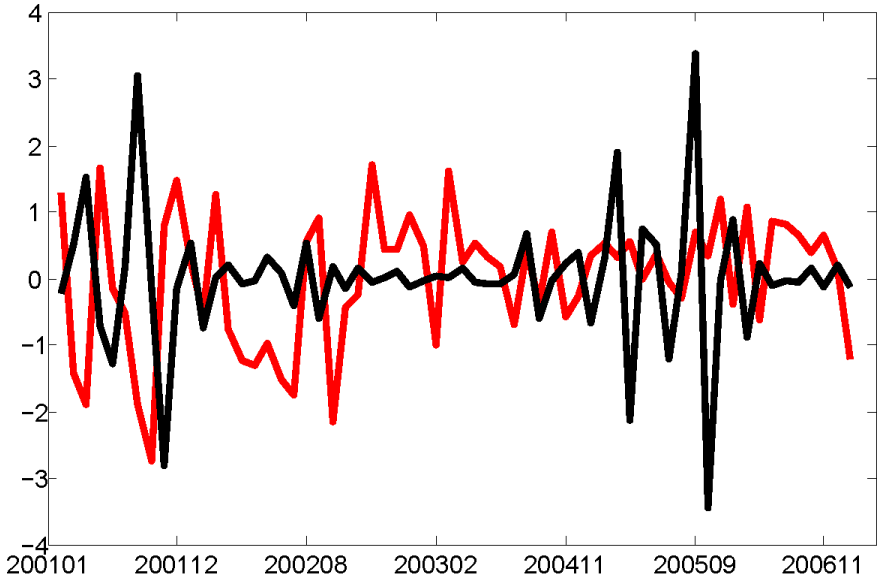


Figure 16: First difference of the area between the bands and the DAX index return (standardized)

10 Conclusions

Pricing Kernels are important elements in understanding investment behavior since they reflect the relative weights given by investors states of nature (Arrow-Debreu securities). Pricing kernels may be deduced in either parametric or nonparametric approaches. Parametric approaches like a simple BS model are too restrictive to account for the dynamics of the risk patterns, which induces the well-known EPK paradox. Nonparametric approaches allow more flexibility and reduce the modeling bias. Simple tools like uniform confidence bands help us to conduct tests against any parametric assumption of the EPKs i.e. shape inspection. Considering the numerical stability, we smooth the IV surface via the Rookley's method, to obtain SPD estimator.

We have studied systematically the methodology of constructing the uniform bands for both semiparametric or nonparametric estimators. Based on the confidence

bands, we explored two indicators to measure risk aversion over time and connected it with DAX index, one is the coverage probability measuring the proportion of the BS fitting curve covered in bands, while another one is the area indicator measuring the variability of the estimator. We have found out that there are strong correlations between DAX index and our indicators with lag effects. The smooth bootstrap is studied without a significant improvement in finite sample performance. One interesting further extension would be employing robust smoothers to improve the bootstrap performance.

11 Appendix

Assumptions:

(A1) $h_{n_q} \to 0$ in such a way that $\{\log n_q/(n_q h_{n_q})\}^{1/2} h_{n_q}^3 \to 0$, and the optimal bandwidth, to guarantee undersmoothing, is of order $\mathcal{O}\{(\log n_q n_q)^{1/4[(p-j)/2]+3+2j+\alpha}\}$, where $j = 0, 1, 2, 3$, and $\alpha > 0$.

(A2) The kernel function $K \in C^{(1)}[-1, 1]$ and takes value 0 on the boundary.

(A3) For the likelihood function $Q \in C^{(1)}(E)$ it holds that $\inf_{x \in E} Q(x) > 0$. $C(x) \in C^{(4)}(E)$. Additionally the third partial derivatives of $Q(Y, C)$ with respect to C exists and is continuous in C for every y. The Fisher information $I(C(x))$ has a continuous derivative and $\inf_{x \in E} I\{C(x)\} > 0$.

(A4) There exists a neighborhood $N(C(x))$ such that

$$\max_{k=1,2} \sup_{x \in E} || \sup_{C \in N\{C(x)\}} \frac{\partial^k}{\partial C^k} Q(y; C)||_\lambda < \infty$$

for some $\lambda \in (2, \infty]$. Furthermore

$$\sup_{x \in E} \mathsf{E}[\sup_{C \in N(C(x))} |\frac{\partial^3}{\partial C^3} Q(y; C)|] < \infty.$$

(A5) The density of underlying $p(x)$ is three times continuously differentiable.

(A6) Let $a_{n_p} = (n_p h_{n_p}/\log n_p)^{-1/2} + h_{n_p}^2$ from (8) and $b_{n_q} = h_{n_q}^{-2}(n_q h_{n_q}/\log n_q)^{-1/2} + h_{n_q}^2$ from Lemma 1. We assume that $n_q/n_p = \mathcal{O}(1)$ and $a_{n_p}/b_{n_q} = o(1)$.

11.1 Proof of Lemma 1

We follow the approach of Zhao (1994) also deployed by Claeskens and Van Keilegom (2003). Consider the disturbed coefficients of the local polynomial

$$C_\varepsilon(x, X_0) = \sum_{j=0}^{p} \{C_j(x) + \varepsilon_{n_q j}\}(X_0 - x)^j,$$

where $\varepsilon_{n_q j}$ are positive-valued sequences for $j = 0, \ldots, p+1$. This leads to the modified system of equations

$$A_{n_q j,\varepsilon}(x) = \frac{1}{n_q} \sum_{i=1}^{n_q} K_{h_{n_q}}(X_i - x) \frac{\partial Q\{Y_i, C_\varepsilon(x, X_i)\}}{\partial C}(X_i - x)^j.$$

Similarly to Lemma 2.2 of Härdle, Janssen and Serfling (1988) and using assumptions (A1)-(A4), we obtain

$$\sup_{x\in E} \sup_{C\in N\{C(x)\}} \sqrt{\frac{n_q h_{n_q}}{\log n_q}} \quad \Big| \frac{1}{n_q h_{n_q}} \sum_{i=1}^{n_q} K_{h_{n_q}}(X_i - x) \frac{\partial Q\{Y_i, C(x, X_i)\}}{\partial C} \frac{(X_i - x)^j}{h_{n_q}^j}$$
$$- \mathsf{E}\Big\{ \frac{1}{n_q h_{n_q}} \sum_{i=1}^{n_q} K_{h_{n_q}}(X_i - x) \frac{\partial Q\{Y_i, C(x, X_i)\}}{\partial C} \frac{(X_i - x)^j}{h_{n_q}^j} \Big\} \Big| = \mathcal{O}(1) \quad (a.s.)$$

This implies that $\sup_{x\in E} |A_{n_q j,\varepsilon}(x) - \mathsf{E}\, A_{n_q j,\varepsilon}(x)| = \mathcal{O}[h_{n_q}^j \{\log n_q/(n_q h_{n_q})\}^{1/2}] = \ell_{n_q j}$ (a.s.). Particularly this implies, that (a.s.) $|A_{n_q j,\varepsilon}(x) - \mathsf{E}\, A_{n_q j,\varepsilon}(x)| \leq \ell_{n_q j}$ for appropriate ε_{n_q}. Consider the first order Taylor expansion of $\partial Q\{Y_i, C_\varepsilon(x, X_i)\}/\partial C$ around the true values of the parameters $C_j(x)$ of the polynomial regression.

$$\frac{\partial Q\{Y, C_\varepsilon(u, X)\}}{\partial C} \approx \mathsf{E}\left[\frac{\partial Q\{Y, C(X)\}}{\partial C}\right]\Big|_{X=u} + \mathsf{E}\left[\frac{\partial^2 Q\{Y, C(X)\}}{\partial C^2}\right]\Big|_{X=u} \cdot \{C_\varepsilon(u, X) - C(u)\}.$$

Noting that $\mathsf{E}\left[\frac{\partial Q\{Y,C(X)\}}{\partial C}\right] = 0$ we obtain

$$\begin{aligned}\mathsf{E}\{A_{n_q j,\varepsilon}(x)\} &= -\mathsf{E}\Big(\frac{1}{n_q}\sum_{i=1}^{n_q} K_{h_{n_q}}(X_i - x)(X_i - x)^j\, \mathsf{E}\Big[-\frac{\partial^2 Q\{Y_i, C(X)\}}{\partial C^2}\Big]\Big|_{X=u} \\ &\quad \times\{C(u, X_i) - C(u) + \sum_{j^*=0}^{p} \varepsilon_{n_q j^*}(X_i - u)^{j^*}\}\Big).\end{aligned}$$

Note that $\sup_{|u-x|\le h_{n_q}} \sum_{j^*=0}^{p}(u-x)^{j^*} \int u^{j^*} K_{h_{n_q}}(u)du = \mathcal{O}(1)$. We decompose the expectation into two parts, one containing $C(u, X_i) - C(u)$ and the second one containing $\sum_{j^*=0}^{p} \varepsilon_{n_q j^*}(X_i - u)^{j^*}\}$. Let $\tilde{p} = 2([p/2] + 1)$ and

$$\omega_j(h_{n_q}) \stackrel{\text{def}}{=} h_{n_q}^{2[(p-j)/2]+1} \int u^{\tilde{p}+j} K_{h_{n_q}}(u)du \cdot \sup_{x\in E} C_{\tilde{p}}(x) = \mathcal{O}(h_{n_q}^{2\{[(p-j)/2]+1\}}).$$

The first part is bounded above by $M\omega_j(h_{n_q})$ for some constant M, while the second part is bounded by $2\ell_{n_q j}$. Choosing $\varepsilon_{n_q j} = \mathcal{O}(\max[2\omega_j(h_{n_q}), 4\ell_{n_q j}/M])$ we obtain that $\mathsf{E}[A_{n_q j,\varepsilon}(x)] \le -\frac{1}{2}M\varepsilon_{n_q j}$. Similarly we can show that $\mathsf{E}[A_{n_q j,-\varepsilon}(x)] \ge \frac{1}{2}M\varepsilon_{n_q j}$. This implies that $A_{n_q j,\varepsilon}(x) \le \ell_{n_q j} - \frac{1}{2}M\varepsilon_{n_q j} < 0$ and $A_{n_q j,-\varepsilon}(x) \ge -\ell_{n_q j} + \frac{1}{2}M\varepsilon_{n_q j} > 0$. Thus there exists a sequence of estimators $\hat{C}_j(x)$ such that $\sup_{x\in E} |\hat{C}_j(x) - C(x)| \le \varepsilon_{n_q j}$. The choice of $\varepsilon_{n_q j}$ implies for $j = 2$

$$\sup_{x\in E} |\hat{q}(x) - q(x)| = \mathcal{O}[h_{n_q}^{-2}\{\log n_q/(n_q h_{n_q})\}^{1/2} + h_{n_q}^2] \quad \text{(a.s.)}$$

11.2 Proof of Lemma 2

Recall from Lemma 1 and (8) that

$$\begin{aligned}&\sup_{x\in E} |\hat{p}(x) - p(x)| = \mathcal{O}\{\{\log n_q/(n_q h_{n_q})\}^{1/2} + h_{n_p}^2\} = \mathcal{O}(a_{n_p}), \\ &\sup_{x\in E} |\hat{q}(x) - q(x)| = \mathcal{O}[h_{n_q}^{-2}\{\log n_q/(n_q h_{n_q})\}^{1/2} + h_{n_q}^2] = \mathcal{O}(b_{n_q}).\end{aligned}$$

To determine the order of the EPK we linearize the ratio $q(x)/p(x)$.

$$\hat{\mathcal{K}}(x) - \mathcal{K}(x) = \frac{\hat{q}(x)}{\hat{p}(x)} - \frac{q(x)}{p(x)} = \frac{\hat{q}(x)p(x) - \hat{p}(x)q(x)}{p^2(x)} \cdot \frac{1}{1 + \frac{\hat{p}(x)-p(x)}{p(x)}}.$$

We decompose the first factor as $\hat{q}(x)p(x) - \hat{p}(x)q(x) = \{\hat{q}(x) - q(x)\}p(x) - \{\hat{p}(x) - p(x)\}q(x)$, while for the second factor we use the first order Taylor expansion. Putting together we obtain

$$\begin{aligned}\sup_{x\in E} |\hat{\mathcal{K}}(x) - \mathcal{K}(x)| &= \sup_{x\in E} |\frac{\hat{q}(x) - q(x)}{p(x)} - \frac{\hat{p}(x) - p(x)}{p(x)} \cdot \frac{q(x)}{p(x)} \\ &\quad - \frac{\{\hat{q}(x) - q(x)\}\{\hat{p}(x) - p(x)\}}{p^2(x)} + \frac{\{\hat{p}(x) - p(x)\}^2}{p^2(x)} \cdot \frac{q(x)}{p(x)}|.\end{aligned}$$

The first two elements are of order $\mathcal{O}(b_{n_q})$ and $\mathcal{O}(a_{n_p})$ respectively, while the last element is of order $o(a_{n_p})$. Summarizing we conclude that

$$\sup_{x\in E} |\hat{\mathcal{K}}(x) - \mathcal{K}(x)| = \mathcal{O}[\max\{a_{n_p}, b_{n_q}\}].$$

11.3 Proof of Theorem 2

The basic idea of the proof is to approximate the process

$$D_{n_q}(x) = n_q^{1/2} h_{n_q}^{5/2} \{\hat{\mathcal{K}}(x) - \mathcal{K}(x)\}/[\widehat{\mathsf{Var}}\{\hat{\mathcal{K}}(x)\}]^{1/2}$$

by a process with a non-stochastic variance term, which is further approximated by a process that can be treated with the tools of Claeskens and Van Keilegom (2003). Here we drop for the simplicity of notation the indices in n_q and h_{n_q}. As first approximation we define

$$D_n^{(1)}(x) \stackrel{\text{def}}{=} n^{1/2} h^{5/2} \{\hat{\mathcal{K}}(x) - \mathcal{K}(x)\}/\{V_{\mathcal{K}}(x)\}^{1/2},$$

where $V_{\mathcal{K}}(x) = \mathsf{Var}\{\hat{\mathcal{K}}(x)\}$ is given in equation (10). Lemma 2 ensures that the approximation by

$$n^{1/2}h^{5/2}\{\hat{q}(x) - q(x)\}/\{p(x)V_{\mathcal{K}}(x)^{1/2}\} \tag{18}$$

is uniformly of order $o_p\{(\log n)^{-1/2}\}$. The process in equation (18) can be approximated as in Claeskens and Van Keilegom (2003) by

$$2!\exp(r\tau)h^2 f_X(x)^{-1/2}V_{\mathcal{K}}(x)^{-1/2}I\{C(x)\}^{-1/2}\sum_{i=0}^{3} f_X(x)^{-1/2}I\{C(x)\}^{-1/2}\{N^{-1}\}_{3,i+1}A_{n,i}(x) \tag{19}$$

For the definition of the local Fisher information, $I\{C(x)\}$, the matrix N and the process $A_{ni}(x)$, we refer to Section 3 and Section 7. Define

$$Y_{ni}(x) \stackrel{\text{def}}{=} (nh)^{1/2}h^{-i}[I\{C(x)\}f_X(x)]^{-1/2}A_{ni}(x).$$

Then equation (19) can be written as

$$F_n(x) \stackrel{\text{def}}{=} 2!\exp(r\tau)h^2\{f_X(x)\}^{-1/2}V_{\mathcal{K}}(x)^{-1/2}I\{C(x)\}^{-1/2}\sum_{i=0}^{3} h^i\{N^{-1}\}_{3,i+1}Y_{ni}(x)$$

Please note that N is not a function of x as Claeskens and Van Keilegom (2003) erroneously write. Following their line of thoughts, we replace $Y_{ni}(x)$ (uniformly) by

$$Y'_{ni}(x) = h^{1/2}\int K_h(z-x)\Big(\frac{z-x}{h}\Big)dz.$$

In order to apply Corollary A1 of Bickel and Rosenblatt (1973), denote the covariance function of the Gaussian process $F_{n_q}(x)$ by $r(x)$, and note that

$$r(x) = \mathsf{Cov}(Y'_{nj}(x), Y'_{nj}(0)) = C_1 - C_2|x|^2 + o(|x|^2)$$

for $x \in E$, where C_1 and C_2 are constants. Since the regularity conditions are satisfied, the result follows.

Finally, we have to show that $\sup_{x\in E} |\widehat{\mathsf{Var}}\{\hat{\mathcal{K}}(x)\} - \mathsf{Var}\{\hat{\mathcal{K}}(x)\}| = o_p(1)$.

$$\begin{aligned}
& \sup_{x\in E} |\widehat{\mathsf{Var}}\{\hat{\mathcal{K}}(x)\} - \mathsf{Var}\{\hat{\mathcal{K}}(x)\}| \\
= & \sup_{x\in E} |\widehat{\mathsf{Var}}\left\{\frac{\hat{q}(x)-q(x)}{p(x)}\right\} - \mathsf{Var}\left\{\frac{\hat{q}(x)-q(x)}{p(x)}\right\}| + o_p\{(nh)^{-(1/2+\alpha)}(\log n)^{1+\alpha}\},
\end{aligned}$$

where $0 < \alpha < 1$.

According to Corollary 2.1 in Claeskens and Van Keilegom (2003), for $j = 3, k = 3$,

$$\sup_{x\in E} |\widehat{\mathsf{Var}}\{\hat{q}(x)\} - \mathsf{Var}\{\hat{q}(x)\}| = o_p\{(nh\log n)^{-1/2}\}.$$

This implies

$$\sup_{x\in E} |\widehat{\mathsf{Var}}\{\hat{\mathcal{K}}(x)\} - \mathsf{Var}\{\hat{\mathcal{K}}(x)\}| = o_p(1).$$

11.4 Expressions for the semiparametric estimator of SPD

Taking the derivatives of $c(M_{it})$ with respect to moneyness (M) and noting that both $d_1(M_{it})$ and $d_2(M_{it})$ depend on M_{it} we obtain

$$\begin{aligned}
\frac{\mathrm{d}c}{\mathrm{d}M} &= \varphi(d_1)\frac{\mathrm{d}d_1}{\mathrm{d}M} - e^{-r\tau}\frac{\varphi(d_2)}{M}\frac{\mathrm{d}d_2}{\mathrm{d}M} + e^{-r\tau}\frac{\Phi(d_2)}{M^2}, \\
\frac{\mathrm{d}^2c}{\mathrm{d}M^2} &= \varphi(d_1)\left\{\frac{\mathrm{d}^2d_1}{\mathrm{d}M^2} - d_1\left(\frac{\mathrm{d}d_1}{\mathrm{d}M}\right)^2\right\} \\
&\quad -\frac{e^{-r\tau}\varphi(d_2)}{M}\left\{\frac{\mathrm{d}^2d_2}{\mathrm{d}M^2} - \frac{2}{M}\frac{\mathrm{d}d_2}{\mathrm{d}M} - d_2\left(\frac{\mathrm{d}d_2}{\mathrm{d}M}\right)^2\right\} - \frac{2e^{-r\tau}\Phi(d_2)}{M^3}
\end{aligned}$$

Computing the first and second order differentials for d_1 and d_2 using the notation $V = \sigma(M)$, $V' = \partial\sigma(M)/\partial M$ and $V'' = \partial^2\sigma(M)/\partial M^2$, we obtain

$$
\begin{aligned}
\frac{\mathrm{d}d_1}{\mathrm{d}M} &= \frac{1}{MV\sqrt{\tau}} + \Big\{ -\frac{\log(M)+r\tau}{V^2\sqrt{\tau}} + \sqrt{\tau}/2 \Big\} V', \\
\frac{\mathrm{d}d_2}{\mathrm{d}M} &= \frac{1}{MV\sqrt{\tau}} + \Big\{ -\frac{\log(M)+r\tau}{V^2\sqrt{\tau}} - \sqrt{\tau}/2 \Big\} V', \\
\frac{\mathrm{d}^2d_1}{\mathrm{d}M^2} &= -\frac{1}{MV\sqrt{\tau}} \Big\{ \frac{1}{M} + \frac{V'}{V} \Big\} + V'' \Big\{ \frac{\sqrt{\tau}}{2} - \frac{\log(M)+r\tau}{V^2\sqrt{\tau}} \Big\} \\
&\quad + V' \Big\{ 2V' \frac{\log(M)+r\tau}{V^3\sqrt{\tau}} - \frac{1}{MV^2\sqrt{\tau}} \Big\}, \\
\frac{\mathrm{d}^2d_2}{\mathrm{d}M^2} &= -\frac{1}{MV\sqrt{\tau}} \Big\{ \frac{1}{M} + \frac{V'}{V} \Big\} + V'' \Big\{ -\frac{\sqrt{\tau}}{2} - \frac{\log(M)+r\tau}{V^2\sqrt{\tau}} \Big\} \\
&\quad + V' \Big\{ 2V' \frac{\log(M)+r\tau}{V^3\sqrt{\tau}} - \frac{1}{MV^2\sqrt{\tau}} \Big\},
\end{aligned}
$$

List of Figures

1 Examples of inter-temporal pricing kernels (as functions of moneyness) with fixed maturity 0.00833 (3 days) in years respectively on 17-Jan-2006 (blue), 18-Apr-2006 (red), 16-May-2006 (magenta), 13-June-2006 (black), see Grith, Härdle and Park (2009). 4

2 Examples of inter-temporal pricing kernels with various maturities in years: 0.02222 (8 days, red) 0.1 (36 days, black) on 12-Jan-2006 and their confidence bands. 5

3 10

4 Examples of sheet for pricing kernels in 060228 17

5 25

6 26

7 27

8 Plot of confidence bands (black), estimated value (red), the BS (blue) SPD with simulated volatility smile, $n_p = 2000$, $n_q = 300$, $h_{n_q} = 0.06$, $h_{n_p} = 0.0106$, $\alpha = 0.05$. 29

9 Plot of confidence bands (black), estimated value (red), the BS (blue) EPK with simulated volatility smile $n_p = 2000$, $n_q = 300$, $h_{n_q} = 0.06$, $h_{n_p} = 0.0106$, $\alpha = 0.05$. 29

10 31

11 31

12 Optional caption for list of figures 32

13 33

14 34

15 35

16 36

List of Tables

1 Coverage probability (area) of the uniform confidence band at 10% with annualized volatility = 0.1878 for SPD 25

2 Coverage probability (area) of the uniform confidence band at 5% with annualized volatility = 0.1878 for SPD 26

3 Coverage probability (area) of the uniform confidence band for the EPK at 5% with volatility(annualized) = 0.1878 26

4 Coverage probability (area) of the uniform confidence band for the EPK at 10% with volatility(annualized) = 0.1878 27

5 Coverage probability (area) of the uniform confidence band for the EPK at 5% 28

6 Coverage probability (area) of the uniform confidence band for the EPK at 10% 28

Declaration of Authorship

I hereby confirm that I have authored this master thesis independently and without use of others than the indicated resources. All passages, which are literally or in general matter taken out of publications or other resources, are marked as such.

Bibliography

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

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

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

Bickel, P. J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates, *The Annals of Statistics* **1**: 1071–1095.

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

Brown, D. P. and Jackwerth, J. C. (2004). The pricing kernel puzzle: Reconciling index option data and economic theory, *Manuscript* .

Carroll, R., Ruppert, D. and Welsh, A. (1998). Local estimating equations, *Journal of American Statistics Association* **93**: 214–227.

Chabi-Yo, Y., Garcia, R. and Renault, E. (2008). State dependence can explain the risk aversion puzzle, *Review of Financial Studies* **21**: 973–1011.

Claeskens, G. and Van Keilegom, I. (2003). Bootstrap confidence bands for regression curves and their derivatives, *The Annals of Statistics* **31**: 1852–1884.

Cleveland, W. (1979). Robust locally weighted regression and smoothing scatterplots, *Journal of American Statistical Association* **74**: 829–836.

Cochrane, J. H. (2001). *Asset pricing*, Princeton University Press.

Fan, J. (1992). Design-adaptive nonparametric regression, *Journal of American Statistical Association* **87**: 998–1004.

Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies, *The Annals of Statistics* **21**: 196–216.

Fengler, M. (2005). *Semiparametric modeling of implied volatility*, Springer.

Fengler, M. (2010). Option data and modeling of implied volatility, *Handbook of Computational Finance, Duan, Gentle, Härdle, eds. Springer Verlag* .

Golubev, Y., Härdle, W. and Timofeev, R. (2009). Testing monotonicity of pricing kernels, *Journal of Applied Econometrics, invited resubmission, 01.12.2009* .

Grith, M., Härdle, W. and Park, J. (2009). Shape invariant modelling pricing kernels and risk aversion, *Journal of Financial Econometrics, submitted, 01.05.2009* .

Hall, P. (1991a). *Edgeworth expansions for nonparametric density estimators, with applications*, Vol. 22, New York: Academic Press.

Härdle, W. (1989). Asymptotic maximal deviation of m-smoothers, *Journal of Multivariate Analysis* **29**: 163–179.

Härdle, W., Janssen, P. and Serfling, R. (1988). Strong uniform consistency rates for estimators of conditional functionals, *Annals of Statistics* **16**: 1428–1449.

Härdle, W. and Marron, J. (1991). Bootstrap simultaneous error bars for nonparametric regression, *Annals of Statistics* **19**: 778–796.

Jackwerth, J. (2000). Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* **13**: 433–451.

Jackwerth, J. and Rubinstein, M. (1996). Recovering probability distributions from option prices, *Journal of Finance* **51**: 1611–1631.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk, *Econometrica* **47**: 263–291.

Liero, H. (1982). On the maximal deviation of the kernel regression function estimate, *Mathematische Operationsforschung und Statistik, Serie Statistics* **13**: 171–182.

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

Rookley, C. (1997). Fully exploiting the information content of intra day option quotes: Applications in option pricing and risk management, *Manuscript* .

Rosenberg, J. and Engle, R. F. (2002). Empirical pricing kernels, *Journal of Financial Economics* **64**: 341–372.

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

Ruppert, D. and Wand, M. (1994). Multivariate locally weighted least squares regression, *The Annals of Statistics* **22**: 1346–1370.

Zhao, P.-L. (1994). Asymptotics of kernal estimators based on local maximum likelihood, *Nonparametric Statistics* **4**: 79–90.