

SFB 649 Discussion Paper 2011-004

A Confidence Corridor for Expectile Functions

Esra Akdeniz Duran*
Mengmeng Guo*
Wolfgang Karl Härdle*



* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

A Confidence Corridor for Expectile Functions*

Esra Akdeniz Duran [†]; Mengmeng Guo [‡]; Wolfgang Karl Härdle [§]

Abstract

Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. rvs and denote $v(x)$ as the unknown τ -expectile regression curve of Y conditional on X . We introduce the expectile-smoother $v_n(x)$ as a localized nonlinear estimator of $v(x)$, and prove the strong uniform consistency rate of $v_n(x)$ under general conditions. The stochastic fluctuation of the process $\{v_n(x) - v(x)\}$ is also studied in our paper. Moreover, using strong approximations of the empirical process and extreme value theory, we consider the asymptotic maximal deviation $\sup_{0 \leq x \leq 1} |v_n(x) - v(x)|$. This paper considers fitting a simultaneous confidence corridor (SCC) around the estimated expectile function of the conditional distribution of Y given x based on the observational data generated according to a nonparametric regression model. Furthermore, we apply it into the temperature analysis. We construct the simultaneous confidence corridors around the expectiles of the residuals from the temperature models to investigate the temperature risk drivers. We find the risk drivers in Berlin and Taipei are different.

Keywords: Expectile Regression; Consistency Rate; Simultaneous confidence corridor; Asymmetric least squares; Kernel Smoothing.

JEL classification: C00; C14; J01; J31

1 Introduction

In regression function estimation, most investigations are concerned with the conditional mean. Geometrically, the observations $\{(X_i, Y_i), i = 1, \dots, n\}$ form a cloud of points in a Euclidean space. The mean regression function focuses on the center

*The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 “Ökonomisches Risiko”, Humboldt-Universität zu Berlin is gratefully acknowledged. China Scholarship Council (CSC) is gratefully acknowledged.

[†]Research fellow at the Institute for Statistics and Econometrics of Humboldt-Universität zu Berlin, Germany and research assistant at Gazi University, Ankara, Turkey.

[‡]Corresponding Author, research associate at the Institute for Statistics and Econometrics of Humboldt-Universität zu Berlin, Germany, Email: guomengm@cms.hu-berlin.de.

[§]Professor at Humboldt-Universität zu Berlin and Director of C.A.S.E - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099, Berlin, Germany.

of the point-cloud, given the covariant X , see Efron (1991). However, more insights about the relation between Y and X can be gained by considering the higher or lower regions of the conditional distribution.

Asymmetric least squares estimation provides a convenient and relatively efficient method of summarizing the conditional distribution of a dependent variable given the regressors. It turns out that similar to conditional percentiles, the conditional expectiles also characterize the distribution. Breckling and Chambers (1988) proposed M -quantiles, which extends this idea by a “quantile-like” generalization of regression based on asymmetric loss functions. Expectile regression, and more generally M -quantile regression, can be used to characterize the relationship between a response variable and explanatory variables when the behaviour of “non-average” individuals is of interest. Jones (1994) described that expectiles and M -quantiles are related to means and quantiles are related to the median, and moreover expectiles are indeed quantiles of a transformed distribution. Expectiles can be generally used in labor market and financial market, which would be as interesting as quantile regression.

The expectile curves can be key aspects of inference in various economic problems and are of great interest in practice. Expectiles have recently been applied in financial and demographic studies. Kuan et al. (2009) considered the conditional autoregressive expectile (CARE) model to calculate the VaR, and expectiles are used to calculate the expected shortfall in Taylor (2008). Schnabel and Eilers (2009a) modelled the relationship between gross domestic product per capita (GDP) and average life expectancy using expectile curves. There are several methods to calculate the expectiles. Schnabel and Eilers (2009b) combined asymmetric least square and P-splines to calculate the smoothing expectile curve. In our paper, we use kernel smoothing method for the expectile curve, and apply it into the temperature studies. As we know, during the last several years, the dynamic of the temperatures is not stable especially in different cities, extreme weather appears occasionally. We investigate the behaviour of the temperature from Berlin and Taipei. We also construct the confidence corridors for the low and high expectile curves of the residuals from the dynamic temperature models, and compare the risk factors between Berlin and Taipei.

Both quantile and expectile can be expressed as minimum contrast parameter estimators. Define $q_\tau(u) = |\mathbf{I}(u \leq 0) - \tau|u$ for $0 < \tau < 1$, then the τ -quantile may be expressed as $\arg \min_\theta \mathbf{E}_{q_\tau}(y - \theta)$. With the interpretation of the contrast function

$\rho_\tau(u)$ as the negative log likelihood of asymmetric Laplace distribution, we can see the τ -quantile as a quasi maximum estimator in the location model. Changing the loss (contrast) function to

$$\rho_\tau(u) = |\mathbf{I}(u \leq 0) - \tau|u^2, \quad \tau \in (0, 1) \quad (1)$$

leads to expectile. Note that for $\tau = \frac{1}{2}$, we obtain the mean respective to the sample average. Putting this into a regression framework, we define the conditional expectile function (to level τ) as:

$$v(x) = \arg \min_{\theta} \mathbf{E}\{\rho_\tau(y - \theta)|X = x\} \quad (2)$$

From now on, we silently assume τ is fixed therefore we suppress the explicit notion. Inserting (1) into (2), we obtain:

$$v(x) = \arg \min_{\theta} (1 - \tau) \int_{-\infty}^{\theta} (y - \theta)^2 dF(y|x) + \tau \int_{\theta}^{\infty} (y - \theta)^2 dF(y|x) \quad (3)$$

$v(x)$ can be equivalent in seen as solving the following equation (w.r.t. v):

$$G(x, v) - \tau = \frac{\int_{-\infty}^v |y - v| dF(y|x)}{\int_{-\infty}^{\infty} |y - v| dF(y|x)} - \tau = 0 \quad (4)$$

Yet another representation of $v(x)$ is given by an average of the conditional upside and downside mean:

$$v(x) = \gamma \mathbf{E}\{Y|Y > v(x)\} + (1 - \gamma) \mathbf{E}\{Y|Y \leq v(x)\}$$

where $\gamma = \tau[1 - F_Y\{v(x)\}]/(\tau[1 - F_Y\{v(x)\}] + (1 - \tau)F_Y\{v(x)\})$ may be interpreted as the weighted probability of $Y > v(x)$. Here $F_Y(\cdot)$ denotes the marginal *cdf* of Y . This property distinguishes the expectile from expected shortfall because the latter is determined only by a conditional downside mean. Newey and Powell (1987) show that $v(x)$ is monotonically increasing in τ and is location and scale equivalent, in the sense that for $\tilde{Y} = aY + b$ and $a > 0$, then $v_{\tilde{Y}}(\tau) = av_Y + b$. In our conditional setting, we need to deal with $v(x)$ from (3) and variation of the RHS of (3) when θ is in a neighborhood of $v(x)$.

Recall conditional quantile $l(x)$ at level τ can be considered as

$$l(x) = \inf\{y \in \mathbb{R}|F(y|x) \geq \tau\}$$

Therefore, the proposed estimate $l_n(x)$ can be expressed :

$$l_n(x) = \inf\{y \in \mathbb{R}|\hat{F}(y|x) \geq \tau\} = \hat{F}_x^{-1}(\tau)$$

where $\widehat{F}(y|x)$ is the kernel estimator of $F(y|x)$:

$$\widehat{F}(y|x) = \frac{\sum_{i=1}^n K_h(x - X_i) \mathbf{I}(Y_i \leq y)}{\sum_{i=1}^n K_h(x - X_i)}$$

With the similar idea, we can treat expectile $v(x)$ as

$$G_{Y|x}(v) = \frac{\int_{-\infty}^{v(x)} |Y - v(x)| dF(Y|x)}{\int_{-\infty}^{\infty} |Y - v(x)| dF(Y|x)} = \tau$$

$$v(x) = G_{Y|x}^{-1}(\tau)$$

τ expectile curve estimator:

$$v_n(x) = \widehat{G}_{Y|x}^{-1}(\tau)$$

where the nonparametric estimate of $G_{Y|x}(v)$ is

$$\widehat{G}_{Y|x}(v) = \frac{\sum_{i=1}^n K_h(x - X_i) \mathbf{I}(Y_i < y) |y - v|}{\sum_{i=1}^n K_h(x - X_i) |y - v|}$$

Quantiles and expectiles both characterize a distribution function although they are different in nature. As an illustration, Figure 1 plots curves of quantiles and expectiles of the standard normal $N(0, 1)$. There is a one-to-one mapping relationship between quantile and expectile, see as Yao and Tong (1996). Fixed x , define $w(\tau)$ such that $v_{w(\tau)}(x) = l(x)$, then $w(\tau)$ is related to $l(x)$ via

$$w(\tau) = \frac{\tau l(x) - \int_{-\infty}^{l(x)} y dF(y|x)}{2 \mathbf{E}(Y|x) - 2 \int_{-\infty}^{l(x)} y dF(y|x) - (1 - 2\tau)l(x)} \quad (5)$$

$l(x)$ is an increasing function of τ , therefore, $w(\tau)$ is also a monotonically increasing function. Expectile corresponds to quantile with transformation w . For example, $Y \sim U(0, 1)$, then $w(\tau) = \tau^2 / (2\tau^2 - 2\tau + 1)$.

In light of the concepts of M -estimation as in Huber (1981), if we define $\psi(u)$ as:

$$\begin{aligned} \psi(u) &= \frac{\partial \rho(u)}{\partial u} \\ &= |\mathbf{I}(u \leq 0) - \tau|u \\ &= \{\tau - \mathbf{I}(u \leq 0)\}|u| \end{aligned}$$

$v_n(x)$ and $v(x)$ can be treated as a zero (w.r.t. θ) of the function:

$$H_n(\theta, x) \stackrel{\text{def}}{=} n^{-1} \sum_{i=1}^n K_h(x - X_i) \psi(Y_i - \theta) \quad (6)$$

$$H(\theta, x) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x, y) \psi(y - \theta) dy \quad (7)$$

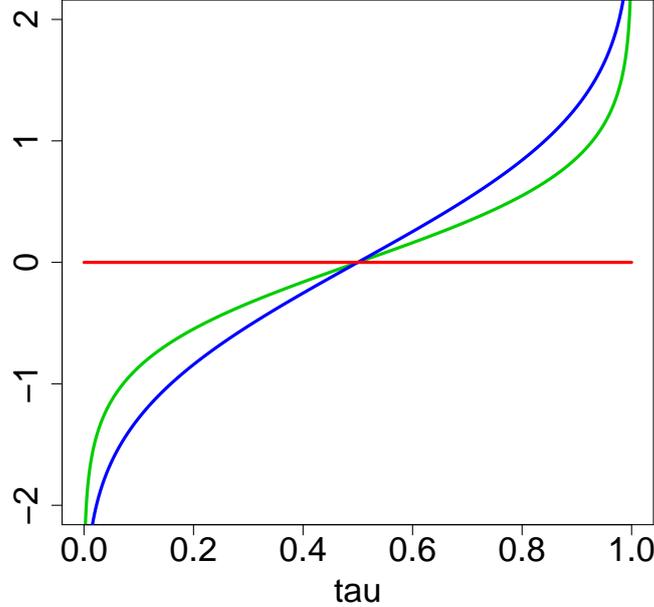


Figure 1: Quantile Curve(blue) and Expectile Curve(green) for Standard Normal Distribution.

correspondingly.

By employing similar methods as those developed in Härdle (1989) it is shown in this paper that

$$\begin{aligned}
 & \mathbb{P} \left((2\delta \log n)^{1/2} \left[\sup_{x \in J} r(x) |\{v_n(x) - v(x)\}| / \lambda(K)^{1/2} - d_n \right] < z \right) \\
 & \longrightarrow \exp\{-2 \exp(-z)\}, \quad \text{as } n \rightarrow \infty.
 \end{aligned} \tag{8}$$

with some adjustment of $v_n(x)$, we can see that the supreme of $v_n(x) - v(x)$ follows the asymptotic Gumbel distribution, where $r(x)$, δ , $\lambda(K)$, d_n are suitable scaling parameters. The asymptotic result (8) therefore allows the construction of simultaneous confidence corridor (SCC) for $v(x)$ based on specifications of the stochastic fluctuation of $v_n(x)$. The strong approximation with Brownian bridge techniques is applied in this paper to prove the asymptotic distribution of $v_n(x)$.

The structure of this paper is as follows. In Section 2, the stochastic fluctuation of the process $\{v_n(x) - v(x)\}$ is studied and the simultaneous confidence corridor (SCC) is presented through the equivalence of several stochastic processes. We get the asymptotic distribution of $v_n(x)$. Further we also get a strong uniform consistency rate of $\{v_n(x) - v(x)\}$. In Section 3, a small Monte Carlo study is studied to investigate the behaviour of $v_n(x)$ when the data is generated with the error terms standard normally distributed. In Section 4, an application considers

the temperature in Berlin and Taipei. Moreover, a simultaneous confidence corridor (SCC) for the residuals after a fitted temperature model will be constructed to detect the risk drivers for temperature. All proofs are attached in Section 5.

2 Results

We make the following assumptions about the distribution of (X, Y) and the score function $\psi(u)$ in addition to the existence of an initial estimator whose error is a.s. uniformly bounded.

(A1) The kernel $K(\cdot)$ is positive, symmetric, has compact support $[-A, A]$ and is Lipschitz continuously differentiable with bounded derivatives;

(A2) $(nh)^{-1/2}(\log n)^{3/2} \rightarrow 0$, $(n \log n)^{1/2}h^{5/2} \rightarrow 0$, $(nh^3)^{-1}(\log n)^2 \leq M$, M is a constant;

(A3) $h^{-3}(\log n) \int_{|y|>a_n} f_Y(y)dy = \mathcal{O}(1)$, $f_Y(y)$ the marginal density of Y , $\{a_n\}_{n=1}^\infty$ a sequence of constants tending to infinity as $n \rightarrow \infty$;

(A4) $\inf_{x \in J} |p(x)| \geq p_0 > 0$, where $p(x) = \partial \mathbf{E}\{\psi(Y - \theta)|x\} / \partial \theta|_{\theta=v(x)} \cdot f_X(x)$, where $f_X(x)$ is the marginal density of X ;

(A5) The expectile function $v(x)$ is Lipschitz twice continuously differentiable, for all $x \in J$.

(A6) $0 < m_1 \leq f_X(x) \leq M_1 < \infty$, $x \in J$, and the conditional density $f(\cdot|y)$, $y \in \mathbb{R}$, is uniform locally Lipschitz continuous of order $\tilde{\alpha}$ (uLL- $\tilde{\alpha}$) on J , uniformly in $y \in \mathbb{R}$, with $0 < \tilde{\alpha} \leq 1$, and $\psi(x)$ is piecewise twice continuously differentiable.

Define also

$$\begin{aligned} \sigma^2(x) &= \mathbf{E}[\psi^2\{Y - v(x)\}|x] \\ H_n(x) &= (nh)^{-1} \sum_{i=1}^n K\{(x - X_i)/h\} \psi\{Y_i - v(x)\} \\ D_n(x) &= \partial(nh)^{-1} \sum_{i=1}^n K\{(x - X_i)/h\} \psi\{Y_i - \theta\} / \partial \theta|_{\theta=v(x)} \end{aligned}$$

and assume that $\sigma^2(x)$ and $f_X(x)$ are differentiable.

Assumption (A1) on the compact support of the kernel could possibly be relaxed by introducing a cutoff technique as in Csörgö and Hall (1982) for density estimators. Assumption (A2) has purely technical reasons: to keep the bias at a lower rate than the variance and to ensure the vanishing of some non-linear remainder terms. Assumption (A3) appears in a somewhat modified form also in Johnston

(1982). Assumptions (A5) and (A6) are common assumptions in robust estimation as in Huber (1981), Härdle et al. (1988) that are satisfied by exponential, and generalized hyperbolic distributions.

Zhang (1994) has proved the asymptotic normality of the nonparametric expectile. Under the Assumptions (A1) to (A4), we have:

$$\sqrt{nh}\{v_n(x) - v(x)\} \xrightarrow{L} N\{0, V(x)\} \quad (9)$$

with

$$V(x) = \lambda(K)f_X(x)\sigma^2(x)/p(x)^2$$

where we can denote

$$\begin{aligned} \sigma^2(x) &= E[\psi^2\{Y - v(x)\}|x] \\ &= \int \psi^2\{y - v(x)\}dF(y|x) \\ &= \tau^2 \int_{v(x)}^{\infty} \{y - v(x)\}^2 dF(y|x) + (1 - \tau)^2 \int_{-\infty}^{v(x)} \{y - v(x)\}^2 dF(y|x) \end{aligned} \quad (10)$$

$$\begin{aligned} p(x) &= E[\psi'\{Y - v(x)\}|x] \cdot f_X(x) \\ &= \left\{ \tau \int_{v(x)}^{\infty} dF(y|x) + (1 - \tau) \int_{-\infty}^{v(x)} dF(y|x) \right\} \cdot f_X(x) \end{aligned} \quad (11)$$

For the uniform strong consistency rate of $v_n(x) - v(x)$, we apply the result of Härdle et al. (1988) by taking $\beta(y) = \psi(y - \theta)$, $y \in \mathbb{R}$, for $\theta \in I = \mathbb{R}$, $q_1 = q_2 = -1$, $\gamma_1(y) = \max\{0, -\psi(y - \theta)\}$, $\gamma_2(y) = \min\{0, -\psi(y - \theta)\}$ and $\lambda = \infty$ to satisfy the representations for the parameters there. We have the following lemma under some specified assumptions:

Lemma 1 *Let $H_n(\theta, x)$ and $H(\theta, x)$ be given by (6) and (7). Under Assumption (A6) and $(nh/\log n)^{1/2} \rightarrow \infty$ through Assumption (A2), for some constant A^* not depending on n , we have a.s. as $n \rightarrow \infty$*

$$\sup_{\theta \in I} \sup_{x \in J} |H_n(\theta, x) - H(\theta, x)| \leq A^* \max\{(nh/\log n)^{-1/2}, h^{\tilde{\alpha}}\} \quad (12)$$

For our result on $v_n(\cdot)$, we shall also require

$$\inf_{x \in J} \left| \int \psi\{y - v(x) + \varepsilon\}dF(y|x) \right| \geq \tilde{q}|\varepsilon|, \quad \text{for } |\varepsilon| \leq \delta_1, \quad (13)$$

where δ_1 and \tilde{q} are some positive constants, see also Härdle and Luckhaus (1984). This assumption is satisfied if there exists a constant \tilde{q} such that $f(v(x)|x) > \tilde{q}/p$, $x \in J$.

Theorem 1 Under the conditions of Lemma 1 and also assuming (13) holds, we have a.s. as $n \rightarrow \infty$

$$\sup_{x \in J} |v_n(x) - v(x)| \leq B^* \max\{(nh/\log n)^{-1/2}, h^{\tilde{\alpha}}\} \quad (14)$$

with $B^* = A^*/m_1\tilde{q}$ not depending on n and m_1 a lower bound of $f_X(x)$. If additionally $\tilde{\alpha} \geq \{\log(\sqrt{\log n}) - \log(\sqrt{nh})\}/\log h$, it can be further simplified to

$$\sup_{x \in J} |v_n(x) - v(x)| \leq B^* \{(nh/\log n)^{-1/2}\}.$$

Theorem 2 Let $h = n^{-\delta}$, $\frac{1}{5} < \delta < \frac{1}{3}$, $\lambda(K) = \int_{-A}^A K^2(u)du$ and

$$\begin{aligned} d_n &= (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2} [\log\{c_1(K)/\pi^{1/2}\} + \frac{1}{2}\{\log \delta + \log \log n\}], \\ &\text{if } c_1(K) = \{K^2(A) + K^2(-A)\}/\{2\lambda(K)\} > 0 \\ d_n &= (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2} \log\{c_2(K)/2\pi\} \\ &\text{otherwise with } c_2(K) = \int_{-A}^A \{K'(u)\}^2 du / \{2\lambda(K)\}. \end{aligned}$$

Then (8) holds with

$$r(x) = (nh)^{-\frac{1}{2}} p(x) \{f_X(x)/\sigma^2(x)\}^{\frac{1}{2}}$$

This theorem can be used to construct uniform confidence intervals for the regression function as stated in the following corollary.

Corollary 1 Under the assumptions of the theorem above, an approximate $(1-\alpha) \times 100\%$ confidence band over $[0, 1]$ is

$$v_n(x) \pm (nh)^{-1/2} \{\hat{\sigma}^2(x)\lambda(K)/\hat{f}_X(x)\}^{1/2} \hat{p}^{-1}(x) \{d_n + c(\alpha)(2\delta \log n)^{-1/2}\}$$

where $c(\alpha) = \log 2 - \log |\log(1-\alpha)|$ and $\hat{f}_X(x)$, $\hat{\sigma}^2(x)$ and $\hat{p}(x)$ are consistent estimates for $f_X(x)$, $\sigma^2(x)$ and $p(x)$.

With $\sqrt{\hat{V}(x)}$ introduced, we can further write Corollary 1 as:

$$v_n(x) \pm (nh)^{-1/2} \{d_n + c(\alpha)(2\delta \log n)^{-1/2}\} \sqrt{\hat{V}(x)}$$

where $\hat{V}(x)$ is the nonparametric estimator of $V(x)$. The proof is essentially based on a liberalization argument after a Taylor series expansion. The leading linear term will then be approximated in a similar way as in Johnston (1982), Bickel and

Rosenblatt (1973). The main idea behind the proof is a strong approximation of the empirical process of $\{(X_i, Y_i)_{i=1}^n\}$ by a sequence of Brownian bridges as proved by Tusnady (1977).

As $v_n(x)$ is the zero (w.r.t. θ) of $H_n(\theta, x)$, it follows by applying 2nd-order Taylor expansions to $H_n(\theta, x)$ around $v(x)$ that

$$v_n(x) - v(x) = \{H_n(x) - \mathbf{E} H_n(x)\}/q(x) + R_n(x) \quad (15)$$

where $\{H_n(x) - \mathbf{E} H_n(x)\}/q(x)$ is the leading linear term and the remainder term is written as:

$$\begin{aligned} R_n(x) &= H_n(x)\{q(x) - D_n(x)\}/\{D_n(x) \cdot q(x)\} + \mathbf{E} H_n(x)/q(x) \\ &\quad + \frac{1}{2}\{v_n(x) - v(x)\}^2 \cdot \{D_n(x)\}^{-1} \end{aligned} \quad (16)$$

$$\cdot (nh)^{-1} \sum_{i=1}^n K\{(x - X_i)/h\} \psi''\{Y_i - v(x) + r_n(x)\}, \quad (17)$$

$$|r_n(x)| < |v_n(x) - v(x)|.$$

We show in Section 5 that (Lemma 4) that $\|R_n\| = \sup_{x \in J} |R_n(x)| = \mathcal{O}_p\{(nh \log n)^{-1/2}\}$.

Furthermore, the rescaled linear part

$$Y_n(x) = (nh)^{1/2} \{\sigma^2(x) f_X(x)\}^{-1/2} \{H_n(x) - \mathbf{E} H_n(x)\}$$

is approximated by a sequence of Gaussian processes, leading finally to the Gaussian process

$$Y_{5,n}(x) = h^{-1/2} \int K\{(x - t)/h\} dW(x). \quad (18)$$

Drawing upon the result of Bickel and Rosenblatt (1973), we finally obtain asymptotically the Gumbel distribution.

We also need the Rosenblatt (1952) transformation,

$$T(x, y) = \{F_{X|y}(x|y), F_Y(y)\},$$

which transforms (X_i, Y_i) into $T(X_i, Y_i) = (X'_i, Y'_i)$ mutually independent uniform rv's. In the event that x is a d -dimension covariate, the transformation becomes:

$$\begin{aligned} T(x_1, x_2, \dots, x_d, y) &= \{F_{X_1|y}(x_1|y), F_{X_2|y}(x_2|x_1, y), \dots, \\ &\quad F_{X_k|x_{d-1}, \dots, x_1, y}(x_k|x_{d-1}, \dots, x_1, y), F_Y(y)\}. \end{aligned} \quad (19)$$

With the aid of this transformation, Theorem 1 of Tusnady (1977) may be applied to obtain the following lemma.

Lemma 2 *On a suitable probability space a sequence of Brownian bridges B_n exists that*

$$\sup_{x \in J, y \in \mathbb{R}} |Z_n(x, y) - B_n\{T(x, y)\}| = \mathcal{O}\{n^{-1/2}(\log n)^2\} \quad a.s.,$$

where $Z_n(x, y) = n^{1/2}\{F_n(x, y) - F(x, y)\}$ denotes the empirical process of $\{(X_i, Y_i)\}_{i=1}^n$.

For $d > 2$, it is still an open problem which deserves further research.

Before we define the different approximating processes, let us first rewrite (18) as a stochastic integral w.r.t. the empirical process $Z_n(x, y)$,

$$Y_n(x) = \{hg'(x)\}^{-1/2} \iint K\{(x-t)/h\}\psi\{y-v(x)\}dZ_n(t, y),$$

$$g'(x) = \sigma^2(x)f_X(x).$$

The approximating processes are now:

$$Y_{0,n}(x) = \{hg(x)\}^{-1/2} \iint_{\Gamma_n} K\{(x-t)/h\} \psi\{y-v(x)\} dZ_n(t,y) \quad (20)$$

where $\Gamma_n = \{|y| \leq a_n\}$, $g(t) = \mathbb{E}[\psi^2\{y-v(x)\} \cdot \mathbf{I}(|y| \leq a_n) | X=x] \cdot f_X(x)$

$$Y_{1,n}(x) = \{hg(x)\}^{-1/2} \iint_{\Gamma_n} K\{(x-t)/h\} \psi\{y-v(x)\} dB_n\{T(t,y)\} \quad (21)$$

$\{B_n\}$ being the sequence of Brownian bridges from Lemma 2.

$$Y_{2,n}(x) = \{hg(x)\}^{-1/2} \iint_{\Gamma_n} K\{(x-t)/h\} \psi\{y-v(x)\} dW_n\{T(t,y)\} \quad (22)$$

$\{W_n\}$ being the sequence of Wiener processes satisfying

$$B_n(t', y') = W_n(t', y') - t'y'W_n(1, 1)$$

$$Y_{3,n}(x) = \{hg(x)\}^{-1/2} \iint_{\Gamma_n} K\{(x-t)/h\} \psi\{y-v(t)\} dW_n\{T(t,y)\} \quad (23)$$

$$Y_{4,n}(x) = \{hg(x)\}^{-1/2} \int g(t)^{1/2} K\{(x-t)/h\} dW(t) \quad (24)$$

$$Y_{5,n}(x) = h^{-1/2} \int K\{(x-t)/h\} dW(t) \quad (25)$$

$\{W(\cdot)\}$ being the Wiener process.

Lemmas 5 to 10 ensure that all these processes have the same limit distributions.

The result then follows from

Lemma 3 (*Theorem 3.1 in Bickel and Rosenblatt (1973)*) *Let d_n , $\lambda(K)$, δ as in Theorem 2. Let*

$$Y_{5,n}(x) = h^{-1/2} \int K\{(x-t)/h\} dW(t).$$

Then, as $n \rightarrow \infty$, the supremum of $Y_{5,n}(x)$ has a Gumbel distribution.

$$\mathbb{P} \left\{ (2\delta \log n)^{1/2} \left[\sup_{x \in J} |Y_{5,n}(x)| / \{\lambda(K)\}^{1/2} - d_n \right] < z \right\} \rightarrow \exp\{-2 \exp(-z)\}.$$

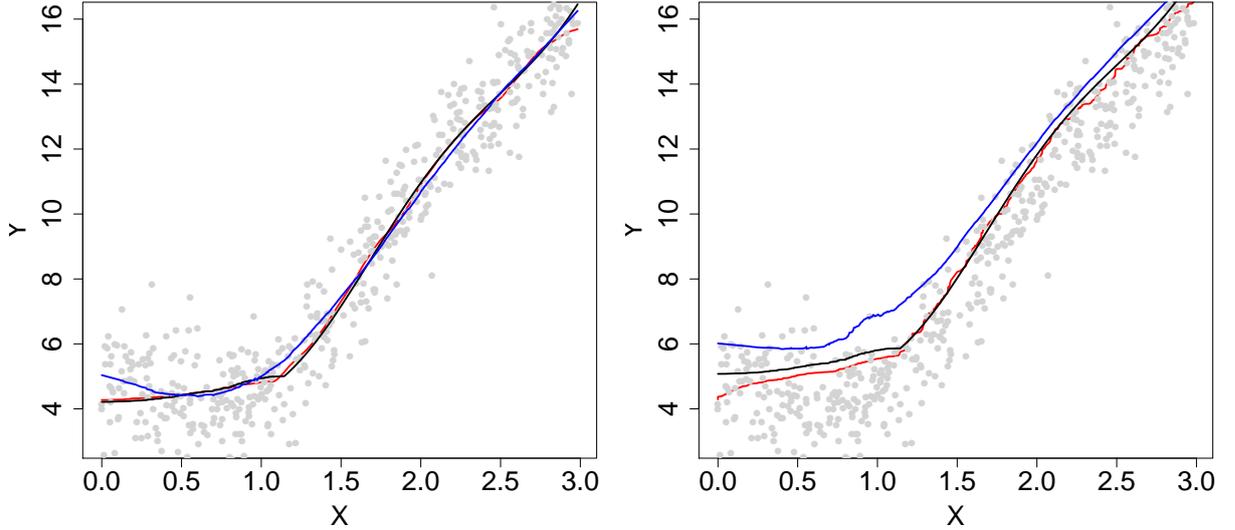


Figure 2: $\tau = 0.5$ (left) and $\tau = 0.9$ (right) Estimated Quantile and Expectile Plot. **Quantile Curve**, Theoretical Expectile Curve, **Estimated Expectile Curve**

3 A Monte Carlo Study

In the design of the simulation, we follow the same idea as in Schnabel and Eilers (2009b), since the result is very intuitive to explain and to compare with quantiles. We generate the bivariate random variables $\{(X_i, Y_i)\}_{i=1}^n$ with the sample size $n = 500$, and X is uniformly distributed on $[0, 3]$

$$Y = 1.5X^2 + 4 + \cos(3X) + \varepsilon \quad (26)$$

where $\varepsilon \sim N(0, 1)$.

Obviously, the theoretical expectiles (fixed τ) are determined by

$$v(x) = 1.5x^2 + 4 + \cos(3x) + v_N(\tau) \quad (27)$$

where $v_N(\tau)$ is the τ th-expectile of the standard Normal distribution.

Figure 2 (in the left part) describes the simulated data (the grey points), together with the 0.5 estimated quantile and estimated expectile and theoretical expectile curves, which represents respectively the conditional median and conditional mean. The conditional mean and conditional median coincide with each other, since the error term is symmetrically distributed, which can be found very obviously in Figure 2. In the right part of the figure, we consider the conditional 0.9 quantile and expectile curves. Since the expectile curve contains more information than the

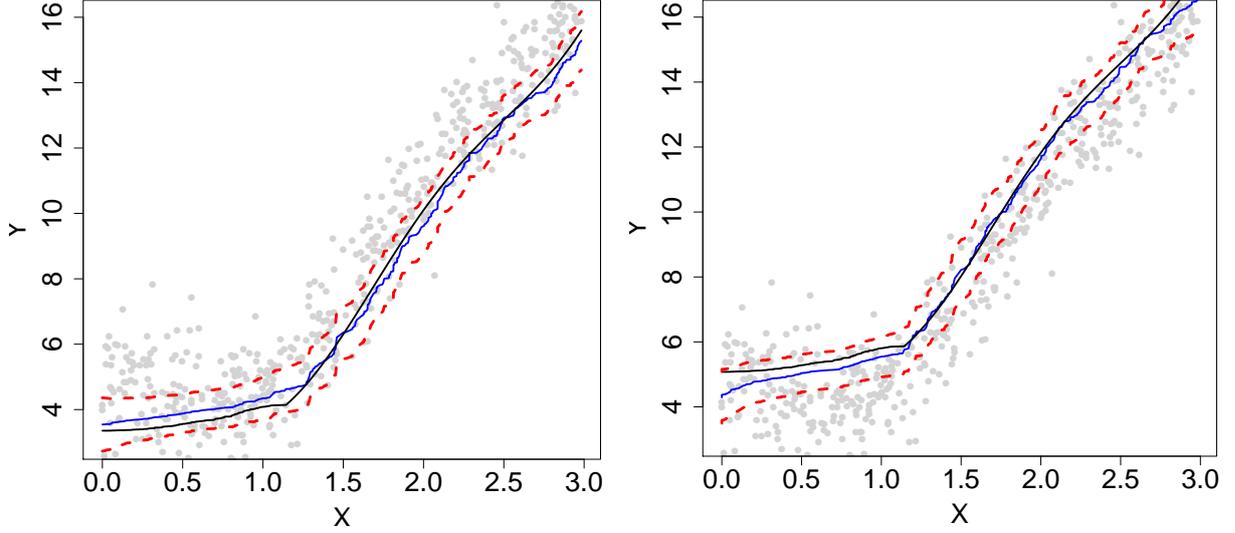


Figure 3: Uniform Confidence Bands for Expectile Curve for $\tau = 0.1$ (Left) and $\tau = 0.9$ (Right). Theoretical Expectile Curve, **Estimated Expectile Curve** and **5% – 95% Uniform Confidence Bands**

quantile curve, there is a gap between the quantile curve and the expectile curve, which can be interpreted by the transformation $w(\tau)$. As we have checked, for the standard normal distribution, the 0.9 quantile can be expressed by the around 0.96 expectile. Moreover, the expectile curve is smoother than the corresponding quantile curve.

Figure 3 shows the 5% – 95% uniform confidence bands for expectile curve, which are represented by the two red dot lines. We calculate both 0.1 (left) and 0.9 (right) expectile curves. The black lines stand for the corresponding 0.1 and 0.9 theoretical expectile curves, and the blue lines are the corresponding estimated expectile curves. Obviously, the theoretical expectile curves locate in the confidence bands.

4 Application

In this part, we apply the expectile into the temperature study. We consider the daily temperature both of Berlin and Taipei, ranging from 19480101 to 20071231, together 21900 observations. The statistical properties of the temperature are summarized in Table 4. The Berlin temperature data was obtained from Deutscher Wetterdienst, and the Taipei temperature data was obtained from the center for adaptive data

	<i>Mean</i>	<i>SD</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>Max</i>	<i>Min</i>
<i>Berlin</i>	9.66	7.89	-0.315	2.38	30.4	-18.5
<i>Taipei</i>	22.61	5.43	-0.349	2.13	33.0	6.5

Table 1: Statistical summary of the temperature in Berlin and Taipei

analysis in National Central University.

Weather risk is the uncertainty in cash flow and earnings caused by weather volatility. Many energy companies have a natural position in weather which is their largest source of financial uncertainty. However, it is a local phenomenon for each city, since the location, the atmosphere and the human activities are quite different. It is well documented that seasonal volatility in the regression residuals appears highest during the winter months where the temperature shows high volatility. To assess the potential for hedging against weather surprises, and to formulate the appropriate hedging strategies, we need to determine how much “weather noise” exists. Time series modeling reveals a wealth of information about both conditional mean dynamics and conditional variance dynamics of daily average temperature, and it provides insights into both the distributions of temperature and temperature surprises, and the differences between them.

Before proceeding to detailed modeling and forecasting results, it is useful to get an overall feel for the daily average temperature data. Figure 4 displays the average temperature series for the last five years of the sample. The black line stands for the temperature in Taipei, and the blue line describes for the temperature in Berlin. The time series plots reveal strong and unsurprising seasonality in average temperature: in each city, the daily average temperature moves repeatedly and regularly through periods of high temperature (summer) and low temperature (winter). Importantly, however, the seasonal fluctuations differ noticeably across cities both in terms of amplitude and detail of pattern.

Based on the pattern of the temperature we observed, we apply a conventional model for temperature dynamics, which is a stochastic model with seasonality and inter temporal autocorrelation. To understand the model clearly, Let us introduce the time series decomposition of the temperature, with $t = 1, \dots, \tau = 365$ days,

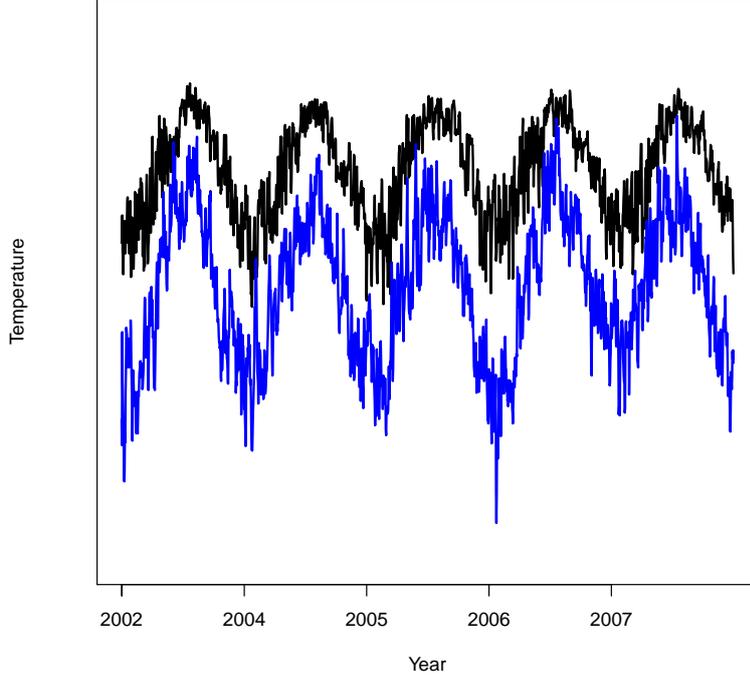


Figure 4: The time series plot of the temperature in Berlin and Taipei from 2002-2007. The black line stands for the temperature in Taipei, and the blue line is in Berlin.

and $j = 0, \dots, J$ years:

$$\begin{aligned}
 X_{365j+t} &= T_{t,j} - \Lambda_t \\
 X_{365j+t} &= \sum_{l=1}^L \beta_{lj} X_{365j+t-l} + \varepsilon_{t,j} \\
 \Lambda_t &= a + bt + \sum_{m=1}^M c_l \cos\left\{\frac{2\pi(t - d_m)}{l \cdot 365}\right\}
 \end{aligned} \tag{28}$$

where $T_{t,j}$ is the temperature at day t in year j , Λ_t denotes the seasonality effect. Motivation of this modeling approach can be found in Diebold and Inoue (2001). Later studies like e.g. Campbell and Diebold (2005) and Benth et al. (2007) have provided evidence that the parameters β_{lj} are likely to be j independent and hence estimated consistently from a global autoregressive process $AR(L_j)$ model with $L_j = L$. The analysis of the partial autocorrelations and Akaike's Information Criterion (AIC) suggests that a simple $AR(3)$ model fits the temperature evolution both in Berlin and Taipei well.

In this part, we consider the residuals of temperature from the fitted model from Equation (28). Since the temperatures have seasonal effects, and also AR

effects, we use the residuals after taking out such effects. We intend to construct the confidence corridors for the 0.01 and 0.9-expectile curves for the residuals from the fitted temperature model. We would like to investigate whether the performance of the extreme values are same in different cities, further to analyze the risk factors in different places.

The left part of the pictures describes the expectile curves for Berlin, and the right part is for Taipei. In each figure, the thick black line depicts the average expectile with the data from 1948 to 2007. The red line is the expectile for the residuals from the model with the first 20 years temperature, i.e. we use the data from 1948 to 1967. The 0.9 expectile for the second 20 years (1968-1987) residuals is described by the green line, and the blue line stands for the expectile curve in the latest 20 years (1988-2007). The dot lines are the 5% – 95% confidence corridors corresponding to the expectile curve with the same color. Figures 4 – 7 describe the 0.9 expectile curve for the residuals from the the models for Berlin and Taipei, as well as their corresponding confidence corridors. Obviously, the variance is higher in winter-earlier summer both in Berlin and Taipei.

It is obvious to detect that the expectile curves in Berlin and Taipei are quite different. Firstly, the variation of the temperature in Berlin is smaller than that of Taipei. All the expectile curves cross with each other during the last 100 observations for Berlin, and the variance in this period is smaller, seen from the expectile curves. Moreover, all of them nearly locate in the corresponding three confidence corridors. However, the performance of the temperature in Taipei is quite different. The expectile curves for Taipei have similar trends for each 20 years. They have highest volatilities in January, and lowest volatility in July. More interestingly, the expectile curve for the latest 20 years does not locate in the confidence corridor constructed using the data from the first 20 years and second 20 years, see Figure 4 and Figure 7. Similarly, the expectile curve for the first 20 years does not locate in the confidence corridor constructed by the latest 20 years.

Further, let us move to study the low expectile for the residuals from the models for Berlin and Taipei. As we know, it is very hard to get the very low quantile curve, due to the character of quantiles. However, it is not a problem for expectiles. One can calculate very low or very high expectiles. Therefore, we also calculate the 0.01 expectiles for the residuals and their corresponding confidence corridors, which are described in Figures 8 – 10. One can tell easily that the shapes of the 0.01 expectile for Berlin and Taipei are different. It does not fluctuate a lot during the whole year

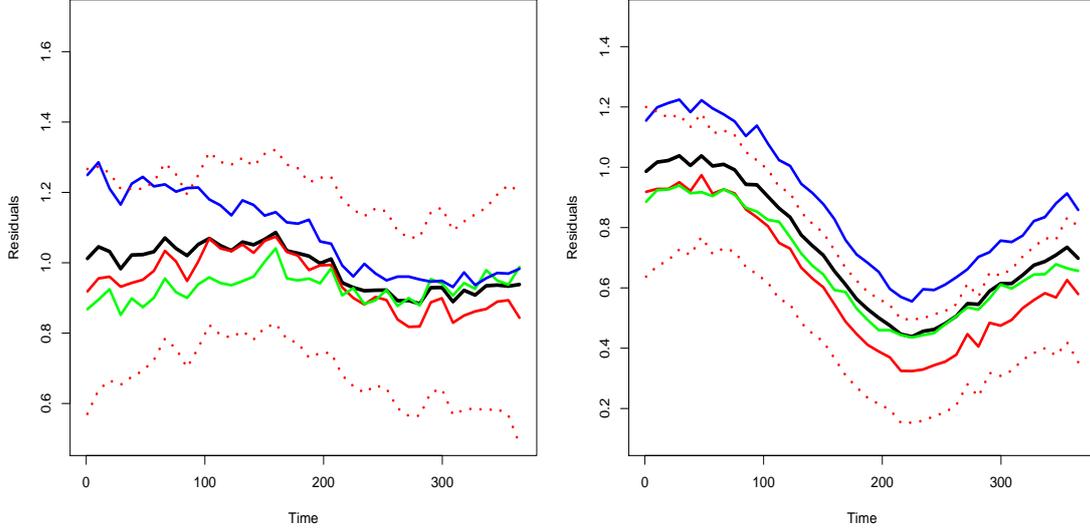


Figure 5: 0.9-expectile curves for Berlin (left) and Taipei (right) daily temperature residuals from 1948-2007 with the 5% – 95% confidence corridors for the first 20 years expectile.

in Berlin, while the variation in Taipei is much bigger. However, all of these curves both for Berlin and Taipei locate in their corresponding confidence corridors.

Obviously, one can say that the performance of the residuals are quite different from Berlin and Taipei, after we take out the regular seasonal effect and AR effect, especially for the high expectiles. Since the temperature can be influenced by the human factors and other natural factors, which have been well documented in literature. We find the variation of the temperature in Taipei is more volatile. As one interpretation, as we know in the last 60 years, Taiwan has been experiencing a fast developing period, such as the industrial expansion, the burning of fossil fuel and deforestation and other sectors, which would be an important factor to induce the more volatility in the temperature of Taipei. However, Germany is well-developed in this period, especially in Berlin, where there is no intensive industries. Therefore, one may say the residuals reveals the influence of the human activities and we conclude that the risk drivers for temperature are localized.

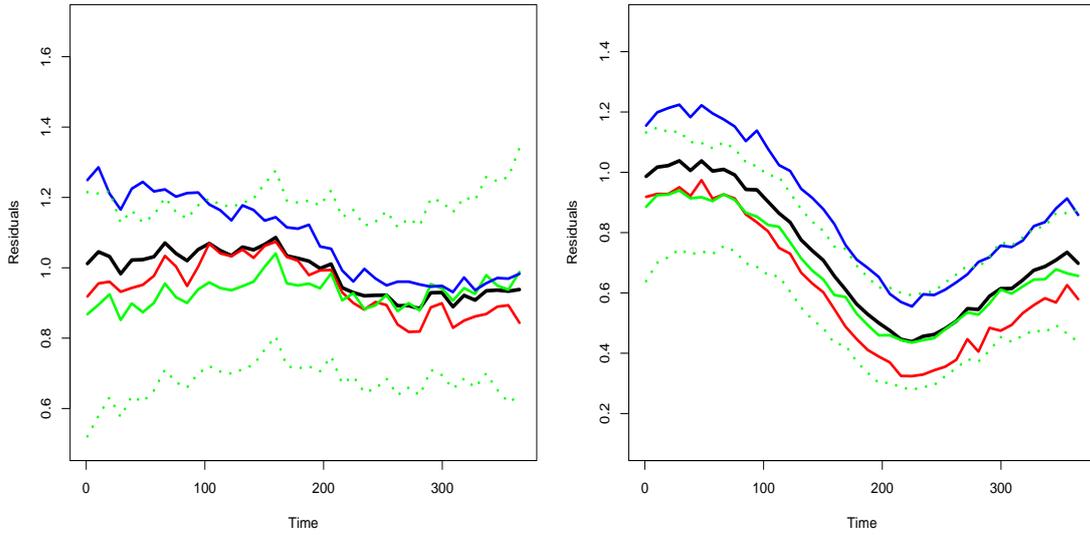


Figure 6: 0.9-expectile curves for Berlin (left) and Taipei (right) daily temperature residuals from 1948-2007 with the 5% – 95% confidence corridors for the second 20 years expectile.

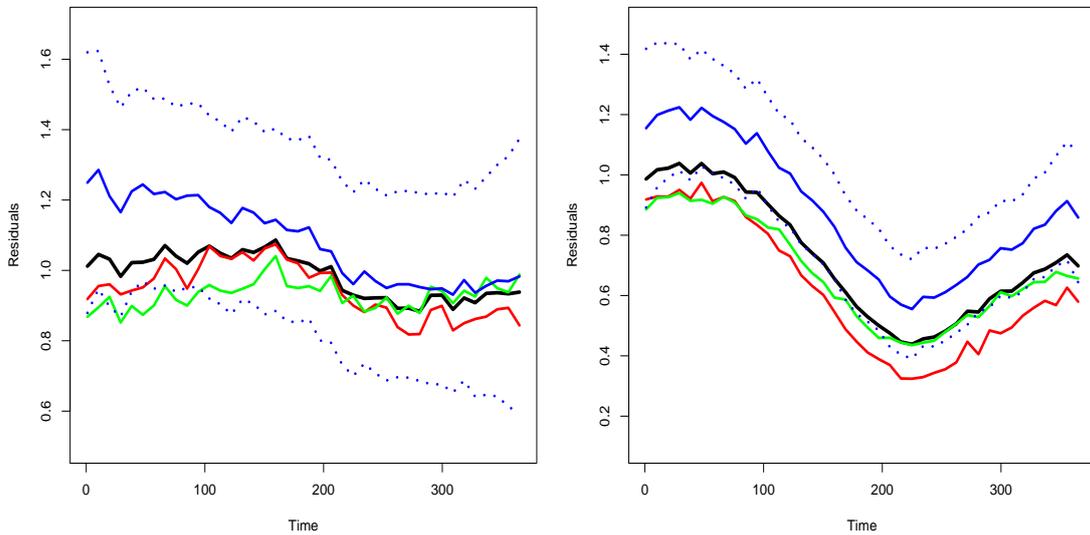


Figure 7: 0.9-expectile curves for Berlin (left) and Taipei (right) daily temperature residuals from 1948-2007 with the 5% – 95% confidence corridors for the latest 20 years expectile

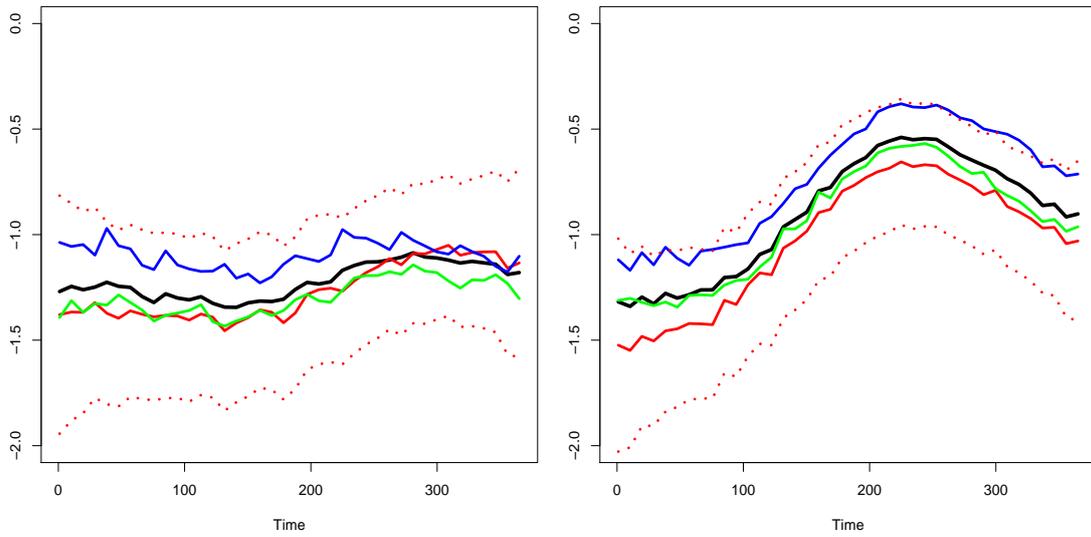


Figure 8: 0.01-expectile curves for Berlin (left) and Taipei (right) daily temperature residuals from 1948-2007 with the 5% – 95% confidence corridors for the first 20 years expectile

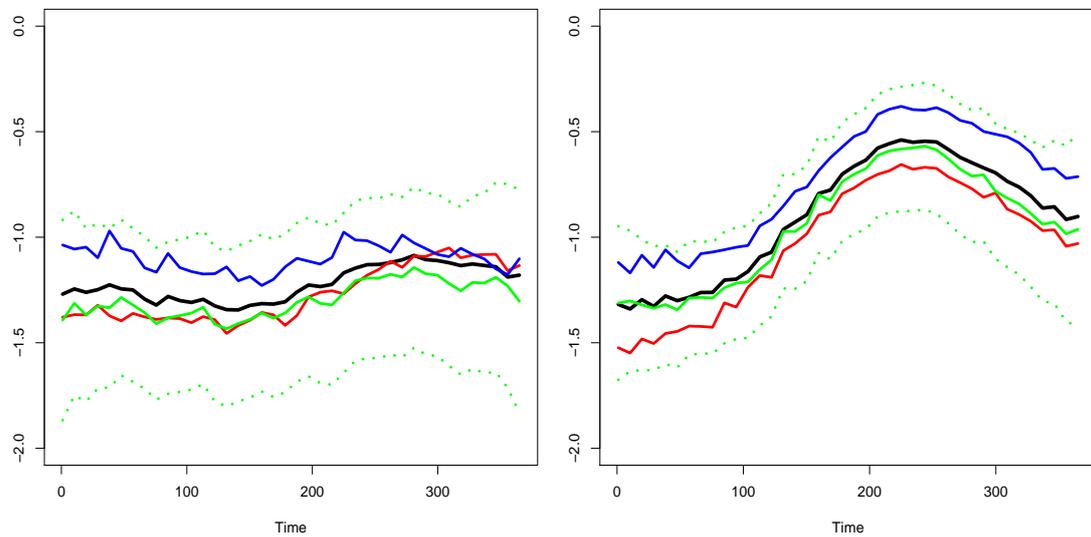


Figure 9: 0.01-expectile curves for Berlin (left) and Taipei (right) daily temperature residuals from 1948-2007 with the 5% – 95% confidence corridors for the second 20 years expectile

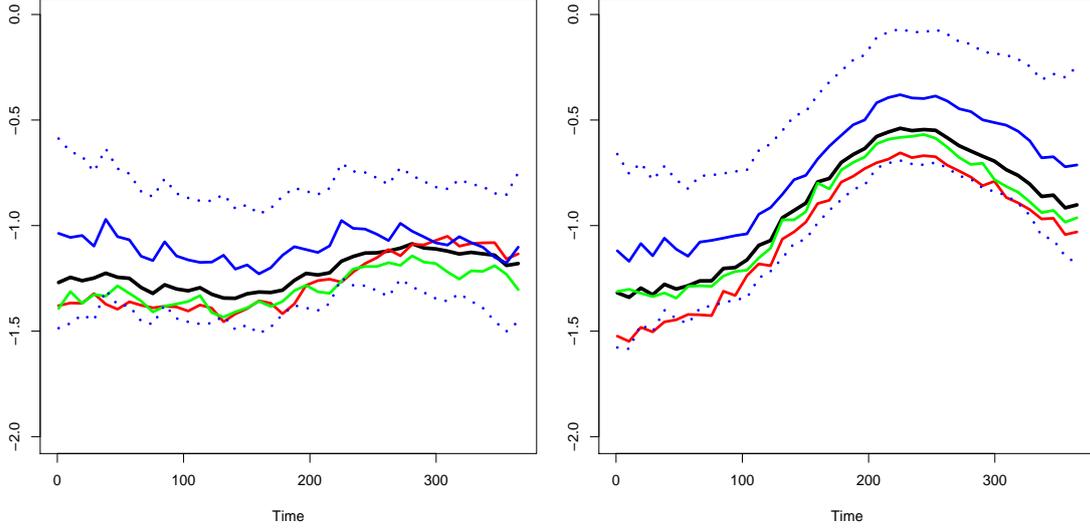


Figure 10: 0.01-expectile curves for Berlin (left) and Taipei (right) daily temperature residuals from 1948-2007 with the 5% – 95% confidence corridors for the latest 20 years expectile

5 Appendix

Proof of Theorem 1. By the definition of $v_n(x)$ as a zero of (6), we have, for $\varepsilon > 0$,

$$\text{if } v_n(x) > v(x) + \varepsilon, \text{ and then } H_n\{v(x) + \varepsilon, x\} > 0. \quad (29)$$

Now

$$H_n\{v(x) + \varepsilon, x\} \leq H\{v(x) + \varepsilon, x\} + \sup_{\theta \in I} |H_n(\theta, x) - H(\theta, x)|. \quad (30)$$

Also, by the identity $H\{v(x), x\} = 0$, the function $H\{v(x) + \varepsilon, x\}$ is not positive and has a magnitude $\geq m_1 \tilde{q} \varepsilon$ by assumption (A6) and (13), for $0 < \varepsilon < \delta_1$. That is, for $0 < \varepsilon < \delta_1$,

$$H\{v(x) + \varepsilon, x\} \leq -m_1 \tilde{q} \varepsilon. \quad (31)$$

Combining (29), (30) and (31), we have, for $0 < \varepsilon < \delta_1$:

$$\text{if } v_n(x) > v(x) + \varepsilon, \text{ and then } \sup_{\theta \in I} \sup_{x \in J} |H_n(\theta, x) - H(\theta, x)| > m_1 \tilde{q} \varepsilon.$$

With a similar inequality proved for the case $v_n(x) < v(x) + \varepsilon$, we obtain, for $0 < \varepsilon < \delta_1$:

$$\text{if } \sup_{x \in J} |v_n(x) - v(x)| > \varepsilon, \text{ and then } \sup_{\theta \in I} \sup_{x \in J} |H_n(\theta, x) - H(\theta, x)| > m_1 \tilde{q} \varepsilon. \quad (32)$$

It readily follows that (32), and (12) imply (14). \blacksquare

Below we first show that $\|R_n\|_\infty = \sup_{x \in J} |R_n(x)|$ vanishes asymptotically faster than the rate $(nh \log n)^{-1/2}$; for simplicity we will just use $\|\cdot\|$ to indicate the sup-norm.

Lemma 4 *For the remainder term $R_n(t)$ defined in (16) we have*

$$\|R_n\| = \mathcal{O}_p\{(nh \log n)^{-1/2}\}. \quad (33)$$

PROOF. First we have by the positivity of the kernel K ,

$$\begin{aligned} \|R_n\| &\leq \left[\inf_{0 \leq x \leq 1} \{|D_n(x)| \cdot q(x)\} \right]^{-1} \{ \|H_n\| \cdot \|q - D_n\| + \|D_n\| \cdot \|\mathbf{E} H_n\| \} \\ &\quad + C_1 \cdot \|v_n - l\|^2 \cdot \left\{ \inf_{0 \leq t \leq 1} |D_n(x)| \right\}^{-1} \cdot \|f_n\|_\infty, \end{aligned}$$

where $f_n(x) = (nh)^{-1} \sum_{i=1}^n K\{(x - X_i)/h\}$.

The desired result (4) will then follow if we prove

$$\|H_n\| = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{1/2}\} \quad (34)$$

$$\|q - D_n\| = \mathcal{O}_p\{(nh)^{-1/4}(\log n)^{-1/2}\} \quad (35)$$

$$\|\mathbf{E} H_n\| = \mathcal{O}(h^2) \quad (36)$$

$$\|v_n - v\|^2 = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{-1/2}\} \quad (37)$$

Since (36) follows from the well-known bias calculation

$$\mathbf{E} H_n(x) = h^{-1} \int K\{(x - u)/h\} \mathbf{E}[\psi\{y - v(x)\} | X = u] f_X(u) du = \mathcal{O}(h^2),$$

where $\mathcal{O}(h^2)$ is independent of x in Parzen (1962), we have from assumption (A2) that $\|\mathbf{E} H_n\| = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{-1/2}\}$.

According to Lemma A.3 in Franke and Mwita (2003),

$$\sup_{x \in J} |H_n(x) - \mathbf{E} H_n(x)| = \mathcal{O}\{(nh)^{-1/2}(\log n)^{1/2}\}.$$

and the following inequality

$$\begin{aligned} \|H_n\| &\leq \|H_n - \mathbf{E} H_n\| + \|\mathbf{E} H_n\|. \\ &= \mathcal{O}\{(nh)^{-1/2}(\log n)^{1/2}\} + \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{-1/2}\} \\ &= \mathcal{O}\{(nh)^{-1/2}(\log n)^{1/2}\} \end{aligned}$$

Statement (34) thus is obtained.

Statement (35) follows in the same way as (34) using assumption (A2) and the Lipschitz continuity properties of K , ψ' , l .

According to the uniform consistency of $v_n(x) - v(x)$ shown before, we have

$$\|v_n - v\| = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{1/2}\}$$

which implies (37).

Now the assertion of the lemma follows, since by tightness of $D_n(x)$, $\inf_{0 \leq t \leq 1} |D_n(x)| \geq q_0$ *a.s.* and thus

$$\|R_n\| = \mathcal{O}_p\{(nh \log n)^{-1/2}\}(1 + \|f_n\|).$$

Finally, by Theorem 3.1 of Bickel and Rosenblatt (1973), $\|f_n\| = \mathcal{O}_p(1)$; thus the desired result $\|R_n\| = \mathcal{O}_p\{(nh \log n)^{-1/2}\}$ follows. ■

We now begin with the subsequent approximations of the processes $Y_{0,n}$ to $Y_{5,n}$.

Lemma 5

$$\|Y_{0,n} - Y_{1,n}\| = \mathcal{O}\{(nh)^{-1/2}(\log n)^2\} \quad a.s.$$

PROOF. Let x be fixed and put $L(y) = \psi\{y - v(x)\}$ still depending on x . Using

integration by parts, we obtain

$$\begin{aligned}
& \iint_{\Gamma_n} L(y)K\{(x-t)/h\}dZ_n(t,y) \\
&= \int_{u=-A}^A \int_{y=-a_n}^{a_n} L(y)K(u)dZ_n(x-h \cdot u,y) \\
&= - \int_{-A}^A \int_{-a_n}^{a_n} Z_n(x-h \cdot u,y)d\{L(y)K(u)\} \\
&\quad +L(a_n)(a_n) \int_{-A}^A Z_n(x-h \cdot u,a_n)dK(u) \\
&\quad -L(-a_n)(-a_n) \int_{-A}^A Z_n(x-h \cdot u,-a_n)dK(u) \\
&\quad +K(A)\left\{ \int_{-a_n}^{a_n} Z_n(x-h \cdot A,y)dL(y) \right. \\
&\quad \left. +L(a_n)(a_n)Z_n(x-h \cdot A,a_n) - L(-a_n)(-a_n)Z_n(x-h \cdot A,-a_n) \right\} \\
&\quad -K(-A)\left\{ \int_{-a_n}^{a_n} Z_n(x+h \cdot A,y)dL(y) + L(a_n)(a_n)Z_n(x+h \cdot A,a_n) \right. \\
&\quad \left. -L(-a_n)(-a_n)Z_n(x+h \cdot A,-a_n) \right\}.
\end{aligned}$$

If we apply the same operation to $Y_{1,n}$ with $B_n\{T(x,y)\}$ instead of $Z_n(x,y)$ and use Lemma 2, we finally obtain

$$\sup_{0 \leq x \leq 1} h^{1/2}g(x)^{1/2}|Y_{0,n}(x) - Y_{1,n}(x)| = \mathcal{O}\{n^{-1/2}(\log n)^2\} \quad a.s..$$

■

Lemma 6 $\|Y_{1,n} - Y_{2,n}\| = \mathcal{O}_p(h^{1/2})$.

PROOF. Note that the Jacobian of $T(x,y)$ is $f(x,y)$. Hence

$$\begin{aligned}
& Y_{1,n}(x) - Y_{2,n}(x) \\
&= \left| \{g(x)h\}^{-1/2} \iint_{\Gamma_n} \psi\{y-v(x)\}K\{(x-t)/h\}f(t,y)dt dy \right| \cdot |W_n(1,1)|.
\end{aligned}$$

It follows that

$$\begin{aligned}
h^{-1/2}\|Y_{1,n} - Y_{2,n}\| &\leq |W_n(1,1)| \cdot \|g^{-1/2}\| \\
&\quad \cdot \sup_{0 \leq t \leq 1} h^{-1} \iint_{\Gamma_n} |\psi\{y-v(x)\}K\{(x-t)/h\}|f(t,y)dt dy.
\end{aligned}$$

Since $\|g^{-1/2}\|$ is bounded by assumption, we have

$$h^{-1/2}\|Y_{1,n} - Y_{2,n}\| \leq |W_n(1,1)| \cdot C_4 \cdot h^{-1} \int K\{(x-t)/h\}dx = \mathcal{O}_p(1).$$

■

Lemma 7 $\|Y_{2,n} - Y_{3,n}\| = \mathcal{O}_p(h^{1/2})$.

PROOF. The difference $|Y_{2,n}(x) - Y_{3,n}(x)|$ may be written as

$$\left| \{g(x)h\}^{-1/2} \iint_{\Gamma_n} [\psi\{y - v(x)\} - \psi\{y - v(t)\}] K\{(x-t)/h\} dW_n\{T(t,y)\} \right|.$$

If we use the fact that l is uniformly continuous, this is smaller than

$$h^{-1/2}|g(x)|^{-1/2} \cdot \mathcal{O}_p(h)$$

and the lemma thus follows. ■

Lemma 8 $\|Y_{4,n} - Y_{5,n}\| = \mathcal{O}_p(h^{1/2})$.

PROOF.

$$\begin{aligned} |Y_{4,n}(x) - Y_{5,n}(x)| &= h^{-1/2} \left| \int \left[\left\{ \frac{g(t)}{g(x)} \right\}^{1/2} - 1 \right] K\{(x-t)/h\} dW(x) \right| \\ &\leq h^{-1/2} \left| \int_{-A}^A W(x-hu) \frac{\partial}{\partial u} \left[\left\{ \frac{g(x-hu)}{g(x)} \right\}^{1/2} - 1 \right] K(u) du \right| \\ &\quad + h^{-1/2} \left| K(A)W(t-hA) \left[\left\{ \frac{g(x-Ah)}{g(x)} \right\}^{1/2} - 1 \right] \right| \\ &\quad + h^{-1/2} \left| K(-A)W(x+hA) \left[\left\{ \frac{g(x+Ah)}{g(x)} \right\}^{1/2} - 1 \right] \right| \\ &= S_{1,n}(x) + S_{2,n}(x) + S_{3,n}(x), \quad \text{say.} \end{aligned}$$

The second term can be estimated by

$$h^{-1/2}\|S_{2,n}\| \leq K(A) \cdot \sup_{0 \leq x \leq 1} |W(x-Ah)| \cdot \sup_{0 \leq x \leq 1} h^{-1} \left| \left[\left\{ \frac{g(x-Ah)}{g(x)} \right\}^{1/2} - 1 \right] \right|;$$

by the mean value theorem it follows that

$$h^{-1/2}\|S_{2,n}\| = \mathcal{O}_p(1).$$

The first term $S_{1,n}$ is estimated as

$$\begin{aligned} h^{-1/2}S_{1,n}(x) &= \left| h^{-1} \int_{-A}^A W(x-uh)K'(u) \left[\left\{ \frac{g(x-uh)}{g(x)} \right\}^{1/2} - 1 \right] du \right. \\ &\quad \left. - \frac{1}{2} \int_{-A}^A W(x-uh)K(u) \left\{ \frac{g(x-uh)}{g(x)} \right\}^{1/2} \left\{ \frac{g'(x-uh)}{g(x)} \right\} du \right| \\ &= |T_{1,n}(x) - T_{2,n}(x)|, \quad \text{say}; \end{aligned}$$

$\|T_{2,n}\| \leq C_5 \cdot \int_{-A}^A |W(t-hu)| du = \mathcal{O}_p(1)$ by assumption on $g(x) = \sigma^2(x) \cdot f_X(x)$. To estimate $T_{1,n}$ we again use the mean value theorem to conclude that

$$\sup_{0 \leq x \leq 1} h^{-1} \left| \left\{ \frac{g(x-uh)}{g(x)} \right\}^{1/2} - 1 \right| < C_6 \cdot |u|;$$

hence

$$\|T_{1,n}\| \leq C_6 \cdot \sup_{0 \leq x \leq 1} \int_{-A}^A |W(x-hu)| K'(u) u du = \mathcal{O}_p(1).$$

Since $S_{3,n}(x)$ is estimated as $S_{2,n}(x)$, we finally obtain the desired result. \blacksquare

The next lemma shows that the truncation introduced through $\{a_n\}$ does not affect the limiting distribution.

Lemma 9 $\|Y_n - Y_{0,n}\| = \mathcal{O}_p\{(\log n)^{-1/2}\}$.

PROOF. We shall only show that $g'(x)^{-1/2}h^{-1/2} \iint_{\mathbb{R}-\Gamma_n} \psi\{y-v(x)\}K\{(x-t)/h\}dZ_n(t,y)$ fulfills the lemma. The replacement of $g'(x)$ by $g(x)$ may be proved as in Lemma A.4 of Johnston (1982). The quantity above is less than $h^{-1/2}\|g^{-1/2}\| \cdot \|\iint_{\{|y|>a_n\}} \psi\{y-v(\cdot)\}K\{(\cdot-t)/h\}dZ(t,y)\|$. It remains to be shown that the last factor tends to zero at a rate $\mathcal{O}_p\{(\log n)^{-1/2}\}$. We show first that

$$\begin{aligned} V_n(x) &= (\log n)^{1/2}h^{-1/2} \iint_{\{|y|>a_n\}} \psi\{y-v(x)\}K\{(x-t)/h\}dZ_n(t,y) \\ &\xrightarrow{p} 0 \quad \text{for all } x \end{aligned}$$

and then we show tightness of $V_n(x)$, the result then follows:

$$\begin{aligned} V_n(x) &= (\log n)^{1/2}(nh)^{-1/2} \sum_{i=1}^n [\psi\{Y_i - v(x)\} \mathbf{I}(|Y_i| > a_n) K\{(x - X_i)/h\} \\ &\quad - \mathbf{E} \psi\{Y_i - v(x)\} \mathbf{I}(|Y_i| > a_n) K\{(x - X_i)/h\}] \\ &= \sum_{i=1}^n X_{n,x}(x), \end{aligned}$$

where $\{X_{n,x}(x)\}_{i=1}^n$ are i.i.d. for each n with $\mathbb{E} X_{n,x}(x) = 0$ for all $x \in [0, 1]$. We then have

$$\begin{aligned} \mathbb{E} X_{n,x}^2(x) &\leq (\log n)(nh)^{-1} \mathbb{E} \psi^2\{Y_i - v(x)\} \mathbf{I}(|Y_i| > a_n) K^2\{(x - X_i)/h\} \\ &\leq \sup_{-A \leq u \leq A} K^2(u) \cdot (\log n)(nh)^{-1} \mathbb{E} \psi^2\{Y_i - v(x)\} \mathbf{I}(|Y_i| > a_n); \end{aligned}$$

hence

$$\begin{aligned} \text{Var}\{V_n(x)\} &= \mathbb{E} \left\{ \sum_{i=1}^n X_{n,x}(x) \right\}^2 = n \cdot \mathbb{E} X_{n,x}^2(x) \\ &\leq \sup_{-A \leq u \leq A} K^2(u) h^{-1} (\log n) \int_{\{|y| > a_n\}} f_y(y) dy \cdot M_\psi. \end{aligned}$$

where M_ψ denotes an upper bound for ψ^2 . This term tends to zero by assumption (A3). Thus by Markov's inequality we conclude that

$$V_n(x) \xrightarrow{p} 0 \quad \text{for all } x \in [0, 1].$$

To prove tightness of $\{V_n(x)\}$ we refer again to the following moment condition as stated in Lemma 4:

$$\begin{aligned} \mathbb{E}\{|V_n(x) - V_n(x_1)| \cdot |V_n(x_2) - V_n(x)|\} &\leq C' \cdot (x_2 - x_1)^2 \\ C' \text{ denoting a constant, } &x \in [x_1, x_2]. \end{aligned}$$

We again estimate the left-hand side by Schwarz's inequality and estimate each factor separately,

$$\begin{aligned} \mathbb{E}\{V_n(x) - V_n(x_1)\}^2 &= (\log n)(nh)^{-1} \mathbb{E} \left[\sum_{i=1}^n \Psi_n(x, x_1, X_i, Y_i) \cdot \mathbf{I}(|Y_i| > a_n) \right. \\ &\quad \left. - \mathbb{E}\{\Psi_n(x, x_1, X_i, Y_i) \cdot \mathbf{1}(|Y_i| > a_n)\} \right]^2, \end{aligned}$$

where $\Psi_n(x, x_1, X_i, Y_i) = \psi\{Y_i - v(x)\} K\{(x - X_i)/h\} - \psi\{Y_i - v(x_1)\} K\{(x_1 - X_i)/h\}$. Since ψ, K are Lipschitz continuous except at one point and the expectation is taken afterwards, it follows that

$$\begin{aligned} &[\mathbb{E}\{V_n(x) - V_n(x_1)\}^2]^{1/2} \\ &\leq C_7 \cdot (\log n)^{1/2} h^{-3/2} |x - x_1| \cdot \left\{ \int_{\{|y| > a_n\}} f_y(y) dy \right\}^{1/2}. \end{aligned}$$

If we apply the same estimation to $V_n(x_2) - V_n(x_1)$ we finally have

$$\begin{aligned} &\mathbb{E}\{|V_n(x) - V_n(x_1)| \cdot |V_n(x_2) - V_n(x_1)|\} \\ &\leq C_7^2 (\log n) h^{-3} |x - x_1| |x_2 - x_1| \times \int_{\{|y| > a_n\}} f_y(y) dy \\ &\leq C' \cdot |x_2 - x_1|^2 \text{ since } x \in [x_1, x_2] \text{ by (A3)}. \end{aligned}$$

■

Lemma 10 Let $\lambda(K) = \int K^2(u)du$ and let $\{d_n\}$ be as in the theorem. Then

$$(2\delta \log n)^{1/2} [\|Y_{3,n}\|/\{\lambda(K)\}^{1/2} - d_n]$$

has the same asymptotic distribution as

$$(2\delta \log n)^{1/2} [\|Y_{4,n}\|/\{\lambda(K)\}^{1/2} - d_n].$$

PROOF. $Y_{3,n}(x)$ is a Gaussian process with

$$\mathbb{E} Y_{3,n}(x) = 0$$

and covariance function

$$\begin{aligned} r_3(x_1, x_2) &= \mathbb{E} Y_{3,n}(x_1)Y_{3,n}(x_2) \\ &= \{g(x_1)g(x_2)\}^{-1/2}h^{-1} \iint_{\Gamma_n} \psi^2\{y - v(x)\}K\{(x_1 - x)/h\} \\ &\quad \times K\{(x_2 - x)/h\}f(t, y)dtdy \\ &= \{g(x_1)g(x_2)\}^{-1/2}h^{-1} \iint_{\Gamma_n} \psi^2\{y - v(x)\}f(y|x)dyK\{(x_1 - x)/h\} \\ &\quad \times K\{(x_2 - x)/h\}f_X(x)dx \\ &= \{g(x_1)g(x_2)\}^{-1/2}h^{-1} \int g(x)K\{(x_1 - x)/h\}K\{(x_2 - x)/h\}dx \\ &= r_4(x_1, x_2) \end{aligned}$$

where $r_4(x_1, x_2)$ is the covariance function of the Gaussian process $Y_{4,n}(x)$, which proves the lemma. ■

References

- Benth, F., Benth, J., and Koekebakker, S. (2007). Putting a price on temperature. *Scandinavian Journal of Statistics*, 34(4):746–767.
- Bickel, P. and Rosenblatt, M. (1973). On some global measures of the deviation of density function estimators. *Annals of Statistics*, 1:1071–1095.
- Breckling, J. and Chambers, R. (1988). m -quantiles. *Biometrika*, 74(4):761–772.

- Campbell, S. and Diebold, F. (2005). Weather forecasting for weather derivatives. *Journal of the American Statistical Association*, 100:6–16.
- Csörgö, S. and Hall, P. (1982). Upper and lower classes for triangular arrays. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 61:207–222.
- Diebold, F. and Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics*, 105:131–159.
- Efron, B. (1991). Regression percentiles using asymmetric squared loss. *Statistica Sinica*, 1:93–125.
- Franke, J. and Mwita, P. (2003). Nonparametric estimates for conditional quantiles of time series. *Report in Wirtschaftsmathematik 87, University of Kaiserslautern*.
- 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*, 16:1428–1429.
- Härdle, W. and Luckhaus, S. (1984). Uniform consistency of a class of regression function estimators. *Annals of Statistics*, 12:612–623.
- Huber, P. (1981). *Robust Statistics*. Wiley, New York.
- Johnston, G. (1982). Probabilities of maximal deviations of nonparametric regression function estimates. *Journal of Multivariate Analysis*, 12:402–414.
- Jones, M. (1994). Expectiles and m -quantiles are quantiles. *Statistics Probability Letters*, 20:149–153.
- Kuan, C. M., Yeh, Y. H., and Hsu, Y. C. (2009). Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics*, 150:261–270.
- Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, 55:819–847.
- Parzen, M. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 32:1065–1076.

- Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 23:470–472.
- Schnabel, S. and Eilers, P. (2009a). An analysis of life expectancy and economic production using expectile frontier zones. *Demographic Research*, 21:109–134.
- Schnabel, S. and Eilers, P. (2009b). Optimal expectile smoothing. *Computational Statistics and Data Analysis*, 53:4168–4177.
- Taylor, J. (2008). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics*, 6:231–252.
- Tusnady, G. (1977). A remark on the approximation of the sample distribution function in the multidimensional case. *Periodica Mathematica Hungarica*, 8:53–55.
- Yao, Q. and Tong, H. (1996). Asymmetric least squares regression estimation: a nonparametric approach. *Journal of Nonparametric Statistics*, 6 (2-3):273–292.
- Zhang, B. (1994). Nonparametric regression expectiles. *Nonparametric Statistics*, 3:255–275.

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Localising temperature risk" by Wolfgang Karl Härdle, Brenda López Cabrera, Ostap Okhrin and Weining Wang, January 2011.
- 002 "A Confidence Corridor for Sparse Longitudinal Data Curves" by Shuzhuan Zheng, Lijian Yang and Wolfgang Karl Härdle, January 2011.
- 003 "Mean Volatility Regressions" by Lu Lin, Feng Li, Lixing Zhu and Wolfgang Karl Härdle, January 2011.
- 004 "A Confidence Corridor for Expectile Functions" by Esra Akdeniz Duran, Mengmeng Guo and Wolfgang Karl Härdle, January 2011.

SFB 649, Ziegelstraße 13a, D-10117 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

