

Expectile asymptotics

Hajo Holzmann and Bernhard Klar

Fachbereich Mathematik und Informatik

Philipps-Universität Marburg

and

Institut für Stochastik

Karlsruher Institut für Technologie (KIT)

e-mail: holzmann@mathematik.uni-marburg.de; bernhard.klar@kit.edu

Abstract: We discuss in detail the asymptotic distribution of sample expectiles. First, we show uniform consistency under the assumption of a finite mean. In case of a finite second moment, we show that for expectiles other than the mean, only the additional assumption of continuity of the distribution function at the expectile implies asymptotic normality, otherwise, the limit is non-normal. For a continuous distribution function we show the uniform central limit theorem for the expectile process. If, in contrast, the distribution is heavy-tailed, and contained in the domain of attraction of a stable law with $1 < \alpha < 2$, then we show that the expectile is also asymptotically stable distributed. Our findings are illustrated in a simulation section.

Keywords and phrases: M-estimator, expectiles, convergence to stable distributions, asymptotic normality, uniform central limit theorem.

Received April 2016.

1. Introduction

Expectile regression, that is, regression on a parameter that generalizes the mean and characterizes the tail behaviour of a distribution, has been introduced by Newey and Powell (1987) as an alternative to more standard quantile regression; Breckling and Chambers (1988) considered regression based on more general asymmetric M-estimators. For a recent comparison between quantile and expectile regression and references see Schulze-Waltrup et al. (2014).

Let Y be a random variable with distribution function F and finite mean $E|Y| < \infty$. For a fixed $\tau \in (0, 1)$, the τ -expectile $\mu_\tau = \mu_\tau(F)$ of Y has been introduced by Newey and Powell (1987) as the minimizer of an asymmetric quadratic loss

$$\begin{aligned} \mu_\tau(F) &= \arg \min_{x \in \mathbb{R}} E S_\tau(x, Y), \\ S_\tau(x, y) &= \tau/2 [((y-x)^+)^2 - (y^+)^2] + (1-\tau)/2 [((y-x)^-)^2 - (y^-)^2]. \end{aligned} \tag{1}$$

Apparently, for $\tau = 1/2$ one obtains the mean. Alternatively to $S_\tau(x, y)$ in (1), one may use other scoring functions for the expectile; these were recently characterized by Gneiting (2011, Theorem 10).

Compared to quantiles, expectiles require the existence of a first moment and hence lack robustness. On the other hand, for any distribution with finite mean, the expectile is unique for each τ , and the expectile curve is always strictly increasing and continuous.

More importantly, as a risk measure it has been shown recently that expectiles have the attractive property of coherence (see Bellini et al. (2014)), while quantiles suffer from the lack of subadditivity. Indeed, expectiles were shown to be the only coherent, elicitable risk measures in Ziegel (2014); for a discussion and comparison between value at risk (quantiles), expectiles and expected shortfall see Emmer et al. (2015). Further discussion and application of expectiles as risk measures are given in Delbaen (2013) and Bellini and Di Bernardino (2015).

In this note we study in detail the statistical, that is, asymptotic properties of the sample expectiles. Somewhat surprisingly and in contrast to the mean, for $\tau \neq 1/2$ we find that even under the assumption of a finite second moment, the sample expectile is only asymptotically normal if the distribution function F is continuous at $\mu_\tau(F)$, otherwise, the limit distribution is non-normal.

First, in Section 2.1 we show uniform consistency under the assumption of a finite mean. Next, in Section 2.2 we show that if the distribution function F is continuous at its τ -expectile $\mu_\tau(F)$, there is an asymptotic linearization of the sample expectile for this τ . In case of finite second moments, this implies asymptotic normality, but if F is in the domain of attraction of a stable law, the sample expectile is also asymptotically stable distributed. If F has a jump at $\mu_\tau(F)$, we show in Section 2.3 that also under the assumption of a finite second moment, the asymptotic distribution of the sample expectile is non-normal. Finally, for a continuous distribution function with second moments, we show the uniform central limit theorem for the expectile process. We illustrate our findings in a simulation in Section 3, using the t -distribution with low degrees of freedom as a prototypical example for heavy-tailed distributions. Based on an explicit representation of the expectile for discrete distributions, we exemplify the nonstandard asymptotic behavior of the empirical expectile by a three-point distribution. Proofs are deferred to Section 4.

In a recent paper, Krättschmer and Zähle (2016) obtained results on the asymptotics of expectiles which are to some extent complementary to our results. Using a non-standard version of the functional delta-method allows them to treat both the case of dependent data as well as expectiles of parametric estimates of the distribution. However, they only consider the case of a finite second moment (they even assume slightly more) and a distribution which is continuous at the expectiles, and further do not investigate properties of the expectile process.

2. Asymptotic properties of sample expectiles

Newey and Powell (1987) state a number of useful properties of expectiles, mainly for absolutely continuous distributions F . Below we state an extension,

and in particular point out the assumptions on F which are actually required. Introducing the identification function

$$I_\tau(x, y) = \tau(y - x)1_{\{y \geq x\}} - (1 - \tau)(x - y)1_{\{y < x\}}$$

of the expectile, it is well-known that $\mu_\tau(F)$ can equivalently be defined as unique solution of the first-order condition

$$EI_\tau(x, Y) = 0, \quad x \in \mathbb{R}. \tag{2}$$

The following identity, obtained by a partial integration, is important for us:

$$I_\tau(x, F) := EI_\tau(x, Y) = \tau \int_x^\infty (1 - F(y)) dy - (1 - \tau) \int_{-\infty}^x F(y) dy. \tag{3}$$

Proposition 1. *Let F be a distribution function with finite mean.*

- (i) *For each $\tau \in (0, 1)$ there is a unique solution $\mu_\tau(F)$ to (1) or, equivalently, to (2).*
- (ii) *The function $\mu_\tau(F) : (0, 1) \rightarrow \mathbb{R}$, $\tau \mapsto \mu_\tau(F)$, is continuous, strictly increasing, and has range $\{y \in \mathbb{R} : 0 < F(y) < 1\}$.*
- (iii) *If F is continuous in a neighborhood of $\mu_\tau(F)$ for a given $\tau \in (0, 1)$, then $\mu_\tau(F)$ is continuously differentiable in a neighborhood of τ with derivative*

$$\partial_\tau \mu_\tau(F) = \frac{\int_{\mu_\tau}^\infty (1 - F(y)) dy + \int_{-\infty}^{\mu_\tau} F(y) dy}{\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau)}.$$

2.1. Sample expectiles and uniform consistency

In this section we show strong uniform consistency of sample expectiles. Let Y have distribution function F , with finite first moment $E_F|Y| = E|Y| < \infty$, and let Y_1, Y_2, \dots be i.i.d. copies of Y , and let \hat{F}_n be the empirical distribution function. The empirical τ -expectile

$$\hat{\mu}_{\tau,n} = \mu_\tau(\hat{F}_n)$$

can be defined as solution of the equation

$$I_\tau(x, \hat{F}_n) = \frac{1}{n} \sum_{k=1}^n I_\tau(x, Y_k) = 0. \tag{4}$$

This type of estimator is often termed Z-estimator, and a large amount of theory is available to obtain asymptotic properties for this type of estimators. Alternatively, asymptotic results can be derived using the representation as an M-estimator, that is,

$$\hat{\mu}_{\tau,n} = \operatorname{argmin}_{x \in \mathbb{R}} \hat{S}_n(x), \quad \hat{S}_n(x) = \frac{1}{n} \sum_{k=1}^n S_\tau(x, Y_k) = \int_{\mathbb{R}} S_\tau(x, y) d\hat{F}_n(y). \tag{5}$$

Here, any other scoring function for the expectile (Gneiting, 2011) could be used instead, they all result in the same estimator, the expectile of the empirical distribution function.

The measurability of $\hat{\mu}_{\tau,n}$ follows from Theorem (1.9) in Pfanzagl (1969), who studied M-estimators under the heading of minimum contrast estimation. More directly, measurability follows from the explicit representation of $\hat{\mu}_{\tau,n}$ in Subsection 3.2.

Theorem 2. *Let Y, Y_1, Y_2, \dots be i.i.d. with distribution function F , and assume $E_F|Y| < \infty$. For any $\tau_l, \tau_u \in (0, 1)$, $\tau_l < \tau_u$, we have*

$$\sup_{\tau_l \leq \tau \leq \tau_u} |\hat{\mu}_{\tau,n} - \mu_\tau(F)| \rightarrow 0 \quad a.s.$$

2.2. Asymptotic linearization and convergence to stable distributions

Let us consider the representation (5) of the sample expectile as an M-estimator. Asymptotic normality or, more generally, asymptotic linearization, requires that the asymptotic contrast function has a second order Taylor expansion at the true parameter. Since $|\partial_x S_\tau(x, y)| = |I_\tau(x, y)| \leq c(|x| + |y|)$ for a suitable constant c , we may differentiate the *asymptotic contrast function*

$$\psi_\tau(x) = ES_\tau(x, Y) = \int S_\tau(x, y) dF(y) \quad (6)$$

under the integral sign to obtain

$$\psi'_\tau(x) = -EI_\tau(x, Y) =: -I_\tau(x, F).$$

We see from (3) that $\psi'_\tau(x)$ has

$$\begin{aligned} \text{right derivative } \psi''^+(x) &= \tau(1 - F(x)) + (1 - \tau)F(x) \\ \text{left derivative } \psi''^-(x) &= \tau(1 - F(x-)) + (1 - \tau)F(x-) \end{aligned} \quad (7)$$

at x , where $F(x-) = P(Y < x)$ is the left limit of F at x . For $\tau = 1/2$ (i.e. the mean), these are always equal, but generally only coincide at $\mu_\tau(F)$ if F has no point mass in its τ -expectile. From Theorems 1 and 10 in Arcones (2000) we deduce the following linearization.

Theorem 3 (Asymptotic linearization). *Let Y, Y_1, Y_2, \dots be i.i.d. with distribution function F . Assume that $E_F|Y| < \infty$ and that F is continuous at $\mu_\tau = \mu_\tau(F)$ for a given $\tau \in (0, 1)$. Let $\{a_n\}$ be a sequence of positive numbers which converges to infinity with $\sup_{n \geq 1} n^{-1}a_n^2 < \infty$, such that*

$$\frac{a_n}{n} \sum_{k=1}^n I_\tau(\mu_\tau, Y_k) = O_P(1). \quad (8)$$

Then

$$a_n (\hat{\mu}_{\tau,n} - \mu_\tau) = \frac{a_n}{n} (\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau))^{-1} \sum_{k=1}^n I_\tau(\mu_\tau, Y_k) + o_P(1).$$

Asymptotic normality

In case of finite second moments, (8) is satisfied with $a_n = \sqrt{n}$ by the central limit theorem, and we obtain asymptotic normality for a finite number of expectiles. In the following, we write $Y_n \xrightarrow{\mathcal{L}} F$ as a short-hand notation for $Y_n \xrightarrow{\mathcal{L}} Y \sim F$, where F denotes the distribution function of Y .

Corollary 4. *Suppose that $EY^2 < \infty$. Let $\tau_i \in (0, 1)$, $i = 1, \dots, m$ be such that F does not have a point mass at any of the μ_{τ_i} , $i = 1, \dots, m$. Then*

$$\sqrt{n}(\hat{\mu}_{\tau_1, n} - \mu_{\tau_1}, \dots, \hat{\mu}_{\tau_m, n} - \mu_{\tau_m})' \xrightarrow{\mathcal{L}} N(\mathbf{0}, \Sigma),$$

where

$$\Sigma_{i,j} = \frac{E [I_{\tau_i}(\mu_{\tau_i}, Y) I_{\tau_j}(\mu_{\tau_j}, Y)]}{(\tau_i(1 - F(\mu_{\tau_i})) + (1 - \tau_i)F(\mu_{\tau_i})) (\tau_j(1 - F(\mu_{\tau_j})) + (1 - \tau_j)F(\mu_{\tau_j}))} \tag{9}$$

for $i, j = 1, \dots, m$.

Convergence to stable distributions

A random variable X has an α -stable distribution if its characteristic function is given by

$$E [e^{iuX}] = \begin{cases} \exp(-|u|^\alpha [1 - i\beta \tan(\frac{\pi\alpha}{2}) \text{sign}(u)]), & \alpha \neq 1, \\ \exp(-|u| [1 + i\beta \frac{2}{\pi} \text{sign}(u) \log |u|]), & \alpha = 1, \end{cases}$$

where $0 < \alpha \leq 2$, $\beta \in [-1, 1]$. Assume that Y belongs to the domain of attraction of an α -stable distribution ($Y \in DA(\alpha)$) with $0 < \alpha < 2$ (see, e.g., Embrechts et al. (1997, Def. 2.2.7)). This is the case if and only if Y has tail probabilities that satisfy

$$P(Y > y) = \frac{c^+ + o(1)}{y^\alpha} L(y) \quad \text{and} \quad P(Y < -y) = \frac{c^- + o(1)}{y^\alpha} L(y), \quad y \rightarrow \infty, \tag{10}$$

where L is slowly varying and $c^+, c^- \geq 0$ with $c^+ + c^- > 0$ (Embrechts et al., 1997, Th. 2.2.8). In the following, we assume $1 < \alpha < 2$ to ensure that $E|Y| < \infty$.

Corollary 5. *Let Y, Y_1, Y_2, \dots be i.i.d. r.v. with distribution function $F \in DA(\alpha)$, where $1 < \alpha < 2$. Assume further that F has no point mass in μ_τ . Then,*

$$\frac{n^{1-1/\alpha}}{L_1(n)} (\hat{\mu}_{\tau, n} - \mu_\tau(F)) \xrightarrow{\mathcal{L}} \frac{\tilde{Z}}{\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau)}.$$

Here, \tilde{Z} follows an α -stable distribution, and L_1 is an appropriate slowly varying function.

Proof. Since $F \in DA(\alpha)$, from (10) we obtain that

$$\begin{aligned} P(I(\mu_\tau, Y) > y) &= \tau^\alpha \frac{c^+ + o(1)}{y^\alpha} L(y) \quad \text{and} \\ P(I(\mu_\tau, Y) < -y) &= (1 - \tau)^\alpha \frac{c^- + o(1)}{y^\alpha} L(y) \quad \text{as } y \rightarrow \infty. \end{aligned}$$

Consequently, $I(\mu_\tau, Y) \in DA(\alpha)$, and the general CLT (Embrechts et al., 1997, Th. 2.2.15) yields

$$\left(n^{1/\alpha} L_1(n) \right)^{-1} \left(\sum_{k=1}^n I(\mu_\tau, Y_k) - nEI(\mu_\tau, Y) \right) \xrightarrow{\mathcal{L}} \tilde{Z} \quad \text{as } n \rightarrow \infty,$$

where \tilde{Z} follows an α -stable distribution and L_1 is an appropriate slowly varying function. This implies that (8) is satisfied, and an application of Theorem 3 together with the general CLT yields the statement of the corollary. \square

Instead of using the assumptions $Y \in DA(\alpha)$, suppose more specifically that Y belongs to the *domain of normal attraction* of some α -stable distribution with $1 < \alpha < 2$, i.e. Y has tail probabilities that satisfy

$$y^\alpha P(Y > y) \rightarrow c^+ \quad \text{and} \quad y^\alpha P(Y < -y) \rightarrow c^-, \quad y \rightarrow \infty, \quad (11)$$

with $c^+ + c^- > 0$ and $1 < \alpha < 2$.

Corollary 6. *Let Y, Y_1, Y_2, \dots be i.i.d. r.v. with distribution function F that belongs to the normal domain of attraction of an α -stable distribution, where $1 < \alpha < 2$, that is, satisfies (11). Assume further that F has no point mass in μ_τ . Then*

$$n^{1-1/\alpha} \tilde{c} (\hat{\mu}_{\tau,n} - \mu_\tau) \xrightarrow{\mathcal{L}} \frac{S(\alpha, \tilde{\beta})}{\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau)},$$

where

$$\tilde{c} = \left(\frac{2\Gamma(\alpha) \sin(\pi\alpha/2)}{\pi(\tau^\alpha c^+ + (1 - \tau)^\alpha c^-)} \right)^{1/\alpha}, \quad \tilde{\beta} = \frac{\tau^\alpha c^+ - (1 - \tau)^\alpha c^-}{\tau^\alpha c^+ + (1 - \tau)^\alpha c^-}.$$

Proof. Since

$$y^\alpha P(I(\mu_\tau, Y) > y) \rightarrow \tau^\alpha c^+ \quad \text{and} \quad y^\alpha P(I(\mu_\tau, Y) < -y) \rightarrow (1 - \tau)^\alpha c^-$$

as $y \rightarrow \infty$, this follows from the general CLT for distributions in the normal domain of attraction of a corresponding stable law (Nolan, 2015, p. 22). \square

2.3. Further asymptotics under finite second moments

Suppose that $Y \sim F$ with $EY^2 < \infty$ and $\text{Var} Y > 0$. In contrast to the mean, asymptotic normality of general expectiles as in Corollary 4 actually requires the additional assumption that Y has no point mass at $\mu_\tau(F)$, otherwise, the limit distribution is non-normal, as the following result shows.

Theorem 7. *Let Y, Y_1, Y_2, \dots be i.i.d. with distribution function F with $EY^2 < \infty$. Let $\tau \in (0, 1)$ and denote $\mu_\tau = \mu_\tau(F)$. Then*

$$\sqrt{n}(\hat{\mu}_{\tau,n} - \mu_\tau) \xrightarrow{\mathcal{L}} \sigma_1 W 1_{W>0} + \sigma_2 W 1_{W<0},$$

where $W \sim N(0, E[I_\tau(\mu_\tau, Y)^2])$,

$$\begin{aligned} \sigma_1 &= [\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau)]^{-1}, \\ \sigma_2 &= [\tau(1 - F(\mu_\tau -)) + (1 - \tau)F(\mu_\tau -)]^{-1}, \end{aligned}$$

and $F(x-) = P(Y < x)$ denotes the left limit of F at x .

We prove Theorem 7 by using empirical process methods and the argmax continuity theorem as presented in Van der Vaart (1998). Alternatively one could exploit the convexity of the contrast and modify the assumptions and the proof in Hjort and Pollard (1993, Theorem 2.1) to give an alternative argument.

In case of a continuous distribution function, we also have convergence of the expectile process.

Theorem 8. *Let Y, Y_1, Y_2, \dots be i.i.d. with distribution function F with $EY^2 < \infty$. Let $0 < \tau_l < \tau_u < 1$ and suppose that F is continuous in a neighborhood of $[\mu_{\tau_l}, \mu_{\tau_u}]$. Then the sequence of processes*

$$\tau \mapsto (\sqrt{n}(\hat{\mu}_{\tau,n} - \mu_\tau))_{n \geq 1}, \quad \tau \in [\tau_l, \tau_u], \quad (12)$$

converges weakly in $C[\tau_l, \tau_u]$ to a Gaussian process with continuous sample paths and covariance function given in (9).

Tran et. al. (2014) also show convergence of the expectile process. They argue via convergence of an associated quantile process, and therefore require that F has a density, further, they do not specify the covariance function of the limit process.

Theorem 7 shows that process convergence, at least in $C[\tau_l, \tau_u]$ or even in $l^\infty[\tau_l, \tau_u]$, cannot be expected if F has a discontinuity in $[\tau_l, \tau_u]$, since in this case the limit process would be discontinuous as well.

3. Some Simulations

3.1. Illustration of convergence to a stable distribution

As an example for a distribution with finite expectation but infinite variance we consider Student's t -distribution t_α , $1 < \alpha < 2$, with symmetric density

$$f_\alpha(x) = \frac{\Gamma((\alpha+1)/2)}{\Gamma(\alpha/2)\sqrt{\alpha\pi}} \left(1 + \frac{x^2}{\alpha}\right)^{-\frac{\alpha+1}{2}}, \quad x \in \mathbb{R}.$$

For $Y_\alpha \sim t_\alpha$,

$$\lim_{y \rightarrow \infty} y^\alpha P(Y_\alpha > y) = \lim_{y \rightarrow \infty} y^\alpha P(Y_\alpha < -y) = \frac{\Gamma((\alpha+1)/2)}{\Gamma(\alpha/2)} \frac{\alpha^{\alpha/2-1}}{\sqrt{\pi}}.$$

Accordingly, t_α belongs to the domain of normal attraction of some α -stable distribution. To compute the theoretical τ -expectile, which is the unique solution of

$$\mu_\tau - EY = \frac{2\tau - 1}{1 - \tau} E[(Y - \mu_\tau)^+], \quad (13)$$

one can use the identity

$$E[(Y_\alpha - \mu_\tau)^+] = \frac{\sqrt{\alpha}\Gamma((\alpha+1)/2)}{\sqrt{\pi}(\alpha-1)\Gamma(\alpha/2)} \left(1 + \frac{\mu_\tau^2}{\alpha}\right)^{\frac{1-\alpha}{2}} - \mu_\tau(1 - F_\alpha(\mu_\tau)),$$

where $F_\alpha(\cdot)$ denotes the distribution function of t_α . The limiting behavior of the empirical τ -expectile then follows directly from Corollary 5. Figure 1 shows the distribution function of $n^{1-1/\alpha} \tilde{c}(\hat{\mu}_{\tau,n} - \mu_\tau)$ (more precisely the empirical distribution function based on 10000 replications) for sample sizes of 20, 200 and 2000 for several values of τ and α . It can be observed that the quality of the approximation by the corresponding limiting stable law depends on both τ and α : the approximation improves for decreasing α (see Figure 1 (a)–(c)) and for τ approaching the value 0.5 (see Figure 1 (d)–(f)).

3.2. Illustration of nonstandard asymptotics under finite second moments

To illustrate the convergence to a non-normal distribution stated in Theorem 7, we first give an explicit formula for the empirical expectile which is interesting in itself. From (13), it follows directly that the τ -expectile satisfies the equivalent conditions

$$\tau = \frac{E[(Y - \mu_\tau)^-]}{E[|Y - \mu_\tau|]}, \quad (14)$$

$$\mu_\tau = \frac{(1 - \tau)E[Y1_{\{Y \leq \mu_\tau\}}] + \tau E[Y1_{\{Y > \mu_\tau\}}]}{(1 - \tau)P(Y \leq \mu_\tau) + \tau P(Y > \mu_\tau)}. \quad (15)$$

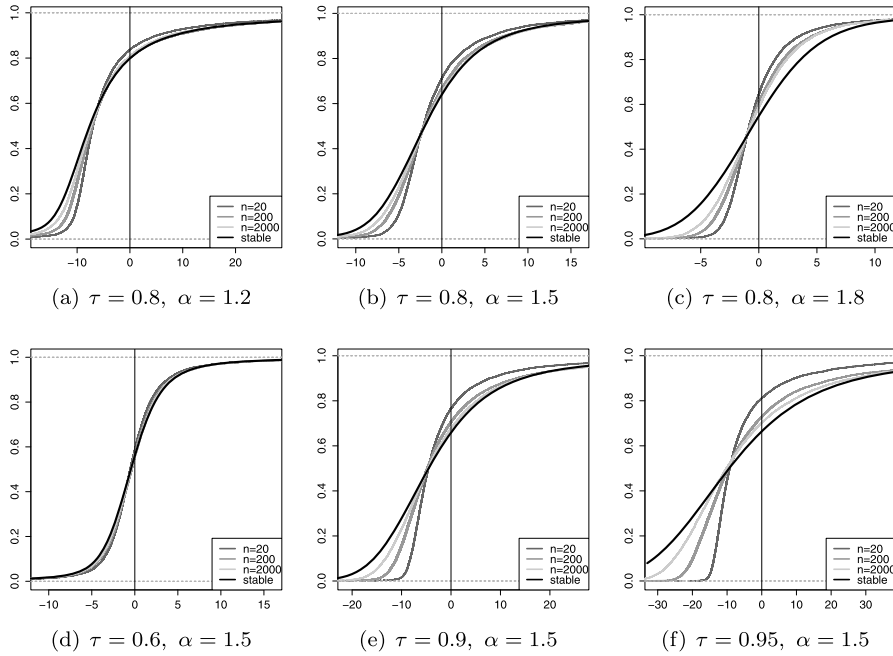


FIG 1. Convergence of the cumulative distribution function (cdf) of the empirical expectile to the corresponding limiting stable cdf. Upper row: Data follow t_α -distribution with different α , $\tau = 0.8$ fixed. Lower row: Data follow t_α -distribution with $\alpha = 1.5$, different values of τ .

The subsequent representation follows Bellini (2012), but formulated for the empirical distribution, and allowing for ties. Let $Y_{(1)} \leq \dots \leq Y_{(n)}$ denote the order statistics of Y_1, \dots, Y_n . From (15), the empirical expectile satisfies

$$\hat{\mu}_{\tau,n} = \frac{(1 - \tau) \sum_k Y_{(k)} 1_{\{Y_{(k)} \leq \hat{\mu}_{\tau,n}\}} + \tau \sum_k Y_{(k)} 1_{\{Y_{(k)} > \hat{\mu}_{\tau,n}\}}}{(1 - \tau) \sum_k 1_{\{Y_{(k)} \leq \hat{\mu}_{\tau,n}\}} + \tau \sum_k 1_{\{Y_{(k)} > \hat{\mu}_{\tau,n}\}}}.$$

Hence, for $\hat{\mu}_{\tau,n} \in [Y_{(i)}, Y_{(i+1)})$, where $Y_{(i)} < Y_{(i+1)}$, one has

$$\hat{\mu}_{\tau,n} = \frac{(1 - \tau) \sum_{k=1}^i Y_{(k)} + \tau \sum_{k=i+1}^n Y_{(k)}}{(1 - \tau)i + \tau(n - i)}. \tag{16}$$

Defining

$$\tau_i^* := \frac{iY_{(i)} - \sum_{k=1}^i Y_{(k)}}{\sum_{k=1}^n |Y_{(k)} - Y_{(i)}|}, \quad i = 1, \dots, n, \tag{17}$$

we have $\hat{\mu}_{\tau,n} = Y_{(i)}$ iff $\tau = \tau_i^*$ for $i = 1, \dots, n$ (and then, (17) is the empirical counterpart of (14)). Note that $\tau_0^* = 0, \tau_n^* = 1$, and since $\hat{\mu}_{\tau,n}$ is nondecreasing in τ , we obtain that $\tau_i^* \leq \tau_{i+1}^*, i = 1, \dots, n - 1$. As a consequence,

$$\hat{\mu}_{\tau,n} \in [Y_{(i)}, Y_{(i+1)}) \Leftrightarrow \tau \in [\tau_i^*, \tau_{i+1}^*), \quad i = 1, \dots, n - 1.$$

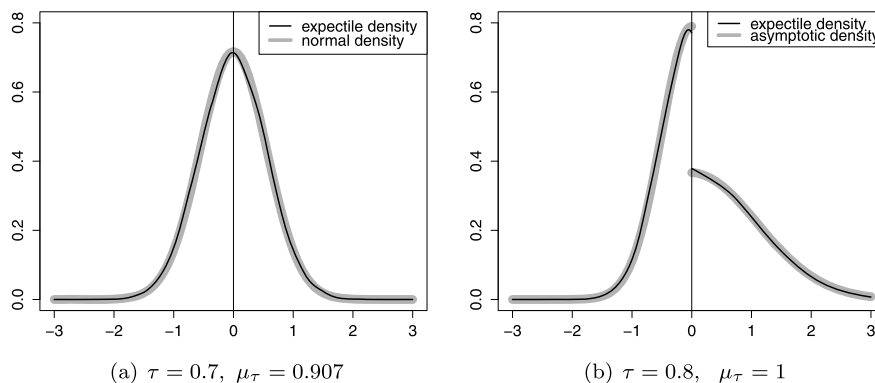


FIG 2. Density function of the standardized empirical expectile for $n = 500$ and of the corresponding limiting distribution. Data follow a three point distribution in $0, 1, 2$. (a) $\tau = 0.7$, normal limiting distribution. (b) $\tau = 0.8$, non-normal limiting distribution.

Remark. 1. Formulas (17) and (16) are especially well-suited for plotting purposes without the need of any numerical root-finding.
2. From (16), $\hat{\mu}_{\tau,n}$ is piecewise differentiable in τ with

$$\frac{d\hat{\mu}_{\tau,n}}{d\tau} = \frac{i \sum_{k=i+1}^n Y(k) - (n-i) \sum_{k=1}^i Y(k)}{((1-\tau)i + \tau(n-i))^2} \quad \text{for } \tau \in (\tau_i^*, \tau_{i+1}^*).$$

If Y has a discrete distribution on $0, 1, 2, \dots$ (say), an analogous reasoning leads to the following explicit formula for the theoretical expectiles μ_τ . Define

$$\tau_i^* := \frac{\sum_{k=0}^{i-1} (i-k)P(Y=k)}{\sum_{k \geq 0} |i-k|P(Y=k)}, \quad i = 0, 1, 2, \dots \quad (18)$$

For $\tau \in [\tau_i^*, \tau_{i+1}^*]$, and accordingly $\mu_\tau \in [i, i+1)$, one has

$$\mu_\tau = \frac{(1-\tau) \sum_{k \leq i} kP(Y=k) + \tau \sum_{k > i} kP(Y=k)}{(1-\tau)P(Y \leq i) + \tau P(Y > i)}. \quad (19)$$

Now, assume that Y follows a three-point distribution with $P(Y=i) = p_i$, $i = 0, 1, 2$, with $p_0, p_1, p_2 > 0$, $p_0 + p_1 + p_2 = 1$. Then, from (18) and (19), we get $\tau_0^* = 0$, $\tau_1^* = p_0/(p_0 + p_2)$, $\tau_2^* = 1$ and

$$\mu_\tau = \begin{cases} \frac{\tau(p_1 + 2p_2)}{(1-\tau)p_0 + \tau(p_1 + p_2)}, & 0 < \tau < \tau_1^*, \\ \frac{(1-\tau)p_1 + 2\tau p_2}{(1-\tau)(p_0 + p_1) + \tau p_2}, & \tau_1^* \leq \tau < 1. \end{cases}$$

Next, we make the choice $p_0 = 4/10$, $p_1 = 5/10$, $p_2 = 1/10$. Then, $\mu_{0.8} = 1$, i.e. the distribution of Y has a point mass in μ_τ for $\tau = 0.8$, but not for other values of τ . Figure 2 (a) shows the density of $\sqrt{n}(\hat{\mu}_{\tau,n} - \mu_\tau)$ (estimated by a nonparametric density estimator based on 20000 replications) for sample size 500

and $\tau = 0.7$ (hence, $\mu_\tau = 49/54$) together with the limiting normal distribution given in Corollary 4. Figure 2 (b) shows the corresponding plot for $\tau = 0.8$ together with the limiting non-normal distribution given in Corollary 7.

4. Proofs

Proof of Proposition 1. Parts (i) and (ii) are from Newey and Powell (1987) except for the general continuity of $\mu_\tau(F)$ in τ . From (3) we see that $I_\tau(x, F)$ is a continuous function of (τ, x) . To show continuity of the expectile, first let $\tau_n \downarrow \tau$, and let $\tilde{\mu}_\tau = \lim_n \mu_{\tau_n}(F)$ for which by monotonicity $\mu_\tau(F) \leq \tilde{\mu}_\tau$. By continuity of $I_\tau(x, F)$ we have

$$0 = \lim_n I_{\tau_n}(\mu_{\tau_n}(F), F) = I_\tau(\tilde{\mu}_\tau, F),$$

but since $\mu_\tau(F)$ is the unique zero, it follows that $\mu_\tau(F) = \tilde{\mu}_\tau$, that is, right-continuity. The argument for left-continuity is the same.

(iii) From (3) we see that if F is continuous in a neighborhood of x , then $I_\tau(\cdot, F)$ is continuously differentiable at x with derivative $-\tau(1 - F(x)) - (1 - \tau)F(x)$. The conclusion follows from the implicit function theorem. \square

Proof of Theorem 2. We start with strong consistency of individual expectiles, that is,

$$\hat{\mu}_{\tau,n} \rightarrow \mu_\tau(F) \quad a.s. \tag{20}$$

We may use the representation (4) of the empirical expectile as a Z-estimator and strengthen Van der Vaart (1998, Lemma 5.10) to almost sure convergence. Since $x \mapsto I_\tau(x, F)$ is strictly decreasing, we have for every $\varepsilon > 0$ that

$$I_\tau(\mu_\tau - \varepsilon, F) > 0 > I_\tau(\mu_\tau + \varepsilon, F).$$

Since $I(\mu_\tau \pm \varepsilon, \hat{F}_n) \rightarrow I_\tau(\mu_\tau \pm \varepsilon, F)$ a.s. as $n \rightarrow \infty$, we have a.s. that $I(\mu_\tau - \varepsilon, \hat{F}_n) > 0 > I(\mu_\tau + \varepsilon, \hat{F}_n)$ for large $n \in \mathbb{N}$. Since each map $x \rightarrow I(x, \hat{F}_n)$, $n \in \mathbb{N}$, is continuous and has exactly one zero $\hat{\mu}_{\tau,n}$, this zero must a.s. lie between $\mu_\tau \pm \varepsilon$ for large $n \in \mathbb{N}$, that is,

$$\limsup_{n \rightarrow \infty} |\hat{\mu}_{\tau,n} - \mu_\tau(F)| \leq \varepsilon \quad a.s. \quad \forall \varepsilon > 0,$$

showing (20).

Using Proposition 1 (ii) and individual consistency, the classical Glivenco-Cantelli argument may be applied. Let $d = \mu_{\tau_u}(F) - \mu_{\tau_l}(F)$, $m \in \mathbb{N}$, and choose by continuity $\tau_l = \tau_0 \leq \tau_1 \leq \dots \leq \tau_m = \tau_u$ such that $\mu_{\tau_k}(F) = \mu_{\tau_l}(F) + kd/m$, $k = 1, \dots, m$. By monotonicity, for $\tau_k \leq \tau \leq \tau_{k+1}$,

$$\hat{\mu}_{\tau,n} - \mu_\tau(F) \leq \hat{\mu}_{\tau_{k+1},n} - \mu_{\tau_{k+1}}(F) + \mu_{\tau_{k+1}}(F) - \mu_{\tau_k}(F).$$

Therefore

$$\sup_{\tau_l \leq \tau \leq \tau_u} (\hat{\mu}_{\tau,n} - \mu_\tau(F)) \leq \max_{0 \leq k \leq m} |\hat{\mu}_{\tau_k,n} - \mu_{\tau_k}(F)| + d/m.$$

Similarly,

$$\sup_{\tau_l \leq \tau \leq \tau_u} (\mu_\tau(F) - \hat{\mu}_{\tau,n}) \leq \max_{0 \leq k \leq m} |\hat{\mu}_{\tau_k,n} - \mu_{\tau_k}(F)| + d/m.$$

Since

$$\sup_{\tau_l \leq \tau \leq \tau_u} |\hat{\mu}_{\tau,n} - \mu_\tau(F)| = \max \left(\sup_{\tau_l \leq \tau \leq \tau_u} (\hat{\mu}_{\tau,n} - \mu_\tau(F)), \sup_{\tau_l \leq \tau \leq \tau_u} (\mu_\tau(F) - \hat{\mu}_{\tau,n}) \right),$$

we have for any $m \in \mathbb{N}$ that

$$\begin{aligned} \limsup_n \sup_{\tau_l \leq \tau \leq \tau_u} |\hat{\mu}_{\tau,n} - \mu_\tau(F)| &\leq \limsup_n \max_{0 \leq k \leq m} |\hat{\mu}_{\tau_k,n} - \mu_{\tau_k}(F)| + d/m \\ &= d/m \quad \text{a.s.} \end{aligned}$$

□

We shall derive Theorem 3 from Theorems 1 and 10 in Arcones (2000). For convenience, we state a version of these results, tailored to our needs.

Theorem (Theorems 1 and 10 in Arcones (2000)). *Let Y, Y_1, Y_2, \dots be i.i.d. with distribution function F . Let $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a function such that $g(\cdot, \vartheta) : \mathbb{R} \rightarrow \mathbb{R}$ is measurable for each $\vartheta \in \mathbb{R}$. Let $\hat{\vartheta}_n$ be a sequence of r.v.'s satisfying*

$$n^{-1} \sum_{k=1}^n g(Y_k, \hat{\vartheta}_n) = \inf_{\vartheta \in \mathbb{R}} n^{-1} \sum_{k=1}^n g(Y_k, \vartheta).$$

Suppose that:

$$(A.1) \quad \hat{\vartheta}_n \xrightarrow{P} \vartheta_0, \quad \vartheta_0 \in \mathbb{R}.$$

(A.2) *There is a positive constant V such that*

$$E[g(Y, \vartheta) - g(Y, \vartheta_0)] = V(\vartheta - \vartheta_0)^2 + o(|\vartheta - \vartheta_0|^2),$$

as $\vartheta \rightarrow \vartheta_0$.

(A.3) *Let $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ and let $\{a_n\}$ be a sequence of positive numbers which converges to infinity with $\sup_{n \geq 1} n^{-1} a_n^2 < \infty$ such that*

$$a_n \left(n^{-1} \sum_{j=1}^n \varphi(Y_j) - E[\varphi(Y)] \right) = O_P(1).$$

(A.4) *There is a function $\zeta : \mathbb{R} \rightarrow \mathbb{R}$ with $E|\zeta(Y)| < \infty$ such that*

$$\lim_{\delta \rightarrow 0} E \left[\sup_{|\vartheta| \leq \delta} \frac{|r(Y, \vartheta) - \vartheta^2 \zeta(Y)|}{\vartheta^2} \right] = 0,$$

where $r(y, \vartheta) = g(y, \vartheta_0 + \vartheta) - g(y, \vartheta_0) - \vartheta \varphi(y)$.

Then,

$$a_n(\hat{\vartheta}_n - \vartheta_0) + \frac{a_n}{2V} \left(\frac{1}{n} \sum_{j=1}^n \varphi(Y_j) - E[\varphi(Y)] \right) \xrightarrow{P} 0.$$

Proof of Theorem 3. We verify the conditions of the above theorem for $\vartheta_0 = \mu_\tau(F)$, $g(y, \vartheta) = S_\tau(\vartheta, y)$, $\varphi(y) = -I_\tau(\mu_\tau(F), y)$ and

$$\zeta(y) = \frac{\tau}{2} 1_{\{y > \mu_\tau\}} + \frac{1 - \tau}{2} 1_{\{y < \mu_\tau\}}.$$

(A1) follows from Theorem 2.

(A2) follows from (6), (7), the assumption of continuity of F at $\mu_\tau(F)$, and Taylor's theorem, which holds under the minimal assumption of an existing second derivative.

(A3) is (8). Finally, for (A4) we compute that for $x > 0$,

$$\begin{aligned} & S_\tau(\mu_\tau + x, y) - S_\tau(\mu_\tau, y) + xI_\tau(\mu_\tau, y) \\ &= -\frac{\tau}{2} (y - \mu_\tau - x)^2 1_{\{\mu_\tau < y \leq \mu_\tau + x\}} + \frac{\tau}{2} x^2 1_{\{y > \mu_\tau\}} \\ & \quad + \frac{1 - \tau}{2} (y - \mu_\tau - x)^2 1_{\{\mu_\tau \leq y < \mu_\tau + x\}} + \frac{1 - \tau}{2} x^2 1_{\{y < \mu_\tau\}} \end{aligned}$$

and similarly for $x < 0$. Therefore for some $c > 0$ we may estimate

$$\begin{aligned} & |S_\tau(\mu_\tau + x, y) - S_\tau(\mu_\tau, y) + xI_\tau(\mu_\tau, y) - x^2\zeta(y)| \\ & \leq c (y - \mu_\tau - x)^2 1_{\{\mu_\tau - |x| \leq y \leq \mu_\tau + |x|\}} \\ & \leq c x^2 1_{\{\mu_\tau - |x| \leq y \leq \mu_\tau + |x|\}}, \end{aligned}$$

and therefore

$$\begin{aligned} & E \left[\sup_{|x| \leq \delta} \frac{|S_\tau(\mu_\tau + x, Y) - S_\tau(\mu_\tau, Y) + xI_\tau(\mu_\tau, Y) - x^2\zeta(Y)|}{x^2} \right] \\ & \leq c P(\mu_\tau - \delta \leq Y \leq \mu_\tau + \delta) \rightarrow 0, \quad \delta \rightarrow 0, \end{aligned}$$

since Y does not have a point mass at μ_τ . □

Proof of Theorem 7. We start by establishing Lipschitz continuity of $S_\tau(x, y)$ as a function of x with square-integrable Lipschitz constant. Since $\partial_x S_\tau(x, y) = -I_\tau(x, y)$, we have for $x_1, x_2 \in B_\delta(\mu_\tau)$

$$|S_\tau(x_1, y) - S_\tau(x_2, y)| \leq c m(y) |x_1 - x_2|, \quad m(y) := \sup_{x \in B_\delta(\mu_\tau)} |I_\tau(x, y)|. \quad (21)$$

Then, the inequality

$$m(y) \leq \sup_{x \in B_\delta(\mu_\tau)} |x - y| \leq \sup_{x \in B_\delta(\mu_\tau)} |x| + |y|$$

yields $E[m(Y)^2] < \infty$ if $EY^2 < \infty$, that is, the Lipschitz constant has finite second moment.

Next, the asymptotic contrast in (6) is continuously differentiable with left and right derivatives in μ_τ given in (7). From Taylor's formula, we obtain

$$\begin{aligned}\psi_\tau(x) - \psi_\tau(\mu_\tau) &= (x - \mu_\tau)^2 \psi_\tau''^+(\mu_\tau)/2 + o(|x - \mu_\tau|^2), & x > \mu_\tau, \\ \psi_\tau(x) - \psi_\tau(\mu_\tau) &= (x - \mu_\tau)^2 \psi_\tau''^-(\mu_\tau)/2 + o(|x - \mu_\tau|^2), & x < \mu_\tau,\end{aligned}\quad (22)$$

where $\psi_\tau''^\pm(\mu_\tau)$ are right/left second derivatives.

Therefore, the assumptions of Theorem 5.52 in Van der Vaart (1998) are satisfied with $\alpha = 2$ and $\beta = 1$ (see the argument in Corollary 5.53, that the Lipschitz property (21) implies the concentration inequality), and we obtain the \sqrt{n} -rate of convergence:

$$\sqrt{n}(\hat{\mu}_{\tau,n} - \mu_\tau) = O_P(1).$$

To obtain the asymptotic distribution, we apply the argmax-continuity theorem, Corollary 5.58 in Van der Vaart (1998). To this end, for a measurable function f with $Ef^2(Y) < \infty$, denote

$$\mathbb{P}_n f = \frac{1}{n} \sum_{k=1}^n f(Y_k), \quad Pf = EF(Y), \quad \text{and} \quad \mathbb{G}_n(f) = \sqrt{n}(\mathbb{P}_n - P)f.$$

By the Lipschitz property (21), from the proof of Lemma 19.31 in Van der Vaart (1998) we obtain for any $M > 0$ that

$$\sup_{|h| \leq M} \mathbb{G}_n[\sqrt{n}(S_\tau(\mu_\tau + h/\sqrt{n}, \cdot) - S_\tau(\mu_\tau, \cdot)) + hI_\tau(\mu_\tau, \cdot)] \xrightarrow{n \rightarrow \infty} 0 \quad (\text{P}).$$

Therefore, for any $M > 0$, the difference between the processes

$$h \mapsto \sqrt{n} \mathbb{P}_n[\sqrt{n}(S_\tau(\mu_\tau + h/\sqrt{n}, \cdot) - S_\tau(\mu_\tau, \cdot))], \quad |h| \leq M,$$

and

$$h \mapsto n[\psi_\tau(\mu_\tau + h/\sqrt{n}) - \psi_\tau(\mu_\tau)] - h \mathbb{G}_n I_\tau(\mu_\tau, \cdot), \quad |h| \leq M,$$

tends to 0 in probability in sup-norm. Using (22), the second process converges to the Gaussian process

$$h \mapsto \frac{1}{2\sigma_1} h^2 1_{h>0} + \frac{1}{2\sigma_2} h^2 1_{h<0} - hW, \quad (23)$$

where W is normally distributed as in the theorem, hence so does the first. From the argmax-continuity theorem, we obtain weak convergence of the minimizers $\sqrt{n}(\hat{\mu}_{\tau,n} - \mu_\tau)$ to the minimizer of the limit process. Now, a parabola $h \mapsto -hW + h^2/(2\sigma)$ for some $\sigma > 0$ is minimized at $h = \sigma W$, yielding the negative value $-\sigma W^2/2$. Therefore, the minimizer of (23) is at $h = \sigma_1 W$ for $W > 0$ and at $h = \sigma_2 W$ for $W < 0$, which gives the statement of the theorem. \square

Proof of Theorem 8. We shall apply Van der Vaart (1995, Theorem 1), which gives asymptotic normality of functional Z-estimators; see also Kosorok (2008, Theorem 13.4), which additionally implies validity of the bootstrap. First, Theorem 2 gives the uniform consistency. Given $\nu \in C[\tau_l, \tau_u] \subset l^\infty[\tau_l, \tau_u]$, the functions $\tau \mapsto I_\tau(\nu(\tau), F)$ and $\tau \mapsto I_\tau(\nu(\tau), \hat{F}_n)$ are also in $C[\tau_l, \tau_u]$, and

$$\tau \mapsto I_\tau(\mu_\tau, F) = 0, \quad \tau \mapsto I_\tau(\hat{\mu}_{\tau,n}, \hat{F}_n) = 0.$$

Next, we check the conditions (2), (3) and (4) in Van der Vaart (1995). Suppose that $\nu \in C[\tau_l, \tau_u]$ is such that F is continuous on the image of ν (this is true by our assumption if $\|\nu - \mu\|_{[\tau_l, \tau_u]}$ is small enough). Then we apply the mean value theorem for each $\tau \in [\tau_l, \tau_u]$ to obtain

$$\begin{aligned} & |I_\tau(\nu(\tau), F) - I_\tau(\mu_\tau, F) + [\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau)](\nu(\tau) - \mu_\tau)| \\ & \leq |F(\xi_\tau) - F(\mu_\tau)| |\nu(\tau) - \mu_\tau|, \end{aligned}$$

where ξ_τ is between $\nu(\tau)$ and μ_τ . Since F is uniformly continuous in a compact neighborhood of $[\tau_l, \tau_u]$, we obtain

$$\begin{aligned} & \sup_{\tau \in [\tau_l, \tau_u]} |I_\tau(\nu(\tau), F) - I_\tau(\mu_\tau, F) + [\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau)](\nu(\tau) - \mu_\tau)| \\ & = o(\|\nu - \mu\|_{[\tau_l, \tau_u]}), \end{aligned} \tag{24}$$

showing Fréchet differentiability, that is, (4) in Van der Vaart (1995). Note that the derivative, multiplication with the function

$$\tau \mapsto -[\tau(1 - F(\mu_\tau)) + (1 - \tau)F(\mu_\tau)],$$

is continuously invertible.

Since by Proposition 1, (iii), $\mu_\tau(F)$ is continuously differentiable in τ , we have for an appropriate constant $c > 0$ that

$$|I_{\tau_1}(\mu_{\tau_1}(F), y) - I_{\tau_2}(\mu_{\tau_2}(F), y)| \leq c|y| |\tau_1 - \tau_2|, \quad \tau_1, \tau_2 \in [\tau_l, \tau_u], \quad y \in \mathbb{R}, \tag{25}$$

so that

$$(I_\tau(\mu_\tau(F), \cdot))_{\tau \in [\tau_l, \tau_u]} \tag{26}$$

is a Donsker class of functions, see Van der Vaart (1998, Example 19.7.), taking care of (2) in Van der Vaart (1995).

Finally, to show (3) in Van der Vaart (1995), we choose $\delta_n \downarrow 0$, and estimate in the first step

$$\begin{aligned} & \sup_{\|\nu - \mu\|_{[\tau_l, \tau_u]} \leq \delta_n} \sup_{\tau \in [\tau_l, \tau_u]} \sqrt{n} |I_\tau(\nu(\tau), \hat{F}_n) - I_\tau(\nu(\tau), F) \\ & \quad - [I_\tau(\mu_\tau, \hat{F}_n) - I_\tau(\mu_\tau, F)]| \\ & \leq \sup_{|x| \leq \delta_n} \sup_{\tau \in [\tau_l, \tau_u]} \sqrt{n} |I_\tau(\mu_\tau + x, \hat{F}_n) - I_\tau(\mu_\tau + x, F) \\ & \quad - [I_\tau(\mu_\tau, \hat{F}_n) - I_\tau(\mu_\tau, F)]|. \end{aligned} \tag{27}$$

Now, for a constant $C > 0$, $y \in \mathbb{R}$, $\tau, \tau_1 \in [\tau_l, \tau_u]$, $|x|, |x_1| \leq 1$,

$$\begin{aligned} & |I_{\tau_1}(\mu_{\tau_1} + x_1, y) - I_{\tau_1}(\mu_{\tau_1}, y) - [I_{\tau}(\mu_{\tau} + x, y) - I_{\tau}(\mu_{\tau}, y)]| \\ & \leq C|y|(|x_1 - x| + |\tau_1 - \tau|), \end{aligned}$$

and

$$|I_{\tau}(\mu_{\tau} + x, y) - I_{\tau}(\mu_{\tau}, y)| \leq C|y||x|.$$

Therefore,

$$\sup_{|x| \leq \delta_n} \sup_{\tau \in [\tau_l, \tau_u]} E(I_{\tau}(\mu_{\tau} + x, Y) - I_{\tau}(\mu_{\tau}, Y))^2 \leq C EY^2 \delta_n,$$

and each

$$\mathcal{F}_n = \left\{ (x, \tau) \mapsto I_{\tau}(\mu_{\tau} + x, y) - I_{\tau}(\mu_{\tau}, y), \quad |x| \leq \delta_n, \tau \in [\tau_l, \tau_u] \right\}$$

is a Lipschitz-class of functions. Therefore we may estimate (27) by the bracketing integral $J_{[]}(\delta_n, \mathcal{F}_n, L_2(F))$ and an additional sequence converging to zero, by Van der Vaart (1998, Lemma 19.34 and Example 19.7), which together $\rightarrow 0$ as $n \rightarrow \infty$ and $\delta_n \downarrow 0$.

Next, we show that weak convergence is actually in $C[\tau_l, \tau_u]$. The expectile processes (12) have continuous sample paths. As for the limiting Gaussian process, it suffices to show continuity of the sample paths of the limit Gaussian process of the empirical process corresponding to the function class (26), since the inverse of the Fréchet derivative in (24) is simply multiplication by a fixed continuous function. By Van der Vaart (1998, Lemma 18.15), the limit process can be constructed to have continuous sample paths w.r.t. its standard deviation semimetric. In order to check that continuity also holds w.r.t. the ordinary distance on $[\tau_l, \tau_u]$, we show that

$$E(I_{\tau_2}(\mu_{\tau_2}, Y) - I_{\tau_1}(\mu_{\tau_1}, Y))^2 \leq C^2(\tau_2 - \tau_1)^2, \quad \tau_j \in [\mu_{\tau_l}, \mu_{\tau_u}], \quad j = 1, 2 \quad (28)$$

for some $C > 0$. But this follows immediately from (25) upon squaring and integrating. This concludes the proof of the theorem. \square

References

- ARCONES, M.A., 2000. *M-Estimators Converging to a Stable Limit*. Journal of Multivariate Analysis 74, 193–221. [MR1794001](#)
- BELLINI, F., 2012. *Isotonicity properties of generalized quantiles*. Statistics and Probability Letters 82, 2017–2024. [MR2970307](#)
- BELLINI, F., KLAR, B., MÜLLER, A., ROSAZZA GIANIN, E., 2014. *Generalized quantiles as risk measures*. Insurance: Mathematics and Economics 54, 41–48. [MR3145849](#)
- BELLINI, F., DI BERNARDINO, E., 2015. *Risk management with expectiles*. The European Journal of Finance, DOI: [10.1080/1351847X.2015.1052150](#)

- BRECKLING, J., CHAMBERS, R., 1988. *M-Quantiles*. *Biometrika* 75, 761–771. [MR0995118](#)
- DELBAEN, F. 2013. *A remark on the structure of expectiles*. Preprint, [arXiv:1307.5881v1](#)
- EMBRECHTS, P., KLÜPPELBERG, C., and MIKOSCH, T., 1997. *Modelling Extremal Events for Insurance and Finance*. Springer. [MR1458613](#)
- EMMER, S., KRATZ, M., TASCHE, D., 2015. *What is the best risk measure in practice? A comparison of standard measures*. *Journal of Risk* 18, 31–60.
- GNEITING, T., 2011. *Making and evaluating point forecasts*. *Journal of the American Statistical Association*, 106, 746–762. [MR2847988](#)
- HJORT, N.L. and POLLARD, D., 1993. *Asymptotics for minimizers of convex processes*. Working paper, [arXiv:1107.3806](#) [math.ST]
- KOSOROK, M.R., 2008. *Introduction to Empirical Processes and Semiparametric Inference*. Springer, New York. [MR2724368](#)
- KRÄTSCHEMER, V., ZÄHLE, H., 2016. *Statistical inference for expectile-based risk measures*. Preprint, [arXiv:1601.05261](#)
- NEWBY, W., POWELL, J., 1987. *Asymmetric least squares estimation and testing*. *Econometrica* 55, 819–847. [MR0906565](#)
- NOLAN, J.P., 2015. *Stable Distributions – Models for Heavy Tailed Data*. Birkhauser. Chapter 1 online at <http://academic2.american.edu/~jpnolan>.
- PFANZAGL, J., 1969. *On the Measurability and Consistency of Minimum Contrast Estimates*. *Metrika*, 14, 249–272.
- SCHULZE-WALTRUP, L., SOBOTKA, F., KNEIB, T., KAUEMANN, G., (2014) *Expectile and quantile regression - David and Goliath?* *Statistical Modeling*, DOI: [10.1177/1471082X14561155](#)
- TRAN, N.M., OSIPENKO, M., HAERDLE, W.K., 2014. *Principal Component Analysis in an Asymmetric Norm*. Preprint, [arXiv:1401.3229](#)
- VAN DER VAART, A.W., 1998. *Asymptotic Statistics*. Cambridge University Press. [MR1652247](#)
- VAN DER VAART, A.W., 1995. *Efficiency of infinite dimensional M-estimators*. *Statistica Neerlandica* 49, 9–30. [MR1333176](#)
- ZIEGEL, J., 2014. *Coherence and elicibility*. *Mathematical Finance*, DOI: [10.1111/mafi.12080](#)