Risk Management with Expectiles

Fabio Bellini* Elena Di Bernardino†

21st December 2014

Abstract

Expectiles (EVaR) are a one-parameter family of coherent risk measures that have been recently suggested as an alternative to quantiles (VaR) and to Expected Shortfall (ES). In this work we review their known properties, we discuss their financial meaning, we compare them with VaR and ES and we study their asymptotic behaviour, refining some of the results in Bellini et al. (2014). Moreover, we present a numerical example of computation of expectiles by means of simple Garch(1,1) models and we assess the accuracy of the forecasts by means of a consistent loss function, as suggested by Gneiting (2011). Theoretical and numerical results indicate that expectiles are perfectly reasonable alternatives to VaR and ES.

1 Introduction

It is well known that the left and right quantiles x_α^- and x_α^+ of a random variable X can be defined through the minimization of an asymmetric, piecewise linear loss function:

$$[x_\alpha^-(X), x_\alpha^+(X)] = \underset{x \in \mathbb{R}}{\operatorname{argmin}}\, \alpha\, \mathbb{E}[(X-x)_+] + (1-\alpha)\, \mathbb{E}[(X-x)_-], \text{ for } \alpha \in (0,1),$$

where $x_+ = \max(x, 0)$ and $x_- = \max(-x, 0)$; see for example Koenker (2005). The expectiles e_q have been introduced by Newey and Powell (1987) as the minimizers of an asymmetric quadratic loss:

$$e_q(X) = \underset{x \in \mathbb{R}}{\operatorname{argmin}}\, q\, \mathbb{E}[(X-x)_+^2] + (1-q)\, \mathbb{E}[(X-x)_-^2], \text{ for } q \in (0,1). \tag{1}$$

When $q = \frac{1}{2}$, it is well known that $e_q(X) = \mathbb{E}[X]$, thus expectiles can be seen as an asymmetric generalization of the mean. The term "expectiles" has been probably suggested by the union of "expectation" and "quantiles". The expectiles are uniquely identified by the first order condition

$$q\, \mathbb{E}[(X-e_q)_+] = (1-q)\, \mathbb{E}[(X-e_q)_-]. \tag{2}$$

*Dipartimento di Statistica e Metodi Quantitativi, Università di Milano-Bicocca (email: fabio.bellini@unimib.it)

†CNAM, Paris, Département IMATH, Laboratoire Cédric EA4629 (email:elena.di_bernardino@cnam.fr)

Electronic copy available at: http://ssrn.com/abstract=2475106

Since Equation (5) is well defined for each $X \in L^1$, which is the natural domain of definition of the expectiles, we take it as the definition of e_q. Letting

$$\ell_q(x) := qx_+ - (1-q)x_-,$$

we see that Equation (5) can be rewritten as

$$\mathbb{E}[\ell_q(X - e_q)] = 0.$$

Hence expectiles are an example of *shortfall risk measures* in the sense of Föllmer and Schied (2002), also known as *zero utility premia* in the actuarial literature. From this point of view, they had been considered in Weber (2006) and by Ben Tal and Teboulle (2007), although the connection with the minimization problem (1) and with the statistical notion of expectile emerged only in the more recent literature. In general, a statistical functional that can be defined as the minimizer of a suitable expected loss function as in (1) is said to be *elicitable*; we refer to Gneiting (2011), Bellini and Bignozzi (2013), Ziegel (2014), Embrechts et al. (2014), Davis (2013), Acerbi and Szekely (2014) for further informations about the elicitability property and the debate about its financial relevance. See also the discussion in Section 4 on the relationship between elicitability and backtesting. In this paper we compare expectiles with the more common financial risk measures, that are Value at Risk (VaR_α) and Expected Shortfall (ES_α). We define

$$\begin{aligned} VaR_\alpha(X) &= -x_\alpha^+(X), \\ ES_\alpha(X) &= -\frac{1}{\alpha}\int_0^\alpha x_u^+(X)du. \end{aligned}$$

To be consistent with these sign conventions, following Kuan et al. (2009), we define the expectile-VaR ($EVaR_q$) as follows:

$$EVaR_q(X) = -e_q(X).$$

$EVaR_q$ is the financial risk measure associated to the expectiles, in the same way as VaR_α is the financial risk measure associated to the quantiles. For $q \le \frac{1}{2}$, $EVaR_q$ is a coherent risk measure, since it satisfies the well known axioms introduced by Artzner et al. (1999). Indeed, it is easy to see that

- $EVaR_q(X+h) = EVaR_q(X) - h$ (translation invariance)
- $X \le Y \Rightarrow EVaR_q(X) \ge EVaR_q(X)$ (monotonicity)
- $EVaR_q(\lambda X) = \lambda EVaR_q(X)$ (positive homogeneity)
- $EVaR_q(X+Y) \le EVaR_q(X) + EVaR_q(Y)$ (subadditivity).

Moreover, it has been shown in several papers, albeit starting from different angles, that $EVaR_q$ with $q \le \frac{1}{2}$ is the only coherent risk measure that is also elicitable (see Weber (2006), Ben Tal and Teboulle (2007), Ziegel (2014), Bellini et al. (2014), Bellini and Bignozzi (2013) and Delbaen et al. (2014)). We refer the interested reader to these works and to Delbaen (2012) and Delbaen (2013) for the properties of $EVaR_q$ as a coherent risk measure, in particular for its

Electronic copy available at: http://ssrn.com/abstract=2475106

dual representation, Kusuoka representation and for the identification of the optimal scenario in its dual representation. In order to better understand the financial meaning of $EVaR_q$, it is interesting to compare its acceptance set with VaR_α and with ES_α. Recall that the acceptance set of a translation invariant risk measure ρ is defined as

$$\mathcal{A}_\rho = \{X \mid \rho(X) \leq 0\},$$

and that ρ can be recovered by $\mathcal{A}_\rho$ by the formula

$$\rho(X) = \inf\{m \in \mathbb{R} \mid X + m \in \mathcal{A}_\rho\}.$$

We refer to Delbaen (2012), Föllmer and Schied (2011) or Pflug and Romisch (2007) for textbook treatments. In the case of VaR_α,

$$\mathcal{A}_{VaR_\alpha} = \{X \mid P(X < 0) \leq \alpha\};$$

notice that we can equivalently write

$$\mathcal{A}_{VaR_\alpha} = \left\{X \;\middle|\; \frac{P(X \geq 0)}{P(X < 0)} \geq \frac{1-\alpha}{\alpha}\right\}. \tag{3}$$

In the case of ES_α, we have

$$\mathcal{A}_{ES_\alpha} = \left\{X \;\middle|\; \frac{1}{\alpha}\int_0^\alpha x_u(X)du \geq 0\right\}.$$

In the case of $EVaR_q$, the acceptance set can be written as

$$\mathcal{A}_{EVaR_q} = \left\{X \;\middle|\; \frac{\mathbb{E}[X_+]}{\mathbb{E}[X_-]} \geq \frac{1-q}{q}\right\} \tag{4}$$

(see for example Delbaen (2013)). The $EVaR_q$ is then the amount of money that should be added to a position in order to have a prespecified, sufficiently high *gain-loss ratio*. We recall that the gain-loss ratio or Ω-ratio is a popular performance measure in portfolio management (see i.e. Shadwick and Keating (2002)) and is also well known in literature on no good deal valuation in incomplete markets (see i.e. Biagini and Pinar (2013) and the references therein). It is sometimes argued that $EVaR_q$ is "difficult to explain" to the financial community, but this is probably due to the fact (1) is usually taken as starting definition instead of (4), which has a transparent financial meaning: in the case of VaR_α, a position is acceptable if the ratio of the *probability* of a gain with respect to the probability of a loss is sufficiently high (3); in the case of $EVaR_q$, a position is acceptable if the ratio between the *expected value* of the gain and the expected value of the loss is sufficiently high (4). In Section 4, we provide a numerical example of computation of expectiles by means of a normal i.i.d. model, an historical method, a Garch(1,1) model with normal innovations and a Garch(1,1) model with Student t innovations. Choosing $q = 0.00145$, the magnitude of $VaR_{0.01}$, $ES_{0.025}$ and $EVaR_{0.00145}$ are closely comparable. In conclusion, we believe that $EVaR_q$ is a perfectly reasonable risk measure, displaying many similarities with VaR_α and ES_α, surely worth of deeper study and practical experimentations by risk managers, regulators and portfolio managers.

Electronic copy available at: https://ssrn.com/abstract=2475106

The paper is structured as follows: in Section 2 we review the basic properties of $EVaR_q$, we discuss the comparison between $EVaR_q$ and VaR_α and the asymptotic behaviour of e_q for $q \to 1$. Similar results have been independently found by Mao et al. (2014). In Section 3 we provide several examples. In Section 4 we compute EVaR, compare the results with VaR and with ES and assess the accuracy of the forecasts by means of the realized losses. Proofs and auxiliary results are postponed to the Appendix.

2 Properties of Expectiles

As mentioned in Section 1, we take as definition of expectiles the following equation, valid for each $X \in L^1$:

$$q\,\mathbb{E}[(X-e_q)_+] = (1-q)\,\mathbb{E}[(X-e_q)_-], \tag{5}$$

which can also be written as

$$q = \frac{\mathbb{E}[(X-e_q)_-]}{\mathbb{E}[|X-e_q|]},$$

which shows that the expectiles e_q can be seen as the quantiles of a transformed distribution with distribution function

$$G(x) := \frac{\mathbb{E}[(X-e_q)_-]}{\mathbb{E}[|X-e_q|]},$$

as noted by Jones (1994). We collect in the following Proposition further properties of expectiles (see i.e. Newey and Powell (1987) and Bellini et al. (2014)) :

Proposition 2.1. *Let $X \in L^1$ and let e_q be the unique solution of (5). Then*

a) e_q is strictly monotonic in q, for $q \in (0,1)$

b) e_q is strictly monotonic in X, in the sense that

$$X \geq Y \, a.s. \text{ and } P(X > Y) > 0 \Rightarrow e_q(X) > e_q(Y)$$

c) $e_q(-X) = -e_{1-q}(X)$

d) if X is symmetric with respect to x_0, then

$$\frac{e_q(X) + e_{1-q}(X)}{2} = x_0$$

e) if X has a C^1 density, then e_q is a C^1 function of q, with

$$\frac{de_q}{dq} = \frac{\mathbb{E}[|X-e_q|]}{(1-q)F(e_q) + q\overline{F}(e_q)}.$$

More refined symmetry properties have been considered in Abdous and Remillard (1995).

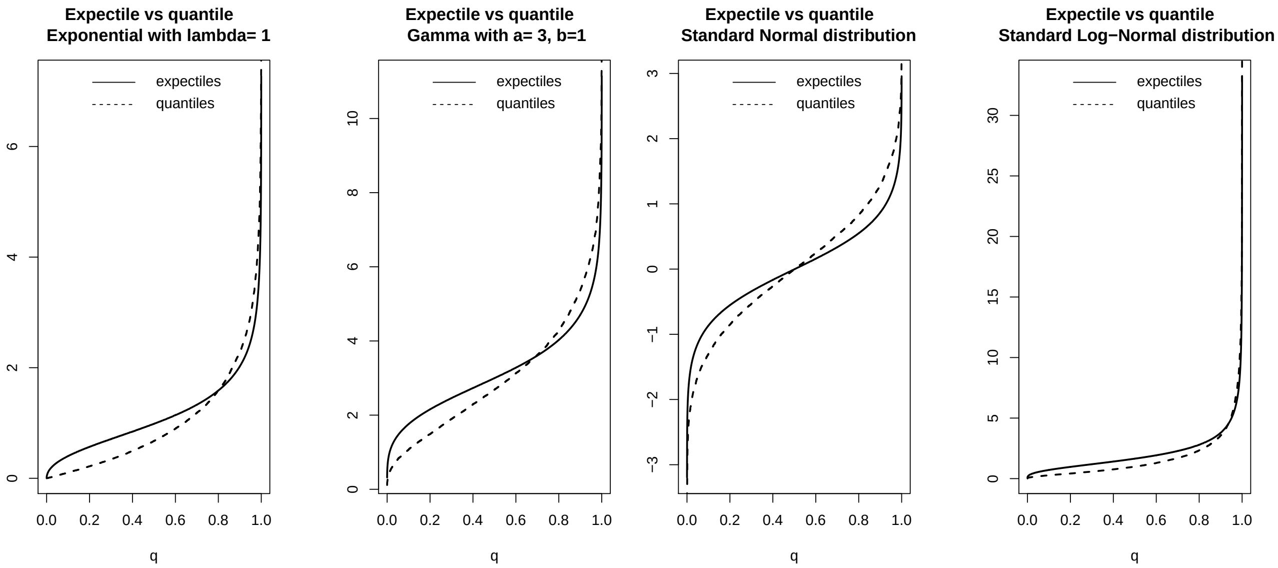

Figure 1: Comparison between expectiles (full line) and quantiles (dashed line) in the Exponential ($\lambda = 1$), Gamma ($a = 3$, $b = 1$), Normal ($\mu = 0, \sigma = 1$) and in the Lognormal ($\mu = 0, \sigma = 1$) cases.

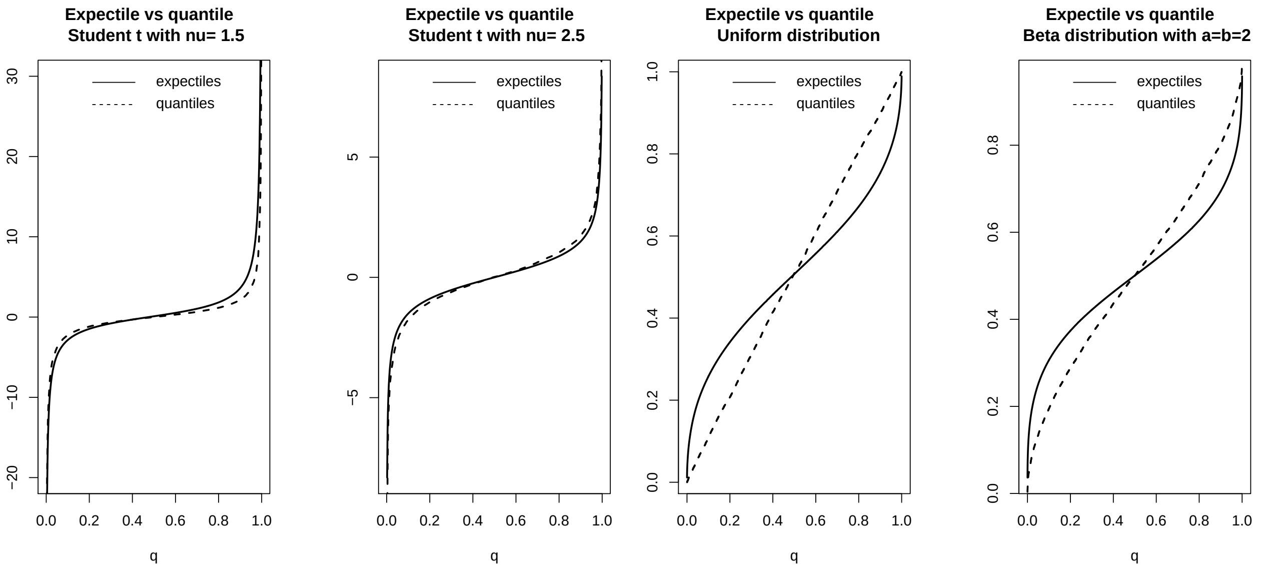

Figure 2: Comparison between expectiles (full line) and quantiles (dashed line) in the Student t with $\nu = 1.5$, $\nu = 2.5$, Uniform and Beta ($a = 2$ and $b = 2$) cases.

2.1 Comparison between expectiles and quantiles

For the most common distributions, expectiles are typically closer to the center of the distribution than the corresponding quantiles. Moreover, the quantile and the expectile curve intersect in a unique point, which corresponds to the center of simmetry in the case of a symmetric distribution (see Examples 3.1, 3.2, 3.3 3.4 3.5 and Figures 1-2).

Koenker (1993) found that for a distribution with quantile function

$$x_\alpha = \frac{2\alpha - 1}{\sqrt{\alpha(1-\alpha)}},$$

which is a rescaled Student t distribution with $\nu = 2$, it holds that $e_\alpha(X) = x_\alpha(X)$ for each $\alpha \in (0,1)$, that is the quantile and the expectile curve coincide. In Zou (2014) the Koenker's argument was generalized to show that for every nondecreasing function $q\colon [0,1] \to [0,1]$ with $q(0) = 0$, $q(1) = 1$ and $q(\alpha_0) = \frac{1}{2}$, the distribution with quantile function

$$x_\alpha = -S(\alpha) \exp\left\{ \int_{\alpha_0}^{\alpha} S(t)\, dt \right\},$$

with

$$S(\alpha) = \frac{2q(\alpha) - 1}{2\alpha q(\alpha) - \alpha - q(\alpha)},$$

satisfies $e_{q(\alpha)}(X) = x_\alpha(X)$. Hence by a proper choice of the function $q(\alpha)$ any arbitrary number of intersections between the quantile curve and the expectile curve is possible.

2.2 Asymptotic behaviour of high expectiles

When X belongs to the domain of attraction of a Generalized Extreme Value distributions, it is possible to derive the asymptotic behaviour of e_q for $q \to 1$ by using techniques from Extreme Value Theory. We refer to Hua and Joe (2011), Tang and Yang (2012), Mao and Hu (2012) for similar results for other one-parameter families of coherent risk measures. Some of the results of this subsection have been independently derived in Mao et al. (2014). We refer to Appendix 5.2 for notations and basic definitions and to Appendix 5.1 for proofs.
In Bellini et al. (2014) it was shown that when $X \in MDA(\Phi_\beta)$, with $\beta > 2$, there exists $\bar{q} < 1$ such that $e_q < x_q$ for each $q \in (\bar{q}, 1)$. That is, for a paretian tail with tail index $\beta > 2$, expectiles are definitively less conservative than the corresponding quantiles. On the contrary, when $\beta < 2$ the opposite inequality holds. Our first result is the extension of this asymptotic comparison between quantiles and expectiles to the Gumbel and Weibull cases, in which expectiles are definitely less conservative than quantiles.

Proposition 2.2. *Let $F \in MDA(\Phi_\beta)$, with $\beta > 1$. If $\beta > 2$ definitely $e_q < x_q$; if $\beta < 2$ definitely $e_q > x_q$. If $F \in MDA(\Psi_\beta)$ or $F \in MDA(\Lambda)$, then definitely $e_q < x_q$.*

As an illustration, see Figure 2 (left and center panel).
When X belongs to the domain of attraction of a Fréchet distribution, that is when X has a Paretian right tail, then it is possible to provide an explicit expression for the asymptotic behaviour of e_q.

Proposition 2.3. *Let $F \in MDA(\Phi_\beta)$, with $\overline{F}(x) \sim x^{-\beta} L(x)$, $L \in RV_0(+\infty)$ and $\beta > 1$. Then, for $q \to 1$,*

$$e_q \sim (\beta - 1)^{-\frac{1}{\beta}} x_q. \tag{6}$$

Moreover, if for some $\lambda_0 > 0$ it holds that

$$\left\{ \frac{L(\lambda_0 t)}{L(t)} - 1 \right\} \log L(t) \to 0 \quad \textit{for } t \to +\infty, \tag{7}$$

then

$$e_q \sim (\beta-1)^{-\frac{1}{\beta}} \frac{(1-q)^{-\frac{1}{\beta}}}{L^{-\frac{1}{\beta}}\left((1-q)^{-\frac{1}{\beta}}\right)}.$$

In particular if $L = C \in (0, +\infty)$, *then*

$$e_q \sim \left(\frac{\beta-1}{C}\right)^{-\frac{1}{\beta}} (1-q)^{-\frac{1}{\beta}}. \tag{8}$$

A similar result can be found in Mao et al. (2014), where also the possible refinements under a second order regular variation condition are discussed in depth.
In the case of the Gumbel domain of attraction the situation is more complicated. For several common distributions (see Examples 3.2, 3.3, 3.4, 3.5), it holds that $\frac{e_q}{x_q} \to 1$, as it happens for ES and for other one parameter families of coherent risk measures (see i.e. Tang and Yang (2012), Mao and Hu (2012)). Typically, the asymptotic f.o.c. is a transcendental equation which can be explicitly solved by means of the so called *Lambert W function* (see Appendix 5.3). We provide a general result for the class of the so called *Weibull-type* distributions (that must not be confused with the distributions in the Weibull domain of attraction; see Beirlant et al. (1995), Dierckx et al. (2009) and the references therein).

Proposition 2.4. *Let* $F \in MDA(\Lambda)$, *with the additional requirement that* $\overline{F} = \exp(-x^\tau L(x))$, *with* $L \in RV_0(+\infty)$ *and* $\tau > 0$. *Then, for* $q \to 1$,

$$e_q^\tau L(e_q) \sim -\log(1-q) \quad \textit{and} \ \log e_q \sim \log x_q. \tag{9}$$

Moreover, if for some $\lambda_0 > 0$ *it holds that*

$$\left\{\frac{L(\lambda_0 t)}{L(t)} - 1\right\} \log L(t) \to 0 \quad \textit{for } t \to +\infty, \tag{10}$$

then

$$e_q \sim x_q.$$

Finally, for distribution in the Weibull domain of attraction we have the following (see also Mao et al. (2014)).

Proposition 2.5. *Let* $F \in MDA(\Psi_\beta)$, *with* $\overline{F}(x) \sim (\widehat{x}-x)^\beta L(\widehat{x}-x)$, $L \in RV_0(0)$ *and* $\beta > 0$. *Then, for* $q \to 1$,

$$(\widehat{x}-e_q)^{\beta+1} L(\widehat{x}-e_q) \sim (\beta+1)(\widehat{x}-\mathbb{E}[X])(1-q).$$

In the particular case $L(x) = C \in (0, +\infty)$, *then*

$$\widehat{x} - e_q \sim \left[\frac{(\widehat{x}-\mathbb{E}[X])(\beta+1)}{C}\right]^{\frac{1}{\beta+1}} (1-q)^{\frac{1}{\beta+1}}. \tag{11}$$

The quality of the first order approximations is graphically assessed in Figures 3 and 4.

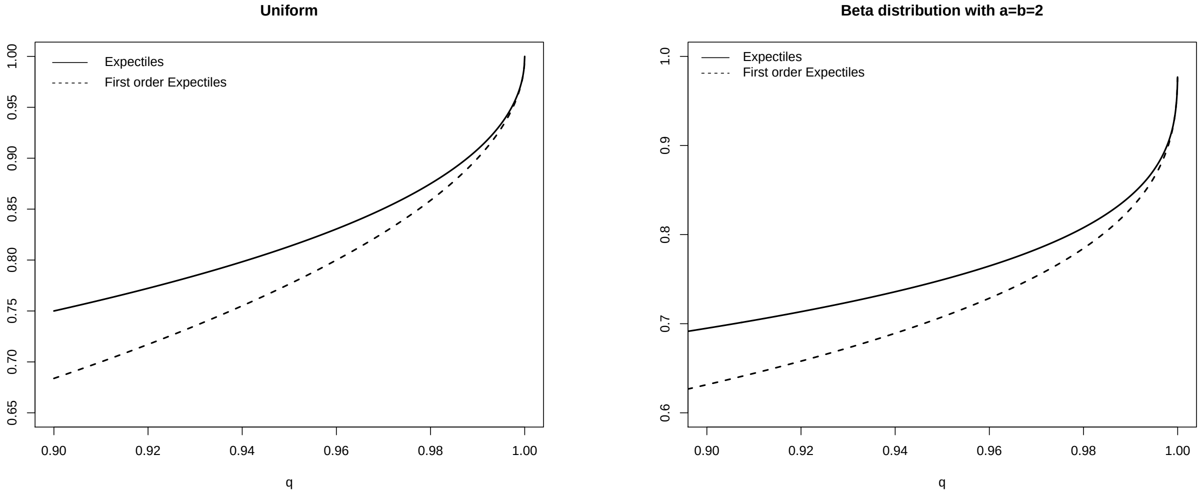

Figure 3: Comparison between expectiles (solid line) and the first order asymptotic approximation (dashed line) in the Uniform and Beta ($a = 2$ and $b = 2$) cases.

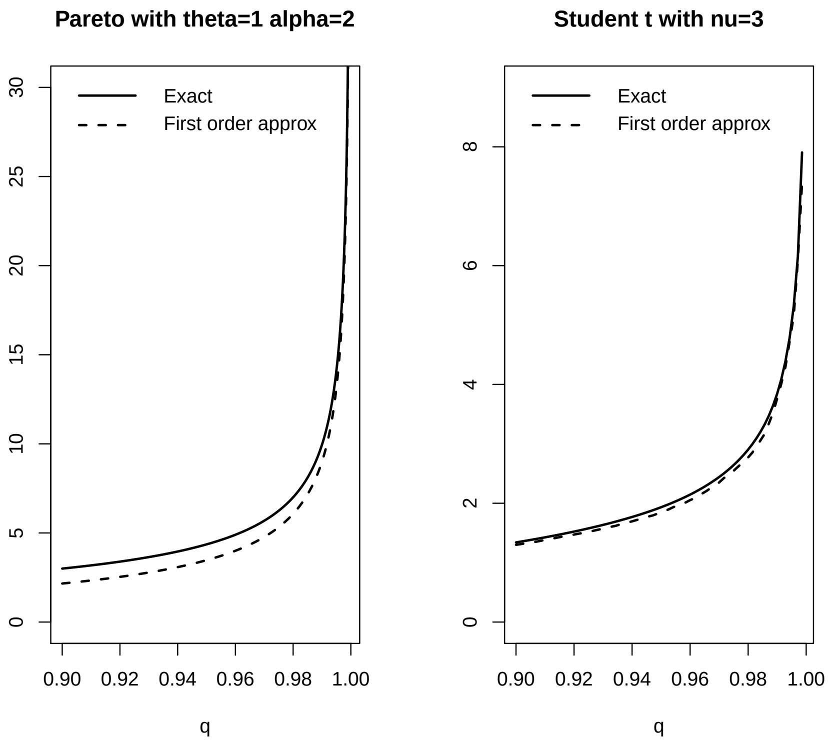

Figure 4: Comparison between expectiles (solid line) and the first order asymptotic approximation (dashed line) in the Pareto ($\theta = 1, \alpha = 2$) and Student t ($\nu = 3$) cases.

3 Examples and Illustrations

Example 3.1 (Uniform distribution)**.** *Let $F(t) = t$, $t \in [0, 1]$. Then the f.o.c. is given by*

$$e_q - \frac{1}{2} = \frac{2q-1}{1-q}\left[\frac{e_q^2 - 2e_q + 1}{2}\right],$$

that gives the explicit solution

$$e_q = \frac{q - \sqrt{q - q^2}}{2q - 1}.$$

In this case $C = 1$ and $\beta = 1$, and indeed

$$1 - e_q \sim \sqrt{1 - q},$$

in accordance with Proposition 2.5.

Example 3.2 (Exponential distribution). *Let $\overline{F} = \exp(-\lambda x)$, $x > 0$, $\lambda > 0$. Then the f.o.c. is given by*

$$e_q - \frac{1}{\lambda} = \frac{(2q-1)\exp(-\lambda x)}{\lambda(1-q)},$$

that can be written as

$$ze^z = \frac{2q-1}{(1-q)e},$$

with $z = \lambda e_q - 1$. Recalling the definition of Lambert's W function (see Appendix 5.3), we find that

$$z = W\left(\frac{2q-1}{(1-q)e}\right)$$

that gives the exact expression

$$e_q = \frac{1}{\lambda}\left\{1 + W\left(\frac{2q-1}{(1-q)e}\right)\right\}.$$

Since $W(x) \sim \log(x)$ for $x \to +\infty$ (see Lemma 5.1), we find

$$e_q \sim \frac{1}{\lambda}\left\{-\log(1-q)\right\} = x_q.$$

Example 3.3 (Logistic distribution). *Let $\overline{F}(x) = \frac{2}{1+\exp(x)}$. Then $m_X(x) \sim 1$ (see Beirlant and Teugels (1992)), and the f.o.c. is*

$$e_q\left(1 + \exp(e_q)\right) = \frac{2}{1-q},$$

and asymptotically we obtain

$$e_q \exp(e_q) = \frac{2}{1-q},$$

that as in the exponential case gives

$$e_q \sim -\log(1-q) \sim x_q.$$

Example 3.4 (Weibull distribution). *Let $\overline{F}(x) = \exp(-\lambda\, x^{\tau})$, with $\tau > 0$. Then $m_X(x) \sim \frac{x^{1-\tau}}{\lambda\tau}$, hence the asymptotic f.o.c. becomes*

$$e_q \sim \frac{e_q^{1-\tau}\exp(-\lambda e_q^{\tau})}{\lambda\tau(1-q)},$$

or

$$\lambda e_q^\tau \exp \lambda e_q \tau \sim \frac{1}{\tau(1-q)},$$

which from Lemma 5.1 gives

$$e_q \sim \left(-\frac{\log(1-q)}{\lambda}\right)^{\frac{1}{\tau}} = x_q.$$

Example 3.5 (Normal distribution). *If $X \sim N(0,1)$, it is well known that*

$$m_X(x) \sim \frac{\overline{F}(x)}{f(x)} \sim \frac{1}{x}, \tag{12}$$

where $f = F'$ is the standard normal density. It follows that

$$\mathbb{E}[(X-x)_+] \sim \frac{\overline{F}(x)}{x} \sim \frac{f(x)}{x^2}.$$

The asymptotic f.o.c. becomes

$$e_q^3 \exp(\frac{e_q^2}{2}) \sim \frac{1}{\sqrt{2\pi}(1-q)},$$

and from Lemma 5.1 it follows that

$$e_q \sim \sqrt{-2\log(1-q)} \sim x_q.$$

Example 3.6 (Pareto distribution). *Let $\overline{F}(t) = \left(\frac{\theta}{t+\theta}\right)^\alpha$, with $t \geq 0$ and $\alpha > 1$. Then $\mathbb{E}[X] = \frac{\theta}{\alpha-1}$ and*

$$\mathbb{E}[(X-x)_+] = \frac{\theta^\alpha (x+\theta)^{-\alpha+1}}{\alpha-1}$$

so the f.o.c. becomes

$$e_q - \frac{\theta}{\alpha-1} = \frac{(2q-1)\theta^\alpha (e_q+\theta)^{-\alpha+1}}{(1-q)(\alpha-1)}.$$

If $\alpha = 2$ the equation can be solved explicitly; we get

$$e_q = \theta\,\frac{\sqrt{q}}{\sqrt{1-q}} > \theta\,\frac{1-\sqrt{1-q}}{\sqrt{1-q}} = x_q.$$

It is easy to check that $\frac{e_q}{x_q} \to 1$, in accordance with Proposition 2.3.

4 Numerical examples

In this section we provide a straightforward example of computation of $EVaR_q$ by standard econometric models. We compare four different methods: a purely historical method, a normal model with fixed volatility, a Garch(1,1) model with normal innovations and a Garch(1,1) model with Student t innovations; all models are estimated on rolling windows of length N =

500. The aim of this simple experiment is just to show that $EVaR_q$ is a perfectly reasonable alternative to VaR_α and ES_α. More sophisticated econometric modelling of expectiles has been pursued in Taylor (2008), Kuan et al. (2009) and De Rossi and Harvey (2009). The first question that has to be addressed is the choice of q. In the case of VaR_α, it is costumary to choose $\alpha = 0.01$. In the case of ES_α, the latest revisions of the Basel Accords suggest $\alpha = 0.025$, in order not to change too much the order of magnitude of the capital requirements. In the case of $EVaR_q$, there seems to be no natural, a priori psychologically acceptable level of gain-loss ratio, so we suggest $q = 0.00145$, which satisfies $EVaR_q(X) = VaR_{0.01}(X)$ for a normally distributed X. See also Rroji (2013) for empirical investigations on the choice of q. The portfolio under consideration is represented by the SP500 Index from 2 November 1994 to 31 December 2009, which correspond to $T = 3818$ trading days. The corresponding logarithmic returns are reported in Figure 5.

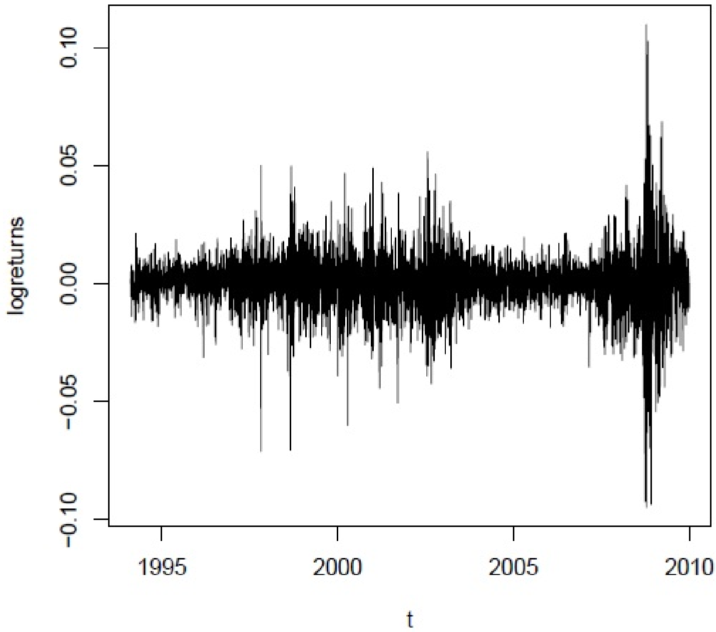

Figure 5: Logreturns of the SP500 Index.

The values of $VaR_{0.01}$ for the four different models are displayed in Figure 6. The number of violations and the p-values of the binomial test are reported in the table below.

Method	Violations of VaR	p-value
Normal model (i)	85	6.9211e-13
Historical model (ii)	57	5.4592e-05
Garch(1,1) model with normal innovations (iii)	61	4.4538e-06
Garch(1,1) model with Student t innovations (iv)	34	0.39849

In accordance with the literature, models with normal innovations are not conservative enough and have a too high number of violations, while the Garch (1,1) model with Student t innovations seems able to capture at least the right frequency of violations. Since VaR_α minimizes a

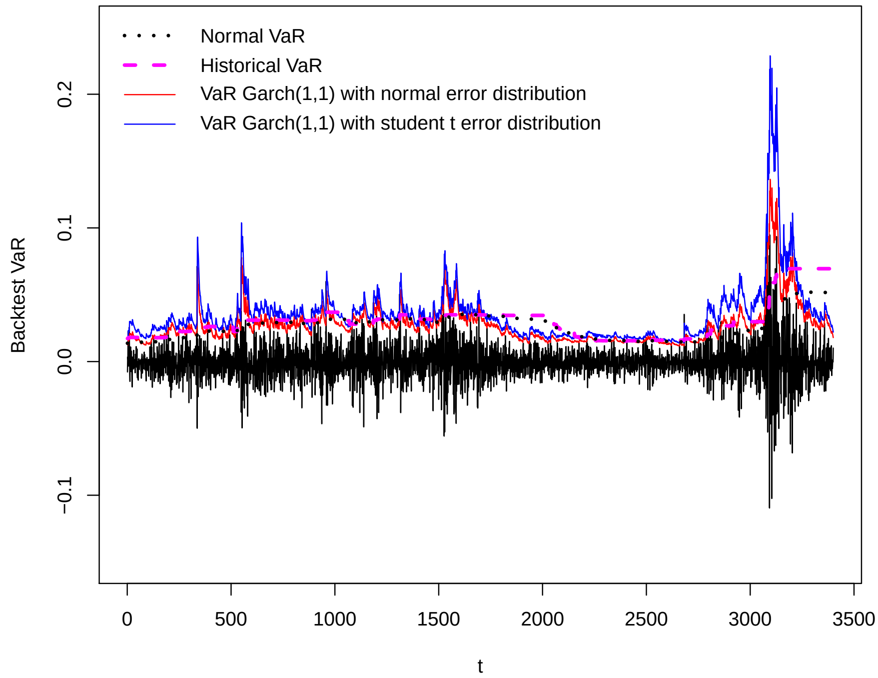

Figure 6: $VaR_{0.01}$ for the SP500 log-return series, evaluated using the four considered methods: a purely historical methods based on rolling windows of length N = 500 (dashed magenta line), a normal model with fixed volatility (dotted-dashed black line), a Garch(1,1) model with normal innovations (red line) and a Garch(1,1) model with Student t innovations (blue line).

piecewise linear, asymmetric scoring function, we consider the *realized loss* given by

$$\mathcal{L}^{VaR}(\alpha) = \frac{1}{T}\sum_{t=1}^{T} L(VaR_t(\alpha), r_t),$$

with

$$L(VaR_t(\alpha), r_t) = \begin{cases} (1-\alpha)\cdot \mid r_t + VaR_t(\alpha) \mid, & \text{if } r_t \le -VaR_t(\alpha) \\ \alpha\cdot \mid r_t + VaR_t(\alpha) \mid, & \text{if } r_t > -VaR_t(\alpha). \end{cases}$$

The values of the realized loss for the different models reported in the following table confirm the superiority of the Garch(1,1) model with Student t innovations. For more sophisticated uses of the realized loss in backtesting and in model selection we refer to Campbell (2005) and Bernardi et al. (2014), and the references therein.

Method	$\mathcal{L}^{VaR}(0.01)$
Normal model (i)	5.6071e-04
Historical model (ii)	4.8657e-04
Garch(1,1) model with normal innovations (iii)	4.0404e-04
Garch(1,1) model with Student t innovations (iv)	3.9078e-04

Our next step is to use the same four models to compute $EVaR_{0.00145}$; the results are displayed in Figure 7. Notice that for $EVaR$ the notion of "violation" is not meaningful, since its definition is related to gain-loss ratios. Theoretically, we know that it should hold

$$\frac{\mathbb{E}(X + EVaR_{0.00145})_+}{\mathbb{E}(X + EVaR_{0.00145})_-} = \frac{\mathbb{E}(X - e_{0.00145})_+}{\mathbb{E}(X - e_{0.00145})_-} = \frac{1 - 0.00145}{0.00145} \simeq 689,$$

so a first assessment of the accuracy of the different models can be obtained by means of the realized gain-loss ratios in the table below.

Method	**Gain-loss ratio**
Normal model (i)	91.27
Historical model (ii)	202.88
Garch(1,1) model with normal innovations (iii)	292.87
Garch(1,1) model with Student t innovations (iv)	705.11

We see that only in the case of the expectiles computed with the Garch (1,1) model with Student t innovations the level of realized gain-loss ratio is satisfactory.

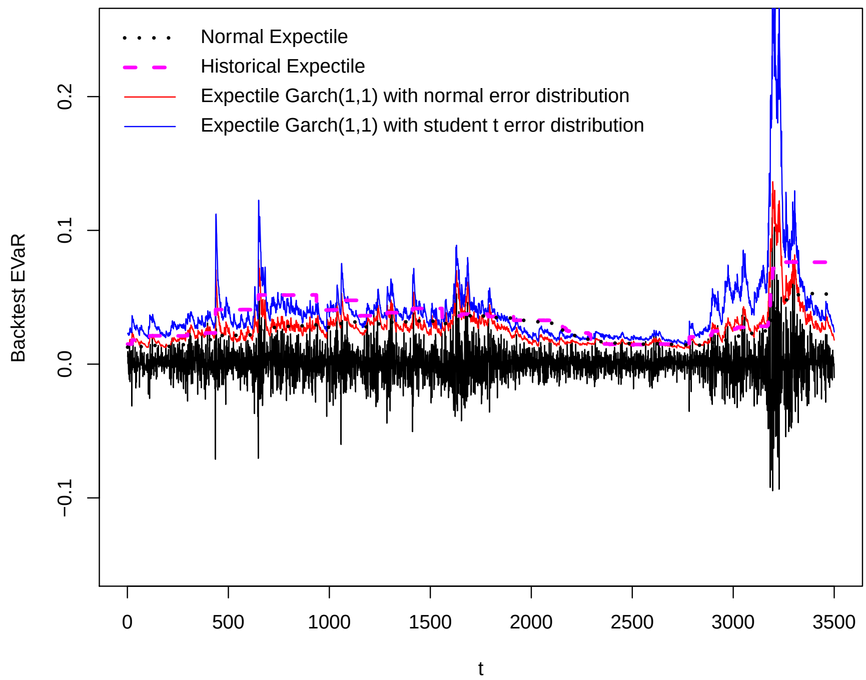

Figure 7: $EVaR_{0.00145}$ for the SP500 log-return series, evaluated using the four considered methods: a purely historical methods based on rolling windows of length N = 500 (dashed magenta line), a normal model with fixed volatility (dotted-dashed black line), a Garch(1,1) model with normal innovations (red line) and a Garch(1,1) model with Student t innovations (blue line).

As we have already done in the case of VaR, exploiting the elicitability property of expectiles, we compute the realized losses

$$\mathcal{L}^{EVaR}(\alpha) = \frac{1}{T}\sum_{t=1}^{T} L(EVaR_t(\alpha), r_t),$$

where the loss function is given this time by

$$L(EVaR_t(\alpha), r_t) = \begin{cases} (1-\alpha)\cdot(r_t + EVaR_t(\alpha))^2, & \text{if } r_t \leq -EVaR_t(\alpha) \\ \alpha\cdot(r_t + EVaR_t(\alpha))^2, & \text{if } r_t > -EVaR_t(\alpha). \end{cases}$$

The results are reported in the following table, where also the square root of $\mathcal{L}^{EVaR}(0.00145)$ is displayed, in order to allow comparisons with the case of VaR.

Method	$\mathcal{L}^{EVaR}(0.00145)$	$\mathcal{L}^{EVaR}(0.00145)^{1/2}$
Normal model (i)	9.2109e-06	0.00303
Historical model (ii)	5.5134e-06	0.00235
Garch(1,1) model with normal innovations (iii)	4.6263e-06	0.00215
Garch(1,1) model with Student t innovations (iv)	4.1235e-06	0.00203

As expected, the best model for forecasting expectiles is Garch(1,1) with Student t innovations, although the differences in the realized losses do not seem very significant; further analysis could be carried out by means of the model selection procedure outlined in Bernardi et al. (2014). Due to the exploratory nature of this example, in this work we limit ourselves to a bootstrap analysis of the distribution of the realized loss under the Garch (1,1) Student t model, that is reported in Figure 8.

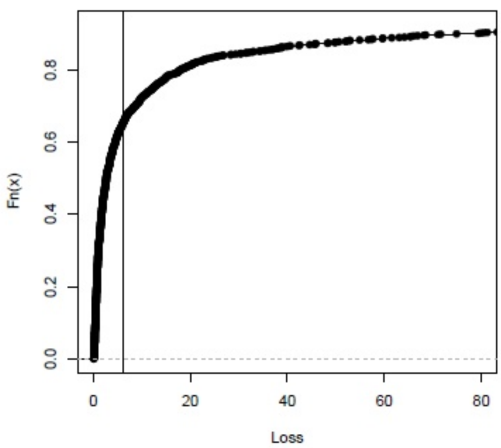

Figure 8: Resimulated distribution of the expectile realized loss using 1000 risimulations of the Garch (1,1) model with Student t innovations. The vertical line represents the observed expectile realized loss.

Finally, in Figure 9 we compare $VaR_{0.01}$, $ES_{0.025}$ and $EVaR_{0.00145}$ computed with the Garch (1,1) model with Student t innovations. The three risk measures are very similar, since the forecasting distributions differs only in a single shape parameter, the number of degrees of

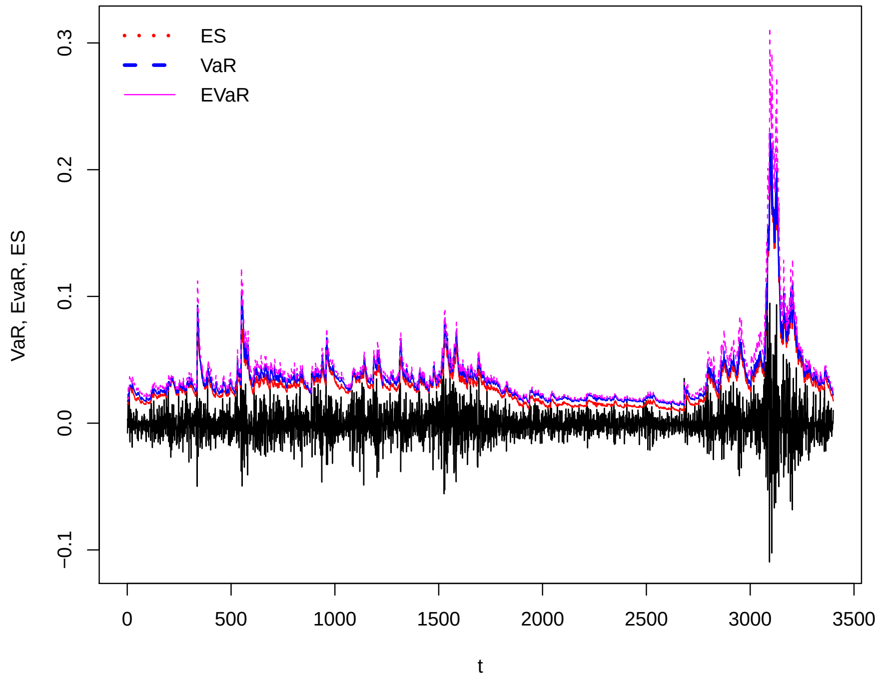

Figure 9: $VaR_{0.01}$, $ES_{0.025}$ and $EVaR_{0.00145}$ for the S&P500 log-return, evaluated using Garch(1,1) risk model with Student t error distribution.

freedom ν. In conclusion, we showed that by a proper choice of q and simple econometric models it is not difficult to forecast EVaR and obtain capital requirements in line with those of VaR and ES. Moreover, the elicitability property of expectiles enables the risk manager to adopt a standard set of techniques mutuated from the forecasting literature to assess the accuracy of the model, to compare between different models and, if necessary, to provide a statistical test of the model by means of the realized loss.

5 Appendix

5.1 Proofs

Proof of Proposition 2.2. Let us first write the f.o.c. (5) in the following equivalent from:

$$e_q - \mathbb{E}[x] = \frac{2q-1}{1-q}\,\mathbb{E}[(X-x)_+].$$

Let us consider separately the three cases. In the Fréchet case $\overline{F} \in RV_{-\beta}(+\infty)$, and since $e_q \to +\infty$ for $q \to 1$, the asymptotic f.o.c. becomes

$$e_q(X) \sim \frac{1}{1-q}\,\mathbb{E}[(X-e_q)_+],$$

and from Proposition 5.3 we obtain

$$e_q \sim \frac{e_q \overline{F}(e_q)}{(1-q)(\beta-1)}.$$

Since $\overline{F}(x_q) \sim 1-q$, it follows that $\overline{F}(e_q)/\overline{F}(x_q) \sim \beta - 1$, from which the thesis follows, as it was already shown in Bellini et al. (2014). In the Gumbel case, we have to distinguish between the two further subcases $\widehat{x} = +\infty$ and $\widehat{x} < +\infty$. When $\widehat{x} = +\infty$, from Proposition 5.3 the asymptotic f.o.c. becomes

$$e_q \sim \frac{m_X(e_q)\overline{F}(e_q)}{1-q}$$

and since in this case $m_X(e_q) = o(e_q)$ (see Proposition 3.3.24 in Embrechts et al. (1997)), it follows that $\overline{F}(e_q)/\overline{F}(x_q) \to +\infty$. When $\widehat{x} < +\infty$, we have that

$$e_q - \mathbb{E}[X] \sim \frac{m_X(e_q)\overline{F}(e_q)}{1-q}$$

and since $m_X(e_q) = o(\widehat{x} - e_q) \to 0$ (see Proposition 3.3.24 in Embrechts et al. (1997)), it follows that again $\overline{F}(e_q)/\overline{F}(x_q) \to +\infty$. Finally, in the Weibull case from Proposition 5.3 the asymptotic f.o.c. is given by

$$\widehat{x} - \mathbb{E}[X] \sim \frac{(\widehat{x} - e_q)\overline{F}(e_q)}{(1-q)(\beta+1)},$$

that gives again

$$\frac{\overline{F}(e_q)}{\overline{F}(x_q)} \sim \frac{(\beta+1)(\widehat{x} - E[X])}{\widehat{x} - e_q} \to +\infty,$$

from which the thesis follows. □

Proof of Proposition 2.3. In the Fréchet case, the previously derived asymptotic relationship

$$\overline{F}(e_q) \sim (\beta-1)(1-q),$$

can also be written as

$$e_q = F^{\leftarrow}(1 - (\beta-1)(1-q+o(1))).$$

By Theorem 1.5.12 in Bingham et al. (1989) if $\overline{F} \in RV_{-\beta}(+\infty)$ then $F^{-1}(1-\cdot) \in RV_{-\frac{1}{\beta}}(0)$ (see also Yang (2013)). Then it follows that for $q \to 1$

$$e_q \sim (\beta-1)^{-\frac{1}{\beta}} x_q.$$

Under condition (7), from Theorem 2.3.3 and Corollary 2.3.4 in Bingham et al. (1989) it follows that

$$x_q \sim \left(\frac{1-q}{L\left((1-q)^{-\frac{1}{\beta}}\right)}\right)^{-\frac{1}{\beta}},$$

from which we have the thesis. □

Proof of Proposition 2.4. Under the given asymptotic hypothesis on $\overline{F}$, Beirlant et al. (1995) prove that

$$m_X(x) \sim \frac{x^{1-\tau}}{\tau\, L(x)}. \tag{13}$$

The asymptotic f.o.c. becomes

$$e_q \sim \frac{e_q^{1-\tau} \exp(-e_q^\tau L(e_q))}{\tau(1-q)L(e_q)},$$

that by setting $z = e_q^\tau\, L(e_q)$ and applying Lemma 5.1 gives

$$e_q^\tau\, L(e_q) \sim -\log(1-q).$$

Beirlant et al. (1995) showed also that in this case $x_q = (-\log(1-q))^{1/\tau} L^*(-\log(1-p))$, for a slowly varying L^*, that implies $\log x_q \sim \frac{1}{\tau} \log(-\log(1-q))$.
Since for a slowly varying L it holds that $\log L(x) = o(\log(x))$, we get $\log e_q \sim \frac{1}{\tau} \log(-\log(1-q)) \sim \log x_q$.
Under (10), it also hold that

$$x_q^\tau\, L(x_q) \sim -\log(1-q),$$

that by asymptotic inversion gives $x_q \sim e_q$. Hence the result. $\square$

Proof of Proposition 2.5. By Corollary 5.3 we have that

$$\widehat{x} - \mathbb{E}[X] \sim \frac{1}{1-q}(\widehat{x} - e_q)\frac{\overline{F}(e_q)}{\beta+1}.$$

Since

$$\overline{F}(e_q) = \overline{F}(\widehat{x} - (\widehat{x} - e_q)) \sim (\widehat{x} - e_q)^\beta L(\widehat{x} - e_q),$$

we get the first thesis. When $L = C$, the result follows by asymptotic inversion. $\square$

5.2 Basic definitions and results from EVT

We always denote with $\widehat{x}$ the right endpoint of F, that is $\widehat{x} := \sup\{x : F(x) < 1\}$. The possible limiting distributions of properly normalized maxima of i.i.d. random variables belong to the class of Extreme Value distributions, that can take three possible shapes: Fréchet, Weibull and Gumbel.

Definition 5.1 (Extreme value distributions)**.**

$$\Phi_\beta(x) = \begin{cases} 0 & x \le 0 \\ \exp(-x^{-\beta}) & x > 0 \end{cases}, \quad \beta > 0$$

$$\Psi_\beta(x) = \begin{cases} \exp(-(-x)^{\beta}) & x \le 0 \\ 0 & x > 0 \end{cases}, \quad \beta > 0$$

$$\Lambda(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}$$

A unifying parametrization due to Jenkinson and von Mises is the following:

$$G_\gamma := \begin{cases} \Phi_{\frac{1}{\gamma}} & \gamma > 0 \\ \Psi_{-\frac{1}{\gamma}} & \gamma < 0 \\ \Lambda & \gamma = 0 \end{cases}$$

A convenient characterization of the maximum domain of attractions of the distributions G_γ can be given by means of the notions of *regular variation* and *extended regular variation*:

Definition 5.2 (Regular variation). *Let $\beta \in \mathbb{R}$. A measurable function $h : \mathbb{R} \to \mathbb{R}$ is of regular variation of index β in x_0 if*

$$\lim_{t \to x_0} \frac{h(tx)}{h(t)} = x^\beta,$$

for each $x \in \mathbb{R}$. The class of regularly varying functions of index β in x_0 is denoted by $RV_\beta(x_0)$. A regularly varying function with $\beta = 0$ is called slowly varying.

Definition 5.3 (Extended regular variation). *Let $\gamma \in \mathbb{R}$. A measurable function $h : \mathbb{R}^+ \to \mathbb{R}$ is of extended regular variation with index γ at $+\infty$ if there exists a function $a : \mathbb{R}^+ \to \mathbb{R}^+$ such that for all $x > 0$,*

$$\lim_{t \to +\infty} \frac{h(tx) - h(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma},$$

where in the case $\gamma = 0$ the right hand side has to be interpreted as $\log x$. The function $a(t)$ is referred as an auxiliary function for h. The class of extended regularly varying functions at $+\infty$ is denoted by $ERV_\gamma(+\infty)$.

When $\gamma \neq 0$, it is easy to see that the notion of extended regular variation boils down to that of regular variation. The advantage of Definition 5.3 is that it enables to give a unified characterization of the maximum domain of attractions of the three limiting laws.
Let us denote with $f^{\leftarrow}$ the left continuous inverse of a nondecreasing function $f : \mathbb{R} \to \mathbb{R}$, that is

$$f^{\leftarrow}(y) := \inf\{x : f(x) \geq y\},$$

and with U the so called *tail quantile function*, given by

$$U := \left(\frac{1}{\overline{F}}\right)^{\leftarrow}, \quad \text{with } \overline{F} = 1 - F.$$

Then the following result holds.

Proposition 5.1 (Theorem 1.1.6 in de Haan and Ferreira (2006)).

$$F \in MDA(G_\gamma) \textit{ if and only if } U \in ERV_\gamma.$$

In the case $\gamma \neq 0$, the notion of regular variation is sufficient to completely characterize $MDA(G_\gamma)$ as stated by the following result. The interested reader is also referred to Example 3.3.7 and 3.3.12 in Embrechts et al. (1997).

Proposition 5.2 (Theorem 3.3.7 and 3.3.12 in Embrechts et al. (1997)).

$$F \in MDA(\Phi_\beta) \textit{ if and only if } \overline{F}(x) \sim x^{-\beta} L(x),$$
$$F \in MDA(\Psi_\beta) \textit{ if and only if } \widehat{x} < +\infty \textit{ and } \overline{F}\left(\widehat{x} - \frac{1}{x}\right) \sim x^{-\beta} L(x),$$

where L is a slowly varying function and $\beta > 0$.

When $F \in MDA(G_\gamma)$, that is when $U \in ERV_\gamma$, we have the following (see Tang and Yang (2012), Mao and Hu (2012), Hua and Joe (2011)):

Proposition 5.3. *Let $F \in MDA(\Phi_\beta)$, with $\beta > 1$. Then*

$$\lim_{x \to \widehat{x}} \frac{\mathbb{E}[(X-x)_+]}{x\overline{F}(x)} = \frac{1}{\beta - 1}$$

Let $F \in MDA(\Psi_\beta)$; then

$$\lim_{x \to \widehat{x}} \frac{\mathbb{E}[(X-x)_+]}{(\widehat{x} - x)\overline{F}(x)} = \frac{1}{\beta + 1}$$

Let $F \in MDA(\Lambda)$; then

$$\lim_{x \to \widehat{x}} \frac{\mathbb{E}[(X-x)_+]}{m_X(x)\overline{F}(x)} = 1,$$

where $m_X(t) := \mathbb{E}[X - t | X > t]$ is the mean excess function of X.

5.3 Lambert W function

In the following we recall the basic properties of the *Lambert W function*, that is defined implicitly by means of the equation

$$W(z) \exp(W(z)) = z.$$

When $z \in \mathbb{R}$, $z > 0$, the solution is unique. For example, it is easy to check that $W(e) = 1$. The Lambert W function arises in many applied problems, typically in connection with the solution of transcendental equations; for a detailed review of its properties see for example Corless et al. (1996). In the case of a real z, the asymptotic behaviour of $W(z)$ for $z \to +\infty$ is the same as $\log z$. For completeness, we report the argument in a slightly more general form:

Lemma 5.1. *Let $W_\alpha(t)$ be the unique solution of the equation*

$$W^\alpha \exp(W_\alpha) = t, \tag{14}$$

with $\alpha > 0$ and $t > 0$. Then for $t \to +\infty$,

$$W_\alpha(t) = \log t + O(\log \log t).$$

Proof. The argument in De Bruijn (1954) applies with almost no modifications. Since the function $w \mapsto w^\alpha \exp w$ is strictly increasing for $w > 0$, the solution of Equation (14) is unique, and $t > e \Rightarrow W(t) > 1$. Rewriting Equation (14) as

$$W = \log t - \alpha \log W, \tag{15}$$

we get that $t > e \Rightarrow W < \log t \Rightarrow \log W < \log \log t$, which inserted in Equation (15) gives the thesis. □

References

Abdous, B. and Remillard, B. (1995). Relating quantiles and expectiles under weighted-symmetry. *Ann. Inst. Statist. Math*, 47(2):371–384.

Acerbi, C. and Szekely, B. (2014). Backtesting Expected Shortfall. *Working paper*.

Artzner, P., Delbaen, F., Eber, J. M., and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, 9(3):203–228.

Beirlant, J., Broniatowski, M., Teugels, J., and Vynckier, P. (1995). The mean residual life function at great age: applications to tail estimation. *Journal of Statistical Planning and Inference*, 45:21–48.

Beirlant, J. and Teugels, J. L. (1992). Modeling large claims in non-life insurance. *Insurance: Mathematics and Economics*, 11(1):17–29.

Bellini, F. and Bignozzi, V. (2013). On elicitable risk measures. *Preprint, to appear in Quantitative Finance, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2334746*.

Bellini, F., Klar, B., Müller, A., and Rosazza-Gianin, E. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, 54:41–48.

Ben Tal, A. and Teboulle, M. (2007). An old new concept of convex risk measures: the optimized certainty equivalent. *Mathematical Finance*, pages 449–476.

Bernardi, M., L., C., and L., P. (2014). Are news important to predict large losses ? *Preprint, arXiv:1410.6898v2*.

Biagini, S. and Pinar, M. (2013). The best gain-loss ratio is a poor performance measure. *SIAM Journal on Financial Mathematics*, 4(1):228–242.

Bingham, N., Goldie, C., and Teugels, J. (1989). *Regular variation*. Cambridge University Press.

Bosi, I. (2014). Risk management with expectiles. Master's thesis, University of Milano Bicocca.

Campbell, S. D. (2005). A review of backtesting and backtesting procedures. *Federal Reserve Finance and Economics Discussion Series*.

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., and Knuth, D. E. (1996). On the Lambert W Function. pages 329–359.

Davis, M. (2013). Consistency of Risk Measure Estimates. *Preprint*.

De Bruijn, N. G. (1954). *Asymptotic Methods in Analysis*. Dover.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer.

De Rossi, G. and Harvey, A. (2009). Quantiles, expectiles and splines. *Journal of Econometrics*, 152:179–185.

Delbaen, F. (2012). *Monetary Utility Functions*. Osaka University Press.

Delbaen, F. (2013). A remark on the structure of expectiles. *Preprint*, arXiv:1307.5881v1.

Delbaen, F., Bellini, F., Bignozzi, V., and Ziegel, J. (2014). Risk measures with the CxLS property. *Working paper*.

Dierckx, G., Beirlant, J., De Waal, D., and Guillou, A. (2009). A new estimaton method for Weibull-type tails based on the mean excess function. *Journal of Statistical Planning and Inference*, 139:1905–1920.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R., and Beleraj, A. (2014). An Academic Response to Basel 3.5. *Risks*, 2(1):25–48.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, 6:429–447.

Föllmer, H. and Schied, A. (2011). *Stochastic Finance: an introduction in discrete time*. Walter De Gruyter.

Gneiting, T. (2011). Making and evaluating point forecast. *Journal of the American Statistical Association*, 106(494):746–762.

Hua, L. and Joe, H. (2011). Second order regular variation and conditional tail expectation of multiple risks. *Insurance: Mathematics and Economics*, 49:537–546.

Jones, M. C. (1994). Expectiles and M-quantiles are quantiles. *Statistics and Probability Letters*, 20:149–153.

Koenker, R. (1993). When are expectiles percentiles (solution). *Econometric Theory*, 9:526–527.

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

Kuan, C., Yeh, J., and Hsu, Y. (2009). Assessing value at risk with CARE, the Conditional Autoregressive Expectile models. *Journal of Econometrics*, 150:261–270.

Mao, T. and Hu, T. (2012). Second-order properties of the Haezendonck-Goovaerts risk measure for extreme risks. *Insurance: Mathematics and Economics*, 51:333–343.

Mao, T., Ng, K., and Hu, T. (2014). Asymptotic Expansions of Generalized Quantiles and Expectiles for Extreme Risks. Working Paper.

Newey, W. and Powell, J. (1987). Asymmetric least squares estimation and testing. *Econometrica*, 55(4):819–847.

Pflug, G. and Romisch, W. (2007). *Modeling, Measuring and Managing Risk.* World Scientific Publishing Company.

Rroji, E. (2013). *Risk attribution and semi-heavy tailed distributions.* PhD thesis, University of Milano Bicocca.

Shadwick, W. and Keating, C. (2002). A universal performance measure. *Journal of Performance Measurement*, Spring Issue:59–84.

Tang, Q. and Yang, F. (2012). On the Haezendonck-Goovaerts risk measure for extreme risks. *Insurance: Mathematics and Economics*, 50:217–227.

Taylor, J. W. (2008). Estimating Value at Risk and Expected Shortfall using Expectiles. *Journal of Financial Econometrics*, 6:231–252.

Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, 16(2):419–441.

Yang, F. (2013). *Asymptotics for Risk Measures of Extreme Risks.* PhD thesis, University of Iowa.

Ziegel, J. (2014). Coherency and elicitability. arXiv:1303.1690v3, to appear in *Mathematical Finance*.

Zou, H. (2014). Generalizing Koenker's distribution. *Journal of Statistical Planning and Inference*, 148:123–127.