Expected Shortfall is not elicitable – so what?

Dirk Tasche

Bank of England – Prudential Regulation Authority[1]
dirk.tasche@gmx.net

Modern Risk Management of Insurance Firms
Hannover, January 23, 2014

[1] The opinions expressed in this presentation are those of the author and do not necessarily reflect views of the Bank of England.

Outline

Outline

Motive of this presentation

- For more than 10 years, academics have been suggesting Expected Shortfall (ES) as a coherent alternative to Value-at-Risk (VaR).
- Recently, the Basel Committee (BCBS, 2013) has confirmed that ES will replace VaR for regulatory capital purposes in the trading book.
- Gneiting (2011) points out that *elicitability* is a desirable property when it comes to "making and evaluating point forecasts". He finds that "conditional value-at-risk [ES] is not [elicitable], despite its popularity in quantitative finance."
- *Expectiles* are coherent and elicitable.
- That is why several authors have suggested to drop both VaR and ES and use expectiles instead.

Outline

What risk do we measure?

- Rockafellar and Uryasev (2013) distinguish 4 approaches to the measurement of risk:
  - Risk measures – aggregated values of random cost.
  - Deviation measures – deviations from benchmarks or targets.
  - Measures of regret – utilities in the context of losses. They 'generate' risk measures.
  - Error measures – quantifications of 'non-zeroness'. They 'generate' deviation measures.
- Risk measures may be understood as measures of solvency $\Rightarrow$ Use by creditors and regulators.
- Deviation measures may be interpreted as measures of uncertainty $\Rightarrow$ Use by investors of own funds (no leverage).

Solvency measures

- There are many papers on desirable properties of risk measures. Most influential: Artzner et al. (1999)
- **Coherent risk measures:** How much capital is needed to make position[2] L acceptable to regulators?
  - Homogeneity ("double exposure ⇒ double risk"):

$$\rho(h\,L) = h\,\rho(L), \quad h \geq 0. \tag{1a}$$

  - Subadditivity ("reward diversification"):

$$\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2). \tag{1b}$$

  - Monotonicity ("higher losses imply higher risk"):

$$L_1 \leq L_2 \quad \Rightarrow \quad \rho(L_1) \leq \rho(L_2). \tag{1c}$$

  - Translation invariance ("reserves reduce requirements")

$$\rho(L - a) = \rho(L) - a, \quad a \in \mathbb{R}. \tag{1d}$$

[2] Convention: Losses are positive numbers, gains are negative.

Important and less important properties

- **Characterisation:** A risk measure ρ is coherent if and only there is a set of probability measures $\mathcal{Q}$ such that

$$\rho(L) = max_{Q\in\mathcal{Q}}\mathrm{E}_Q[L], \quad \text{for all } L. \tag{2}$$

$\Rightarrow$ Interpretation of coherent measures as expectations in stress scenarios.

- **Duality:** $\rho(L)$ solvency risk measure $\Rightarrow$
  $\delta(L) = \rho(L) - \mathrm{E}[L]$ deviation measure
- Homogeneity and subadditivity are preserved in δ.
  Monotonicity and translation invariance are not preserved.
- **Conclusion:** Monotonicity and translation invariance are less important properties.

Other important properties

- **Comonotonic additivity** ("No diversification for total dependence"):

$$L_1 = f_1 \circ X,\ L_2 = f_2 \circ X \ \Rightarrow\ \rho(L_1 + L_2) = \rho(L_1) + \rho(L_2). \qquad \text{(3a)}$$

  X common risk factor, f_1, f_2 increasing functions.

- **Law-invariance** ("context independence"[3]):

$$\mathrm{P}[L_1 \le \ell] = \mathrm{P}[L_2 \le \ell], \ell \in \mathbb{R} \ \Rightarrow\ \rho(L_1) = \rho(L_2). \qquad \text{(3b)}$$

- **Proposition:** Coherent risk measures ρ that are also law-invariant and comonotonically additive are **spectral measures**, i.e. there is a convex distribution function F_ρ on $[0, 1]$ such that

$$\rho(L) = \int_0^1 q_u(L)\, F_\rho(du), \quad \text{for all } L. \qquad \text{(3c)}$$

  $q_u(L) = \min\{\mathrm{P}[L \le \ell] \ge u\}$ denotes the u-quantile of L.

[3] Identical observations in a downturn and a recovery imply the same risk.

Risk contributions

- Generic one-period **loss model:**

$$L = \sum_{i=1}^{m} L_i. \tag{4}$$

L portfolio-wide loss, m number of risky positions in portfolio, L_i loss with i-th position.

- **Risk sensitivities** $\rho(L_i \,|\, L) = \frac{d\,\rho(L+h\,L_i)}{d\,h}\,\Big|_{h=0}$ are of interest for risk management and optimisation.
- ρ homogeneous and differentiable $\Rightarrow$

$$\sum_{i=1}^{m} \rho(L_i \,|\, L) = \rho(L). \tag{5}$$

$\Rightarrow$ Interpretation of sensitivities as risk contributions[4].

[4] This approach to contributions is called *Euler allocation*.

Some properties of risk contributions

- $\rho(L)$ positively homogeneous $\Rightarrow$

$$\rho(L_i \,|\, L) \le \rho(L_i) \quad \Longleftrightarrow \quad \rho \text{ subadditive}$$

For subadditive risk measures, the risk contributions of positions do never exceed their stand-alone risks.

- $\rho(L)$ positively homogeneous and subadditive $\Rightarrow$

$$\rho(L) - \rho(L - L_i) \le \rho(L_i \,|\, L) \tag{6}$$

So-called 'with – without' risk contributions underestimate the Euler contributions.

- ρ spectral risk measure, smooth loss distribution $\Rightarrow$

$$\rho(L_i \,|\, L) = \int_0^1 \mathrm{E}\big[L_i \mid L = q_u(L)\big]\, F_\rho(du). \tag{7}$$

Outline

Shortfall probability risk measures

- Special case of solvency risk measures.
- **Construction principle:** For a given confidence level γ, the *risk measure* $\rho(L)$ specifies a level of loss that is exceeded only with probability less than $1 - \gamma$.
- Formally, $\rho(L)$ should satisfy

$$\mathrm{P}[L > \rho(L)] \;\leq\; 1 - \gamma. \tag{8}$$

- γ is often chosen on the basis of a target rating, for example for a target A rating with long-run average default rate[5] of 0.07%:

$$1 - \gamma = 0.07\%$$

- Popular examples: (Scaled) standard deviation, Value-at-Risk (VaR), Expected Shortfall (ES).

[5] Source: S&P (2013), table 21.

Standard deviation

- Scaled **standard deviation** (with constant $a > 0$):

$$\sigma_a(L) \;=\; \mathrm{E}[L] + a\sqrt{\mathrm{var}[L]} \;=\; \mathrm{E}[L] + a\sqrt{\mathrm{E}\big[(L-\mathrm{E}[L])^2\big]}. \qquad (9a)$$

- By Chebychev's inequality:

$$\mathrm{P}[L > \sigma_a(L)] \;\le\; \mathrm{P}\big[|L-\mathrm{E}[L]| > a\sqrt{\mathrm{var}[L]}\big] \;\le\; a^{-2}. \qquad (9b)$$

- Hence, choosing $a = \frac{1}{\sqrt{\gamma}}$ (e.g. $\gamma = 0.001$) yields

$$\mathrm{P}[L > \sigma_a(L)] \;\le\; \gamma. \qquad (9c)$$

- Alternative: Choose a such that (9c) holds for, e.g., normally distributed L. Underestimates risk for skewed loss distributions.

Properties of standard deviation

- Homogeneous, subadditive and law-invariant
- Not comonotonically additive, but additive for risks with correlation 1
- Not monotonic, hence not coherent
- Easy to estimate – moderately sensitive to 'outliers' in sample
- **Overly expensive** if calibrated (by Chebychev's inequality) to be a shortfall measure
- Risk contributions:

$$\sigma_a(L_i \,|\, L) \;=\; a\,\frac{\mathrm{cov}(L_i, L)}{\sqrt{\mathrm{var}(L)}} + \mathrm{E}[L_i]. \tag{10}$$

Value-at-Risk

- For $\alpha \in (0,1)$: α-quantile $q_\alpha(L) = \min\{\ell : \mathrm{P}[L \le \ell] \ge \alpha\}$.
- In finance, $q_\alpha(L)$ is called **Value-at-Risk** (VaR).
- If L has a continuous distribution (i.e. $\mathrm{P}[L = \ell] = 0$, $\ell \in \mathbb{R}$), then $q_\alpha(L)$ is a solution of $\mathrm{P}[L \le \ell] = \alpha$.
- **Quantile** / **VaR-based** risk measure:

$$\mathrm{VaR}_\alpha(L) \;=\; q_\alpha(L). \tag{11a}$$

- By definition $\mathrm{VaR}_\alpha(L)$ satisfies

$$\mathrm{P}\big[L > \mathrm{VaR}_\alpha(L)\big] \le 1 - \alpha. \tag{11b}$$

Properties of Value-at-Risk

- Homogeneous, comonotonically additive and law-invariant
- **Not subadditive**, hence not coherent
- Easy to estimate by sorting sample – not sensitive to extreme 'outliers'
- Provides least loss in worst case scenario – may be misleading.
- Risk contributions:

$$\mathrm{VaR}_\alpha(L_i \,|\, L) \;=\; \mathrm{E}\big[L_i \;|\; L = q_\alpha(L)\big]. \tag{12}$$

- Estimation of risk contributions is difficult in continuous case.

Expected Shortfall

- **Expected Shortfall** (ES, Conditional VaR, superquantile). Spectral risk measure with $F_\rho(u) = (1-\alpha)^{-1} \max(u, \alpha)$:

$$\begin{aligned} \mathrm{ES}_\alpha(L) &= \tfrac{1}{1-\alpha} \int_\alpha^1 q_u(L)\, du \\ &= \mathrm{E}[L \,|\, L \geq q_\alpha(L)] \\ &\quad + \big(\mathrm{E}[L \,|\, L \geq q_\alpha(L)] - q_\alpha(L)\big) \left(\tfrac{\mathrm{P}[L \geq q_\alpha(L)]}{1-\alpha} - 1\right). \end{aligned} \tag{13}$$

- If $\mathrm{P}[L = q_\alpha(L)] = 0$ (in particular, if L has a density),

$$\mathrm{ES}_\alpha(L) \;=\; \mathrm{E}[L \,|\, L \geq q_\alpha(L)].$$

- ES dominates VaR: $\mathrm{ES}_\alpha(L) \;\geq\; \mathrm{VaR}_\alpha(L)$.

Properties of Expected Shortfall

- **Coherent, comonotonically additive and law-invariant**
- Easy to estimate by sorting. Provides average loss in worst case scenario
- Least coherent law-invariant risk measure that dominates VaR
- Risk contributions (continuous case):

$$\mathrm{ES}_\alpha(L_i \,|\, L) \;=\; \mathrm{E}\big[L_i \,|\, L \geq q_\alpha(L)\big]. \tag{14}$$

- Very sensitive to extreme ‘outliers’. For same accuracy, many more observations than for VaR at same confidence level might be required.
- Big gap between VaR and ES indicates heavy tail loss distribution.

Outline

Related definitions

- A **scoring function** is a function

$$s : \mathbb{R} \times \mathbb{R} \to [0, \infty),\ (x, y) \mapsto s(x, y), \tag{15a}$$

  where x and y are the *point forecasts* and *observations* respectively.
- Let ν be a functional on a class of probability measures $\mathcal{P}$ on $\mathbb{R}$:

$$\nu : \mathcal{P} \to 2^{\mathbb{R}},\ P \mapsto \nu(P) \subset \mathbb{R}.$$

  A scoring function $s : \mathbb{R} \times \mathbb{R} \to [0, \infty)$ is **consistent** for the functional ν relative to $\mathcal{P}$ if and only if

$$\mathrm{E}_P\left[s(t, Y)\right] \le \mathrm{E}_P\left[s(x, Y)\right] \tag{15b}$$

  for all $Y \sim P \in \mathcal{P}$, $t \in \nu(P)$ and $x \in \mathbb{R}$.
- s is **strictly consistent** if it is consistent and

$$\mathrm{E}_P\left[s(t, Y)\right] = \mathrm{E}_P\left[s(x, Y)\right] \Rightarrow x \in \nu(P). \tag{15c}$$

Elicitability

- The functional ν is **elicitable** relative to $\mathcal{P}$ if and only if there is a scoring function s which is strictly consistent for ν relative to $\mathcal{P}$.
- Examples:

$$\text{Expectation: } \nu(P) = \int x\, P(dx), \quad s(x, y) = (y - x)^2. \tag{16a}$$

$$\text{Quantiles: } \nu(P) = \big\{x : P[(-\infty, x)] \le \alpha \le P[(-\infty, x]]\big\}, \tag{16b}$$
$$s(x, y) = \tfrac{\alpha}{1-\alpha} \max(y - x, 0) + \max(x - y, 0).$$

- Interpretation:
  - Point estimates of elicitable functionals can be determined by means of regression:
  $$\nu(P) = \arg\min_x \mathrm{E}_P[s(x, Y)], \quad Y \sim P. \tag{16c}$$
  - Point estimation methods of elicitable functionals can be compared by means of the related scoring functions (interesting for backtesting).

Standard deviation and ES are not elicitable

- **Necessary** for ν being elicitable ("convex level sets"):

$$\begin{aligned} 0 < \pi < 1, \quad & t \in \nu(P_1) \cap \nu(P_2) \\ \Rightarrow \quad & t \in \nu(\pi\, P_1 + (1-\pi)\, P_2) \end{aligned} \tag{17a}$$

- By counter-examples: Standard deviation and ES violate (17a).
  $\Rightarrow$ Standard deviation and ES are not elicitable.
- But standard deviation and ES can be calculated by means of regression, with s as in (16a) and (16b):

$$\mathrm{var}(P) = \min_x \mathrm{E}_P\big[(Y-x)^2\big] \tag{17b}$$

$$\begin{aligned} \mathrm{ES}_\alpha(P) = \min_x \Big\{ & \mathrm{E}_P\big[\tfrac{\alpha}{1-\alpha}\, \max(Y-x, 0) \\ & + \max(x-Y, 0)\big] + \mathrm{E}_P[Y]\Big\}. \end{aligned} \tag{17c}$$

Outline

Expectiles

- For $0 < \tau < 1$ the τ**-expectile** of square-integrable Y is defined by

$$e_\tau(Y) = \arg\min_x \mathrm{E}\big[\tau \max(Y - x, 0)^2 + (1-\tau) \max(x - Y, 0)^2\big] \quad (18a)$$

- e_τ is elicitable with scoring function

$$s(x, y) = \tau \max(y - x, 0)^2 + (1-\tau) \max(x - y, 0)^2. \quad (18b)$$

- $e_\tau(Y)$ is the unique solution of

$$\tau\,\mathrm{E}[\max(Y - x, 0)] = (1-\tau)\,\mathrm{E}[\max(x - Y, 0)] \quad (18c)$$

- e_τ is law-invariant and coherent for $\tau \geq 1/2$ (Bellini et al., 2013).

Properties of expectiles

- $e_{1/2}[Y] = \mathrm{E}[Y]$.
- e_τ is sensitive to extreme 'outliers'.
- $\mathrm{corr}[Y_1, Y_2] = 1 \;\Rightarrow\; e_\tau(Y_1 + Y_2) = e_\tau(Y_1) + e_\tau(Y_2)$
- But e_τ is **not comonotonically additive** for $\tau > 1/2$.
  - If e_τ were comonotonically additive then it would be a spectral measure.
  - By Corollary 4.3 of Ziegel (2013) the only elicitable spectral measure is the expectation. Hence $\tau = 1/2$ – contradiction!
- Hence, for non-linear dependence expectiles may see diversification where there is none.
- Risk contributions (conceptually easy to estimate):

$$e_\tau(L_i \,|\, L) \;=\; \frac{\tau\,\mathrm{E}[L_i\,\mathbf{1}_{\{L\ge e_\tau(L)\}}] + (1-\tau)\,\mathrm{E}[L_i\,\mathbf{1}_{\{L< e_\tau(L)\}}]}{\tau\,\mathrm{P}[L \ge e_\tau(L)] + (1-\tau)\,\mathrm{P}[L < e_\tau(L)]}. \qquad (19)$$

Comparison

- Expectiles:
  - Coherent, law-invariant and elicitable.
  - No obvious interpretation in terms of solvency.
  - May see diversification where there is none.
- Expected Shortfall:
  - Coherent, law-invariant and comonotonically additive.
  - Clearly related to solvency probability (via confidence level).
  - Not elicitable but composition of elicitable conditional expectation and quantile.
  - From (13):

$$\begin{aligned}\mathrm{ES}_\gamma(L) \approx 1/4\ \big(q_\gamma(L) + q_{0.75\,\gamma+0.25}(L) \\ + q_{0.5\,\gamma+0.5}(L) + q_{0.25\,\gamma+0.75}(L)\big)\end{aligned} \tag{20}$$

  - Hence backtest $q_\gamma(L)$, $q_{0.75\,\gamma+0.25}(L)$, $q_{0.5\,\gamma+0.5}(L)$, and $q_{0.25\,\gamma+0.75}(L)$ to backtest ES.

Outline

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

BCBS. *Fundamental review of the trading book: A revised market risk framework*. Basel Committee on Banking Supervision, October 2013.

F. Bellini, B. Klar, A. Müller, and E. Rosazza Gianin. Generalized quantiles as risk measures. Preprint, 2013. URL http://ssrn.com/abstract=2225751.

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

R.T. Rockafellar and S. Uryasev. The fundamental risk quadrangle in risk management, optimization and statistical estimation. *Surveys in Operations Research and Management Science*, 18(1):33–53, 2013.

S&P. Default, Transition, and Recovery: 2012 Annual Global Corporate Default Study And Rating Transitions. Report, Standard & Poor's, March 2013.

J. F. Ziegel. Coherence and elicitability. *ArXiv e-prints*, March 2013.