Expected Shortfall is not elicitable – so what?

Dirk Tasche

Bank of England – Prudential Regulation Authority

dirk.tasche@gmx.net

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

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

Background

Risk measures

Value-at-Risk and Expected Shortfall

Elicitability

Expected Shortfall and Expectiles

References
Outline

Background

Risk measures

Value-at-Risk and Expected Shortfall

Elicitability

Expected Shortfall and Expectiles

References
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

Background

Risk measures

Value-at-Risk and Expected Shortfall

Elicitability

Expected Shortfall and Expectiles

References
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
⇒ Use by creditors and regulators.

Deviation measures may be interpreted as measures of uncertainty
⇒ 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 \(\rightarrow\) double risk”):

\[ \rho(hL) = h \rho(L), \quad h \geq 0. \] (1a)

▶ Subadditivity (“reward diversification”):

\[ \rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2). \] (1b)

▶ Monotonicity (“higher losses imply higher risk”):

\[ L_1 \leq L_2 \quad \Rightarrow \quad \rho(L_1) \leq \rho(L_2). \] (1c)

▶ Translation invariance (“reserves reduce requirements”)

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

\(^2\)Convention: Losses are positive numbers, gains are negative.
Important and less important properties

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

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

⇒ Interpretation of coherent measures as expectations in stress scenarios.

- **Duality:** $\rho(L)$ solvency risk measure ⇒

$$\delta(L) = \rho(L) - \mathbb{E}[L] \text{ deviation measure}$$

- Homogeneity and subadditivity are preserved in $\delta$.
  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"):  
  \[ L_1 = f_1 \circ X, \ L_2 = f_2 \circ X \Rightarrow \rho(L_1 + L_2) = \rho(L_1) + \rho(L_2). \]  
  \[ (3a) \]
  \[ X \] common risk factor, \( f_1, f_2 \) increasing functions.

- **Law-invariance** ("context independence"):  
  \[ P[L_1 \leq \ell] = P[L_2 \leq \ell], \ \ell \in \mathbb{R} \Rightarrow \rho(L_1) = \rho(L_2). \]  
  \[ (3b) \]

- **Proposition**: Coherent risk measures \( \rho \) that are also law-invariant and comonotonically additive are **spectral measures**, i.e. there is a convex distribution function \( F_\rho \) on \([0, 1]\) such that  
  \[ \rho(L) = \int_0^1 q_u(L) F_\rho(du), \text{ for all } L. \]  
  \[ (3c) \]
  \[ q_u(L) = \min \{ P[L \leq \ell] \geq u \} \] 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:**

\[
L = \sum_{i=1}^{m} L_i.
\]  

(4)

*L*: portfolio-wide loss, *m*: number of risky positions in portfolio, \(L_i\): loss with \(i\)-th position.

- **Risk sensitivities**

\[
\rho(L_i \mid L) = \frac{d}{dh} \rho(L + h L_i) \bigg|_{h=0}
\]

are of interest for risk management and optimisation.

- \(\rho\) homogeneous and differentiable \(\Rightarrow\)

\[
\sum_{i=1}^{m} \rho(L_i \mid L) = \rho(L).
\]  

(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 $\Rightarrow$

$$\rho(L_i \mid L) \leq \rho(L_i) \iff \rho \text{ subadditive}$$

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

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

$$\rho(L) - \rho(L - L_i) \leq \rho(L_i \mid L) \quad (6)$$

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

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

$$\rho(L_i \mid L) = \int_0^1 \mathbb{E}[L_i \mid L = q_u(L)] \, F_\rho(du). \quad (7)$$
Outline

Background

Risk measures

Value-at-Risk and Expected Shortfall

Elicitability

Expected Shortfall and Expectiles

References
Shortfall probability risk measures

- Special case of solvency risk measures.
- **Construction principle:** For a given confidence level $\gamma$, 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

$$P[L > \rho(L)] \leq 1 - \gamma.$$  \hspace{1cm} (8)

- $\gamma$ 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).

Standard deviation

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

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

- By Chebychev’s inequality:

$$P[L > \sigma_a(L)] \leq P[|L - E[L]| > a \sqrt{\text{var}[L]}] \leq a^{-2}. (9b)$$

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

$$P[L > \sigma_a(L)] \leq \gamma. (9c)$$

- 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:

\[
\sigma_a(L_i \mid L) = a \frac{\text{cov}(L_i, L)}{\sqrt{\text{var}(L)}} + \text{E}[L_i].
\] (10)
Value-at-Risk

- For $\alpha \in (0, 1)$: $\alpha$-quantile $q_\alpha(L) = \min\{\ell : P[L \leq \ell] \geq \alpha\}$.
- In finance, $q_\alpha(L)$ is called **Value-at-Risk** (VaR).
- If $L$ has a continuous distribution (i.e. $P[L = \ell] = 0$, $\ell \in \mathbb{R}$), then $q_\alpha(L)$ is a solution of $P[L \leq \ell] = \alpha$.
- **Quantile / VaR-based** risk measure:
  \[ \text{VaR}_\alpha(L) = q_\alpha(L). \] (11a)
- By definition $\text{VaR}_\alpha(L)$ satisfies
  \[ P[L > \text{VaR}_\alpha(L)] \leq 1 - \alpha. \] (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:
  \[
  \text{VaR}_\alpha(L_i \mid L) = \mathbb{E}[L_i \mid L = q_\alpha(L)].
  \] (12)
- Estimation of risk contributions is difficult in continuous case.
Expected Shortfall

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

$$\text{ES}_\alpha(L) = \frac{1}{1-\alpha} \int_0^1 q_u(L) \, du$$

$$= \mathbb{E}[L \mid L \geq q_\alpha(L)]$$

$$+ \left( \mathbb{E}[L \mid L \geq q_\alpha(L)] - q_\alpha(L) \right) \left( \frac{\mathbb{P}[L \geq q_\alpha(L)]}{1-\alpha} - 1 \right).$$

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

$$\text{ES}_\alpha(L) = \mathbb{E}[L \mid L \geq q_\alpha(L)].$$

- ES dominates VaR: $\text{ES}_\alpha(L) \geq \text{VaR}_\alpha(L)$. 
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):
  \[ \text{ES}_\alpha(L_i | L) = E[L_i | L \geq q_\alpha(L)]. \] (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

Background

Risk measures

Value-at-Risk and Expected Shortfall

Elicitability

Expected Shortfall and Expectiles

References
Related definitions

- A **scoring function** is a function
  \[ s : \mathbb{R} \times \mathbb{R} \to [0, \infty), \ (x, y) \mapsto s(x, y), \quad (15a) \]
  where \( x \) and \( y \) are the *point forecasts* and *observations* respectively.

- Let \( \nu \) 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 \( \nu \) relative to \( \mathcal{P} \) if and only if
  \[ E_P [s(t, Y)] \leq E_P [s(x, Y)] \quad (15b) \]
  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
  \[ E_P [s(t, Y)] = E_P [s(x, Y)] \Rightarrow x \in \nu(P). \quad (15c) \]
Elicitability

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

**Examples:**

Expectation: $\nu(P) = \int x \, P(dx), \quad s(x, y) = (y - x)^2$.  
Quantiles: $\nu(P) = \{ x : P((-\infty, x]) \leq \alpha \leq P((-\infty, x]) \}$,  
$s(x, y) = \frac{\alpha}{1-\alpha} \max(y - x, 0) + \max(x - y, 0)$.  

**Interpretation:**

Point estimates of elicitable functionals can be determined by means of regression:

$$\nu(P) = \arg \min_x E_P[s(x, Y)], \quad Y \sim P.$$  

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 $\nu$ being elicitable (“convex level sets”):

$$0 < \pi < 1, \quad t \in \nu(P_1) \cap \nu(P_2) \Rightarrow t \in \nu(\pi P_1 + (1 - \pi) P_2) \quad (17a)$$

- By counter-examples: Standard deviation and ES violate (17a).
  $$\Rightarrow \text{Standard deviation and ES are not elicitable.}$$

- But standard deviation and ES can be calculated by means of regression, with $s$ as in (16a) and (16b):

$$\text{var}(P) = \min_x E_P[(Y - x)^2] \quad (17b)$$

$$\text{ES}_\alpha(P) = \min_x \left\{ E_P\left[\frac{\alpha}{1-\alpha} \max(Y - x, 0) \right] + \max(x - Y, 0) \right\} + E_P[Y] \quad (17c)$$
Outline

Background

Risk measures

Value-at-Risk and Expected Shortfall

Elicitability

Expected Shortfall and Expectiles

References
Expectiles

- For $0 < \tau < 1$ the $\tau$-expectile of square-integrable $Y$ is defined by
  \[
e_{\tau}(Y) = \arg \min_x E\left[\tau \max(Y - x, 0)^2 + (1 - \tau) \max(x - Y, 0)^2\right]\]  \hspace{1cm} (18a)

- $e_{\tau}$ is elicitable with scoring function
  \[
s(x, y) = \tau \max(y - x, 0)^2 + (1 - \tau) \max(x - y, 0)^2. \]  \hspace{1cm} (18b)

- $e_{\tau}(Y)$ is the unique solution of
  \[
  \tau E[\max(Y - x, 0)] = (1 - \tau) E[\max(x - Y, 0)] \]  \hspace{1cm} (18c)

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

Dirk Tasche (PRA)   ES is not elicitable – so what? 25 / 29
Properties of expectiles

- $e_{1/2}[Y] = \mathbb{E}[Y]$.
- $e_\tau$ is sensitive to extreme ‘outliers’.
- $\text{corr}[Y_1, Y_2] = 1 \Rightarrow e_\tau(Y_1 + Y_2) = e_\tau(Y_1) + e_\tau(Y_2)$
- But $e_\tau$ is **not comonotonically additive** for $\tau > 1/2$.
  - If $e_\tau$ 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):

$$e_\tau(L_i | L) = \frac{\tau \mathbb{E}[L_i \mathbf{1}_{\{L \geq e_\tau(L)\}}] + (1 - \tau) \mathbb{E}[L_i \mathbf{1}_{\{L < e_\tau(L)\}}]}{\tau \mathbb{P}[L \geq e_\tau(L)] + (1 - \tau) \mathbb{P}[L < e_\tau(L)]}. \quad (19)$$
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):
    \[
    \text{ES}_\gamma(L) \approx \frac{1}{4} (q_\gamma(L) + q_{0.75} \gamma + 0.25(L) \\
    + q_{0.5} \gamma + 0.5(L) + q_{0.25} \gamma + 0.75(L))
    \]
  - 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

Background

Risk measures

Value-at-Risk and Expected Shortfall

Elicitability

Expected Shortfall and Expectiles

References


