Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization

Yasuhiro Yamai and Toshinao Yoshiba

We compare expected shortfall with value-at-risk (VaR) in three aspects: estimation errors, decomposition into risk factors, and optimization.

We describe the advantages and the disadvantages of expected shortfall over VaR. We show that expected shortfall is easily decomposed and optimized while VaR is not. We also show that expected shortfall needs a larger size of sample than VaR for the same level of accuracy.

Key words: Expected shortfall; Value-at-risk; Optimization

Research Division I, Institute for Monetary and Economic Studies, Bank of Japan (E-mail: yasuhiro.yamai@boj.or.jp, toshinao.yoshiba@boj.or.jp)

The authors would like to thank Tsukasa Yamashita (BNP Paribas Tokyo) for his helpful comments.

I. Introduction

Artzner *et al.* (1997) proposed the use of expected shortfall to alleviate the problems inherent in value-at-risk (VaR). Expected shortfall considers loss beyond the VaR level and is shown to be sub-additive, while VaR disregards loss beyond the percentile and is not sub-additive.[1]

In this paper, we compare expected shortfall with VaR in three aspects: their estimation errors, their decomposition into risk factors, and their optimization. These comparative analyses can help risk managers consider the usefulness of expected shortfall as a risk management tool.

The rest of the paper is organized as follows. Section II gives our definitions and concepts of VaR and expected shortfall. Section III examines the estimation error of expected shortfall using Monte Carlo simulations. Section IV provides an example of estimating expected shortfall with sample portfolios. Section V describes a method of decomposing VaR and expected shortfall developed by Hallerbach (1999) and Tasche (2000). Section VI describes a method of optimizing portfolios based on expected shortfall developed by Rockafeller and Uryasev (2000). Section VII concludes the paper.

II. Definitions and Concepts of Expected Shortfall

Artzner *et al.* (1997) have proposed the use of expected shortfall (also called "conditional VaR," "mean excess loss," "beyond VaR," or "tail VaR") to alleviate the problems inherent in VaR. The expected shortfall is defined as follows.

Definition of expected shortfall
Suppose X is a random variable denoting the loss of a given portfolio and $VaR_\alpha(X)$ is the VaR at the $100(1-\alpha)$ percent confidence level.[2] $ES_\alpha(X)$ is defined by the following equation.[3]

$$ES_\alpha(X) = E[X \mid X \geq VaR_\alpha(X)]. \tag{1}$$

Expected shortfall measures how much one can lose on average in states beyond the VaR level. When the loss distribution is not normal, VaR disregards the loss beyond

1. A risk measure ρ is sub-additive when the risk of the total position is less than or equal to the sum of the risk of individual portfolios. Intuitively, sub-additivity requires that "risk measures should consider risk reduction by portfolio diversification effects."
Sub-additivity is defined as follows. Let X and Y be random variables denoting the losses of two individual positions. A risk measure ρ is sub-additive if the following equation holds.

$$\rho(X+Y) \leq \rho(X) + \rho(Y).$$

2. In this paper, VaR is defined as the upper 100α percentile of the loss distribution.
3. $E[x|B]$ is the conditional expectation of the random variable x given event B. Since X is defined as the loss, X is positive in loss and negative in profit.

the VaR level and fails to be sub-additive. Expected shortfall considers the loss beyond the VaR level and is shown to be sub-additive.

III. Estimation Error of Expected Shortfall

A. Concepts of Estimation Error

Estimates of VaR and expected shortfall are affected by estimation error, the natural sampling variability due to limited sample size. For example, consider a situation where we estimate the VaR of a given portfolio by Monte Carlo simulations. The VaR estimates vary according to the realizations of random numbers. To reduce estimation error, risk managers must increase the sample size of the simulations.

This section compares the estimation errors of expected shortfall and VaR, and considers whether more calculation time is needed when estimating expected shortfall than when estimating VaR.

B. Estimation Error under Stable Distribution

In this subsection, we compare the estimation errors of VaR and expected shortfall by simulating random variables with stable distributions.

When a random variable X obeys the stable distribution,[4] there exist constants α and γ_n such that

$$S_n \overset{d}{=} n^{1/\alpha}X + \gamma_n, \tag{2}$$

where S_n is the sum of independently and identically distributed n copies of X.[5] α is the index of stability. The smaller α is, the heavier the tail of the distribution. If α = 2, the stable distribution reduces to the normal, and it reduces to Cauchy[6] if α = 1 (Figure 1). The stable distribution is a generalization of the normal in that a sum of stable random variables is also a stable random variable.

We evaluate the estimation errors of VaR and expected shortfall as follows. First, we run 10,000 sets of Monte Carlo simulations with a sample size of 1,000, assuming that the underlying loss distributions are stable with α = 2.0, 1.9, $\cdots$, 1.2, 1.1.[7,8]

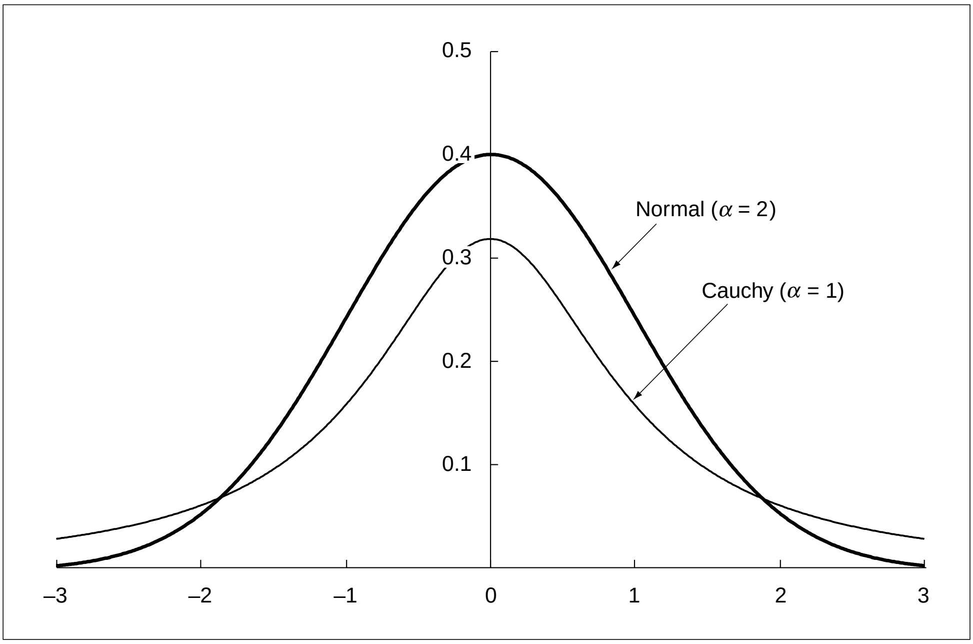
4. For details of the stable distribution, see Feller (1969) and Shiryaev (1999).
5. $\overset{d}{=}$ denotes equality in distribution.
6. The first moment of Cauchy distribution (α = 1) is infinite. Therefore, when the loss obeys Cauchy distribution, one cannot define expected shortfall since it is the conditional expectation of loss given that the loss is in the right tail of the loss distribution.
7. Stable random variables are commonly described by the following characteristic functions:

$$\Phi(\theta) = \begin{cases} \exp\{-\sigma^{\alpha}|\theta|^{\alpha}(1 - i\beta(\operatorname{sgn}\theta)\tan\frac{\pi\alpha}{2}) + i\mu\theta\}, & \text{if } \alpha \neq 1, \\ \exp\{-\sigma|\theta|(1 + i\beta\frac{2}{\pi}(\operatorname{sgn}\theta)\ln\theta) + i\mu\theta\}, & \text{if } \alpha = 1, \end{cases}$$

where α is the index of stability, β is the skewness parameter, σ is the scale parameter, and μ is the location parameter.

In this section, we set β = 0, μ = 0, and $\sigma = 1/\sqrt{2}$. We set $\sigma = 1/\sqrt{2}$ so that the loss distribution reduces to the standard normal when α = 2.
8. We obtained uniform random numbers with Mersenne Twister, and transformed them into stable random numbers with the algorithm developed by Chambers, Mallows, and Stuck (1976).

Figure 1 Comparison of Normal and Cauchy Density Functions



Second, we obtain 10,000 estimates[9] of VaR and expected shortfall from those sets of simulations, and calculate the average, the standard deviation, and the 95 percent confidence level of those estimates.[10] The estimation errors of VaR and expected shortfall are compared by the relative standard deviation (the standard deviation divided by the average). Tables 1–2 and Figures 2–5 show the results.

The estimation error of expected shortfall is larger than that of VaR when the underlying loss distribution is fat-tailed. As α gets closer to one (i.e., as the underlying loss distribution becomes fat-tailed), the relative standard deviation of the expected shortfall estimate becomes much larger than that of the VaR estimate. On the other hand, when α is close to two (i.e., when the underlying loss distribution is approximately normal), the relative standard deviation of VaR and expected shortfall estimates are almost equal.

9. The estimator of VaR at the 100(1 – α) percent confidence interval is the upper 100α percent quantile of the empirical loss distribution. We take the VaR estimator as the ($n\alpha$ + 1)th largest sample of loss, where n is the sample size. That is, we take $X_{(n\alpha+1)}$ as the VaR estimator where the sequence $X_{(n)}, X_{(n-1)}, \cdots, X_{(n\alpha+1)}, X_{(n\alpha)}, \cdots, X_{(1)}$ is the loss sample rearranged in increasing order. We take the following as the expected shortfall estimator.

$$ES_\alpha = \frac{X_{(1)} + X_{(2)} + \cdots + X_{(n\alpha+1)}}{n\alpha + 1}.$$

10. The asymptotic standard deviation of the VaR estimate can be obtained in closed form. Furthermore, there is a closed-form formula that approximates the standard deviation of the expected shortfall estimate (see Appendix 1 for details). When the underlying loss distribution is relatively thin-tailed (such as the normal and t distributions), those closed-form formulas give almost equal numbers to those calculated by Monte Carlo simulation. On the other hand, when the underlying distribution is fat-tailed (such as in a Pareto distribution), they give substantially different numbers.

Table 1 Estimates of VaR and Expected Shortfall with Stable Distribution (Confidence Level: 95 Percent)

α	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
2.0 (normal)	VaR	1.64	0.07	0.04	[1.51 1.77]
	Expected shortfall	2.05	0.08	0.04	[1.90 2.21]
1.9	VaR	1.70	0.08	0.04	[1.55 1.85]
	Expected shortfall	2.42	0.80	0.33	[2.06 3.14]
1.8	VaR	1.77	0.09	0.05	[1.60 1.95]
	Expected shortfall	2.90	1.81	0.63	[2.28 4.20]
1.7	VaR	1.86	0.11	0.06	[1.67 2.08]
	Expected shortfall	3.53	3.84	1.09	[2.58 5.60]
1.6	VaR	1.98	0.13	0.07	[1.75 2.26]
	Expected shortfall	4.39	8.34	1.90	[2.96 7.62]
1.5	VaR	2.15	0.16	0.08	[1.86 2.50]
	Expected shortfall	5.67	19.31	3.41	[3.48 10.71]
1.4	VaR	2.38	0.21	0.09	[2.02 2.82]
	Expected shortfall	7.71	48.95	6.35	[4.16 15.76]
1.3	VaR	2.68	0.26	0.10	[2.22 3.25]
	Expected shortfall	11.46	139.60	12.19	[5.10 25.13]
1.2	VaR	3.08	0.34	0.11	[2.49 3.85]
	Expected shortfall	19.79	463.10	23.40	[6.48 42.45]
1.1	VaR	3.65	0.46	0.13	[2.86 4.67]
	Expected shortfall	44.41	1,866.40	42.03	[8.59 81.44]

Table 2 Estimates of VaR and Expected Shortfall with Stable Distribution (Confidence Level: 99 Percent)

α	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
2.0 (normal)	VaR	2.30	0.12	0.05	[2.09 2.54]
	Expected shortfall	2.62	0.14	0.05	[2.36 2.90]
1.9	VaR	2.57	0.20	0.08	[2.25 3.03]
	Expected shortfall	3.94	3.68	0.93	[2.70 7.02]
1.8	VaR	3.00	0.35	0.12	[2.47 3.86]
	Expected shortfall	5.58	8.36	1.50	[3.27 11.25]
1.7	VaR	3.61	0.55	0.15	[2.78 4.94]
	Expected shortfall	7.70	17.74	2.30	[4.05 16.84]
1.6	VaR	4.40	0.78	0.18	[3.23 6.29]
	Expected shortfall	10.66	38.62	3.62	[5.01 25.03]
1.5	VaR	5.41	1.08	0.20	[3.81 8.00]
	Expected shortfall	15.16	89.50	5.91	[6.31 37.93]
1.4	VaR	6.76	1.49	0.22	[4.56 10.37]
	Expected shortfall	22.76	226.92	9.97	[8.02 60.08]
1.3	VaR	8.63	2.10	0.24	[5.58 13.64]
	Expected shortfall	37.59	647.21	17.22	[10.39 100.13]
1.2	VaR	11.34	3.04	0.27	[7.00 18.77]
	Expected shortfall	72.74	2,147.04	29.52	[13.90 176.21]
1.1	VaR	15.53	4.63	0.30	[9.09 26.85]
	Expected shortfall	181.77	8,653.26	47.61	[19.63 351.63]

Figure 2 Relative Standard Deviation of Estimates (Confidence Level: 95 Percent)

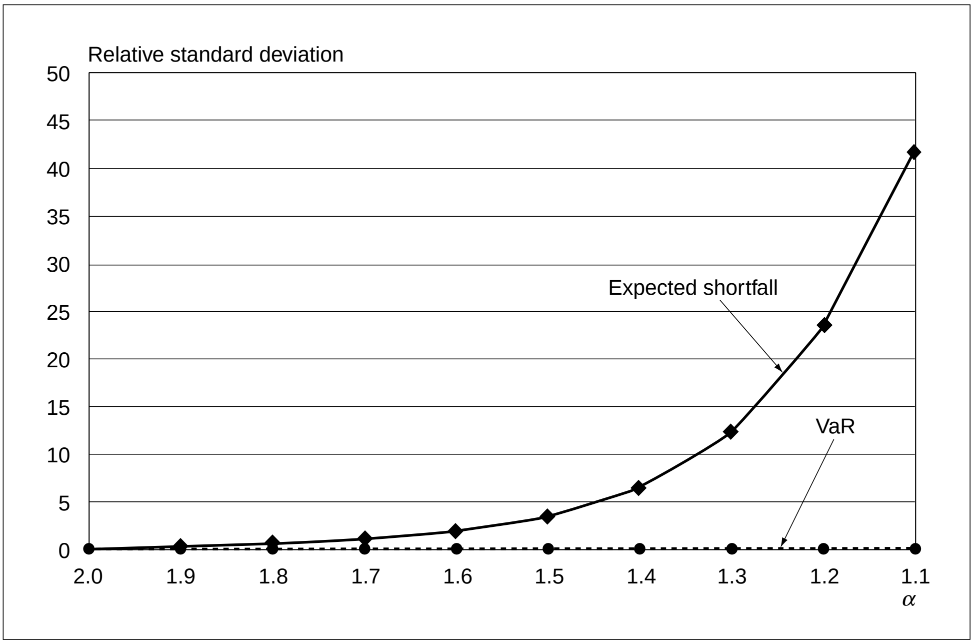


Figure 3 Relative Standard Deviation of Estimates (Confidence Level: 95 Percent, Enlarged 1.9 ≤ α ≤ 2)

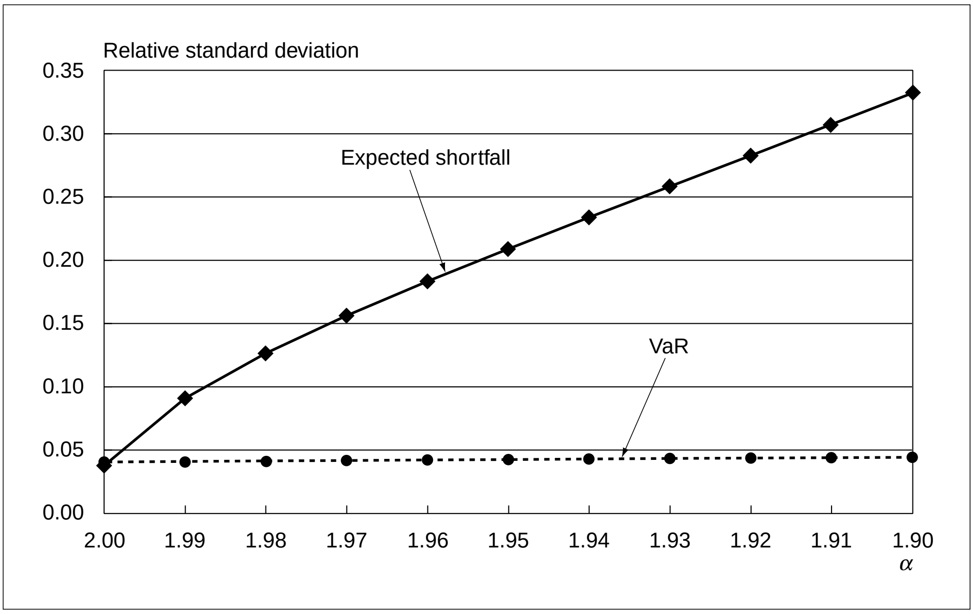


Figure 4 Relative Standard Deviation of Estimates (Confidence Level: 99 Percent)

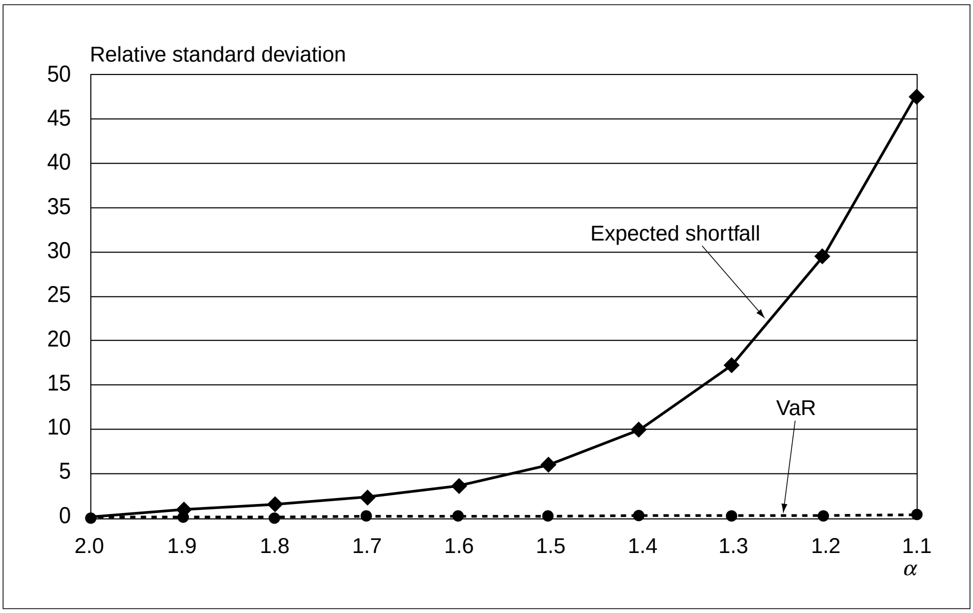
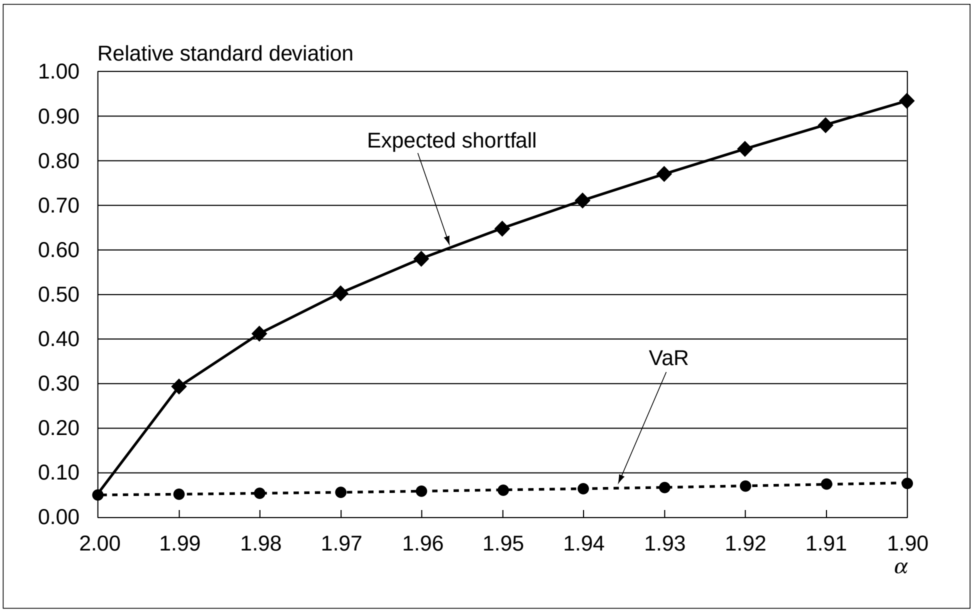


Figure 5 Relative Standard Deviation of Estimates (Confidence Level: 99 Percent, Enlarged 1.9 ≤ α ≤ 2)



This result can be explained as follows. When the underlying distribution is fat-tailed, the probability of infrequent and large loss is high. The expected shortfall estimates are affected by whether large and infrequent loss is realized in the obtained sample, since expected shortfall considers the right tail of the loss distribution. On the other hand, the VaR estimates are less affected by large and infrequent loss than the expected shortfall estimates, since the VaR method disregards loss beyond the VaR level. Therefore, when the underlying loss distribution becomes more fat-tailed, the expected shortfall estimates become more varied due to infrequent and large loss, and their estimation error becomes larger than the estimation error of VaR.

Furthermore, we investigate whether the increase in sample size reduces the estimation error of expected shortfall. We run 10,000 sets of Monte Carlo simulations with sample sizes of 1,000, 10,000, and 100,000 each, assuming that the underlying loss distributions are stable with α = 2.0, 1.5, 1.1. We calculate the average, the standard deviation, and the 95 percent confidence interval of those 10,000 estimates. Tables 3–4 and Figure 6 show the results.

The increase in sample size from 1,000 to 100,000 reduces the relative standard deviations (the standard deviation divided by the average) of the expected shortfall estimates.[11] Therefore, we are able to reduce the estimation error of expected shortfall by increasing sample size.[12]

Table 3 Convergence of Expected Shortfall Estimates under Stable Distributions (Confidence Level: 95 Percent)

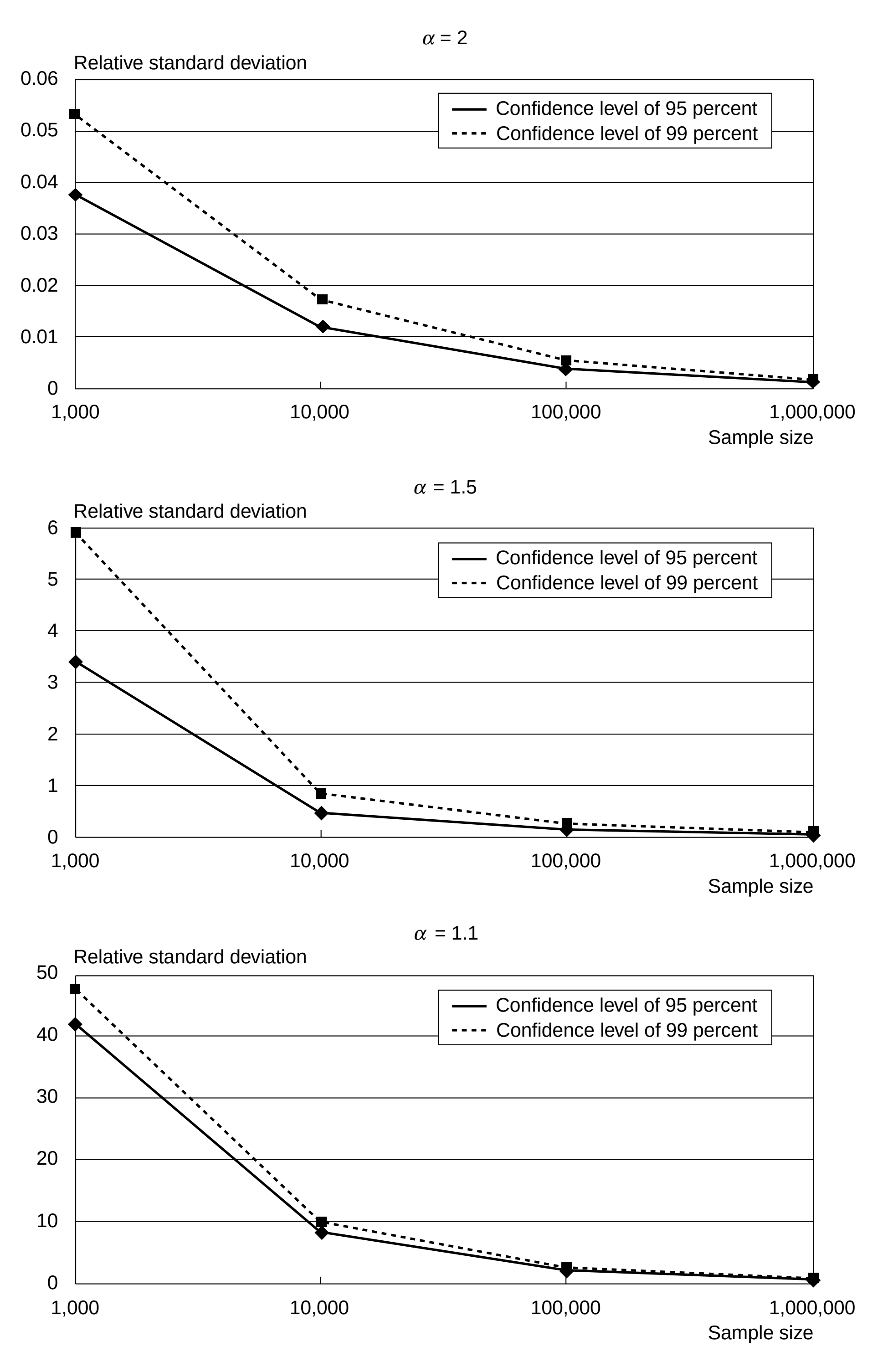
Sample size	α = 2.0		α = 1.5		α = 1.1	
	Relative standard deviation	Confidence interval (95 percent)	Relative standard deviation	Confidence interval (95 percent)	Relative standard deviation	Confidence interval (95 percent)
1,000	0.04	[1.90 2.21]	3.41	[3.48 10.71]	42.03	[8.59 81.44]
10,000	0.01	[2.01 2.11]	0.47	[4.51 8.01]	8.29	[14.02 75.20]
100,000	0.00	[2.05 2.08]	0.15	[5.09 6.91]	2.07	[18.80 76.69]
1,000,000	0.00	[2.06 2.07]	0.05	[5.41 6.28]	0.64	[22.64 71.65]

Table 4 Convergence of Expected Shortfall Estimates under Stable Distributions (Confidence Level: 99 Percent)

Sample size	α = 2.0		α = 1.5		α = 1.1	
	Relative standard deviation	Confidence interval (95 percent)	Relative standard deviation	Confidence interval (95 percent)	Relative standard deviation	Confidence interval (95 percent)
1,000	0.05	[2.36 2.90]	5.91	[6.31 37.93]	47.61	[19.63 351.63]
10,000	0.02	[2.57 2.75]	0.84	[10.40 27.33]	9.91	[43.86 346.70]
100,000	0.01	[2.64 2.69]	0.26	[13.07 22.08]	2.50	[66.87 356.04]
1,000,000	0.00	[2.67 2.67]	0.10	[14.58 18.96]	0.78	[85.99 330.92]

11. Table 3 shows that, when the underlying loss distribution is stable with α = 1.5, we must have a sample size of somewhere between several hundred thousand and one million to ensure the same level of relative standard deviation as occurs when we estimate VaR with a sample size of 1,000 (0.08).
12. This result is consistent with proposition 3.1 of Acerbi and Tasche (2001), which says that the expected shortfall estimate converges with probability one as sample size tends to infinity.

Figure 6 Convergence of Expected Shortfall Estimates



IV. Examples of Estimating Expected Shortfall

This section gives examples of estimating expected shortfall, and compares the estimation error of expected shortfall with that of VaR. We consider two cases: an option portfolio and a credit portfolio.

A. Equity Option Portfolio

This subsection treats a sample portfolio consisting of three issues of U.S. stocks (General Electric, McDonald's, and Intel) and short positions on options whose underlying securities are those U.S. stocks (Table 5).

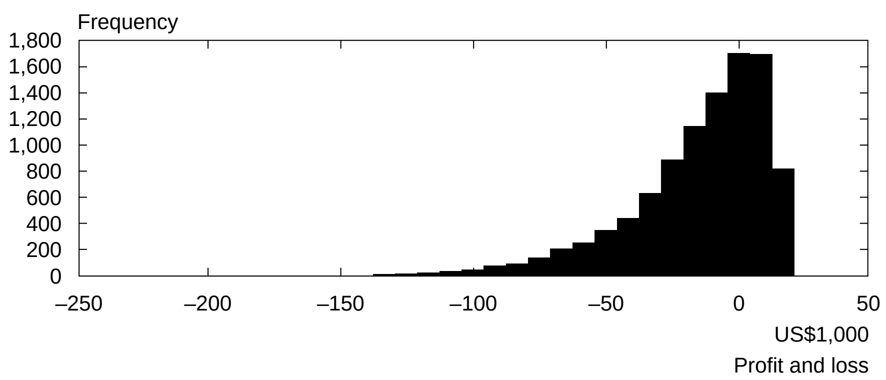
Table 5 Sample Option Portfolio

	Days to maturity	Strike price (US$)	Price (premium) (US$)	Dollar amount (US$1,000)	Position (1,000 units)
General Electric stock	—	—	49-13/16	1,000	20.1
McDonald's stock	—	—	31-1/4	1,000	32.0
Intel stock	—	—	42-1/32	1,000	23.8
Call option on GE stock	18	50.00	2	–100	–50.0
Call option on McDonald's stock	18	30.00	1-3/4	–100	–57.1
Call option on Intel stock	18	40.00	3-1/2	–100	–28.6

Source: Bloomberg L.P. as of November 28, 2000.

We estimate the VaR and expected shortfall of this sample portfolio. We assume that the log returns of the stocks obey the multivariate normal distribution, and estimate the variance-covariance matrix of those log returns from historical data. We set the holding period to be one day, and assume that implied volatility is constant throughout this period (see Appendix 2 for details). Figure 7 shows the profit and loss distribution of the sample portfolio. The distribution is skewed to the left because of the substantial short positions on call options.

Figure 7 Profit and Loss Distribution of Sample Option Portfolio (Sample Size: 10,000)



Note: The profit (loss) is shown as a positive (negative) number.

We evaluate the estimation errors of VaR and expected shortfall as follows. We run 10,000 sets of Monte Carlo simulations with a sample size of 1,000, and calculate the average, the standard deviation, and the 95 percent confidence interval of the VaR and expected shortfall estimates. Tables 6–7 show the result.

Table 6 Estimates of VaR and Expected Shortfall of Sample Portfolio (Confidence Level: 95 Percent)

Sample size	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
1,000	VaR	68.33	3.34	0.0489	[61.72 75.03]
	Expected shortfall	91.20	4.63	0.0508	[82.44 100.30]
10,000	VaR	68.18	1.05	0.0154	[66.10 70.22]
	Expected shortfall	91.56	1.38	0.0151	[88.97 94.26]
100,000	VaR	68.15	0.33	0.0049	[67.51 68.81]
	Expected shortfall	91.57	0.46	0.0050	[90.73 92.50]

Table 7 Estimates of VaR and Expected Shortfall of Sample Portfolio (Confidence Level: 99 Percent)

Sample size	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
1,000	VaR	107.00	7.50	0.0701	[93.16 123.31]
	Expected shortfall	127.10	9.47	0.0745	[109.94 146.46]
10,000	VaR	106.23	2.28	0.0215	[101.82 110.81]
	Expected shortfall	128.35	3.10	0.0242	[122.46 134.68]
100,000	VaR	105.97	0.71	0.0067	[104.53 107.37]
	Expected shortfall	128.28	0.99	0.0078	[126.30 130.31]

The estimation errors of VaR and expected shortfall are almost equal. This result is similar to the result in Section III in which the underlying distribution is found to be normal. The right tail of the loss distribution of this sample option portfolio is similar to the normal, since the strike prices of options are close to at-the-money.

Next, we consider a sample portfolio with far-out-of-the-money options (Table 8). The only difference between this portfolio and the sample portfolio in Table 5 is that

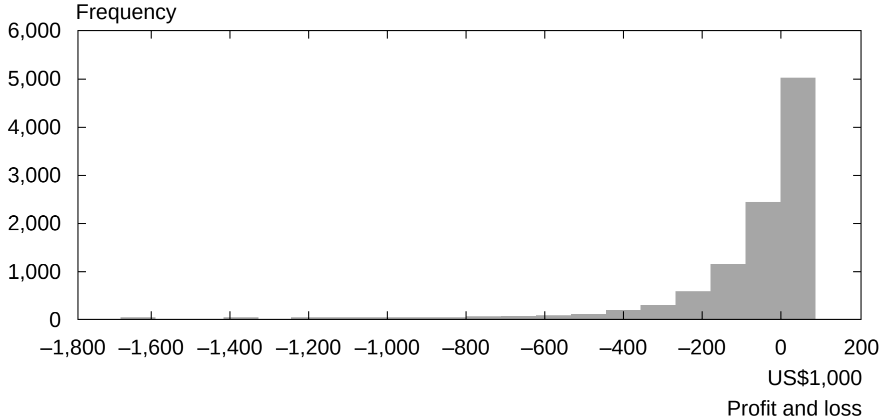
Table 8 Sample Portfolio with Far-Out-of-the-Money Options

	Days to maturity	Strike price (US$)	Price (premium) (US$)	Dollar amount (US$1,000)	Position (1,000 units)
General Electric stock	—	—	49-13/16	1,000	20.1
McDonald's stock	—	—	31-1/4	1,000	32.0
Intel stock	—	—	42-1/32	1,000	23.8
Call option on GE stock	18	58-3/8	3/16	–100	–533.3
Call option on McDonald's stock	18	40	1/16	–100	–1,600.0
Call option on Intel stock	18	55	1/16	–100	–1,600.0

Source: Bloomberg L.P. as of November 28, 2000.

the strike prices of options are far-out-of-the-money[13] in this portfolio. Figure 8 shows the profit and loss distribution of the portfolio. The distribution is more skewed to the left than the profit and loss distribution in Figure 7 because it includes large positions in far-out-of-the-money options.

Figure 8 Profit and Loss Distribution of Sample Portfolio with Far-Out-of-the-Money Options (Sample Size: 10,000)



Note: The profit (loss) is shown as a positive (negative) number.

We estimate the VaR and expected shortfall of this sample portfolio by following the same steps that we used when we estimated the VaR and expected shortfall of the sample portfolio in Table 5. Tables 9–10 show the results.

The estimation error of expected shortfall is larger than that of VaR when the strike prices of the options are far-out-of-the-money. This is because the underlying loss distribution becomes fat-tailed when the strike prices of the options are made far-out-of-the-money.

Table 9 Estimates of VaR and Expected Shortfall of Sample Portfolio with Far-Out-of-the-Money Options (Confidence Level: 95 Percent)

Sample size	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
1,000	VaR	334.62	22.01	0.0658	[293.70 380.86]
	Expected shortfall	500.46	35.00	0.0699	[433.38 571.32]
10,000	VaR	332.56	7.06	0.0212	[318.53 346.88]
	Expected shortfall	502.01	11.17	0.0222	[481.57 524.05]
100,000	VaR	332.26	2.15	0.0065	[328.14 336.70]
	Expected shortfall	502.19	3.45	0.0069	[495.53 509.05]

13. As we assume that the log returns of the stocks obey the multivariate normal distribution and that the volatility is constant, the probability that the stock prices rise beyond the strike prices of options during the holding period is less than 0.01 percent.

Table 10 Estimates of VaR and Expected Shortfall of Sample Portfolio with Far-Out-of-the-Money Options (Confidence Level: 99 Percent)

Sample size	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
1,000	VaR	612.56	58.08	0.0948	[506.46 739.03]
	Expected shortfall	781.31	85.48	0.1094	[625.85 964.29]
10,000	VaR	602.79	18.42	0.0306	[566.84 639.76]
	Expected shortfall	790.73	27.46	0.0347	[738.77 849.00]
100,000	VaR	602.25	5.57	0.0093	[591.23 613.17]
	Expected shortfall	792.14	9.13	0.0115	[774.75 810.97]

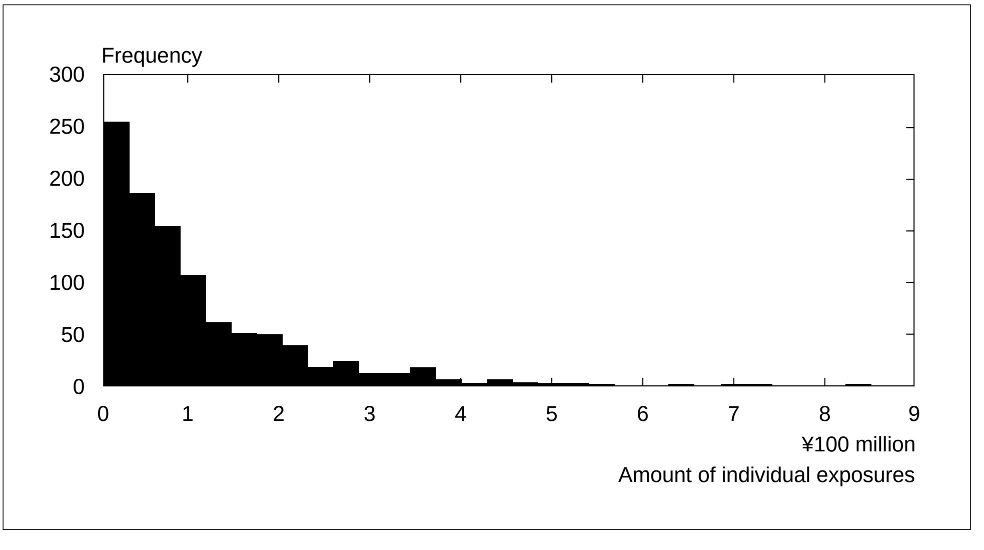
B. Loan Portfolio

This subsection deals with a sample loan portfolio. The sample portfolio (Table 11) consists of 1,000 loans with homogeneous default rates of 1 percent or 0.1 percent. The individual loan amounts obey the exponential distribution with an average of ¥100 million (see Figure 9 for the distribution of the loan amount). The correlation

Table 11 Sample Loan Portfolio

Amount of exposures	¥100 billion
Number of exposures	1,000
Amount of individual exposures	Obeys exponential distribution with average of ¥100 million
Definition of loss	Default mode (recognizes loss only if the borrower defaults during the risk evaluation period)
Recovery rate	Zero
Correlation of default events	The correlation coefficients between default events are assumed to be homogeneous at 0.00, 0.03, and 0.05.

Figure 9 Distribution of Amount of Individual Exposures



coefficients between default events are homogeneous at 0.00, 0.03, and 0.05. We adopt "default mode" as a loss definition, that is, we recognize loss only if the borrower defaults during the risk evaluation period. We estimate VaR and expected shortfall by following the algorithm developed by Ieda, Marumo, and Yoshiba (2000) (see Appendix 3 for details).

We run 1,000 sets of Monte Carlo simulations with a sample size of 1,000, and calculate the average, the standard deviation, and the 95 percent confidence interval of the VaR and expected shortfall estimates. We evaluate the estimation errors of VaR and expected shortfall by using the relative standard deviation. Tables 12–15 show the results.

Table 12 Estimates of VaR and Expected Shortfall of Sample Loan Portfolio (Confidence Level: 95 Percent, Default Rate: 1 Percent, Sample Size: 1,000)

Correlation coefficients of default events	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
0.00	VaR	18.28	0.43	0.0233	[17.39 19.10]
	Expected shortfall	20.99	0.52	0.0248	[20.03 22.02]
0.03	VaR	41.03	3.13	0.0763	[35.05 47.45]
	Expected shortfall	69.09	5.98	0.0865	[57.71 81.80]
0.05	VaR	45.79	4.37	0.0955	[37.95 54.62]
	Expected shortfall	86.16	8.79	0.1021	[71.37 104.95]

Table 13 Estimates of VaR and Expected Shortfall of Sample Loan Portfolio (Confidence Level: 99 Percent, Default Rate: 1 Percent, Sample Size: 1,000)

Correlation coefficients of default events	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
0.00	VaR	22.65	0.79	0.0350	[21.21 24.33]
	Expected shortfall	24.90	1.02	0.0408	[22.99 26.99]
0.03	VaR	85.03	9.79	0.1151	[67.39 106.34]
	Expected shortfall	117.55	15.90	0.1353	[89.42 151.08]
0.05	VaR	108.34	14.31	0.1321	[83.53 141.34]
	Expected shortfall	158.03	23.53	0.1489	[118.87 208.67]

Table 14 Estimates of VaR and Expected Shortfall of Sample Loan Portfolio (Confidence Level: 95 Percent, Default Rate: 0.1 Percent, Sample Size: 1,000)

Correlation coefficients of default events	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
0.00	VaR	3.99	0.20	0.0508	[3.62 4.38]
	Expected shortfall	5.45	0.28	0.0522	[4.93 6.04]
0.03	VaR	4.72	0.71	0.1500	[3.50 6.37]
	Expected shortfall	15.54	3.15	0.2029	[10.26 22.52]
0.05	VaR	3.92	0.72	0.1826	[2.74 5.53]
	Expected shortfall	17.32	4.23	0.2444	[10.74 27.30]

Table 15 Estimates of VaR and Expected Shortfall of Sample Loan Portfolio (Confidence Level: 99 Percent, Default Rate: 0.1 Percent, Sample Size: 1,000)

Correlation coefficients of default events	Risk measures	Average (a)	Standard deviation (b)	Relative standard deviation (c) = (b)/(a)	Confidence interval (95 percent)
0.00	VaR	6.37	0.48	0.0757	[5.53 7.37]
	Expected shortfall	7.70	0.59	0.0769	[6.59 8.89]
0.03	VaR	19.40	4.25	0.2189	[12.74 29.76]
	Expected shortfall	39.62	11.45	0.2891	[22.20 66.87]
0.05	VaR	21.11	5.43	0.2574	[12.83 33.21]
	Expected shortfall	49.32	16.06	0.3257	[25.82 88.95]

The estimation error of expected shortfall is larger than that of VaR when the default rate is low and the default correlation is high. Table 12 shows that, when the default rate is 1 percent, the estimation errors of VaR and expected shortfall at the 95 percent confidence level are almost equal. On the other hand, Table 13 shows that, at the 99 percent confidence level, the estimation error of expected shortfall is larger than that of VaR when the default correlation is high. Table 14 shows that, when the default rate is 0.1 percent and the default correlation is high, the estimation error of expected shortfall is larger than that of VaR at both the 95 percent and 99 percent confidence levels.

The estimates of expected shortfall vary more than those of VaR with low default rates because of the low frequency of portfolio loss and limited sample size. When the loss frequency declines, the estimation of expected shortfall requires a larger sample to ensure the same level of accuracy, since expected shortfall considers the loss in the tail as a conditional expectation. Thus, expected shortfall varies more than VaR at low default rates if we estimate it with the same sample size.

The estimation error of expected shortfall is large when the default correlation is high because of the fat tail of the underlying loss distribution. Ieda, Marumo, and Yoshiba (2000) show that the underlying loss distribution is fat-tailed when the default correlation is high. Furthermore, Subsection III.B shows that the estimation error of expected shortfall is larger than that of VaR when the underlying loss distribution is fat-tailed.

V. Decomposition of VaR and Expected Shortfall

This section describes a method of decomposing portfolio risk into risk factors.[14] The decomposition of risk is a useful tool for managing portfolio risk. For example, risk decomposition enables risk managers to select assets that provide the best risk-return trade-off, or to allocate "economic capital" to individual risk factors.

14. The concept of VaR decomposition was proposed by Garman (1997), whose terminology we followed in using "marginal VaR" and "component VaR."

We describe the method of decomposing VaR and expected shortfall developed by Hallerbach (1999) and Tasche (2000). We also give an example of decomposing VaR and expected shortfall with the sample option portfolio in Subsection V.A.

We show that it is more straightforward to decompose expected shortfall than to decompose VaR.

A. Decomposing VaR

Hallerbach (1999) and Tasche (2000) developed a method of decomposing simulation-based VaR and expected shortfall into individual risk factors. This subsection describes this method following Tasche (2000).

We assume that the portfolio loss X is a linear combination of the losses of individual risk factors X_i (i denotes risk factors):

$$X = \sum_{i=1}^{n} X_i \omega_i. \tag{3}$$

X : portfolio loss
X_i: loss of individual risk factor i[15]
ω_i: sensitivity to individual risk factor i

Since the portfolio VaR is a linearly homogeneous function of sensitivity to individual risk factors, the following equality holds.

$$VaR_\alpha(X) = \sum_{i=1}^{n} \frac{\partial VaR_\alpha(X)}{\partial \omega_i} \cdot \omega_i. \tag{4}$$

Therefore, the portfolio VaR is decomposed[16] into $\partial VaR/\partial\omega_i$ multiplied by the risk factor ω_i.

Tasche (2000) proved the following under certain conditions.[17]

Marginal VaR

The partial derivative of VaR at the $100(1-\alpha)$ percent confidence level with respect to ω_i (we call this "marginal VaR," which is denoted by $M-VaR_i$) is represented as a conditional expectation as follows (Tasche [2000], remark 5.4):

$$M-VaR_i = \frac{\partial VaR_\alpha(X)}{\partial \omega_i} = E[X_i | X = VaR_\alpha(X)]. \tag{5}$$

15. For example, consider the situation where you have 1,000 shares of a stock whose current market price is US$1,200/share and whose original cost is US$1,500/share. The loss of the stock X_i (considered here as one of a number of individual factors) is US$300/share, and sensitivity to this stock ω_i is 1,000 shares. Furthermore, if the market price of this stock is US$2,000/share, X_i is equal to –US$500/share.
16. When X is a nonlinear function of X_i, equation (4) does not hold. To deal with options, we consider option premiums additional risk factors.
17. The condition includes the continuity of the distributions and the integrability of expectations.

Thus, equation (6) provides a method to decompose VaR.

$$VaR_\alpha(X) = \sum_{i=1}^{n} \frac{\partial VaR_\alpha(X)}{\partial \omega_i} \cdot \omega_i = \sum_{i=1}^{n} E[X_i | X = VaR_\alpha(X)] \cdot \omega_i. \qquad (6)$$

Considering $(\partial VaR_\alpha(X)/\partial \omega_i) \cdot \omega_i$ the contribution of risk factor i to the portfolio VaR,[18] we define component VaR as follows.

Component VaR
The contribution of risk factor i to the portfolio VaR (we call this "component VaR," which is denoted by $C - VaR_i$) is defined as follows:

$$C - VaR_i = \frac{\partial VaR_\alpha(X)}{\partial \omega_i} \cdot \omega_i = E[X_i | X = VaR_\alpha(X)] \cdot \omega_i. \qquad (7)$$

It is not straightforward to estimate the right-hand side of equations (5) and (7) when we calculate VaR by simulations. It is difficult to estimate a conditional expectation conditioned by the equality $X = VaR_\alpha(X)$ when the distribution is discrete.

Hallerbach (1999) proposed and evaluated several methods of estimating this conditional expectation approximately. He concluded that the "conditional mean model" provides the best result. This method chooses a data window whose portfolio losses are close to the level of VaR, and takes the mean loss of this window to obtain an estimation of the conditional expectation.

B. An Example of Decomposing VaR

This subsection gives an example of decomposing the VaR of the sample option portfolio in Subsection IV.A. We adopt the "conditional mean model" proposed by Hallerbach (1999) in estimating marginal VaR.

We estimate marginal VaR as follows. Suppose we obtain N samples from a simulation. We choose a data window whose portfolio losses are close to the portfolio VaR level. This window is chosen as

$$X \in [VaR_\alpha(X) - \varepsilon_d,\ VaR_\alpha(X) + \varepsilon_u], \qquad (8)$$

for some small positive ε_d and ε_u. Suppose $X^j (1 \le j \le N)$ denotes the portfolio loss of the j-th sample, $X_i^j (1 \le i \le 6,\ 1 \le j \le N)$ denotes the loss of the i-th risk factor of the j-th sample, and T denotes the number of samples in the chosen data window. The marginal VaR of individual risk factor i is estimated by

18. Component VaR approximates how the portfolio VaR would change if the corresponding risk factor were deleted from the portfolio. This approximation works well when the risk factor makes a relatively small contribution to the portfolio VaR. However, we should note that component VaR is defined using marginal VaR, which is the "marginal" change in VaR with respect to ω_i. This means that this approximation does not work well when the contribution of the risk factor is large.

$$M - VaR_i = \frac{\partial VaR_\alpha(X)}{\partial \omega_i} = E[X_i | X = VaR_\alpha(X)] \cong \frac{1}{T} \sum_j X_i^j, \tag{9}$$

where we take the sum only for the data included in the chosen window.

We calculate VaR at the 95 percent confidence level of the sample option portfolio in Subsection IV.A by using a Monte Carlo simulation with sample size of 10,000, and decompose VaR into risk factors.[19] Table 16 shows the result.

We evaluate the estimation error of this method by comparing marginal VaR estimated by equation (9) with "recalculated marginal VaR" obtained by re-estimating portfolio VaR for a slightly changed portfolio (we use 0.1 percent, 0.5 percent, and 1 percent changes). Table 17 shows the result.[20]

The marginal VaR estimated by equation (9) differs from recalculated marginal VaR. This difference is especially apparent in Intel stock and in the call option on Intel stock, where the signs of those numbers are opposite. Therefore, equation (9) is not necessarily an accurate estimator of marginal VaR.

Table 16 VaR Decomposition of Sample Option Portfolio (Confidence Level: 95 Percent)

	Marginal VaR (US$/unit) (a)	Investment amount (US$1,000)	Position (1,000 units) (b)	Component VaR (US$1,000) (c) = (a) × (b)
General Electric stock	–2.30	1,000	20.1	–46.12
McDonald's stock	–1.67	1,000	32.0	–53.58
Intel stock	0.60	1,000	23.8	14.17
Call option on GE stock	–1.55	–100	–50.0	77.50
Call option on McDonald's stock	–1.53	–100	–57.1	87.43
Call option on Intel stock	0.37	–100	–28.6	–10.71
Total	—	2,700	—	68.70

Table 17 Comparison of Marginal VaR (Confidence Level: 95 Percent)

	Marginal VaR (US$/unit)	Recalculated marginal VaR (US$/unit)		
		Change in position: 0.1 percent	Change in position: 0.5 percent	Change in position: 1 percent
General Electric stock	–2.30	–1.20	–0.44	–1.60
McDonald's stock	–1.67	–2.75	–2.55	–1.86
Intel stock	0.60	–0.37	0.33	0.46
Call option on GE stock	–1.55	–1.05	–0.82	–1.20
Call option on McDonald's stock	–1.53	–2.20	–1.46	–1.64
Call option on Intel stock	0.37	–0.14	0.12	0.50

19. We choose the data window as 51 observations centered around the portfolio VaR level.
20. The difference between marginal VaR and recalculated marginal VaR is the estimation error, since we use the same sample for calculations.

C. Decomposing Expected Shortfall

This subsection describes a method of decomposing expected shortfall developed by Tasche (2000).

Suppose equation (3) holds. The following equation holds, since $ES_\alpha(X)$ is a linearly homogeneous function of ω_i.

$$ES_\alpha(X) = \sum_{i=1}^{n} \frac{\partial ES_\alpha(X)}{\partial \omega_i} \cdot \omega_i. \tag{10}$$

Tasche (2000) proved the following under certain conditions.[21]

> **Marginal expected shortfall**
> The partial derivative of expected shortfall at the $100(1-\alpha)$ percent confidence level with respect to ω_i (we call this "marginal expected shortfall," which is denoted by $M - VaR_i$) is represented as a conditional expectation as follows (Tasche [2000], remark 5.4):
>
> $$M - ES_i = \frac{\partial ES_\alpha(X)}{\partial \omega_i} = E[X_i | X \geq VaR_\alpha(X)]. \tag{11}$$

Thus, equation (12) provides a method of decomposing expected shortfall.[22]

$$ES_\alpha(X) = \sum_{i=1}^{n} \frac{\partial ES_\alpha(X)}{\partial \omega_i} \omega_i = \sum_{i=1}^{n} E[X_i | X \geq VaR_\alpha(X)] \cdot \omega_i. \tag{12}$$

Based on this observation, we define component expected shortfall as follows.

> **Component expected shortfall**
> The contribution of risk factor i to the portfolio expected shortfall (we call this "component expected shortfall," which is denoted by $C - ES_i$) is defined as follows:
>
> $$C - ES_i = \frac{\partial ES_\alpha(X)}{\partial \omega_i} \omega_i = E[X_i | X \geq VaR_\alpha(X)] \cdot \omega_i. \tag{13}$$

It is relatively straightforward to estimate the right-hand side of equation (11). The conditioning event is the inequality $X \geq VaR_\alpha(X)$, and we can take more than one sample as this event.

21. The conditions are the same as in Footnote 17.
22. We should note that component expected shortfall is defined using marginal expected shortfall, which is the "marginal" change in expected shortfall with respect to ω_i (see Footnote 18).

D. An Example of Decomposing Expected Shortfall

This subsection gives an example of decomposing expected shortfall of the sample option portfolio in Subsection IV.A.

Suppose we obtain N samples from a simulation. Let $X^j(1 \leq j \leq N)$ denote the portfolio loss of the j-th sample, and $X_i^j(1 \leq i \leq 6,\ 1 \leq j \leq N)$ denote the loss of the i-th risk factor of the j-th sample. We choose a data window whose portfolio losses are more than or equal to the portfolio VaR level. We let T denote the number of samples in the chosen data window. The marginal VaR of individual risk factor i is estimated by

$$M-ES_i = E[X_i|X \geq VaR_\alpha(X)] = \frac{1}{T}\sum_j X_i^j, \tag{14}$$

where we take the sum only for the data included in the chosen window.

We calculate the expected shortfall at the 95 percent confidence level of the sample option portfolio in Subsection IV.A by a Monte Carlo simulation with a sample size of 10,000, and decompose the expected shortfall into risk factors. Table 18 shows the result.

Table 18 Expected Shortfall Decomposition of Sample Option Portfolio (Confidence Level: 95 Percent)

	Marginal expected shortfall (US$/unit) (a)	Investment amount (US$1,000)	Position (1,000 units) (b)	Component expected shortfall (US$1,000) (c) = (a) × (b)
General Electric stock	–3.33	1,000	20.1	–66.92
McDonald’s stock	–2.10	1,000	32.0	–67.1
Intel stock	0.15	1,000	23.8	3.57
Call option on GE stock	–2.30	–100	–50.0	115.23
Call option on McDonald’s stock	–1.91	–100	–57.1	108.93
Call option on Intel stock	0.08	–100	–28.6	–2.28
Total	—	2,700	—	91.43

We also evaluate the estimation error by comparing the marginal expected shortfall estimated by equation (14) with the “recalculated marginal expected shortfall” obtained by re-estimating the portfolio expected shortfall for a slightly changed portfolio (we take 0.1 percent, 0.5 percent, and 1 percent changes). Table 19 shows the result.[23]

The marginal expected shortfall estimated by equation (14) is almost equal to the recalculated marginal expected shortfall. Therefore, we conclude that equation (14) provides an accurate estimate of the marginal expected shortfall.

23. The difference between marginal expected shortfall and recalculated marginal expected shortfall is the estimation error, since we use the same sample for calculations, as was explained in Footnote 20.

Table 19 Comparison of Marginal Expected Shortfall (Confidence Level: 95 Percent)

	Marginal expected shortfall (US$/unit)	Recalculated marginal expected shortfall (US$/unit)		
		Change in position: 0.1 percent	Change in position: 0.5 percent	Change in position: 1 percent
General Electric stock	–3.33	–3.34	–3.34	–3.34
McDonald's stock	–2.10	–2.10	–2.10	–2.10
Intel stock	0.15	0.15	0.15	0.15
Call option on GE stock	–2.30	–2.31	–2.31	–2.31
Call option on McDonald's stock	–1.91	–1.91	–1.90	–1.91
Call option on Intel stock	0.08	0.08	0.08	0.08

VI. Portfolio Optimization Based on Expected Shortfall

This section provides an overview of methods that can be used to optimize portfolios based on VaR and expected shortfall. We focus in particular on the situation in which the underlying loss distribution is not normal and VaR and expected shortfall are calculated by simulations.

A. Portfolio Optimization Based on VaR by the Variance-Covariance Method

Portfolio optimization based on VaR is straightforward when VaR is calculated by the variance-covariance method.[24] The traditional mean-variance analysis (see Markowitz [1952]) is directly applied to VaR-based portfolio optimization, since VaR is a scalar multiple of the standard deviation of loss when the underlying distribution is normal.[25] Mean variance analysis selects the portfolio with the best mean-variance profile by minimizing variance subject to the constraint of expected portfolio return. This optimization problem is formulated as follows:

$$\min_{\{w\}} \frac{1}{2}\omega'\sum\omega, \qquad (15)$$

$$\text{subject to } \omega'\mu = \mu_X$$
$$\omega' e = 1$$

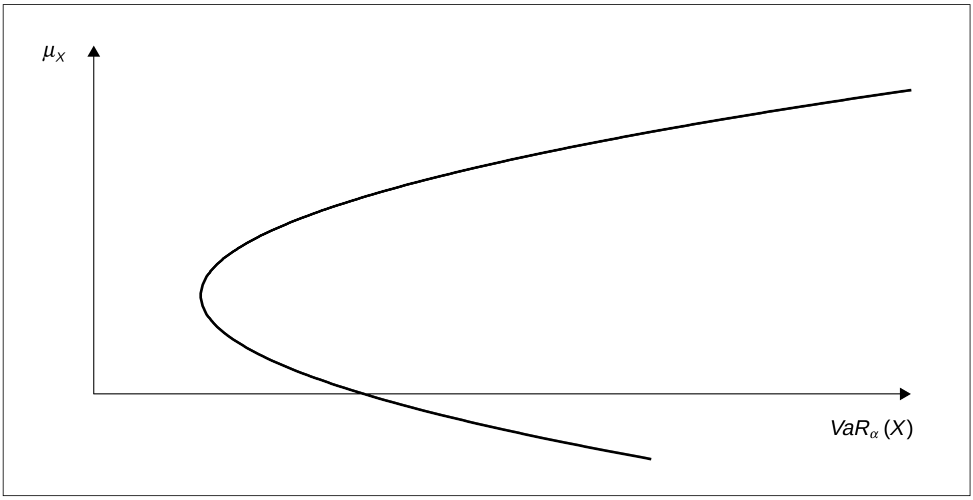
where μ : vector of expected returns of risk factors
μ_X: fixed expected return on portfolio
$\sum$: variance-covariance matrix of risk factors
e : vector of ones
ω : vector of exposures to risk factors
ω': transposed vector of ω

24. The variance-covariance method assumes that portfolios are linear combinations of normally distributed risk factors, and calculates VaR with the variance-covariance matrix of risk factors. This method is also called the "delta-normal" method.
25. To put it more precisely, when the underlying distribution is elliptical (a generalization of the normal), with finite variance, VaR is a scalar multiple of the standard deviation. However, our explanation here assumes normality for the sake of simple illustration.

The solution to this problem is given as ω for each μ_X, from which we obtain an optimized σ_X for each μ_X.[26] This relationship between σ_X and μ_X gives us the efficient frontier in the $\mu_X - \sigma_X$ space. From this efficient frontier, we choose the best portfolio that fits our risk tolerance and return appetite.

Since VaR is a scalar multiple of the standard deviation, we can obtain the efficient frontier in the $\mu_X - VaR$ space as shown in Figure 10. We select the best portfolio using this efficient frontier in the $\mu_X - VaR$ space.

Figure 10 Efficient Frontier of Expected Profit and VaR



B. Portfolio Optimization Based on VaR by Simulation-Based Methods

When VaR is calculated by simulations, it is no longer an efficient tool for optimizing a portfolio, since VaR is no longer a scalar multiple of the standard deviation and is not optimized using equation (15).

Mausser and Rosen (1998) show that it is difficult to optimize simulation-based VaR, since VaR is not generally a convex function of risk factors.[27]

26. The solution ω^* is given as

$$\omega^* = \lambda \sum^{-1} e + \gamma \sum^{-1} \mu$$

where $\lambda = \dfrac{C - \mu_X B}{D}$, $\gamma = \dfrac{\mu_X A - B}{D}$, $A = e' \sum^{-1} e$, $B = e' \sum^{-1} \mu$, $C = \mu' \sum^{-1} \mu$, $D = AC - B^2$.

From this, the efficient frontier is derived as

$$\sigma_X^2 = \frac{A\mu_X^2 - 2B\mu_X + C}{D}.$$

27. When an objective of an optimization problem is not convex, the problem is difficult to solve, since multiple local solutions may exist. See Mausser and Rosen (1998) for the difficulty of optimizing simulation-based VaR.

C. Portfolio Optimization Based on Expected Shortfall by Simulation-Based Methods

Rockafeller and Uryasev (2000) provide a simple algorithm for optimizing portfolios based on a simulation-based expected shortfall. This subsection describes the algorithm they developed, and gives an example of optimizing the sample option portfolio in Subsection IV.A.

We assume that the portfolio loss X is a linear combination of the losses of individual risk factors X_i (i denotes risk factors):

$$X = \sum_{i=1}^{n} X_i \omega_i. \tag{16}$$

X : portfolio loss
X_i: loss of individual risk factor i [28]
ω_i: sensitivity to individual risk factor i

We also assume that the loss of risk factors $(X_1, \ldots, X_n)$ has a probability density function $p(X_1, \ldots, X_n)$.

Suppose $\Psi(\omega, \beta)$ denotes the probability that the portfolio loss X does not exceed some threshold value β.

$$\Psi(\omega, \beta) = \int_{\sum_{i=1}^{n} X_i \omega_i \leq \beta} p(X_1, \ldots, X_n) dX_1 \cdots dX_n. \tag{17}$$

VaR at the 100α percent confidence level is $\beta(\omega, \alpha)$ defined by

$$\beta(\omega, \alpha) = \min\{\beta \in \mathbf{R} | \Psi(\omega, \beta) \geq \alpha\}. \tag{18}$$

We then define the following function denoted by $\Phi(\omega)$.

$$\Phi(\omega) = \int_{\sum_{i=1}^{n} X_i \omega_i \geq \beta(\omega, \alpha)} \left(\sum_{i=1}^{n} X_i \omega_i\right) \cdot p(X_1, \ldots, X_n) dX_1 \cdots dX_n, \tag{19}$$

The expected shortfall is $\Phi(\omega)/(1 - \alpha)$, since it is the conditional expectation given that the portfolio loss $\sum_{i=1}^{n} X_i \omega_i$ is more than $\beta(\omega, \alpha)$.

It is difficult to optimize $\Phi(\omega)$ because $\beta(\omega, \alpha)$ is involved in its definition. Rockafeller and Uryasev (2000) show that optimizing $\Phi(\omega)$ is equivalent to optimizing $F(\omega, \beta)$ (see Appendix 4 for proof).

$$F(\omega, \beta) = (1 - \alpha)\beta + \int_{\omega} \left(\sum_{i=1}^{n} X_i \omega_i - \beta\right)^{+} p(X_1, \ldots, X_n) dX_1 \cdots dX_n. \tag{20}$$

28. See Footnote 15.

Furthermore, the expected shortfall is given as minimized $F(\omega, \beta)/(1-\alpha)$ with respect to β, and VaR is given as corresponding β.

We use this result to minimize the simulation-based expected shortfall. Suppose we sample $X_1, \ldots, X_n$ J times (those samples are denoted by X_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, J$) from the probability density function $p(X_1, \ldots, X_n)$. The integral in equation (20) is calculated approximately as follows:

$$\int_{\omega}\left(\sum_{i=1}^{n} X_i\omega_i - \beta\right)^{+} p(X_1, \ldots, X_n)\,dX_1 \cdots dX_n \approx J^{-1}\sum_{j=1}^{J}\left(\sum_{i=1}^{n} X_{ij}\omega_i - \beta\right)^{+}. \qquad (21)$$

We reduce minimization of $F(\omega, \beta)$ to the following linear programming problem.

$$\min_{\omega\in\mathbf{R}^n,\, \mathbf{z}\in\mathbf{R}^J,\, \beta\in\mathbf{R}} (1-\alpha)\beta + J^{-1}\sum_{j=1}^{J} z_j, \qquad (22)$$

subject to

$$z_j \geq \sum_{i=1}^{n} X_{ij}\omega_i - \beta,\ z_j \geq 0,\ j = 1, \ldots, J. \qquad (23)$$

The constraint on the portfolio expected return is formulated as follows.

$$J^{-1}\sum_{j=1}^{J}\sum_{i=1}^{n} X_{ij}\omega_i = -R. \qquad (24)$$

Furthermore, the constraint on the portfolio investment amount is formulated as follows.

$$\sum_{i=1}^{n} P_i\omega_i = W_0. \qquad (25)$$

P_i : initial value of risk factor i

W_0: initial investment amount in the portfolio

By solving this constrained minimization problem, we can optimize portfolios based on the expected shortfall.

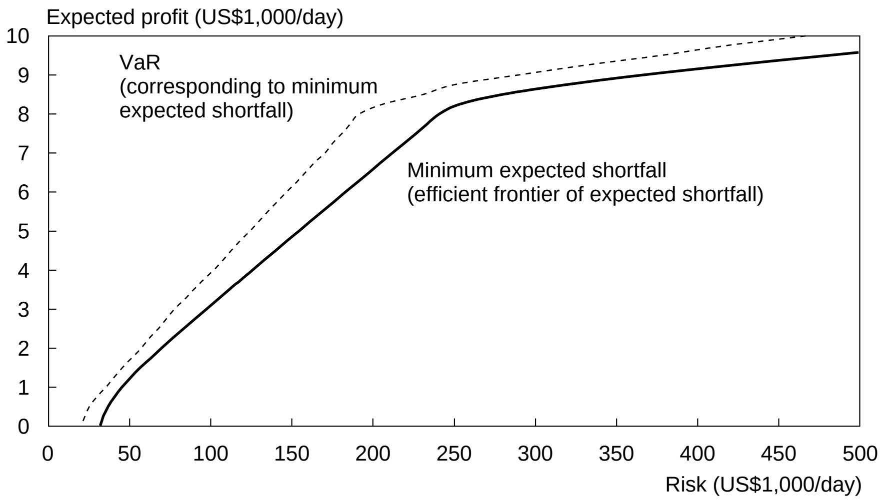
D. An Example of Portfolio Optimization Based on Expected Shortfall

This subsection gives an example of optimizing the sample option portfolio in Subsection IV.A.

We minimize the expected shortfall of this portfolio at the 95 percent confidence level by solving the optimization problem of equation (22) with the constraints of equations (23)–(25). The sample size is 1,000. The portfolio expected return is constrained to be a constant ranging from US\$0 to US\$10,000 in US\$250 increments. To ensure the convergence of the solution, we add the constraint that neither short sales of stocks nor long positions on the call options are allowed.

Figure 11 shows the efficient frontier of the portfolio in the return-expected shortfall space. It is seen that this frontier is convex, and is similar to the one obtained in the return-VaR space when VaR is calculated by the variance-covariance method. Table 20 shows the composition of this efficient frontier when the portfolio expected return is US$5,000.

Figure 11 Efficient Frontier of Expected Shortfall



Note: Sample size is 1,000.

Table 20 Portfolio Composition of Minimum Expected Shortfall

	Investment amount (US$1,000)
General Electric stock	0
McDonald's stock	1,714.45
Intel stock	1,113.05
Call option on GE stock	0
Call option on McDonald's stock	0
Call option on Intel stock	–127.50

Note: Expected profit = US$5,000; expected shortfall = US$154,400.

VII. Concluding Remarks

We compared expected shortfall with VaR in three aspects: their estimation errors, their decomposition into risk factors, and their optimization. We showed that expected shortfall is easily decomposed and optimized, while VaR is not. We also showed that expected shortfall requires a larger size of sample than VaR for the same level of accuracy.

APPENDIX 1: CLOSED-FORM FORMULAS OF ESTIMATION ERRORS

A. Closed-Form Formulas

The estimator of VaR is a quantile of the empirical loss distribution. We take the estimator of VaR at the 100(1 $-\alpha$) percent confidence level to be $X_{(n\alpha+1)}$, where $X_{(n)}$, $X_{(n-1)}, \cdots, X_{(n\alpha+1)}, X_{(n\alpha)}, \cdots, X_{(1)}$ are loss samples arranged in increasing order.

Stuart and Ord ([1994], pp. 356–358) show that the quantiles of distributions asymptotically obey the normal distribution and their asymptotic standard deviation is given in a closed-form formula. From this result, we obtain the following closed-form formula for the asymptotic standard deviation of VaR estimates:

$$\sigma_{VaR_\alpha(X)} = \frac{1}{f(x_\alpha)}\sqrt{\frac{\alpha(1-\alpha)}{n}}, \tag{A.1}$$

where $f(x_\alpha)$ is the probability density function of loss evaluated at the 100(1 $-\alpha$) percent quantile, and n is the sample number.

We take the estimator of the expected shortfall to be

$$ES_\alpha = \frac{X_{(1)} + X_{(2)} + \cdots + X_{(n\alpha+1)}}{n\alpha + 1}. \tag{A.2}$$

When the sample size n is large, the standard deviation of the expected shortfall estimates is approximated by

$$\begin{aligned}\sigma_{ES_\alpha(X)} = \frac{1}{\sqrt{n}}\Bigg[&\frac{(1-\alpha)x_{1-\alpha}^2 + \beta x_{1-\beta}^2}{(\alpha-\beta)^2} + \frac{1}{(\alpha-\beta)^2}\int_{x_{1-\alpha}}^{x_{1-\beta}} x^2 f(x)\,dx \\ &- \frac{1}{(\alpha-\beta)^2}\left\{\beta x_{1-\beta} + (1-\alpha)x_{1-\alpha} + \int_{x_{1-\alpha}}^{x_{1-\beta}} x f(x)\,dx\right\}^2\Bigg]^{\frac{1}{2}},\end{aligned} \tag{A.3}$$

for some β such as $\beta << \alpha$, where $x_{1-\alpha}$ and $x_{1-\beta}$ are $(1-\alpha)$ and $(1-\beta)$ quantiles of the underlying loss distribution.[29]

29. $x_{1-\alpha} = F^{-1}(1-\alpha)$, $x_{1-\beta} = F^{-1}(1-\beta)$, where $F^{-1}(x)$ is the inverse of the distribution function of loss $F(x)$.

B. Comparison between Closed-Form Approximation and Simulation Estimate

We compare the results obtained by the closed-form formulas (A.2) and (A.3) and the results obtained by Monte Carlo simulation. We simulate random numbers from the standard normal, *t* (degrees of freedom: 2),[30] and Pareto ($\beta = 2$)[31] as the loss samples,[32] and compare the standard deviation estimates obtained by those methods. Appendix Table 1 shows the results. When the underlying loss distribution is normal or *t*, those numbers are almost equal. However, when the underlying distribution is Pareto, they are different.

Appendix Table 1 Standard Deviation of VaR and Expected Shortfall Estimates: Comparison between Closed-Form and Simulation Methods

[1] 95 Percent Confidence Interval

	VaR		Expected shortfall	
	Closed-form	Simulation	Closed-form	Simulation
Normal	0.0668	0.0664	0.0780	0.0773
t	0.1080	0.1074	0.1885	0.1872
Pareto	0.3082	0.3090	1.6124	2.0818

[2] 99 Percent Confidence Interval

	VaR		Expected shortfall	
	Closed-form	Simulation	Closed-form	Simulation
Normal	0.1181	0.1153	0.1449	0.1386
t	0.2884	0.2839	0.5346	0.5068
Pareto	1.5732	1.5721	7.0509	8.9681

Note: Sets of simulation = 100,000; sample size = 1,000; β in equation (A.3) = 10^{-5}.

C. Derivation of Closed-Form Formula

Equation (A.3) is derived using the result of robust statistics.[33] We utilize the fact that the estimator of expected shortfall is the L-estimate in robust statistics literature. Under certain conditions, the L-estimate asymptotically obeys the normal distribution, and its asymptotic variance is given in closed-form formula.

30. The probability density function of the *t*-distribution with degrees of freedom of *m* is

$$f(x) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{\pi \cdot m} \cdot \Gamma\left(\frac{m}{2}\right)} \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}} \quad \text{where } \Gamma(\ ) \text{ is the Gamma function.}$$

31. The probability density function of the Pareto distribution is

$$f(x) = \frac{\beta}{x^{\beta+1}}, \quad x \geq 1.$$

32. We used normal, *t*, and Pareto distributions since they have closed-form representation of probability density functions, and are convenient methods of evaluating the standard deviation using equation (A.3).
33. See Huber (1981) for the details of robust statistics.

Suppose $X_{(n)}, X_{(n-1)}, \ldots, X_{(n\alpha)}, \ldots, X_{(n\beta)}, \ldots, X_{(1)}$ are losses of n samples from simulation, rearranged in increasing order. Take some constant β such that $\beta << \alpha$. Consider the following estimator.

$$ES_{\alpha,\beta} = \frac{X_{(n\beta)} + \cdots + X_{(n\alpha)}}{n(\alpha-\beta)} = \frac{1}{n}\sum_{i=n\beta}^{n\alpha}\frac{1}{\alpha-\beta}X_{(i)}. \tag{A.4}$$

This estimator approximates to the estimator of expected shortfall (equation [A.2]) when n is large and β is efficiently small. We define a weighting function h as follows:

$$h(u) = \begin{cases} 1/(\alpha-\beta) & 1-\alpha \le u \le 1-\beta, \\ 0 & \text{otherwise.} \end{cases} \tag{A.5}$$

Using this function, equation (A.4) becomes

$$ES_{\alpha,\beta} = \sum_{i=1}^{n}\frac{1}{n}\cdot h\left(\frac{i}{n+1}\right)X_{(i)}. \tag{A.6}$$

Let $F(x)$ denote the distribution function of the loss, $f(x)$ denote its probability density function, and $T(x) = F^{-1}(x)$ denote the inverse of the distribution function. With equation (3.12) in Huber (1981), the influence function[34] of this estimator is given as follows:

$$\begin{aligned} IC(x, F, T) &= \int_{-\infty}^{x} h(F(y))dy - \int_{-\infty}^{\infty}(1-F(y))h(F(y))dy \\ &= \frac{1}{\alpha-\beta}\left\{(1-\beta)F^{-1}(1-\beta) - (1-\alpha)F^{-1}(1-\alpha)\right. \\ &\quad \left. - \int_{1-\alpha}^{1-\beta}F^{-1}(v)dv\right\} - \int_{F(x)}^{1}h(v)(F^{-1}(v))'dv. \end{aligned} \tag{A.7}$$

Theorem 3.2 of Huber (1981) says that, under certain conditions,[35] L-estimates asymptotically obey the normal distribution with asymptotic variance of $(1/n)\int_{-\infty}^{\infty}IC(x, F, T)^2 f(x)dx$. We apply this theorem to $ES_{\alpha,\beta}$ and obtain the following result.

$$\begin{aligned} \sigma^2_{ES_{\alpha,\beta}} &= \frac{1}{n}\int_{-\infty}^{\infty}IC(x, F, T)^2 f(x)dx \\ &= \frac{1}{n}\left[\frac{(1-\alpha)x_{1-\alpha}^2 + \beta x_{1-\beta}^2}{(\alpha-\beta)^2} + \frac{1}{(\alpha-\beta)^2}\int_{x_{1-\alpha}}^{x_{1-\beta}}x^2 f(x)dx\right. \\ &\quad \left. - \frac{1}{(\alpha-\beta)^2}\left\{\beta x_{1-\beta} + (1-\alpha)x_{1-\alpha} + \int_{x_{1-\alpha}}^{x_{1-\beta}}xf(x)dx\right\}^2\right]. \end{aligned} \tag{A.8}$$

This proves equation (A.3).

34. See Huber (1981) for definitions and concepts of the influence function.

35. The conditions are (1) the support of h is contained in $[\alpha, 1-\alpha]$ for some $\alpha > 0$, and (2) no discontinuity of h coincides with a discontinuity of the inverse of the distribution function, etc. All of those conditions are satisfied here.

APPENDIX 2: METHOD OF SIMULATING LOSS FOR OPTION PORTFOLIO

This appendix describes the method of simulating loss for the sample option portfolio in Subsection IV.A.

A. Stock Price

The log returns on stocks are assumed to obey the multivariate normal distribution. The variance-covariance matrix of the log returns is estimated using historical stock price data for the past three years. Independently and identically distributed standard normal random variables are transformed into correlated normal random numbers with the Cholesky factors of the historical variance-covariance matrix. The average of historical log returns is added to obtain the simulated log returns of stock prices.

B. Option Premium

Option premiums are simulated using the Black-Scholes formula assuming that the implied volatility is constant. The implied volatilities of the options are calibrated from the data in Table 5.

Even though those options are American, they are priced using the Black-Scholes formula, since the dividend payments on the underlying stocks are not expected until the option maturity. Prices of American call options are shown to be equal to those of European call options in such circumstances.[36]

APPENDIX 3: METHOD OF SIMULATING LOSS FOR LOAN PORTFOLIO

This appendix describes the method of simulating loss of the sample loan portfolio in Subsection IV.B. The method is developed by Ieda, Marumo, and Yoshiba (2000), and the description here is totally dependent on them.

We consider the random variable $D_i (i = 1, 2, \ldots, n)$ which has a Bernoulli distribution:

$$D_i = \begin{cases} 1 \text{ (with probability } p\text{)}, \\ 0 \text{ (with probability } 1-p\text{)}. \end{cases} \tag{A.9}$$

In other words, $D_i (i = 1, 2, \ldots, n)$ for exposure i in the portfolio (comprising n exposures) takes the value 1 (default) with probability p and 0 (non-default) with probability $1-p$. Also, the correlation coefficient of each D_i is ρ (constant). The process of generating multivariate Bernoulli random numbers that takes account of the correlation is not a simple application of the Cholesky decomposition. However, the Cholesky decomposition can be used for normal distributions, so one method is to use the normal distribution as a medium for generating Bernoulli random numbers.

36. See Hull ([2000], pp. 175–176).

We first consider a random variable $X_i (i = 1, 2, \ldots, n)$ that follows the standard normal distribution with 0 for its mean and 1 for its variance. (However, individual variables are correlated rather than independent.) At this time, D_i is expressed as

$$D_i = \begin{cases} 1 \; (-\infty < X_i \leq \Phi^{-1}(p)), \\ 0 \; (\Phi^{-1}(p) < X_i < \infty), \end{cases} \tag{A.10}$$

where $\Phi^{-1}(\cdot)$ is the inverse function of the distribution function of the standard normal distribution.

For the correlation coefficient of $D_i (i = 1, 2, \ldots, n)$ to be ρ, one need properly set a correlation coefficient $\tilde{\rho}$ for $X_i (i = 1, 2, \ldots, n)$. ρ can be expressed as

$$\rho = \frac{E[D_i D_j] - p^2}{\sqrt{p(1-p)}\sqrt{p(1-p)}}, \tag{A.11}$$

where

$$D_i D_j = \begin{cases} 1 \; (-\infty < X_i \leq \Phi^{-1}(p), -\infty < X_j \leq \Phi^{-1}(p)) \\ 0 \qquad\qquad (\text{otherwise}). \end{cases} \tag{A.12}$$

Therefore, $E[D_i D_j]$ is the distribution function of a two-dimensional normal distribution with a correlation coefficient of $\tilde{\rho}$.

$$E[D_i D_j] = \int_{-\infty}^{\Phi^{-1}(p)} \int_{-\infty}^{\Phi^{-1}(p)} \frac{1}{2\pi\sqrt{1-\tilde{\rho}^2}} \exp\left(-\frac{1}{2(1-\tilde{\rho}^2)} \{x_i^2 + x_j^2 - 2\tilde{\rho} x_i x_j\} \right) dx_i dx_j. \tag{A.13}$$

This makes it possible to use equations (A.13) and (A.11) to obtain a $\tilde{\rho}$ that will satisfy equation (A.11). (However, numerical calculations will be required to obtain the definite integral above.)

It is, therefore, possible to obtain multivariate Bernoulli random numbers D_i by using equation (A.10) after generating multivariate normal random numbers in the n-th dimension with a mean of 0, a variance of 1, and a constant correlation coefficient of $\tilde{\rho}$.

The portfolio loss L can be expressed as follows:

$$L = \sum_{i=1}^{n} D_i v_i (1 - r_i), \tag{A.14}$$

where v_i is the amount of exposure and $r_i (0 \leq r_i \leq 1)$ is the recovery rate at default of exposure i.

APPENDIX 4: PROOF OF THE THEOREM OF ROCKAFELLER AND URYASEV (2000)

This appendix explains a theorem on expected shortfall established by Rockafeller and Uryasev (2000). This theorem is used to develop an algorithm of efficiently minimizing expected shortfall.

Consider the following function:

$$\Psi(\beta) = \int_{x \leq \beta} dF(x). \tag{A.15}$$

$\Psi(\beta)$ is the probability that the loss x does not exceed some threshold β, where $F(x)$ is the loss distribution function. VaR at the 100α percent confidence level is defined as β_α, where

$$\beta_\alpha = \min\{\beta \in \mathbf{R} | \Psi(\beta) \geq \alpha\}. \tag{A.16}$$

Expected shortfall at the 100α percent confidence level is defined as the following function.

$$\phi_\alpha = (1-\alpha)^{-1} \int_{x \geq \beta_\alpha} x dF(x). \tag{A.17}$$

Theorem 1 below shows that expected shortfall is the minimization of a function $F_\alpha(\beta)$ defined below with respect to β.

$$F_\alpha(\beta) = \beta + (1-\alpha)^{-1} \int_{x \in \mathbf{R}} [x-\beta]^+ dF(x). \tag{A.18}$$

THEOREM 1 (ROCKAFELLER AND URYASEV [2000])

$F_\alpha(\beta)$ is a convex function of β. It is also continuous and differentiable with respect to β. Expected shortfall is given by

$$\phi_\alpha = \min_{\beta \in \mathbf{R}} F_\alpha(\beta), \tag{A.19}$$

where[37]

$$B_\alpha \equiv \arg\min_{\beta \in \mathbf{R}} F_\alpha(\beta). \tag{A.20}$$

VaR is given by

$$\beta_\alpha = \textit{the left end-point of } B_\alpha. \tag{A.21}$$

Furthermore, the following equality holds.

37. $\arg\min_{\beta \in \mathbf{R}} F_\alpha(\beta)$ is the β that minimizes $F_\alpha(\beta)$.

$$\beta_\alpha = \arg\min_{\beta \in \mathbf{R}} F_\alpha(\beta) \text{ and } \phi_\alpha = F_\alpha(\beta_\alpha). \tag{A.22}$$

Before proving Theorem 1, we prove the following Lemma.

LEMMA
Suppose $G(\beta) = \int_{x\in\mathbf{R}}[x - \beta]^+ dF(x)$ is a function of β with fixed x. $G(\beta)$ is convex with respect to β, and $G'(\beta) = \Psi(\beta) - 1$.

Proof:
The convexity of $G(\beta)$ is apparent from the convexity of $\beta \mapsto [x-\beta]^+$.
From $G(\beta) = \int_{-\infty}^{\infty}(x-\beta)1_{\{x-\beta\geq 0\}}dF(x)$,[38]

$$\begin{aligned} G'(\beta) &= -\int_{-\infty}^{\infty} 1_{\{x-\beta\geq 0\}}dF(x) + \int_{-\infty}^{\infty}(x-\beta)\frac{\partial 1_{\{x-\beta\geq 0\}}}{\partial\beta}dF(x) \\ &= -\left\{1 - \int_{x\leq\beta}dF(x)\right\} = \Psi(\beta) - 1. \end{aligned} \tag{A.23}$$

Q.E.D.

We prove Theorem 1 using this lemma.

Proof of Theorem 1:
From the Lemma, we obtain

$$\frac{\partial}{\partial\beta}F_\alpha(\beta) = 1 + (1-\alpha)^{-1}[\Psi(\beta) - 1] = (1-\alpha)^{-1}[\Psi(\beta) - \alpha]. \tag{A.24}$$

Since $F_\alpha(\beta)$ is convex, $F_\alpha(\beta)$ is minimized when the first-order condition $\Psi(\beta) - \alpha = 0$ is satisfied (or when $\beta \in B_\alpha$). Since $\Psi(\beta)$ is continuous and non-increasing with respect to β, β takes the lowest value that satisfies $\Psi(\beta) \geq \alpha$ when $\Psi(\beta) - \alpha = 0$. Therefore, $\beta = \beta_\alpha$ when $\Psi(\beta) - \alpha = 0$, and the following equality holds.

$$\min_{\beta\in\mathbf{R}} F_\alpha(\beta) = F_\alpha(\beta_\alpha) = \beta_\alpha + (1-\alpha)^{-1}\int_{x\in\mathbf{R}}[x-\beta_\alpha]^+ dF(x). \tag{A.25}$$

Thus, the integral in equation (A.25) is

$$\int_{x\geq\beta_\alpha}[x-\beta_\alpha]dF(x) = \int_{x\geq\beta_\alpha}x\,dF(x) - \beta_\alpha\int_{x\geq\beta_\alpha}dF(x). \tag{A.26}$$

The first term of the right-hand side of equation (A.26) is $(1-\alpha)\phi_\alpha$ by the definition of expected shortfall. The second term is $1 - \Psi(\beta_\alpha)$, by the definition of the distribution function. Since $\Psi(\beta_\alpha) = \alpha$, the following equality holds.

$$\min F_\alpha(\beta) = \beta_\alpha + (1-\alpha)^{-1}[(1-\alpha)\phi_\alpha - \beta_\alpha(1-\alpha)] = \phi_\alpha. \tag{A.27}$$

38. 1_A is an indicator function that takes the value of 1 when A is true and takes the value of zero otherwise.

$\beta \in \mathbf{R}$

This concludes the proof of Theorem 1.

Q.E.D.

References

Acerbi, C., and D. Tasche, "On the Coherence of Expected Shortfall," working paper, Center for Mathematical Sciences, Munich University of Technology, 2001.

Artzner, P., F. Delbaen, J. M. Eber, and D. Heath, "Thinking Coherently," *Risk*, 10 (11), November, 1997, pp. 68–71.

———, ———, ———, and ———, "Coherent Measures of Risk," *Mathematical Finance*, 9 (3), June, 1999, pp. 203–228.

Chambers, J. M., C. L. Mallows, and B. W. Stuck, "A Method for Simulating Stable Random Variables," *Journal of the American Statistical Association*, 71 (354), 1976, pp. 340–344.

Feller, W., *An Introduction to Probability Theory and Its Applications*, Volume 2, John Wiley and Sons, 1969.

Garman, M., "Taking VaR to Pieces," *Risk*, 10 (10), October, 1997, pp. 70–71.

Hallerbach, W. G., "Decomposing Portfolio Value at Risk: A General Analysis," Tinbergen Institute Discussion Paper, TI 99-034/2, 1999.

Huber, P. J., *Robust Statistics*, John Wiley & Sons, 1981.

Hull, J., *Options, Futures, and Other Derivatives*, Fourth Edition, Prentice-Hall, 2000.

Ieda, A., K. Marumo, and T. Yoshiba, "A Simplified Method for Calculating the Credit Risk of Lending Portfolios," *Monetary and Economic Studies*, 18 (2), Institute for Monetary and Economic Studies, Bank of Japan, 2000, pp. 49–82.

Markowitz, H., "Portfolio Selection," *The Journal of Finance*, 7 (1), 1952, pp. 77–91.

Mausser, H., and D. Rosen, "Beyond VaR: From Measuring Risk to Managing Risk," *ALGO Research Quarterly*, 1 (2), 1998, pp. 5–20.

Rockafeller, T., and S. Uryasev, "Optimization of Conditional Value-at-Risk," *Journal of Risk*, 2 (3), Spring, 2000, pp. 21–41.

Shiryaev, A. N., *Essentials of Stochastic Finance, Facts, Models, Theory*, World Scientific, 1999.

Stuart, A., and J. K. Ord, *Kendall's Advanced Theory of Statistics: Volume 1, Distribution Theory*, Sixth Edition, London/Melbourne/Auckland: Edward Arnold, 1994.

Tasche, D., "Risk Contributions and Performance Measurement," working paper, Munich University of Technology, 2000.

Yamai, Y., and T. Yoshiba, "On the Validity of Value-at-Risk: Comparative Analyses with Expected Shortfall," *Monetary and Economic Studies*, 20 (1), Institute for Monetary and Economic Studies, Bank of Japan, 2002, pp. 57–86 (this issue).