IMES DISCUSSION PAPER SERIES

Comparative Analyses of Expected Shortfall and Value-at-Risk (2)
: expected utility maximization and tail risk

Toshinao YOSHIBA and Yasuhiro YAMAI

Discussion Paper No. 2001-E-14

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

C.P.O BOX 203 TOKYO
100-8630 JAPAN

IMES Discussion Paper Series 2001-E-14
August 2001

Comparative Analyses of Expected Shortfall and Value-at-Risk (2)
: expected utility maximization and tail risk

Toshinao YOSHIBA and Yasuhiro YAMAI

Abstract

We compare expected shortfall and Value-at-Risk (VaR) in terms of consistency with expected utility maximization and elimination of tail risk. We use the concept of stochastic dominance in studying these two aspects of risk measures.

We conclude that expected shortfall is more applicable than VaR as regards two aspects. Expected shortfall is consistent with expected utility maximization and is free of tail risk, under more lenient conditions than VaR is.

Key words: expected shortfall, Value-at-Risk, tail risk, stochastic dominance, expected utility maximization

JEL classification: G20

* Research Division 1, Institute for Monetary and Economic Studies, Bank of Japan (E-mail: toshinao.yoshiba@boj.or.jp, yasuhiro.yamai@boj.or.jp)

The authors would like to thank Professor Hiroshi Konno (Chuo University) for his helpful comments.

Contents

I. Introduction

In this paper, we compare expected shortfall and VaR from two aspects: consistency with expected utility maximization and elimination of tail risk. We use the concept of stochastic dominance in studying the two aspects of risk measures.

Expected utility maximization is the most widely accepted preference representation in finance and economics literature. It represents the rational investor's preference if we accept the four axioms put forward by von Neumann and Morgenstern [1953].

In this paper, we define the consistency of a risk measure with expected utility maximization. A risk measure is consistent with expected utility maximization if it provides the same ranking of investment opportunities (portfolios) as expected utility maximization does. The use of a risk measure consistent with expected utility maximization leads to rational investment decisions in the sense of von Neumann and Morgenstern [1953].

Also, in this paper, we define tail risk as follows. A risk measure is free of tail risk if it takes into account information about the tail of the underlying distribution. The use of a risk measure free of tail risk avoids extreme loss in the tail of the underlying distribution.

Several studies have discussed the concept of tail risk. The BIS Committee on the Global Financial System [2000] proposes and describes the concept of tail risk with simple illustrations. It shows that a single set of risk measures, including VaR and the standard deviation, disregards the risk of extreme loss in the tail of the underlying distribution. Basak and Shapiro [2000] show that the use of VaR, which disregards the loss beyond the quantile of the underlying distribution, increases the extreme loss in the tail of the distribution. Yamai and Yoshiba [2001a] point out the same problem in the use of VaR for managing options and loan portfolios.

Those studies, however, do not give a definition of tail risk.

A number of comparative studies have been done on expected shortfall and VaR[1]. Those studies describe the advantages and the disadvantages of expected shortfall over VaR in various aspects. For example, Artzner et al. [1997, 1999] say that expected shortfall is sub-additive[2] while VaR is not. Rockafeller and Uryasev [2000] show that expected shortfall is easily optimized using the linear programming approach, while VaR is not. Yamai and Yoshiba [2001b] show that expected shortfall needs a larger sample size than VaR for the same level of accuracy.

The rest of the paper is as follows. Chapter II gives the definition of consistency with expected utility maximization and elimination of tail risk. Chapter III considers whether expected shortfall and VaR are consistent with expected utility maximization and whether they are free of tail risk. Chapter IV provides an example in which expected shortfall is neither consistent with expected utility maximization, nor free of tail risk. Chapter V concludes the paper.

II. Expected Utility Maximization and Tail Risk

In this chapter, we describe the definition of and concept involved in consistency with expected utility maximization and elimination of tail risk. We use the concept of stochastic dominance in defining and studying these two aspects of risk measures.

In this paper, we suppose that investment opportunities (portfolios)

---

1 See, for example, Acerbi and Tasche [2001], Acerbi, Nordio and Sirtori [2001], Artzner et al. [1997, 1999], Basak and Shapiro [2000], Bertsimas, Lauprete and Samarov [2000], Pflug [2000], Rockafeller and Uryasev [2000], and Yamai and Yoshiba [2001a, b].

2 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 can be 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 is satisfied.

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

are described by the set of possible payoffs (profit and loss) and their probabilities. For simplicity, we consider only static investment problems, or one period of investment uncertainty between two dates 0 and 1. We also assume that the distribution functions of the payoffs are continuously differentiable, and thus possess density functions.

A. Consistency with expected utility maximization

Expected utility maximization is one of the most widely accepted preference representations for the analysis of decision under uncertainty. If we accept the axioms put forward by von Neumann and Morgenstern [1953], every rational investor should follow expected utility maximization as his/her decision criterion[3].

Finance and economics literature usually considers the class of utility functions $U(X)$ that satisfy $U'(x) \geq 0$ (non-decreasing) and $U''(x) \leq 0$ (concave) for $\forall x \in R$. This means that investors are nonsatiated and are risk averse.

We study whether expected shortfall and VaR are consistent with expected utility maximization. We say that a risk measure is consistent with expected utility maximization when it provides the same ranking of portfolios as expected utility maximization does. If a risk measure is consistent with expected utility maximization, the use of the risk measure leads to a rational decision.

In order to consider consistency of risk measures with expected utility maximization, we use the concept of stochastic dominance. Stochastic dominance ranks investment opportunities using partial information regarding utility functions. Stochastic dominance is a practical concept since one is able to rank portfolios without specifying the

---

[3] See Ingersoll [1987], and Huang and Litzenberger [1993] for the details of expected utility maximization.

forms of the utility functions used[4,5].

In this section, we describe the definition and the concept[6] of stochastic dominance in order to consider the consistency of risk measures with expected utility maximization.

1. Second order stochastic dominance

We describe the definition and concept of second order stochastic dominance, which employs nonsatiety and risk aversion as partial information about the preferences.

Second order stochastic dominance is defined by the cumulation of distribution functions. Let X be a random variable denoting the profit and loss of a portfolio. Suppose that X has a distribution function $F(x)$ and a density function $f(x)$. We then define the cumulation of the distribution function of X as follows.

$$F^{(2)}(x) = \int_{-\infty}^{x} F(u)du \,. \qquad (1)$$

We call this function "the second order distribution function."

The next theorem shows that the second order distribution function is equal to the first lower partial moment (denoted by $LPM_{1,x}(X)$ below), a risk measure first proposed by Fishburn [1977] (See p.139 of Ingersoll [1987] for the proof).

Theorem 1

$$F^{(2)}(x) = \int_{-\infty}^{x} F(u)du = \int_{-\infty}^{x} (x-u)f(u)du \equiv LPM_{1,x}(X) \qquad (2)$$

---

4 See Levy [1998], Bawa [1975], Ingersoll [1987], and Huang and Litzenberger [1993] for the details of stochastic dominance.

5 Cumperayot et al. [2000], Guthoff, Pfingsten and Wolf [1997], Ogryczak and Ruszczynski [1999, 2001], Pflug [1999, 2000] consider consistency of risk measures with stochastic dominance.

6 We refer to Levy [1998] and Ingersoll [1987] in describing the concept and definition of stochastic dominance.

Second order stochastic dominance is defined as follows.[7]

Definition 1

Let X_1 and X_2 be random variables denoting the profit and loss of two portfolios. We say that X_1 dominates X_2 in the sense of second order stochastic dominance ($X_1 \geq_{SSD} X_2$) if the following holds.

$$F_1^{(2)}(x) \leq F_2^{(2)}(x) \quad \text{for} \quad \forall x \in R \tag{3}$$

where $F_1^{(2)}(x)$ and $F_2^{(2)}(x)$ are the second degree distribution functions of X_1 and X_2 respectively.

Figure 1 shows the distribution functions and the second order distribution functions of two random variables X_1 and X_2. In this figure, X_1 dominates X_2 in the sense of second order stochastic dominance ($X_1 \geq_{SSD} X_2$). Even though the distribution functions cross each other, the two random variables are ranked by second order stochastic dominance as long as the second degree distribution functions do not cross each other.

---

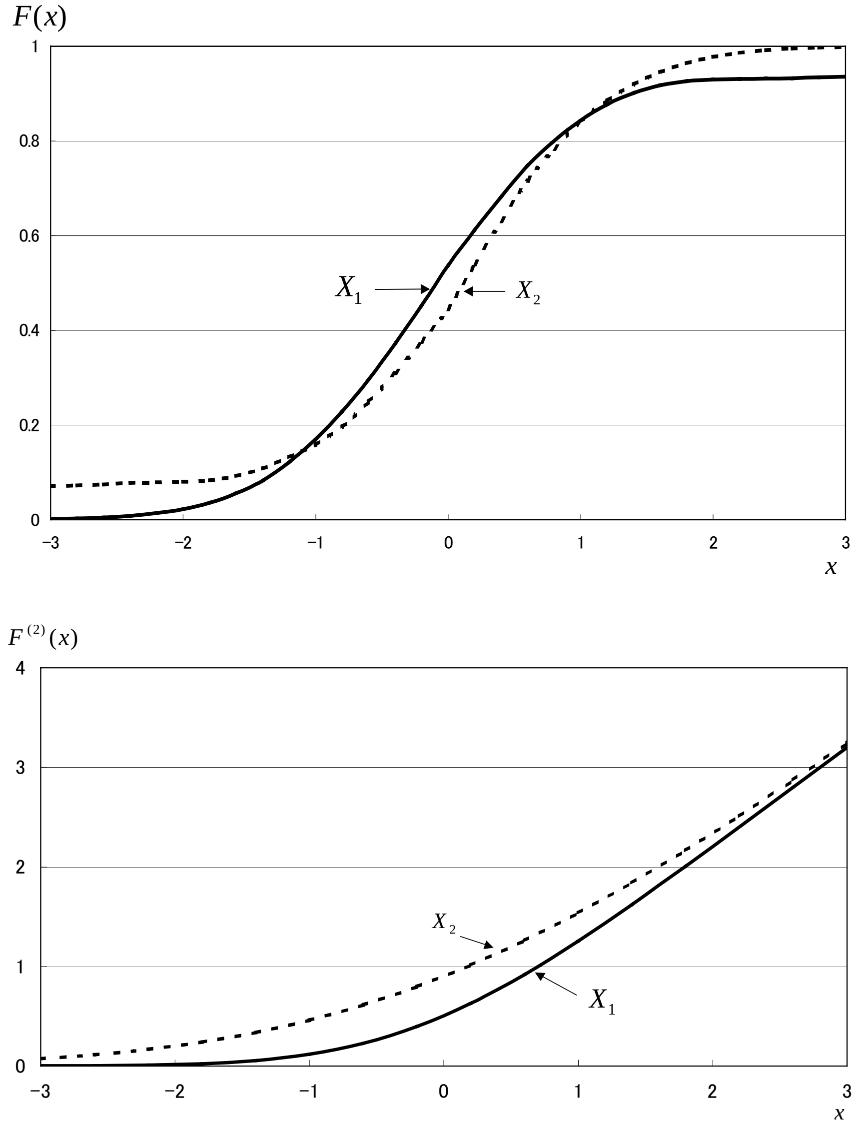
7 First order stochastic dominance is defined as follows.

> A random variable X_1 dominates a random variable X_2 in the sense of first order stochastic dominance ($X_1 \geq_{FSD} X_2$) if $F_1(x) \leq F_2(x)$ for $\forall x \in R$, where $F_1(x)$ and $F_2(x)$ are the distribution functions of X_1 and X_2 respectively.

Then, the following theorem holds (See Theorem 3.1 of Levy[1998] for the proof).

> Let X_1 and X_2 be random variables denoting the profit and loss of two portfolios. $X_1 \geq_{FSD} X_2$ if and only if $E[U(X_1)] \geq E[U(X_2)]$ for all $U(x)$ satisfying $U'(x) \geq 0$ for all x (with at least one $U_0(x)$ satisfying $U_0'(x) > 0$ for some x).

Figure 1　Second order stochastic dominance



Theorem 1 shows that second order stochastic dominance is defined also by the first lower partial moment as follows.

$$LPM_{1,x}(X_1) \leq LPM_{1,x}(X_2)\,, \tag{4}$$

The following theorem shows that second order stochastic dominance employs nonsatiety and risk-aversion as partial information about the preference (See Theorem3.2 of Levy [1998] for the proof).

Theorem 2

$X_1 \geq_{SSD} X_2$ if and only if

$$E[U(X_1)] \geq E[U(X_2)], \qquad (5)$$

for all $U(x)$ satisfying $U'(x) \geq 0$ and $U''(x) \leq 0$ for all x (with at least one $U_0(x)$ satisfying $U_0'(x) > 0$ and $U_0''(x) < 0$ for some x).

The condition that $U(x)$ is non-decreasing and concave for all x means that $U(x)$ represents a nonsatiated and risk-averse preference. Thus, this theorem says that every risk-averse investor chooses X_1 over X_2 if X_1 dominates X_2 in the sense of second order stochastic dominance.

It should be noted that second order stochastic dominance only provides a "partial ordering" of portfolios. This means that second order stochastic dominance is unable to rank all the portfolios. For example, if the second order distribution functions were to cross each other in Figure 1, neither $F_1^{(2)}(x) \leq F_2^{(2)}(x) \quad \forall x \in R$ nor $F_1^{(2)}(x) \geq F_2^{(2)}(x) \quad \forall x \in R$ holds. Thus, one is unable to tell which portfolio dominates the other in the sense of second order stochastic dominance. This corresponds to the situation where one non-decreasing, concave utility function prefers X_1 while another non-decreasing, concave utility function prefers X_2.

When portfolios are not ranked by second order stochastic dominance, one needs to examine third or higher order stochastic dominance in order to rank those portfolios.

2. n-th order stochastic dominance

We now define n-th order stochastic dominance, which is able to rank a larger class of portfolios. N-th order stochastic dominance is defined by n-th order distribution functions defined inductively below.

$$F^{(1)}(x) \equiv F(x), \quad F^{(n)}(x) \equiv \int_{-\infty}^{x} F^{(n-1)}(u)du,$$

where $F(u)$ is the distribution function. (6)

The n-th order distribution function is shown to be equal to the scalar multiple of the $(n-1)$-th lower partial moment (denoted by $LPM_{n-1,x}(X)$ below), a risk measure proposed by Fishburn [1977] (See p.139 of Ingersoll [1987] for proof).

Theorem 3

$$F^{(n)}(x) = \frac{1}{(n-1)!}\int_{-\infty}^{x}(x-u)^{n-1}f(u)du \equiv \frac{1}{(n-1)!}LPM_{n-1,x}(X) \qquad (7)$$

N-th order stochastic dominance is defined as follows.

Definition 2

Let X_1 and X_2 be random variables denoting the profit and loss of two portfolios. We say that X_1 dominates X_2 in the sense of n-th order stochastic dominance ($X_1 \geq_{SD(n)} X_2$) if the following holds.

$$F_1^{(n)}(x) \leq F_2^{(n)}(x) \quad \text{for} \quad \forall x \in R, \qquad (8)$$

where $F_1^{(n)}(x)$ and $F_2^{(n)}(x)$ are the n-th degree distribution functions of X_1 and X_2 respectively.

The following theorem characterizes the relationships between different orders of stochastic dominance.

Theorem 4

If $X_1 \geq_{SD(n)} X_2$, then $X_1 \geq_{SD(n+1)} X_2$.

Proof

If $X_1 \geq_{SD(n)} X_2$ holds, then Equation (8) holds for all x. Thus, the following also holds for all x.

$$\int_{-\infty}^{x} F_1^{(n)}(u)du \leq \int_{-\infty}^{x} F_2^{(n)}(u)du, \qquad (9)$$

From Equation (6) the following holds for all x.

$$F_1^{(n+1)}(x) \le F_2^{(n+1)}(x), \tag{10}$$

Therefore, by definition, $X_1 \ge_{SD(n+1)} X_2$. QED

This theorem shows that, if X_1 dominates X_2 in the sense of n-th order stochastic dominance, X_1 dominates X_2 in the sense of all higher order stochastic dominance.

The following theorem shows how the n-th order stochastic dominance is related to expected utility maximization (see p.139 of Ingersoll [1987] and pp.116-7 of Levy [1998] for the proof).

> Theorem 5
>
> $X_1 \ge_{SD(n)} X_2$ if and only if
>
> $$E[U(X_1)] \ge E[U(X_2)], \tag{11}$$
>
> for all $U(x)$ satisfying $(-1)^k U^{(k)}(x) \le 0$ $(k = 1,2,3,\cdots,n)$ for all x (with at least one $U_0(x)$ satisfying with inequality for some x).

Thus, n-th order stochastic dominance is consistent with expected utility maximization for utility functions $U(x)$ satisfying $(-1)^k U^{(k)}(x) \le 0$ $(k = 1,2,\cdots,n)$.

N-th order stochastic dominance is still a partial ordering, and is unable to rank all the portfolios. However, n-th order stochastic dominance is more applicable than first or second order stochastic dominance in that it is able to rank broader class of portfolios.

3. Consistency of risk measures with stochastic dominance

Following Guthoff, Pfingsten and Wolf [1997], Ogryczak and Ruszczynski [1999, 2001], and Pflug [1999, 2000], we define consistency of risk measures with stochastic dominance as follows.

Definition 3

We say that a risk measure $\rho(X)$ is consistent with n-th order stochastic dominance if the following holds.

$$X_1 \geq_{SD(n)} X_2 \Rightarrow \rho(X_1) \leq \rho(X_2). \quad (12)$$

Taking contraposition of Definition 3, we see that the following holds if a risk measure $\rho(X)$ is consistent with n-th order stochastic dominance.

$$\rho(X_1) > \rho(X_2) \Rightarrow \text{not } (X_1 \geq_{SD(n)} X_2). \quad (13)$$

Thus, when $\rho(X_1) > \rho(X_2)$ holds, either of the following holds.

(i) X_2 dominates X_1 in the sense of n-th order stochastic dominance.

(ii) n-th order stochastic dominance is unable to rank X_1 and X_2.

Theorem 5 shows that, when (i) holds, $\rho(X)$ is consistent with expected utility maximization since it always chooses portfolios whose expected utility is higher. Thus, if portfolios are ranked by n-th order stochastic dominance, a risk measure consistent with n-th order stochastic dominance is also consistent with expected utility maximization.

On the other hand, when (ii) holds, $\rho(X)$ is not necessarily consistent with expected utility maximization. Thus, if portfolios are not ranked by n-th order stochastic dominance, consistency with stochastic dominance is not equivalent to consistency with expected utility maximization.

The following theorem shows the relationship between risk measures and orders of stochastic dominance.

Theorem 6

A risk measure consistent with $(n+1)$-th degree stochastic dominance is also consistent with n-th order stochastic dominance.

Proof

From Theorem 4, the following holds.

$$X_1 \geq_{SD(n)} X_2 \;\Rightarrow\; X_1 \geq_{SD(n+1)} X_2 . \qquad (14)$$

If a risk measure $\rho(X)$ is consistent with $(n+1)$-th order stochastic dominance, then:

$$X_1 \geq_{SD(n+1)} X_2 \;\Rightarrow\; \rho(X_1) \leq \rho(X_2) . \qquad (15)$$

From Equations (14) and (15),

$$X_1 \geq_{SD(n)} X_2 \;\Rightarrow\; \rho(X_1) \leq \rho(X_2) . \qquad (16)$$

Therefore, $\rho(X)$ is consistent with n-th order stochastic dominance.

QED

This theorem shows that, if a risk measure is consistent with n-th order stochastic dominance, the risk measure is consistent with all lower order stochastic dominance. Thus, a risk measure consistent with higher order stochastic dominance is more applicable than a risk measure consistent with lower order stochastic dominance.

B. Tail risk

1. Definition of tail risk

In this section, we provide our definition of tail risk. Our definition is based on our concept of tail risk: a risk measure fails to eliminate tail risk when it fails to summarize the choice between portfolios as a result of its disregard of information on the tail of the distribution. This concept is motivated by the BIS Committee on the Global Financial System [2000], which shows that a single set of risk measures, including VaR and the standard deviation, disregards the risk of extreme loss in the tail of the underlying distributions. Furthermore, Basak and Shapiro [2000] show that the use of VaR, which disregards the loss beyond the quantile of the underlying distribution, increases the extreme loss in the tail of the distribution. Yamai and Yoshiba [2001a] point out the same problem in the use of VaR for managing options and loan portfolios.

Based on this concept, we provide our definition of tail risk according to what kind of partial information about the tail is taken into account by risk measures. We take partial information since a single risk measure is not able to consider all information about the tail.

As a first step, we take the value of the distribution function at some level of loss as partial information on the tail of the profit and loss distributions. Suppose there are two portfolios X_1 and X_2. Also suppose, at some level of loss l, the value of the distribution function of X_1 is larger than the value of the distribution function of X_2. Then, the probability that the loss is larger than l is higher for portfolio X_1 than for portfolio X_2. Thus, any reasonable risk measure should consider X_1 to be the riskier portfolio.

From this observation, we define "first order tail risk" as follows.

Definition 4

We say that a risk measure $\rho(X)$ is free of first order tail risk with a threshold K if the following holds for any two random variables X_1 and X_2 with $\rho(X_1) < \rho(X_2)$.

$$F_1(x) \le F_2(x), \qquad \forall x \quad x \le K \tag{17}$$

where $F_1(x)$ and $F_2(x)$ are the distribution functions of X_1 and X_2.

This definition essentially says that, when a risk measure $\rho(X)$ is free of first order tail risk with a threshold K, the portfolio with the smallest $\rho(X)$ has the lowest probabilities of any loss beyond the threshold K. Thus, a risk measure free of first order tail risk takes into account partial information about the tail.

The following theorem shows the relationship between first order tail risk and first order stochastic dominance.

Theorem 7

When portfolios are ranked by first order stochastic dominance, a risk measure consistent with first order stochastic dominance is free of first order tail risk with any level of threshold.

Proof

Let X_1 and X_2 denote two random variables that are ranked by first order stochastic dominance. Suppose a risk measure $\rho(X)$ is consistent with first order stochastic dominance and $\rho(X_1) < \rho(X_2)$ holds.

Since X_1 and X_2 are ranked by first order stochastic dominance, $X_1 \geq_{FSD} X_2$ holds. From the definition of first order stochastic dominance, Equation (17) holds, with any level of threshold K. QED

Thus, when portfolios are ranked by first order stochastic dominance, the risk measure $\rho(X)$ is free of first order tail risk with any level of threshold K.

On the other hand, when portfolios are not ranked by first order stochastic dominance, one is unable to tell whether a risk measure is free of first order tail risk. We need a more applicable definition of tail risk since the condition that portfolios are ranked by first order stochastic dominance is strict.

As a more applicable definition, we define "second order tail risk" as follows.

Definition 5

A risk measure $\rho(X)$ is free of second order tail risk with a threshold K if the following holds for any two random variables X_1 and X_2 with $\rho(X_1) < \rho(X_2)$.

$$\int_{-\infty}^{x} (x-u) f_1(u) du \leq \int_{-\infty}^{x} (x-u) f_2(u) du, \qquad \forall x \quad x \leq K \qquad (18)$$

where $f_1(x)$ and $f_2(x)$ are the density functions of X_1 and X_2.

This definition uses the expectation as partial information on the tail. This is a more applicable definition than first order tail risk since it penalizes larger losses more than smaller ones.

From Theorem 1, Equation (18) is equivalent to the following.

$$F_1^{(2)}(x) \leq F_2^{(2)}(x) \quad \forall x \quad x \leq K. \qquad (19)$$

The following theorem[8] holds in the same way as Theorem 7 does.

Definition 8

When portfolios are ranked by second order stochastic dominance, a risk measure consistent with second order stochastic dominance is free of second order tail risk with any level of threshold.

The relationship between second order tail risk and first order tail risk is characterized by the following theorem.

Theorem 9

When portfolios are ranked by first order stochastic dominance, a risk measure free of second order tail risk with any level of threshold is also free of first order tail risk with any level of threshold.

This theorem comes from Theorem 6 and the definitions of first and second order tail risk.

We are unable to determine whether a risk measure is free of second order tail risk when portfolios are not ranked by second order stochastic dominance. We may need a more applicable concept of tail risk in this case.

As a more applicable definition, we define n-th order tail risk as follows.

---

[8] This theorem is consistent with a result of Rothschild and Stiglitz [1970]. They say that second order stochastic dominance of portfolio X over portfolio Y is equivalent to "portfolio Y having more weight in the tails than portfolio X."

Definition 6

We say that a risk measure $\rho(X)$ is free of n-th order tail risk with a threshold K if the following holds for any two random variables X_1 and X_2 with $\rho(X_1) < \rho(X_2)$.

$$\int_{-\infty}^{x} (x-u)^{n-1} f_1(u)du \le \int_{-\infty}^{x} (x-u)^{n-1} f_2(u)du \text{ , } \forall x \quad x \le K \tag{20}$$

where $f_1(x)$ and $f_2(x)$ are the density function of X_1 and X_2.

This definition uses the $(n-1)$-th lower partial moment as partial information on the tail. This is a more applicable definition of tail risk than second order tail risk since it penalizes larger losses more than smaller ones because it takes the $(n-1)$-th power of the loss.

From Theorem 3, Equation (20) is equivalent to the following equation.

$$F_1^{(n)}(x) \le F_2^{(n)}(x) \quad \forall x \quad x \le K \text{ .} \tag{21}$$

The following theorem holds in the same way as Theorem 7 does.

Theorem 10

When portfolios are ranked by n-th order stochastic dominance, a risk measure consistent with n-th order stochastic dominance is free of n-th order tail risk with any level of threshold.

The relationship between different orders of tail risk is characterized by the following theorem. This holds in the same way as Theorem 9 does.

Theorem 11

When portfolios are ranked by n-th order stochastic dominance, a risk measure free of $(n+1)$-th order tail risk with any level of threshold is also free of n-th order tail risk with any level of threshold.

III. VaR and Expected Shortfall

In this chapter, we study whether expected shortfall[9] and VaR[10] are consistent with expected utility maximization and whether they are free of tail risk. We show in Chapter II that a risk measure consistent with n-th order stochastic dominance is also consistent with expected utility maximization and free of tail risk, if portfolios are ranked by n-th order stochastic dominance. Thus, we check whether expected shortfall and VaR are consistent with stochastic dominance in order to study their consistency with expected utility maximization and elimination of tail risk.

A. VaR

In this section, we show that VaR is consistent with expected utility maximization and free of tail risk under two conditions. The first is that portfolios are ranked by first order stochastic dominance. The second is that the underlying distributions are elliptical.

1. Consistency with first order stochastic dominance

Levy and Kroll [1978] show that VaR is consistent with first order stochastic dominance as follows (Levy and Kroll [1978] Theorem 1').

---

9 VaR at the 100(1-α)% confidence level, denoted $VaR_\alpha(X)$, is the lower 100α percentile of the profit-loss distribution. This is defined by the following equation.

$$VaR_\alpha(X) = -\inf\{x \mid P[X \le x] > \alpha\},$$

where X is the profit-loss of a given portfolio. $\inf\{x \mid A\}$ is the lower limit of x given event A, and $\inf\{x \mid P[X \le x] > \alpha\}$ indicates the lower 100α percentile of profit-loss distribution.

10 Expected shortfall is the conditional expectation of loss given that the loss is beyond the VaR level. When the underlying distributions are continuous, expected shortfall at the 100(1-α)% confidence level ($ES_\alpha(X)$) is defined by the following equation.

$$ES_\alpha(X) = E[-X \mid -X \ge VaR_\alpha(X)].$$

When the underlying distributions are discrete, we have to adopt the definition by Acerbi and Tasche [2001], so that expected shortfall is sub-additive. See the Definition 2 of Acerbi and Tasche [2001] for details.

Theorem 12

VaR is consistent with first order stochastic dominance. That is, if we let X_1 and X_2 be random variables denoting profit and loss of any two portfolios, the following holds.

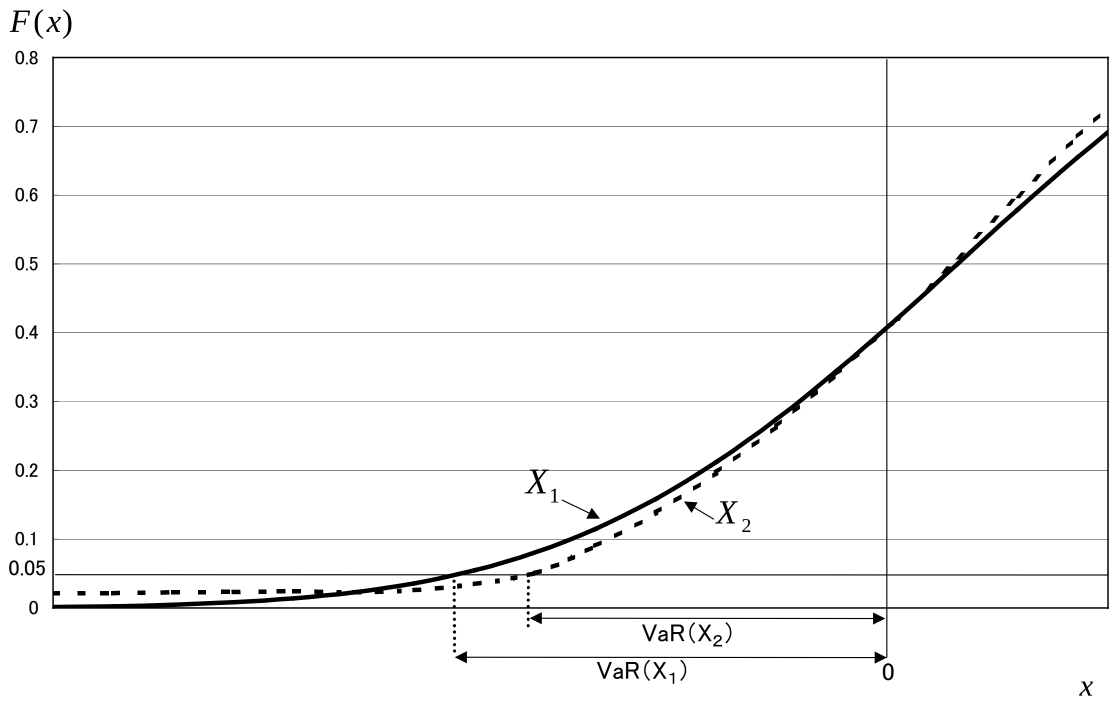
$$X_1 \geq_{FSD} X_2 \Rightarrow VaR_\alpha(X_1) \leq VaR_\alpha(X_2). \qquad (22)$$

Thus, when portfolios are ranked by first order stochastic dominance, VaR is consistent with expected utility maximization and is free of tail risk (first order tail risk).

However, the condition that portfolios are ranked by first order stochastic dominance is too strict to hold in practice. This condition means that the value of the distribution function of one variable is always larger than that of the other.

While VaR is consistent with first order stochastic dominance, it is not generally consistent with second order stochastic dominance, as is shown by Guthoff, Pfingsten and Wolf [1997]. We describe this inconsistency using the illustration in Guthoff, Pfingsten and Wolf [1997]. Figure 2 shows the distribution functions of two random variables X_1 and X_2, where $X_1 \geq_{SSD} X_2$ holds. VaR at the 95% confidence interval, or the 5% quantile of the profit-loss distribution, corresponds to the point where the distribution function and the horizontal line at the cumulative probability of 5% intersect. In this case, $VaR(X_1) > VaR(X_2)$ while $X_1 \geq_{SSD} X_2$. Thus, X_1 is preferred to X_2 based on VaR while X_2 is preferred to X_1 based on second order stochastic dominance. This means that the ranking of portfolios according to VaR contradicts the ranking of portfolios according to second order stochastic dominance.

Figure 2 Inconsistency of VaR and second order stochastic dominance



2. Elliptical Distributions

VaR is consistent with expected utility maximization and is free of tail risk when the underlying profit-loss distribution is an elliptical distribution.

Elliptical distributions are defined as follows.

Definition 7

A n-dimensional random vector $R=[R_1 \cdots R_n]^T$ has an elliptical distribution if the density function of R (denoted by $f(R)$) is represented below with a function $\varphi(\cdot\ ;n)$:

$$f(R;\theta,\Sigma)=\frac{1}{|\Sigma|^{1/2}}\varphi((R-\theta)^T\Sigma^{-1}(R-\theta);n), \qquad (23)$$

where Σ is a n-dimensional positive definite matrix ("scale parameter matrix"), θ is a n-dimensional column vector ("location parameter vector").

Elliptical distributions include the normal distribution as a special case, as well as the Student's t-distribution and the Cauchy distribution. Elliptical distributions are called "elliptical" since the contours of equal density are ellipsoids (See Fang and Anderson [1990] for the concepts and definitions of elliptical distributions).

VaR has useful properties when the underlying distributions are elliptical. The following is the most important property of VaR in an elliptical distribution (See Embrechts, McNeil, and Straumann[1998])[11].

Theorem 13

When a random variable X has an elliptical distribution with finite variance $V[X]$, VaR at the $100(1-\alpha)\%$ confidence level ($VaR_\alpha(X)$) is represented as follows.

$$VaR_\alpha(X) = E[X] + q_\alpha \sqrt{V[X]}, \qquad (24)$$

where q_α is the 100α percentile of the standardized distribution of this type.

This theorem shows that VaR and the standard deviation share the same properties when the underlying distribution is elliptical[12]. In particular, VaR, like the standard deviation, is consistent with second order stochastic dominance in an elliptical distribution.

---

[11] This theorem holds since the elliptical distributions share many properties with the normal distribution: the linear combination of elliptically distributed random vectors is also elliptical; the variance of an elliptically distributed random variable is a scalar multiple of the scale parameter.

[12] This holds only if the underlying distributions are of the same type of elliptical distribution in all portfolios. For example, if one portfolio has a normal distribution and another has the Pareto distribution, VaR does not have the same properties as the standard deviation.

Theorem 14

VaR is consistent with second order stochastic dominance when portfolios' profits and losses have an elliptical distribution with finite variance and the same mean.

Proof

According to the Proposition 6 of Ogryczak and Ruszczynski [1999], the standard deviation is consistent with second order stochastic dominance if the mean of profit and loss is equal across portfolios.

Let X_1 and X_2 denote profit and loss of two portfolios with equal mean. Then,

$$X_1 \geq_{SSD} X_2 \;\Rightarrow\; \sqrt{V[X_1]} \leq \sqrt{V[X_2]}. \tag{25}$$

Therefore, from Equation (24) and $E[X_1] = E[X_2]$,

$$\begin{aligned} X_1 \geq_{SSD} X_2 \;&\Rightarrow\; \sqrt{V[X_1]} \leq \sqrt{V[X_2]} \\ &\Rightarrow E[X_1] + q_\alpha \sqrt{V[X_1]} \leq E[X_2] + q_\alpha \sqrt{V[X_2]} \;\Rightarrow\; VaR_\alpha(X_1) \leq VaR_\alpha(X_2). \end{aligned} \tag{26}$$

This shows that VaR is consistent with second order stochastic dominance.

QED

Thus, VaR is consistent with second order stochastic dominance if the underlying distribution is elliptical and the mean of profit and loss are equal across portfolios[13]. From Theorem 2 and 8, VaR is consistent with expected utility maximization and free of tail risk under this condition.

Elliptical distributions include fat-tailed distributions such as the Student's t-distribution and the Pareto distribution. Thus, the fat tails of the underlying distributions do not necessarily indicate VaR's inconsistency with expected utility maximization and failure to eliminate tail risk[14].

---

[13] Selecting a minimum-risk portfolio within the portfolios of equal mean return is the first step in the mean-risk analysis, which is the most popular approach in financial practice.

[14] This holds only if the underlying distributions are of the same type of elliptical distribution in all portfolios. See Footnote 12.

B. Expected shortfall

In this section, we show that expected shortfall is consistent with expected utility maximization and free of tail risk if portfolios are ranked by second order stochastic dominance. This holds since expected shortfall is consistent with second order stochastic dominance.

The following theorem shows that expected shortfall is consistent with second order stochastic dominance.

Theorem 15

Expected shortfall is consistent with second order stochastic dominance.

Proof

Let X be a random variable denoting the profit and loss of a portfolio. We suppose X has a density function $f(x)$.

Expected shortfall at the $100(1-\alpha)\%$ confidence level is:

$$ES_\alpha(X) = E[-X \mid -X \geq VaR_\alpha(X)] = \frac{E[-X; -X \geq VaR_\alpha(X)]}{P[-X \geq VaR_\alpha(X)]}$$

$$= \frac{1}{\alpha}\int_{-\infty}^{q(\alpha)} (-x) f(x) dx,$$

where $q(\alpha)$ is the α-quantile of X. (27)

Let $F(x)$ denote the distribution function of X and suppose $F(x) = t$. Then, the following equation holds from $f(x)dx = dt$, $F(q(\alpha)) = \alpha$ and $F(-\infty) = 0$.

$$ES_\alpha(X) = \frac{1}{\alpha}\int_{-\infty}^{q(\alpha)} (-x) f(x) dx = -\frac{1}{\alpha}\int_0^\alpha F^{-1}(t) dt = -\frac{1}{\alpha}\int_0^\alpha q(t) dt. \qquad (28)$$

From Theorem 5' of Levy and Kroll [1978][15], the following holds for any two random variables X_1 and X_2.

---

15 Bertsimas, Lauprete and Samarov [2000] first adopted the result of Levy and Kroll [1978] to show the consistency of expected shortfall with second order stochastic dominance. Ogryczak and Ruszczynski [2001] independently prove Theorem 5' of Levy and Kroll [1978] with conjugate convex functions.

$$X_1 \geq_{SSD} X_2 \iff \int_0^{\alpha} q_1(t)dt \geq \int_0^{\alpha} q_2(t)dt \quad \forall\alpha \quad (0 \leq \alpha \leq 1), \tag{29}$$

where $q_1(t)$ and $q_2(t)$ are t-quantiles of X_1 and X_2.

Thus, from Equations (28) and (29), the following holds.

$$X_1 \geq_{SSD} X_2 \Rightarrow ES_{\alpha}(X_1) \leq ES_{\alpha}(X_2). \tag{30}$$

This shows that expected shortfall is consistent with second order stochastic dominance.

QED

From this theorem, expected shortfall is shown to be consistent with expected utility maximization and free of tail risk if portfolios are ranked by second order stochastic dominance.

Thus, expected shortfall is consistent with expected utility maximization and free of tail risk under more lenient conditions than VaR is. In Section A, we show that VaR is consistent with expected utility maximization and free of tail risk if portfolios are ranked by first order stochastic dominance or if the underlying distributions are elliptical. This condition for VaR is more strict than the condition for expected shortfall since portfolios that are ranked by second order stochastic dominance includes portfolios that are ranked by first order stochastic dominance and portfolios whose underlying distributions are elliptical with equal mean.

The condition for expected shortfall, however, is not general. Expected shortfall is neither consistent with expected utility maximization nor free of tail risk, if portfolios are not ranked by second order stochastic dominance. Thus, one may need a risk measure that is consistent with third or higher order stochastic dominance to deal with such portfolios.

C. An alternative: n-th lower partial moment

When portfolios are not ranked by second order stochastic dominance, expected shortfall is no longer consistent with expected utility maximization or free of tail risk.

An alternative to expected shortfall in this case is the lower partial moment with second or higher order. The n-th lower partial moment is defined below.

$$LPM_{n,K}(X) = E[\{(K - X)^+\}^n] = \int_{-\infty}^{K} (K - u)^n f(u)du ,$$

where K is a constant.

From the definition of stochastic dominance, the n-th lower partial moment is consistent with (n+1)-th order stochastic dominance. Thus, it is consistent with expected utility maximization and free of tail risk, as long as portfolios are ranked by (n+1)-th order stochastic dominance.

The n-th lower partial moment, however, has several disadvantages compared to expected shortfall. The n-th lower partial moment may not be comparable across various classes of portfolios since one has to set the same level of constant K across all classes of portfolios[16]. Furthermore, the n-th lower partial moment is not sub-additive while expected shortfall is. This means that the n-th lower partial moment does not consider risk reduction by portfolio diversification effects while expected shortfall does.

IV. Problems with Expected Shortfall

Chapter III shows that, when portfolios are not ranked by second order stochastic dominance, expected shortfall is no longer consistent with expected utility maximization or free of tail risk. This chapter shows a simple example of this kind of situation.

Table 1 shows the payoff of two sample portfolios A and B. The

[16] One way to make the n-th lower partial moment comparable across portfolios is to set K at some “target” or “benchmark” return. However, this may be difficult since K becomes stochastic in this case.

expected payoffs of those portfolios are equal at 97.05. We assume that the initial investment amount in portfolios A and B are equal at 97.05.

Table 1 Payoff of the sample portfolio[17]

Portfolio A			Portfolio B		
payoff	loss	Probability	payoff	Loss	probability
100.00	– 2.95	50.000%	98	– 0.95	50.000%
95.00	2.05	49.000%	97	0.05	49.000%
50.00	47.05	1.000%	90	7.05	0.457%
			20	77.05	0.543%

Most of the time, both portfolios A and B do not incur large losses. The probability that the loss is less than 10 is about 99% for both portfolios. However, there is a very small probability that they may incur extreme loss. The magnitude of extreme loss is higher for portfolio B since portfolio B may lose three quarters of its value while portfolio A never loses more than half of its value. Thus, portfolio B is considered risky when one is worried about extreme loss.

We calculate the expected utility, VaR, expected shortfall, and the second lower partial moment of portfolios A and B. We use a log function ($\ln W$) and a polynomial function with degree three ($-W^3/3+10{,}000W$) as utility functions[18], and take 99% as the confidence level of VaR and expected shortfall. We set a constant K for the second lower partial moment at –1. Table 2 shows the result.

---

[17] The numbers for probability are rounded off to the third decimal place.

[18] Both utility functions satisfy $U'(W) \geq 0$ and $U''(W) \leq 0$ in the range of $0 \leq W \leq 100$. Thus, they represent unsatiated and risk-averse utility, and have consistency with second order stochastic dominance in the sense of Theorem 2. On the other hand, as for $U'''(W)$, the log utility is positive while the polynomial utility is negative. This means that the log utility is consistent with third order stochastic dominance in the sense of Theorem 5 while the polynomial utility is not.

Table 2 Risk profiles of portfolio A and B

	Portfolio A	Portfolio B	Note
Expected Payoff	97.050	97.050	the same
Expected Utility (log function)	4.573	4.571	larger for portfolio A
Expected Utility (polynomial with degree 3)	663,379	663,439	larger for portfolio B
VaR (99% confidence level)	47.050	7.050	larger for portfolio A
Expected Shortfall (99% confidence level)	47.050	45.050	larger for portfolio A
Second Lower Partial Moment ($K=-1$)	21.746	31.564	larger for portfolio B

First of all, portfolios A and B are not ranked by second order stochastic dominance. The two types of utility functions, both of which are increasing and concave, provide conflicting preferences for portfolios A and B.

Second, expected shortfall fails to eliminate tail risk. As we explained above, the magnitude of extreme loss is much higher for portfolio B than for portfolio A. Thus, if a risk measure is free of tail risk, the risk measure should choose portfolio A since its extreme loss is smaller than portfolio B's. However, according to the result in Table 2, expected shortfall chooses portfolio B. This shows that expected shortfall fails to take into account extreme loss.

Third, expected shortfall is not consistent with expected utility maximization. Based on the log utility function, portfolio A is better since the expected utility is higher for portfolio A. On the other hand, based on expected shortfall, portfolio B is better since expected shortfall is lower for portfolio B.

Fourth, the second lower partial moment, which is consistent with third order stochastic dominance, chooses portfolio A, whose extreme loss is smaller than portfolio B's. This means that the lower partial moment with higher order is more effective in eliminating tail risk than expected shortfall.

The example in this chapter shows that expected shortfall is neither consistent with expected utility maximization nor free of tail risk, when portfolios are not ranked by second order stochastic dominance. The example also shows that the second lower partial moment is more effective in eliminating tail risk than expected shortfall.

V. Concluding Remarks

We compare two aspects of expected shortfall and Value-at-Risk (VaR): consistency with expected utility maximization and elimination of tail risk. We use the concept of stochastic dominance in studying the two aspects of risk measures.

We conclude that expected shortfall is more applicable than VaR in regard to both aspects. Expected shortfall is consistent with expected utility maximization and free of tail risk, under more lenient conditions than VaR is.

We show that the condition for expected shortfall is not general. Thus, expected shortfall has problems in certain circumstances.

References

Acerbi, C., C. Nordio, and C. Sirtori, "Expected Shortfall as a Tool for Financial Risk Management," Working Paper, 2001.

Acerbi, C., and D. Tasche, "Expected Shortfall: a natural coherent alternative to Value at Risk," Working Paper, 2001.

Artzner, P., F. Delbaen, J. M. Eber, and D. Heath, "Thinking Coherently," *Risk*, Vol. 10, No. 11, November 1997, pp. 68-71.

Artzner, P., F. Delbaen, J. M. Eber, and D. Heath, "Coherent Measures of Risk," *Mathematical Finance,* Vol. 9, No. 3, July 1999, pp. 203-228.

Basak, S., and A. Shapiro, "Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices," *The Review of Financial Studies*, Vol. 14, No. 2, Summer 2000, pp. 371-405.

Bawa, V. S., "Optimal Rules for Ordering Uncertain Prospects," *Journal of Financial Economics*, Vol. 2, No. 1, March 1975, pp. 95-121.

Bertsimas, D., G. J. Lauprete, and A. Samarov, "Shortfall as a Risk Measure: Properties, Optimization and Applications," Preprint, October 2000.

BIS Committee on the Global Financial System, "Stress Testing by Large Financial Institutions: Current Practice and Aggregation Issues," No. 14, April 2000.

Cumperayot, P. J., J. Danielsson, B. N. Jorgenson, and C. G. de Vries, "On the (Ir)Relevancy of Value-at-Risk Regulation," *Measuring Risk in Complex Stochastic Systems*, Springer Verlag, 2000, pp. 103-119.

Embrechts, P., A. McNeil, and D. Straumann, "Correlation and Dependency in Risk Management: Properties and Pitfalls," Preprint, ETH Zürich, 1998.

Fang, K.T., and T. W. Anderson, *Statistical Inference in Elliptically Contoured and Related Distributions*, Allerton Press, 1990.

Fishburn, P. C., "Mean-Risk Analysis with Risk Associated with Below-

Target Returns," *The American Economic Review*, Vol. 67, No. 2, March 1977, pp. 116-126.

Guthoff, A., A. Pfingsten, and J. Wolf, "On the Compatibility of Value at Risk, Other Risk Concepts, and Expected Utility Maximization," Diskussionsbeitrag 97-01, Westfälishce Wilhelms-Universität Münster, Institut für Kreditwesen, January 1997.

Huang, C. and R. H. Litzenberger, *Foundations for Financial Economics*, Prentice-Hall, 1993.

Ingersoll, J. E. Jr., *Theory of Financial Decision Making*, Rowman & Littlefield Publishers, 1987.

Levy, H., *Stochastic Dominance: Investment Decision Making under Uncertainty*, Kluwer Academic Publishers, 1998.

Levy, H. and Y. Kroll, "Ordering Uncertain Options with Borrowing and Lending," *The Journal of Finance*, Vol. 33, No. 2, 1978, pp. 553-574.

Ogryczak, W. and A. Ruszczynski, "From Stochastic Dominance to Mean-risk Models: Semideviations as Risk Measures," *European Journal of Operational Research*, Vol. 116, No. 1, 1999, pp. 33-50.

Ogryczak, W. and A. Ruszczynski, "Dual Stochastic Dominance and Related Mean-Risk Models," Rutcor Research Report, RRR 10-2001, 2001.

Pflug, G. C., "How to Measure Risk?" *Modelling and Decisions in Economics. Essays in Honor of Franz Ferschl*, Physica-Verlag, 1999.

Pflug, G. C., "Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk," *Probabilistic Optimization: Methodology and Applications*, Kluwer Academic Publishers, 2000, pp. 278-287.

Rockafeller, R. T. and S. Uryasev, "Optimization of Conditional Value-at-Risk," *Journal of Risk*, Vol. 2, No. 3, Spring 2000, pp. 21-41.

Rothschild, M., and J. E. Stiglitz, "Increasing Risk: I. A Definition," *Journal of Economic Theory*, Vol. 2, No. 3, 1970, pp. 225-243.

Von Neumann, J., and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey, 1953.

Yamai, Y., and T. Yoshiba, "On the Validity of Value at Risk: Comparative Analyses with Expected Shortfall," Institute for Monetary and Economic Studies Discussion Paper Series, No. 2001-E-14, Bank of Japan, March 2001a.

Yamai, Y., and T. Yoshiba, "Comparative Analyses of Expected Shortfall and Value-at-Risk: estimation error, decomposition and optimization," Institute for Monetary and Economic Studies Discussion Paper Series, No. 2001-E-21, Bank of Japan, July 2001b.