

# Relevance of Maximum Drawdown in the Investment Fund Selection Problem When Utility is Nonadditive

Daehwan Kim

Assistant Professor, Department of Economics, Konkuk University  
Hwayangdong 1, Gwangjingu, Seoul, 143-701, South Korea  
dkim@konkuk.ac.kr

## Abstract

This paper explores the circumstances under which maximum drawdown becomes relevant to a rational investor's choice of an investment fund. We adopt the framework of the nonadditive expected utility theory of Schmeidler (1989), and consider the consequence of extreme uncertainty over cash flow and also of extreme uncertainty aversion. We show that an investor facing extreme *uncertainty* makes a choice based on maximum drawdown. We also show that an extremely *uncertainty-averse* investor makes a choice based on maximum drawdown even if uncertainty is not extreme. An implication for the mean-variance analysis of Markowitz is discussed as well.

Keywords: maximum drawdown, nonadditive utility, uncertainty over cash flow, uncertainty aversion

JEL Classifications: G11, D81, C44

# 1 Introduction

Consider an investor who wants to select an investment fund among many alternatives. Suppose that the investor has a predictive distribution of returns for each fund. The investor would most likely consider mean and variance of returns as important factors. Would the investor also consider some moments of maximum drawdown as relevant factors?

Maximum drawdown is the biggest peak-to-trough decline in the price of a fund, given a particular realization of price series.<sup>1</sup> It is the worst return that an investor could experience, i.e., the return of an investor who buys the fund at the highest price and sells the fund at the lowest price. Note that, looking forward, maximum drawdown is a random variable, and we may talk about moments of maximum drawdown such as mean of maximum drawdown or variance of maximum drawdown.

Institutional investors often consider maximum drawdown as an important criterion in selecting a fund. It appears that institutional investors consider the maximum drawdown of a fund to be indicative of the fund's risk. The widespread use of maximum drawdown is not hard to document. See, for example, Grossman and Zhou (1993), Acar and James (1997), Harding, Nakou, and Najjir (2003), Leal and Mendes (2005), Alexander and Baptista (2006), and Hayes (2006).<sup>2</sup>

Institutional investors' adoption of maximum drawdown in their decision making is puzzling indeed. Maximum drawdown is a measure of the worst outcome, a near-zero probability event if the distribution of buy and sell timing

---

<sup>1</sup>If a fund pays dividends, maximum drawdown is defined with respect to the total return index of the fund rather than the price of the fund.

<sup>2</sup>Regulators take maximum drawdown seriously as well. The US Commodity Futures Trading Commission requires futures managers to report maximum drawdown in their performance reporting. See Harding, Nakou, and Najjir (2003).

is bell-shaped. Maximum drawdown is relevant only if one buys and sells the fund at the worst possible moments. This is extremely unlikely even for investors with no market timing skills at all.

Despite the widespread use of maximum drawdown among practitioners, financial economists have not paid much attention to this concept. Grossman and Zhou (1993) and Alexander and Baptista (2006) consider the portfolio choice problem with a drawdown constraint, and Magdon-Ismael, Atiya, Pratap, and Abu-Mostafa (2004), Rebonato and Gaspari (2006), and Pospisil and Vecer (2008) study the statistical properties of maximum drawdown. However, none of them explains why maximum drawdown enters their problems in the first place. Cvitanic and Karatzas (1999) and Chekhlov, Uryasev, and Zabarankin (2005) discuss classes of risk measures that include maximum drawdown. However, their discussions do not highlight why investors might or should use it. Our paper is an attempt to fill this gap in the literature. The question we focus on is why investors consider maximum drawdown in making their investment choice.

We start our investigation with the ideas from Dow and Werlang (1992). Dow and Werlang describe the situation where maximum loss becomes relevant to the decision maker. While maximum drawdown and maximum loss are two different concepts, we borrow analytical tools from Dow and Werlang. Adopting the nonadditive expected utility theory of Schmeidler (1986, 1989), Dow and Werlang show that maximum loss can be the only factor in the decision making, under a situation that they described as “extreme uncertainty aversion.” We also adopt the nonadditive expected utility theory of Schmeidler as our framework. However, contrary to the analysis of Dow and Werlang, we make distinction between uncertainty aversion and uncertainty. We show that under extreme uncertainty only maximum drawdown matters, while under ex-

treme uncertainty aversion other factors enter in importance. This distinction is useful, as extreme uncertainty is rather a theoretical possibility while extreme uncertainty aversion is more likely to have practical relevance. The idea of separating uncertainty and uncertainty aversion is influenced by Klibanoff, Marinacci, and Mukerji (2005), who discuss ambiguity aversion and ambiguity as two separate concepts in the context of the maximin utility theory.

Another significant departure from Dow and Werlang is the fact that we model the uncertainty over cash flow as well as uncertainty over returns. We consider an investor who expects cash inflow sometime in the near future, but does not know the exact timing of cash inflow. The investor also expects that there will be cash outflow sometime in the future, but does not know the exact timing of cash outflow either. Having no market timing skills, the investor plans to buy the fund when there is cash inflow and plans to sell it when there is cash outflow. Thus, for this investor, uncertainty over cash flow is also the uncertainty over buy and sell timing. One could characterize our analysis as an extension of Dow and Werlang's analysis to a multi-period setting with uncertainty over cash flow.

While the main goal of the paper is to explain the relevance of maximum drawdown, we hope to contribute to the literature on the nonadditive expected utility theory. Schmeidler (1986, 1989) identifies uncertainty aversion with the convexity of a probability function, which is defined over a sequence of sets (events). We relate uncertainty aversion to the convexity of a function that is defined over a sequence of real numbers. This allows us to think of uncertainty aversion geometrically, in a two-dimensional space, and also to compare the uncertainty aversions of two individuals. We provide a simple geometric interpretation for the statement "individual  $j$  is more uncertainty averse than individual  $i$ ." We also discuss the implication of uncertainty aversion on the

mean-variance analysis of Markowitz (1952). Given the popularity of the mean-variance analysis among practitioners, it is worth exploring the link between maximum drawdown and the mean variance analysis.

We start with the analysis of uncertainty in Section 2. Let  $\Omega$  be the set of states of the world, let  $\sigma_i$  be the nonadditive probability of individual  $i$ , let  $a$  be an act (random variable) defined on  $\Omega$ , and let  $E_1, \dots, E_K$  be events that partition  $\Omega$  such that  $a$  is constant on each  $E_k$ . Nonadditivity means that the sum of the probabilities of these events,  $\sum_{k=1}^K \sigma_i(E_k)$ , can be less than 1. The difference,  $1 - \sum_{k=1}^K \sigma_i(E_k)$ , indicates the existence of uncertainty. It corresponds, according to Schmeidler (1989), to the lack of “decision maker’s confidence in the probability assessment.” The implication of nonadditivity is clearest if we consider an extreme situation. Suppose that individual  $i$  has absolutely no confidence in the probability assessment such that the probability of every event, except the  $\Omega$  event which always has the probability of 1, approaches 0.<sup>3</sup> Here, the only relevant event is the  $\Omega$  event, as all the other events have 0 probability. Therefore, individual  $i$ ’s utility depends only on the  $\Omega$  event. If  $a$  is the return of a fund, the  $\Omega$  event can be described as “the return is better than or equal to the worst possibility.” Thus, we can say that individual  $i$ ’s utility depends only on the worst possibility. As mentioned earlier, this idea appeared first in Dow and Werlang (1992).

In Section 3, we examine the implication of extreme uncertainty aversion. We first clarify the concept of uncertainty and the concept of uncertainty aversion. We consider uncertainty as a characteristic of an act, and uncertainty aversion as a characteristic of an individual. While uncertainty is captured by the non-additivity of probability, uncertainty aversion is captured by the convexity of

---

<sup>3</sup>Any subset of  $\Omega$  can be considered as an event. Thus,  $\Omega$  itself can be considered an event as well. As an event,  $\Omega$  allows any possible states of the world.

probability.<sup>4</sup> Quite often, probability exhibits both nonadditivity and convexity, but these two are not identical. To interpret the convexity of probability geometrically, we consider a sequence of increasing events,  $F_1 \subset \dots \subset F_L = \Omega$ . Given this sequence, we construct a pseudo-probability  $\bar{\sigma}_i^*$ , which is additive and shares certain features with  $\sigma_i$ . We call  $\bar{\sigma}_i^*$  a linearization, or *additive extension*, of  $\sigma_i$ . We show that the convexity of  $\sigma_i$  is equivalent to the convexity of a function from  $\bar{\sigma}_i^*$  to  $\sigma_i$ . We denote the new function by  $g_i$ . The main assumption required for the equivalence is that probability space  $\Omega$  is uniform. One important advantage of considering the convexity of  $g_i$  instead of the convexity of  $\sigma_i$  is that we can easily compare the convexities of two functions, say,  $g_i$  and  $g_{i'}$ . Extreme uncertainty aversion is achieved when the convexity of  $g_i$  is extreme, in which case  $g_i$  is linear everywhere except near  $F_L = \Omega$ .<sup>5</sup> In this case, individual  $i$ 's utility becomes the sum of conventional expected utility (where the expectation is based on the pseudo-probability  $\bar{\sigma}_i^*$ ) and the utility from the worst possibility.

In Section 4 of this paper, we apply the tools we develop in the preceding sections to the investment fund selection problem. We consider an investor who faces uncertainty over cash inflow and outflow timing. To focus on the uncertainty over cash flow, we assume that returns are purely risky, i.e. the randomness of returns is captured by a conventional additive probability. Then we show that if uncertainty is extreme, the investor selects a fund based on maximum drawdown. We also show that if the investor is extremely uncertainty averse, then the investor selects a fund based on maximum drawdown

---

<sup>4</sup>Nonadditivity means that  $\sigma_i(A \cup B) \neq \sigma_i(A) + \sigma_i(B)$  when  $A$  and  $B$  are mutually exclusive events. Convexity means that  $\sigma_i(A \cup B) \geq \sigma_i(A) + \sigma_i(B)$  when  $A$  and  $B$  are mutually exclusive events.

<sup>5</sup> $y = g_i(x)$  is bounded below by  $y = cx$  ( $0 < c \leq 1$ ), and always passes through  $(x, y) = (1, 1)$ . When convexity increases,  $y = g_i(x)$  approaches  $y = cx$ ,  $0 \leq x \leq 1$ , and  $x = 1$ ,  $c \leq y \leq 1$ .

and expected utility (where the expectation is based on the pseudo-probability  $\bar{\sigma}_i^*$ ).

Section 5 summarizes our analysis. Our answer to the question stated at the beginning can be summarized as follows: An investor considers maximum drawdown in the investment fund selection problem if uncertainty over cash flow is extreme or if the investor is extremely uncertainty averse. Section 5 also discusses a possible extension of the mean-variance analysis of Markowitz (1952).

While the current paper is formulated within the framework of nonadditive expected utility, the maximin expected utility of Gilboa and Schmeidler (1989) suggests an alternative formulation. The near equivalence between the maximin expected utility and the nonadditive expected utility is discussed by Gilboa and Schmeidler. Chamberlain (2000) discusses the portfolio choice problem under the maximin expected utility. The fund selection problem of this paper can be considered complementary to Chamberlain's analysis.

The model that we describe in the paper is discrete. That is, both the return and the time are treated as discrete variables. A discrete model's advantage over a continuous model is twofold. First, discrete model formulas are easier to interpret and have more intuitive appeal.<sup>6</sup> Second, the interpretation of maximum drawdown is also easier when the return is a discrete variable with a finite range. In any case, it would be fairly straightforward to present the same idea in a continuous model.

---

<sup>6</sup>The assumption that the probability space is uniform would be unnecessary in a continuous model. In this respect, a continuous model is simpler.

## 2 Consequence of Extreme Uncertainty

Dow and Werlang (1992) discussed a portfolio choice problem in a situation they described as “extreme uncertainty aversion.” We refine their discussion by adopting the concept of uncertainty aversion that is separate from the concept of uncertainty. Following Schmeidler (1989), we identify uncertainty with the nonadditivity of the probability, and uncertainty aversion with the convexity of the probability.

That uncertainty and uncertainty aversion are two distinct concepts (or that such distinction is useful) can be illustrated in the following way. The statement “act (random variable)  $a$  has more uncertainty than act  $b$ ” is probably more intuitive than the statement “act  $a$  is more uncertainty averse than act  $b$ ”. On the other hand, saying “individual  $j$  is more uncertainty averse than individual  $i$ ” makes more sense than “individual  $j$  has more uncertainty than individual  $i$ .” That is, uncertainty is a characteristic of an act given a non-additive probability space, while uncertainty aversion is a characteristic of an individual facing a nonadditive probability space.<sup>7</sup> Most importantly, uncertainty aversion is defined without reference to any particular act. We note that our usage of uncertainty and uncertainty aversion does not correspond to that of Dow and Werlang (1992). What Dow and Werlang describe as “extreme un-

---

<sup>7</sup>Similarly, risk is a characteristic of an act, while risk aversion is a characteristic of an individual. However, the analogy goes only this far. While no consideration of a probability space is necessary to talk about risk aversion, the same thing cannot be said about uncertainty aversion. In the formulation of Schmeidler (1989) and also in the application of Dow and Werlang (1992), an individual’s utility depends only on two things: the felicity function and the nonadditive probability. As we attempt to relate four concepts—risk, risk aversion, uncertainty, and uncertainty aversion—to two functions, some mix-ups are unavoidable. Klibanoff, Marinacci, and Mukerji (2005) avoided this problem by introducing additional functions into the individual utility. Our goal is to stay within the framework of Schmeidler, if possible, and did not follow the strategy of Klibanoff, Marinacci, and Mukerji (2005).

certainty aversion” is called “extreme uncertainty” in this paper. As Dow and Werlang do not consider uncertainty and uncertainty aversion as two separate concepts, we believe that different usages do not indicate disagreement over the concepts. In fact, we confirm Dow and Werlang’s conclusion in this paper.

In this section we characterize uncertainty and extreme uncertainty. Then we discuss the decision making of an individual who faces extreme uncertainty. Uncertainty aversion is discussed in the next section.

Let  $\Omega$  be a set with a finite number of elements. A set function  $\sigma_i$  defined over  $\Omega$  is called *nonadditive probability* if the following three properties are satisfied: (i)  $\sigma_i(\emptyset) = 0$ . (ii)  $\sigma_i(\Omega) = 1$ . (iii)  $\sigma_i$  is monotonic, i.e., for all  $A, B \subset \Omega$ ,  $A \subset B$  implies  $\sigma_i(A) \leq \sigma_i(B)$ . The following lemma from Gilboa and Schmeidler (1994) provides a useful characterization of nonadditive probability.<sup>8</sup>

**Lemma 1.** *Nonadditive probability  $\sigma_i$  has the following Moebius transform:*

$$\sigma_i(B) = \sum_{A \subseteq B} \varphi_{\sigma_i}(A) \quad (1)$$

$\varphi_{\sigma_i}$  is defined as

$$\varphi_{\sigma_i}(A) = \sigma_i(A) - \sum_{\substack{I \subseteq \{1, \dots, L\} \\ I \neq \emptyset}} (-1)^{|I|+1} \sigma_i \left( \bigcap_{l \in I} A_l \right) \quad (2)$$

where  $A_l = A \setminus \{\omega_l\}$  and  $A = \{\omega_1, \dots, \omega_L\}$ .

$\varphi_{\sigma_i}(A)$  represents the *specific weight* that is assigned to  $A$  and cannot be further divided among its subsets. Lemma 1 states that the nonadditive probability of event  $B$  is determined as the total of specific weights of its subsets. Let us separate the subsets of  $B$  into two groups: the subsets of singletons

---

<sup>8</sup>See Theorem 3.3 of Gilboa and Schmeidler (1994). This Lemma is typically presented for the case when  $\sigma_i$  is total monotonic, i.e.  $\varphi_{\sigma_i}$  is non-negative. See Ghirardato (1997).

( $A \subseteq B$ ,  $|A| = 1$ ) and all other subsets ( $A \subseteq B$ ,  $|A| \geq 2$ ). Then the right hand side of (1) is divided into two parts: the additive component based on the subsets of singletons, and the remaining nonadditive component. We denote the *additive component* of  $\sigma_i$  by  $\bar{\sigma}_i$ , and the *nonadditive component* by  $\tilde{\sigma}_i$ . That is,

$$\begin{aligned}\bar{\sigma}_i(B) &= \sum_{\substack{A \subseteq B \\ |A|=1}} \varphi_{\sigma_i}(A) \\ \tilde{\sigma}_i(B) &= \sum_{\substack{A \subseteq B \\ |A|>1}} \varphi_{\sigma_i}(A)\end{aligned}\tag{3}$$

The additive component  $\bar{\sigma}_i$  can be viewed as a “linearization” of  $\sigma_i$ , in the sense that  $\bar{\sigma}_i$  is the maximum linear (i.e. additive) component of  $\sigma_i$ . By dividing  $\bar{\sigma}_i$  by  $\bar{\sigma}_i(\Omega)$ , additive probability is obtained. We call it the *additive extension* of  $\sigma_i$  and denote it by  $\bar{\sigma}_i^*$ , i.e.,

$$\bar{\sigma}_i^*(B) = \frac{\bar{\sigma}_i(B)}{\bar{\sigma}_i(\Omega)}\tag{4}$$

If  $\sigma_i$  is additive,  $\sigma_i$  and  $\bar{\sigma}_i^*$  are identical. If  $\sigma_i$  is very nonlinear, then  $\sigma_i$  and  $\bar{\sigma}_i^*$  are far apart. This motivates characterization of uncertainty as the distance between  $\sigma_i$  and  $\bar{\sigma}_i^*$ .

Before discussing the concept of uncertainty, we make two assumptions regarding nonadditive probabilities. The first assumption says that the probability is uniform over the probability space.<sup>9</sup> Needless to say, even though the probability is uniform over the probability space, an act can have any kind of probability distribution over the event space. The second assumption says that the nonadditive probability measure is characterized by its additive component. If two events have identical additive components, then they have identical spe-

---

<sup>9</sup>This assumption would have been unnecessary in a continuous model. In a sense, this assumption requires that our discrete model is an approximation of a continuous world.

cific weights, identical nonadditive components, and identical probabilities.

**Assumption 1.**  $\sigma_i$  is uniform over  $\Omega$ . That is,  $\sigma_i(\{\omega\})$  is constant for all  $\omega \in \Omega$ .

**Assumption 2.** If  $\bar{\sigma}_i(A) = \bar{\sigma}_i(B)$ ,  $A, B \in \mathcal{A}$ , then  $\varphi_{\sigma_i}(A) = \varphi_{\sigma_i}(B)$ .

Assumption 1 and Assumption 2 together imply that each of  $\sigma_i(B)$ ,  $\bar{\sigma}_i(B)$ ,  $\tilde{\sigma}_i(B)$ , and  $\varphi_{\sigma_i}(B)$  depends only on the number of elements of  $\Omega$  that  $B$  includes.

In formulating these two assumptions, we are thinking of the following situation: An investor wants to select an investment fund among many alternatives. The only information that the investor has about each fund is the historical distribution of returns, say,  $r_1, \dots, r_T$ . Not having any other information, the investor creates histogram out of  $r_1, \dots, r_T$ , and treat it as the predictive distribution of future returns. Assumption 1 captures the aspect that a histogram is used as a predictive distribution. Needless to say, the histogram can have any shape. However, the probability space is uniform, allowing us to equally weight each realization.

Suppose further that the investor is not certain that the historical histogram is identical to the true data generating process. So the investor is not willing to allocate 100% probability to the historical histogram. He sets aside a fraction, say 30%, and reduces the probability of each event by up to 30%. The probabilities of some events are reduced more, while the probabilities of other events are reduced less. Nonetheless, probabilities of two events with the identical frequency remain identical. This is so because the investor does not have any information other than historical histogram, and there is no reason to assign different probabilities if one event occurred as frequently as the other event. Assumption 2 captures the idea that historical frequency determines predictive probability.

We now discuss the concept of uncertainty. The concept of uncertainty is

often illustrated by Ellsberg (1961) paradox, which indicates that a decision maker may prefer one random outcome (“lottery”) to another random outcome if the former has less uncertainty than the latter even though they have identical additive probabilities.<sup>10</sup> If probability is additive, then there is no uncertainty. This leads one to characterize and measure uncertainty as deviation from additivity.

Dow and Werlang (1992) proposed to measure the uncertainty of event  $A \subset \Omega$  as follows<sup>11</sup>:

$$c(\sigma_i, A) = 1 - \sigma_i(A) - \sigma_i(A^c) \quad (5)$$

We may decompose the uncertainty into two parts as follows:

$$c(\sigma_i, A) = c_+(\sigma_i, A) + c_+(\sigma_i, A^c) \quad (6)$$

where

$$c_+(\sigma_i, A) = \bar{\sigma}_i^*(A) - \sigma_i(A) \quad (7)$$

$c_+(\sigma_i, A)$  is the part of uncertainty attributable to  $A$  but not to  $A^c$ . It will be useful in later analysis.

---

<sup>10</sup>Ellsberg paradox goes something like the following. There are two urns, each containing hundred balls. Each ball is either black or red. For the first urn, it is known that 50 balls are black and that the other 50 balls are red. For the second urn, no such thing is known. A ball is drawn from each urn. There are four lotteries whose payoffs depend on the colors of balls drawn. The first lottery pays \$100 if the ball drawn from the first urn is black. Otherwise, it pays nothing. The second lottery pays \$100 only if the ball drawn from the first urn is red. The third lottery pays \$100 only if the ball drawn from the second urn is black. The fourth lottery pays \$100 only if the ball drawn from the second urn is red. Presented with these four lotteries, a decision maker tends to prefer the first two lotteries to the last two lotteries. This preference is not consistent with the expected utility theory of von Neumann and Morgenstern, and violates the “sure thing principle” of Savage (1972).

<sup>11</sup>As noted before, Dow and Werlang called it uncertainty aversion, but in the terminology of this paper, it is uncertainty.

We would like to consider uncertainty as a characteristic of an act, rather than as a characteristic of an event. Given that we have a measure of uncertainty for an event, it is natural to obtain the uncertainty of an act by summing the uncertainty of relevant events. An act is a function from  $\Omega$  to real numbers. Suppose that act  $a$  takes a value from  $\{\alpha_1, \dots, \alpha_K\}$  and that  $\alpha_1 > \dots > \alpha_K$ . We construct a sequence of events  $A_1, \dots, A_K$  such that  $A_k = \{\omega \in \Omega : a(\omega) \geq \alpha_k\}$ ,  $k = 1, \dots, K$ . Then the sum of the uncertainty of these events represents the *uncertainty of act  $a$*  given nonadditive probability  $\sigma_i$ :

$$\gamma(\sigma_i, a) = \sum_{k=1}^{K-1} c_+(\sigma_i, A_k) \bar{\sigma}_i^*(A_{k+1} - A_k) + c_+(\sigma_i, A_K^c) \bar{\sigma}_i^*(A_{K-1}^c - A_K^c) \quad (8)$$

In the above formula,  $\bar{\sigma}_i^*(A_{k+1} - A_k)$  and  $\bar{\sigma}_i^*(A_{k-1}^c - A_k^c)$  are normalization factors. Let us define  $\gamma_+(\sigma_i, a)$  as:

$$\gamma_+(\sigma_i, a) = \sum_{k=1}^{K-1} c_+(\sigma_i, A_k) \bar{\sigma}_i^*(A_{k+1} - A_k) \quad (9)$$

Then we can express  $\gamma(\sigma_i, a)$  as:

$$\gamma(\sigma_i, a) = \gamma_+(\sigma_i, a) + \gamma_+(\sigma_i, -a) \quad (10)$$

We call  $\gamma_+(\sigma_i, a)$  *downside uncertainty* of act  $a$  given  $\sigma_i$ , as it accumulates the uncertainty of an act from the highest payoff event  $A_1$  to the lowest payoff event  $A_K$ .  $\gamma_+(\sigma_i, -a)$  can be called *upside uncertainty* of act  $a$ . It accumulates the uncertainty of an act from the lowest payoff event  $A_{K-1}^c = A_K$  to the highest payoff event  $A_1^c$ .

By construction, uncertainty measure  $\gamma(\sigma_i, a)$  is bounded above by  $1 - \frac{1}{L}$ , where  $L$  is the number of elements in  $\Omega$ . Lemma 2 shows that  $1 - \frac{1}{L}$  is in fact the supremum of  $\gamma(\sigma_i, a)$ . Lemma 3 shows that the supremum of  $\gamma(\sigma_i, a)$  is

obtained if and only if the supremum of  $\gamma_+(\sigma_i, a)$  is obtained. Thus, we may work with downside uncertainty measure instead of uncertainty measure.

**Lemma 2.** *Let  $L$  be the number of elements in  $\Omega$ . Then (i)  $\sup \gamma_+(\sigma_i, a) = \frac{1}{2}(1 - \frac{1}{L})$  and (ii)  $\sup \gamma(\sigma_i, a) = 1 - \frac{1}{L}$ . In both cases, sup is found over all nonadditive probabilities and acts.*

*Proof.* See the appendix. □

**Lemma 3.** *Let  $\sigma_1, \sigma_2, \dots$  be a sequence of nonadditive probabilities and let  $a$  be an act. Then the following statements are equivalent: (i) The sequence of nonadditive probabilities and the act obtain the supremum of downside uncertainty.*

*That is,*

$$\lim_{n \rightarrow \infty} \gamma_+(\sigma_n, a) = \frac{1}{2}(1 - \frac{1}{L}) \quad (11)$$

(ii) *The sequence of events generated by the act satisfies the following:*

$$\bar{\sigma}_n^*(a = \alpha_k) = \frac{1}{L} \quad (12)$$

*for all  $k, 1 \leq k \leq K$ , and*

$$\lim_{n \rightarrow \infty} \sigma_n(a \geq \alpha_k) = 0 \quad (13)$$

*for all  $k, 1 \leq k \leq K - 1$ . (iii) The sequence of nonadditive probabilities and the act obtain the supremum of upside uncertainty.*

$$\lim_{n \rightarrow \infty} \gamma_+(\sigma_n, -a) = \frac{1}{2}(1 - \frac{1}{L}) \quad (14)$$

(iv) *The sequence of events generated by the act satisfies the following:*

$$\bar{\sigma}_n^*(a = \alpha_k) = \frac{1}{L} \quad (15)$$

for all  $k$ ,  $1 \leq k \leq K$ , and

$$\lim_{n \rightarrow \infty} \sigma_n(a \leq \alpha_k) = 0 \quad (16)$$

for all  $k$ ,  $2 \leq k \leq K$ . (v) The sequence of nonadditive probabilities and the act obtain the supremum of uncertainty.

$$\lim_{n \rightarrow \infty} \gamma(\sigma_n, -a) = 0 \quad (17)$$

*Proof.* See the appendix. □

Before describing individuals' decision making, we need to describe individuals' utility in some detail. Given act  $a$  and felicity function  $u_i$ , individual  $i$ 's *nonadditive expected utility* is given by

$$U_i(a) = \int_{\Omega} u_i(a) d\sigma_i \quad (18)$$

The integral in (18) should be interpreted as a Choquet integral (Schmeidler (1986)), i.e.

$$U_i(a) = \int_0^{\infty} \sigma_i(u_i(a) \geq \beta) d\beta \quad (19)$$

To simplify the formula, we assumed above that  $u_i$  is non-negative.<sup>12</sup> Let  $\{\alpha_1, \dots, \alpha_K\}$  be the set of values that  $a$  takes, such that  $\alpha_1 > \dots > \alpha_K$ . Let  $\alpha_0$  be a number greater than  $\alpha_1$ . Then nonadditive expected utility has the following representation:

$$U_i(a) = \sum_{k=1}^K u_i(\alpha_k) [\sigma_i(a \geq \alpha_k) - \sigma_i(a \geq \alpha_{k-1})] \quad (20)$$

---

<sup>12</sup>If we allow  $u_i$  to be negative, then expected utility should be written as  $\int_{-\infty}^0 [\sigma_i(u_i(a) \geq \beta) - 1] d\beta + \int_0^{\infty} \sigma_i(u_i(a) \geq \beta) d\beta$ .

If  $\sigma_i$  is additive, the above formula is reduced to the standard expected utility of von Neumann and Morgenstern (1947).

How do individuals make decision when they face extreme uncertainty? Lemma 3 showed that the supremum of uncertainty is achieved if and only if the supremum of downside uncertainty is achieved. Moreover, the the supremum of downside uncertainty is achieved if and only if

$$\lim_{n \rightarrow \infty} \sigma_n(a \geq \alpha_k) = 0 \quad (21)$$

for all  $k, 1 \leq k \leq K - 1$ . Thus,

$$\begin{aligned} \lim_{n \rightarrow \infty} U_n(a) &= \lim_{n \rightarrow \infty} \sum_{k=1}^K u_n(\alpha_k) [\sigma_n(a \geq \alpha_k) - \sigma_n(a \geq \alpha_{k-1})] \\ &= \lim_{n \rightarrow \infty} u_n(\alpha_K) \sigma_n(a \geq \alpha_K) \\ &= \lim_{n \rightarrow \infty} u_n(\alpha_K) \end{aligned} \quad (22)$$

**Theorem 1.** *As the supremum of uncertainty is approached, an individual's nonadditive expected utility  $U_i(a)$  approaches  $u_i(\min a)$ .*

Intuitively, uncertainty is the largest if the individual cannot assign any probability to any event except the  $\Omega$  event. In that case, the utility depends only on the  $\Omega$  event, as all the other events become irrelevant. This idea was originally discussed by Dow and Werlang (1992).

### 3 Consequence of Extreme Uncertainty Aversion

In the previous section, we discussed the consequence of extreme uncertainty. In this section, we discuss the consequence of extreme uncertainty aversion. To do

so, we first characterize uncertainty aversion and extreme uncertainty aversion.

Schmeidler (1989) established the relationship between uncertainty aversion and the convexity of a probability function. The analysis of Klibanoff, Marinacci, and Mukerji (2005), which is presented for maximin utility, suggests the following: Firstly, the convexity of a probability function (which is a set function) can be captured by the convexity of a function defined on the line.<sup>13</sup> Secondly, the uncertainty aversions of two individuals can be compared using a measure comparable to the absolute risk aversion coefficient of Pratt (1964) and Arrow (1965). We propose a measure of uncertainty aversion based on a function defined on the line. To justify this measure, we show that the convexity of this function is equivalent to the convexity of a probability function. We also show that our measure can be used to indicate “who is more uncertainty averse,” as in Klibanoff, Marinacci, and Mukerji (2005).

Recall that risk aversion is characterized as having lower utility when mean preserving spread is applied. We modify mean preserving spread so that we can characterize uncertainty aversion based on the modified mean preserving spread. Our modification of mean preserving spread is based on what we term additive extension expected utility. We introduce these concepts in turn.

Let  $a$  be an act that takes a value from  $\{\alpha_1, \dots, \alpha_K\}$ , where  $\alpha_1 > \dots > \alpha_K$ . Let  $\alpha_0$  be a number greater than  $\alpha_1$ . Then nonadditive utility in (20) can be rewritten as:

$$\begin{aligned}
 U_i(a) = & \sum_{k=1}^K u_i(\alpha_k) \bar{\sigma}_i(a = \alpha_k) \\
 & + \sum_{k=1}^K u_i(\alpha_k) [\tilde{\sigma}_i(a \geq \alpha_k) - \tilde{\sigma}_i(a \geq \alpha_{k-1})]
 \end{aligned} \tag{23}$$

---

<sup>13</sup>Klibanoff, Marinacci, and Mukerji introduced  $\phi$ , which is a convex transformation of an additive expectation. It is defined on the line.

Another useful representation of utility is:

$$\begin{aligned}
U_i(a) = & \bar{\sigma}_i(\Omega) \sum_{k=1}^K u_i(\alpha_k) \bar{\sigma}_i^*(a = \alpha_k) \\
& + \sum_{k=1}^K u_i(\alpha_k) \bar{\sigma}_i^*(a = \alpha_k) \frac{[\tilde{\sigma}_i(a \geq \alpha_k) - \tilde{\sigma}_i(a \geq \alpha_{k-1})]}{\bar{\sigma}_i^*(a = \alpha_k)}
\end{aligned} \tag{24}$$

We call  $\sum_{k=1}^K u_i(\alpha_k) \bar{\sigma}_i^*(a = \alpha_k)$  the *additive extension expected utility*, and denote it by  $E^*[u_i(a)]$ . refers to .

Let  $b$  be an act whose range is  $\{\beta_1, \dots, \beta_K\}$ . We say that act  $a$  is obtained by applying a *mean preserving spread* to act  $b$  if the following conditions are satisfied: (i) Act  $a$  and act  $b$  are co-monotonic in the sense of Schmeidler(1986). That is,  $b = \beta_k$  if and only if  $a = \alpha_k$  for all  $k, k = 1, \dots, K$ . (ii) Two acts produces identical additive extension expected utilities. That is,  $\sum_{k=1}^K u(\beta_k) \bar{\sigma}^*(b = \beta_k) = \sum_{k=1}^K u(\alpha_k) \bar{\sigma}^*(a = \alpha_k)$ . This is the sense in which “mean” is preserved. (iii) Utility is shifted from high index, low utility events to low index, high utility events. That is,  $\sum_{k'=1}^k u(\beta_{k'}) \bar{\sigma}^*(b = \beta_{k'}) \leq \sum_{k'=1}^k u(\alpha_{k'}) \bar{\sigma}^*(a = \alpha_{k'})$  for all  $k, k = 1, \dots, K$ . This is the sense in which we talk about “spread.”

We now define uncertainty aversion based on our modification of mean-preserving spread. Individual  $i$  is said to be *uncertainty averse* if a mean preserving spread decreases the utility of  $i$ . That is,  $i$  is uncertainty averse if  $U(a) \leq U(b)$  whenever act  $a$  is obtained by applying a mean-preserving spread to act  $b$ .

Lemma 4 shows that our definition of uncertainty aversion is consistent with the idea that uncertainty aversion is equivalent to the convexity of probability. Schmeidler (1989) defined uncertainty aversion in a slightly different way.<sup>14</sup> However, his definition of uncertainty aversion is not consistent with his charac-

<sup>14</sup>For any three acts  $a, b, c$  and any  $\alpha$  in  $[0, 1]$ , if  $U(a) \geq U(c)$  and  $U(b) \geq U(c)$ , then  $U(\alpha a + (1 - \alpha)b) \geq U(c)$

terization of uncertainty aversion being equivalent to the convexity of probability.<sup>15</sup> So it is necessary to either modify the definition of uncertainty aversion, or to give up the equivalence property. We choose the former, and propose a definition of uncertainty aversion that is consistent with the equivalence property.

Nonadditive probability  $\sigma_i$  is said to be *convex* if  $\sigma_i(A) + \sigma_i(B) \leq \sigma_i(A + B) + \sigma_i(A \cap B)$  where  $A$  and  $B$  are two events.  $\sigma_i$  is said to be *strictly convex* if  $\sigma_i(A) + \sigma_i(B) < \sigma_i(A + B) + \sigma_i(A \cap B)$  where  $A$  and  $B$  are two different events.

The convexity of  $\sigma_i$  can be expressed in terms of the convexity of a function defined the line. The analysis later becomes significantly simplified if we work with a function defined on the line. Also, it is easier to provide geometrical interpretation when we work with such function. Let  $\mathcal{A}_1, \dots, \mathcal{A}_L$  be a partition of event space  $\mathcal{A} = 2^\Omega$  such that  $\sigma_i$  is constant on each  $\mathcal{A}_l$ . Such partition exists by Assumption 2. Choose a sequence of events  $A_1, \dots, A_L$  by selecting  $A_l$  from  $\mathcal{A}_l$ . Then we define the *additive-to-nonadditive mapping*  $g_i$  as follows:

$$g_i(0) = 0 \tag{25}$$

and

$$g_i(\bar{\sigma}_i^*(A_l)) = \sigma_i(A_l) \tag{26}$$

for  $l = 1, 2, \dots, L$ .

**Lemma 4.** *The following three statements are equivalent: (i) Individual  $i$  is uncertainty averse, (ii) Nonadditive probability  $\sigma_i$  is convex, (iii)  $g_i$  is convex.*

*Proof.* See the appendix. □

---

<sup>15</sup>The last two lines of the proof of the proposition in Schmeidler (1989) are incorrect. Thus, the proposition does not establish the equivalence between the convexity of probability and uncertainty aversion.

We have defined uncertainty aversion as having lower utility after applying a mean preserving spread. Thus, it is natural to think of “being more uncertainty averse” as having bigger drop in utility after applying a mean preserving spread. In comparing uncertainty aversion, we restrict our attention to a pair of individuals who have identical additive extension expected utility.<sup>16</sup> We formalize the idea in the following way: Let  $a$  be an act that is obtained by applying a mean-preserving spread to act  $b$ . Then there exists constant act  $c$  such that  $U(a + c) = U(b)$ . We call such  $c$  *compensation constant*, and write it as  $c(a, b)$ . Suppose that individuals  $i$  and  $j$  have identical additive extension expected utility. Then individual  $j$  is said to be *more uncertainty averse* than individual  $i$  if  $c_j(a, b) \geq c_i(a, b)$  for any act  $a$  and act  $b$  if act  $a$  is obtained by applying a mean preserving spread to act  $b$ .

Klibanoff, Marinacci, and Mukerji (2005) defined “being more uncertainty averse” in a somewhat different way<sup>17</sup>. If individual  $j$  is more uncertainty averse than individual  $i$  in the sense of our definition, then  $j$  is also more uncertainty averse than  $i$  in the sense of Klibanoff, Marinacci, and Mukerji (2005). However, the opposite is not the case.<sup>18</sup> While related to our definition, the definition of Klibanoff, Marinacci, and Mukerji (2005) is not consistent with the equivalence

---

<sup>16</sup>Property 6 in the appendix shows that  $\tilde{\sigma}_i(\Omega)$  provides an upper bound of uncertainty, and we interpret it as *total uncertainty*. Property 7 shows that when two individuals have identical additive extension expected utilities, they also have identical total uncertainties. Thus, our proposal ensures that we compare uncertainty aversion of two individuals only when they face identical total uncertainty.

<sup>17</sup>Individual  $j$  is more uncertainty averse than individual  $i$  if individual  $i$  prefers uncertain act  $a$  to purely risky act  $a_0$  whenever individual  $j$  prefers  $a$  to  $a_0$ . An act is purely risky if the probability is additive over events defined by the act.

<sup>18</sup>When two individuals have identical additive extension expected utilities,  $j$  is more uncertainty averse than  $i$  in the sense of Klibanoff, Marinacci, and Mukerji (2005) if and only if  $\sigma_j \leq \sigma_i$ . It is easy to show that  $\sigma_j \leq \sigma_i$  if  $j$  is more uncertainty averse than  $i$  in the sense of our definition.

between uncertainty aversion and the convexity of nonadditive probability.

Lemma 5 shows that being more uncertainty averse corresponds to  $\sigma_i$  and  $g_i$  being more convex.

**Lemma 5.** *For two uncertainty averse individuals,  $i$  and  $j$ , who have identical additive extension expected utilities, the following statements are identical: (i)  $j$  is more uncertainty averse than  $i$ , (ii)  $\sigma_j$  is more convex than  $\sigma_i$ , i.e.,*

$$\begin{aligned} & \frac{\sigma_i(a \geq \alpha_{k+1}) - \sigma_i(a \geq \alpha_k)}{\bar{\sigma}_i^*(a = \alpha_{k+1})} - \frac{\sigma_i(a \geq \alpha_k) - \sigma_i(a \geq \alpha_{k-1})}{\bar{\sigma}_i^*(a = \alpha_k)} \\ \leq & \frac{\sigma_j(a \geq \alpha_{k+1}) - \sigma_j(a \geq \alpha_k)}{\bar{\sigma}_j^*(a = \alpha_{k+1})} - \frac{\sigma_j(a \geq \alpha_k) - \sigma_j(a \geq \alpha_{k-1})}{\bar{\sigma}_j^*(a = \alpha_k)} \end{aligned} \quad (27)$$

for any act  $a$  with the range  $\alpha_1 > \dots > \alpha_K$  and any  $k, k = 1, \dots, K - 1$ . (iii)  $g_j$  is more convex than  $g_i$ , i.e.,

$$\begin{aligned} & \frac{g_i(x_{l+1}) - g_i(x_l)}{x_{l+1} - x_l} - \frac{g_i(x_l) - g_i(x_{l-1})}{x_l - x_{l-1}} \\ \leq & \frac{g_j(x_{l+1}) - g_j(x_l)}{x_{l+1} - x_l} - \frac{g_j(x_l) - g_j(x_{l-1})}{x_l - x_{l-1}} \end{aligned} \quad (28)$$

for any  $l, l = 1, \dots, L - 1$ .

*Proof.* See the appendix. □

Lemma 5 corresponds to Theorem 1 of Pratt (1964), which discusses the implication of one individual being more risk averse than another individual. Lemma 5, together with Lemma 4, motivates us to propose the following measure of uncertainty aversion. Let  $\mathcal{A}_1, \dots, \mathcal{A}_L$  be a partition of  $\mathcal{A}$  such that  $\bar{\sigma}_i^*(A_l) = x_l$  for  $A_l \in \mathcal{A}_l$  and  $x_1 < \dots < x_L$ . The *uncertainty aversion* of individual  $i$  is defined as

$$\lambda_i = \sum_{l=2}^L \frac{g_i''(x_l)}{g_i'(x_{l-1})} \quad (29)$$

where  $g_i'(x_l) = \frac{g_i(x_l) - g_i(x_{l-1})}{x_l - x_{l-1}}$ ,  $g_i''(x_l) = \frac{g_i'(x_l) - g_i'(x_{l-1})}{x_l - x_{l-1}}$  and  $g_i(x_0) = g_i(0) = 0$ .

Note the parallelism to risk aversion. Risk aversion is identified with the concavity of utility  $u$ , and the scaled second derivative of  $u$  is used as a measure of absolute risk aversion. Uncertainty aversion is identified with function  $g_i$ . Thus, we propose a measure of uncertainty aversion based on the second derivative of  $g_i$ .

We interpret the maximum of  $\lambda_i$  given  $\tilde{\sigma}_i(\Omega)$ , shown in Lemma 6, as *extreme uncertainty aversion*.

**Lemma 6.** *Given  $\tilde{\sigma}_i(\Omega)$ , the maximum of uncertainty aversion  $\lambda_i$  is  $L^2 \frac{1-\tilde{\sigma}_i(\Omega)}{\tilde{\sigma}_i(\Omega)}$ . This maximum is achieved if and only if*

$$\sigma_i(A) = \begin{cases} \tilde{\sigma}_i(A) & \text{if } A \neq \Omega \\ 1 & \text{if } A = \Omega \end{cases} \quad (30)$$

What happens when uncertainty aversion is extreme? Uncertainty aversion is extreme when  $g_i$  is almost linear except near  $\Omega$ . In this case, nonadditive expected utility will depend only on the additive component of probability and the  $\Omega$  event. Thus,

$$U_i(a) = \tilde{\sigma}_i(\Omega) E^* [u_i(a)] + u_i(\min a) \quad (31)$$

Recall that we call  $E^* [u_i(a)]$  is additive extension expected utility.

**Theorem 2.** *If an individual has extreme uncertainty aversion, the individual's nonadditive expected utility depends only on additive extension expected utilities and the minimum value of the act.*

## 4 Cash Flow Uncertainty And Maximum Draw-down

In this section, we show that uncertainty over cash flow can explain the role of maximum drawdown in the investment fund selection problem. When cash flow timing is uncertain, an uncertainty-averse investor would care more about the possibility that he buys a fund when the fund price is high and that he sells a fund when the fund price is low. If uncertainty is extreme or uncertainty aversion is extreme, the investor's choice will be driven by maximum drawdown.

Consider an investor who wants to select an investment fund among many alternatives. We make the following assumptions about the investor: The investor plans to buy the fund when there is cash inflow and plans to sell the fund when there is cash outflow. He does not know the timing of cash inflow and outflow in advance. He only knows that there will be cash inflow of one unit sometime during the next  $H$  periods, and that there will be cash outflow of one unit also sometime during the next  $H$  periods. Cash outflow can occur only after cash inflow. That is, if  $T$  is the current period and  $T + h$  and  $T + h'$  are the periods of cash inflow and cash outflow, then  $1 \leq h \leq h' \leq H$ .

Some aspects of our characterization of the investor requires clarification. The investor selects only one fund out of many. We are thinking of an investor who does not want to own securities directly, but wants to have exposure to a large number of securities by investing in a diversified fund. In many countries, only a small number of individuals own securities directly. Many individuals prefer to invest through mutual funds. When an investor decides to invest through mutual funds, the investor may not want to buy multiple funds because of the high cost of buying many funds. Instead, the investor may decide to select single fund. Even institutional investors such as corporate treasuries quite often

choose to invest in single fund to lower the transaction costs.

That the investor buys the fund at the time of cash inflow and sells the fund at the time of cash outflow is not unrealistic. Many professional money managers admit that they do not have “timing skills”, while they may boast of their “selection skills”. These managers would not be engaged in timing, i.e. they will buy the fund when they have cash to invest and they will sell the fund when they need to generate cash. The pattern may be even more true for individual investors.

That the investor does not know the timing of cash inflow and outflow is not strange either. A small business owner may not know when there will be cash inflow so that he has extra cash to invest. A retiree may face significant cash outflow anytime, perhaps for medical expense. Institutional investors quite often have significant uncertainty about cash flow as well. A company treasurer does not know the timing of cash flow precisely either.

The investor faces two types of uncertainty: uncertainty over cash flow and uncertainty over returns. We assume that uncertainty over cash flow and uncertainty over returns are independent from each other. That is, the realization of cash flow does not influence the realization of returns, and the realization of returns does not influence the realization of cash flow either. We make a more formal statement below.

Let  $(X, \mathcal{A}_X, \mu_i)$  and  $(Y, \mathcal{A}_Y, \nu_i)$  be two nonadditive probability spaces that satisfy Assumptions 1 and 2.  $X$  and  $Y$  are sets of states,  $\mathcal{A}_X$  and  $\mathcal{A}_Y$  are sets of events, and  $\mu_i$  and  $\nu_i$  are convex nonadditive probabilities. Consider the product space  $(\Omega, \mathcal{A}, \sigma_i)$  where  $\Omega = X \times Y$ . We assume that  $\sigma_i$  is the *Moebius-independent* product of  $\mu_i$  and  $\nu_i$  in the sense that, for every  $A \in \mathcal{A}$ ,

$$\varphi_{\sigma_i}(A) = \begin{cases} \varphi_{\mu_i}(S)\varphi_{\nu_i}(T) & \text{if } A = S \times T \text{ for } S \in \mathcal{A}_X \text{ and } T \in \mathcal{A}_Y \\ 0 & \text{otherwise.} \end{cases} \quad (32)$$

This definition of independence is from Ghirardato (1997). Recall that specific weight  $\varphi_{\sigma_i}$  is the probability that is assigned to  $A$  and not divisible among its subsets. Then Moebius independence means that non-rectangles have no specific weights. It corresponds to the situation where there is no uncertainty about the independence between  $\mu_i$  and  $\nu_i$ .

Let the  $H$ -period return of fund  $j$  be  $r_j = (r_{j,T+1}, \dots, r_{j,T+H})$ . If the investor buys fund  $j$  at  $T+h$  and sells at  $T+h'$ , then the investment return is determined as  $R(r_j, h, h') = r_{j,T+h} + \dots + r_{j,T+h'}$ . We consider a fund return,  $x \equiv r_j$ , as an act in  $(X, \mathcal{A}_X, \mu_i)$ , and cash flow timing,  $y \equiv (h, h')$ , as an act in  $(Y, \mathcal{A}_Y, \nu_i)$ . Then an investment return,  $a \equiv R(r_j, h, h')$ , is an act in the product space  $(\Omega, \mathcal{A}, \sigma_i)$ . Moebius independence indicates that there is no uncertainty regarding the fact that the timing of cash inflow and outflow does not influence the returns of funds.

It is likely that there is more substantial uncertainty over cash flow than over returns. To capture this aspect in a simple analysis, we assume that returns are purely risky. That is, an additive probability adequately describes the distribution of returns. Given that historical returns are available, coming up with a predictive distribution for future returns is not a difficult task. The only question is whether the investor is willing to allocate all of 100% probability to the predictive distribution so that probability remains additive. Even if the investor is not willing to allocate all of 100% probability to the predictive distribution, the uncertainty regarding returns is most likely smaller than uncertainty regarding cash flow. It is more difficult to have a predictive distribution for cash flow. For future returns, investors have a histogram to start from. For cash flow, investors do not have an intuitive approach to generate a histogram. One might consider a uniform distribution as a starting point, i.e. that the probability of cash inflow and outflow is constant across each point in time. But then, this

appears to be exactly the situation where uncertainty concept is applicable.

The following lemma from Ghirardato (1997) shows that, if there is risk but no uncertainty in returns, then nonadditive expected utility has a special form.<sup>19</sup>

**Lemma 7.** *If  $\mu_i$  is additive, then the individual's utility is given as follows:*

$$U_i(a) = \sum_{x \in X} \mu_i(x) \int_Y u_i(x, y) d\nu_i(y) \quad (33)$$

Now we can apply Theorems 1 and 2 to show that the investment fund selection problem is solved with the aid of maximum drawdown. Suppose that cash flow uncertainty is extreme. This is the case when the probability of having the worst cash flow timing is 100%. In that case, the investor's utility will be based on the maximum drawdown of each realization of returns. Theorem 1 implies

$$U_i(a) = \sum_{x \in X} \mu_i(x) u_i(\min a) = E[u_i(\min a)] \quad (34)$$

In the above formula,  $\min a$  indicates the maximum drawdown for a given realization of the return, i.e.  $\min_{h, h'} R(r_j, h, h')$  where  $r_j$  is given. We call  $E[u(\min a)]$  the *expected utility of maximum drawdown*. So we can say that the investor's utility depends only on the expected utility of maximum drawdown.

Suppose that the investor has extreme uncertainty aversion. Now the probability of having the worst cash flow timing can be less than 100%. Uncertainty aversion is extreme when the probability is additive everywhere, except near the worst possibility. Theorem 2 implies

$$\begin{aligned} U_i(a) &= \sum_{x \in X} \mu_i(x) \{ \bar{\nu}_i(Y) E^*[u_i(a)] + u_i(\min a) \} \\ &= \bar{\nu}_i(Y) E\{E^*[u_i(a)]\} + E[u_i(\min a)] \end{aligned} \quad (35)$$

---

<sup>19</sup>See Lemma 5 of Ghirardato (1997).

Recall that we call  $E^*[u_i(a)]$  is additive extension expected utility. So we can say that the investor's utility depends on the additive extension expected utility and the expected utility of maximum drawdown. We summarize this idea in a theorem.

**Theorem 3.** *Suppose that there is cash flow uncertainty, while return is purely risky. Then the following statements are true: (1) If cash flow uncertainty is extreme, the investor's utility depends only on the expected utility of maximum drawdown. (2) If the investor has extreme uncertainty aversion, the investor's nonadditive expected utility depends only on the additive extension expected utility and the expected utility of maximum drawdown.*

## 5 Concluding Remarks

We have shown that a rational investor's investment fund selection problem may involve maximum drawdown in certain situations. Maximum drawdown becomes relevant if uncertainty over cash flow is extreme or the investor's uncertainty aversion is extreme. When uncertainty over cash flow is extreme, the investor discounts all the possibilities to zero probability, except the possibility of buying the fund at the lowest price and selling the fund at the highest price. Thus, the fund selection is based on maximum drawdown. When the investor's uncertainty aversion is extreme, the investor's utility is separated into two parts: additive utility of all the possibilities and the expected utility of the worst possibility.

How likely is the situation of extreme uncertainty and extreme uncertainty aversion? An act is unlikely to have extreme uncertainty, which requires an individual to have absolutely no opinion regarding the likely realizations of the act, except its minimum value. On the other hand, an individual is quite likely to be extremely uncertainty averse. What we require is that the individual's

probability is linear everywhere (except near the worst outcome), but the probability does not sum to one. We are thinking of two situations. Firstly, an individual may construct a predictive distribution out of the history, but, to account for uncertainty, he may just decide to lower the total probability that he assigns to the predictive distribution. Secondly, if no history is available, the individual may just use a uniform distribution as the predictive distribution, but to account for uncertainty, he decides to lower the total probability that he assigns to the uniform distribution. In both situations, the individual's probability is linear almost everywhere, and the probability does not sum to one. We believe that these two situations arise quite often in reality, and therefore, extreme uncertainty aversion is a realistic model of individuals' behavior.

The discussion in this paper suggests a natural extension of the mean-variance analysis of Markowitz (1952). Given the popularity of the mean-variance analysis among practitioners, it is worth discussing the extension. Defining  $\pi_i \equiv \mu_i \bar{\nu}_i^*$ , we may rewrite the individual's utility in (35) as:

$$\begin{aligned}
U_i(a) &= E\{E^*[u_i(a)]\} + E[u_i(\min a)] \\
&= \sum_{x \in X} \sum_{y \in Y} u_i[a(x, y)] \mu_i(x) \bar{\nu}_i^*(y) + E[u_i(\min a)] \\
&= \sum_{x \in X} \sum_{y \in Y} u_i[a(x, y)] \pi_i(x, y) + E[u_i(\min a)] \\
&= \bar{\nu}_i(Y) E^{**}[u_i(a)] + E[u_i(\min a)]
\end{aligned} \tag{36}$$

We used double asterisk \*\* to indicate that the expectation is calculated using additive measure  $\pi_i$ . The idea of Markowitz (1952) and Levy and Markowitz (1979) is that the expected utility can, quite often, be approximated by a function of mean and variance. Applying the idea, let us approximate the expected utility  $E^{**}[u_i(a)]$  by a function  $F_i$  of mean  $E^{**}(a)$  and variance  $V^{**}(a)$ , and also the expected utility  $E[u_i(\min a)]$  by a function  $G_i$  of mean  $E(\min a)$  and

variance  $V(\min a)$ . Thus, we can express the utility approximately as

$$U_i(a) = F_i [E^{**}(a), V^{**}(a)] + G_i [E(\min a), V(\min a)] \quad (37)$$

That is, the individual's utility is expressed as the sum of two functions, where the first function is a function of mean and variance of investment returns and the second function is a function of mean and variance of maximum drawdown.

## Appendix

Proof of Lemma 2 uses the following property:

**Property 1.** Let  $x_1, \dots, x_K$  be a sequence of positive numbers such that

$$\sum_{k=1}^K x_k = 1$$

Then

$$\sum_{k=1}^K x_k^2 \leq \frac{1}{K}$$

and the equality is obtained if and only if  $x_k = \frac{1}{K}$  for all  $k, k = 1, \dots, K$ .

*Proof.* Suppose that  $x_i < x_j$ . Let  $x'_i$  and  $x'_j$  be  $\frac{x_i + x_j}{2}$ . Then

$$x'^2_i + x'^2_j - x^2_i - x^2_j = -\frac{(x_i - x_j)^2}{2} < 0$$

Thus  $\sum_{k=1}^K x_k^2$  can be reduced whenever two  $x$ 's are different. When all  $x$ 's are identical,  $x_k = \frac{1}{K}$  for all  $k, k = 1, \dots, K$ , and  $\sum_{k=1}^K x_k^2 = \frac{1}{K}$ .  $\square$

### Lemma 2

*Proof.* (i) We first show that  $\sup \gamma_+(\sigma_i, a) \leq \frac{1}{2}(1 - \frac{1}{L})$  for any  $\sigma_i$  and any  $a$ .

Then we show that there exist a sequence  $\sigma_1, \sigma_2, \dots$  and  $a$  such that

$$\lim_{n \rightarrow \infty} \gamma_+(\sigma_n, a) = \frac{1}{2}(1 - \frac{1}{L}) \quad (38)$$

From the definition of downside uncertainty,

$$\begin{aligned}
\gamma_+(\sigma_i, a) &= \sum_{k=1}^{K-1} [\bar{\sigma}_i^*(a \geq \alpha_k) - \sigma_i(a \geq \alpha_k)] \bar{\sigma}_i^*(a = \alpha_{k+1}) \\
&\leq \sum_{k=1}^{K-1} \bar{\sigma}_i^*(a \geq \alpha_k) \bar{\sigma}_i^*(a = \alpha_{k+1}) \\
&= \frac{1}{2} \sum_{k=1}^{K-1} \bar{\sigma}_i^*(a \geq \alpha_k) \bar{\sigma}_i^*(a = \alpha_{k+1}) \\
&\quad + \frac{1}{2} \sum_{k=1}^{K-1} \bar{\sigma}_i^*(a \geq \alpha_k) \bar{\sigma}_i^*(a = \alpha_{k+1}) \\
&= \frac{1}{2} \sum_{k=2}^K \bar{\sigma}_i^*(a \geq \alpha_{k-1}) \bar{\sigma}_i^*(a = \alpha_k) \\
&\quad + \frac{1}{2} \sum_{k=1}^{K-1} \bar{\sigma}_i^*(a = \alpha_k) \bar{\sigma}_i^*(a \leq \alpha_{k+1}) \\
&= \frac{1}{2} \sum_{k=1}^K \bar{\sigma}_i^*(a = \alpha_k) - \frac{1}{2} \sum_{k=1}^K \bar{\sigma}_i^*(a = \alpha_k)^2
\end{aligned}$$

From Property 1,

$$\sum_{k=1}^K \bar{\sigma}_i^*(a = \alpha_k)^2 \geq \frac{1}{K}$$

Thus,

$$\frac{1}{2} \sum_{k=1}^K \bar{\sigma}_i^*(a = \alpha_k) - \frac{1}{2} \sum_{k=1}^K \bar{\sigma}_i^*(a = \alpha_1)^2 \leq \frac{1}{2} - \frac{1}{2} \frac{1}{K} \leq \frac{1}{2} - \frac{1}{2} \frac{1}{L} \quad (39)$$

The last inequality comes from the fact that the number of distinct events ( $K$ ) cannot be greater than the number of distinct states ( $L$ ).

We now prove that there exist a sequence  $\sigma_1, \sigma_2, \dots$  and  $a$  such that (38) is

satisfied. Define  $\sigma_n$  as

$$\begin{aligned}\bar{\sigma}_n(A) &= \frac{|A|}{n} \\ \tilde{\sigma}_n(A) &= \begin{cases} 0 & \text{if } A \neq \Omega \\ 1 - \frac{L}{n} & \text{if } A = \Omega \end{cases}\end{aligned}\quad (40)$$

where  $|A|$  refers to the number of elements in event  $A$ . Consider act  $a$  that takes a value from a decreasing sequence of real numbers  $\alpha_1, \dots, \alpha_L$  such that

$$a(\{\omega_l\}) = \alpha_l \quad (41)$$

where  $\{\omega_1, \dots, \omega_L\}$  is the set of states,  $\Omega$ . Then

$$\lim_{n \rightarrow \infty} \gamma_+(\sigma_n, a) = \lim_{n \rightarrow \infty} \sum_{k=1}^{L-1} \left( \frac{k}{L} - \frac{k}{n} \right) \frac{1}{L} = \frac{1}{2} \left( 1 - \frac{1}{L} \right)$$

(ii) Note that

$$\gamma(\sigma_i, a) = \gamma_+(\sigma_i, a) + \gamma_+(\sigma_i, -a) \leq 1 - \frac{1}{L}$$

Thus, it is enough to prove that there exist a sequence  $\sigma_1, \sigma_2, \dots$  and  $a$  such that

$$\lim_{n \rightarrow \infty} \gamma(\sigma_n, a) = 1 - \frac{1}{L}$$

It is easy to verify that  $\sigma_1, \sigma_2, \dots$  and  $a$  defined in (40) and (41) satisfy this.  $\square$

**Lemma 3**

*Proof.* From (i) to (ii). Note that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \gamma_+(\sigma_n, a) \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^{K-1} \bar{\sigma}_n^*(a \geq \alpha_k) \bar{\sigma}_n^*(a = \alpha_{k+1}) - \lim_{n \rightarrow \infty} \sum_{k=1}^{K-1} \sigma_n(a \geq \alpha_k) \bar{\sigma}_n^*(a = \alpha_{k+1}) \end{aligned}$$

Since the second term is non-negative, (i) implies

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{k=1}^{K-1} \bar{\sigma}_n^*(a \geq \alpha_k) \bar{\sigma}_n^*(a = \alpha_{k+1}) &= \frac{1}{2} \left(1 - \frac{1}{L}\right) \\ \lim_{n \rightarrow \infty} \sum_{k=1}^{K-1} \sigma_n(a \geq \alpha_k) \bar{\sigma}_n^*(a = \alpha_{k+1}) &= 0 \end{aligned} \tag{42}$$

The first of these two equations implies

$$\bar{\sigma}_n^*(a = \alpha_k) = \frac{1}{L} \tag{43}$$

for all  $k$ ,  $1 \leq k \leq K$ . See (39). Given (43), the second of (42) implies

$$\lim_{n \rightarrow \infty} \sigma_n(a \geq \alpha_k) = 0$$

for all  $k$ ,  $1 \leq k \leq K - 1$ .

The proof from (ii) to (i) is the reverse of the proof from (i) to (ii). The proof from (iii) to (iv) is similar to the proof from (i) to (ii). The proof from (iv) to (iii) is similar to the proof from (ii) to (i).

Note that

$$\bar{\sigma}_n^*(a \leq \alpha_k) = \frac{K - k + 1}{L} = \bar{\sigma}_n^*(a \geq \alpha_{K-k+1})$$

Thus,

$$\sigma_n(a \leq \alpha_k) = \sigma_n(a \geq \alpha_{K-k+1})$$

which implies

$$\sum_{k=2}^K \sigma_n(a \leq \alpha_k) = \sum_{k=2}^K \sigma_n(a \geq \alpha_k)$$

This equality ensures the equivalence of (ii) and (iv).

From (i) to (v). (i) is equivalent to (iii). (i) and (iii) together imply (v).

From (v) to (i). Downside uncertainty is non-negative. Thus, (v) implies (i) and (iii).  $\square$

The proof of Lemma 4 uses the following properties.

**Property 2.** Let  $x_1, x_2, \dots, x_K$  and  $y_1, y_2, \dots, y_K$  be two sequences of numbers with the following properties: (i)  $\sum_{k=1}^K x_k = \sum_{k=1}^K y_k$ . (ii)  $\sum_{k'=1}^k x_{k'} \geq \sum_{k'=1}^k y_{k'}$  for all  $k, k = 1, \dots, K$ . Let  $w_1, w_2, \dots, w_K$  be an increasing sequence of numbers. Then  $\sum_{k=1}^K w_k x_k \leq \sum_{k=1}^K w_k y_k$ .

*Proof.* Let  $I$  be the set of index  $k$  for which  $x_k < y_k$ . Denote the elements of  $I$  as  $i_1, \dots, i_M$  where  $i_1 < \dots < i_M$ . Define  $I_m = \{i_1, \dots, i_m\}$  for any  $m, 1 \leq m \leq M$ . Let  $S$  be a constant defined as  $S = \sum_{k \in I} y_k - x_k$ . Let  $f$  be a function from interval  $[0, S]$  to  $I$  such that  $s \in [0, S]$  is mapped to  $i_m \in I$  such that

$$\sum_{k \in I_{m-1}} y_k - x_k < s \leq \sum_{k \in I_m} y_k - x_k$$

Similarly, let  $J = \{j_1, \dots, j_N\}$  be the set of index  $k$  for which  $x_k > y_k$  and  $j_1 < \dots < j_N$ . Define  $J_n = \{j_1, \dots, j_n\}$  for any  $n, 1 \leq n \leq N$ . From (i),  $\sum_{k \in J} x_k - y_k = S$ . Define function  $g : [0, S] \rightarrow J$  such that  $s \in [0, S]$  is mapped to  $j_n \in J$  such that

$$\sum_{k \in J_{n-1}} x_k - y_k < s \leq \sum_{k \in J_n} x_k - y_k$$

We first show that  $f(s) > g(s)$ . Suppose not. That is,  $f(s) < g(s)$ . Consider

$$\sum_{k < g(s)} y_k - x_k = \sum_{\substack{k < g(s) \\ k \in I}} y_k - x_k + \sum_{\substack{k < g(s) \\ k \in J}} y_k - x_k$$

The first term in the right hand side is greater than or equal to  $s$  since  $f(s) < g(s)$ . The second term is greater than  $-s$  as we excluded  $g(s)$ . Thus,  $\sum_{k < g(s)} y_k - x_k > 0$ , which contradicts (ii). Thus,  $f(s) > g(s)$  for all  $s$ .

Note that

$$\sum_{k \in I} (y_k - x_k) w_k = \int_0^S w_{f(s)} ds$$

and

$$\sum_{k \in J} (x_k - y_k) w_k = \int_0^S w_{g(s)} ds$$

As  $w$  is increasing,  $f(s) > g(s)$  implies  $w_{f(s)} > w_{g(s)}$ , and

$$\sum_{k=1}^K (y_k - x_k) w_k = \int_0^S w_{f(s)} - w_{g(s)} ds \geq 0$$

□

**Property 3.** Let  $P_K$  be a set of pairs of  $K$ -element sequences with the following properties: For any  $(x, y) \in P_K$ , (i)  $\sum_{k=1}^K x_k = \sum_{k=1}^K y_k$ , and (ii)  $\sum_{k'=1}^k x_{k'} \geq \sum_{k'=1}^k y_{k'}$  for all  $k$ ,  $1 \leq k \leq K$ . Let  $w_1, w_2, \dots, w_K$  be a sequence of numbers. If  $\sum_{k=1}^K w_k x_k \leq \sum_{k=1}^K w_k y_k$  for any  $(x, y) \in P$ , then  $w_1, w_2, \dots, w_K$  is an increasing sequence.

*Proof.* Suppose that  $w_k > w_{k+1}$ . Choose  $(x, y) \in P_K$  such that  $y_{k'} = x_{k'}$  for  $k' \neq k, k+1$ ,  $y_k = x_k - 1$ , and  $y_{k+1} = x_{k+1} + 1$ . Then

$$\sum_{k=1}^K w_k (y_k - x_k) = -w_k + w_{k+1} < 0$$

This is a contradiction.  $\square$

**Property 4.** (Convexity of  $\sigma_i$ ) Let  $\sigma_i$  be a nonadditive probability and let  $\tilde{\sigma}_i$  be its nonadditive component. Then the following statements are equivalent: (i)  $\sigma_i$  is convex. (ii) Let  $a$  be an act, which takes a value from a decreasing sequence of numbers,  $\alpha_1, \dots, \alpha_K$ . Let  $\alpha_0$  be a number larger than  $\alpha_1$ . Then

$$\frac{\sigma_i(a \geq \alpha_k) - \sigma_i(a \geq \alpha_{k-1})}{\bar{\sigma}_i^*(a = \alpha_k)} \leq \frac{\sigma_i(a \geq \alpha_{k+1}) - \sigma_i(a \geq \alpha_k)}{\bar{\sigma}_i^*(a = \alpha_{k+1})} \quad (44)$$

for any  $k, k = 1, \dots, K-1$ . (iii)  $\tilde{\sigma}_i$  is convex. (iv) Let  $a$  be an act, which takes a value from a decreasing sequence of numbers,  $\alpha_1, \dots, \alpha_K$ . Let  $\alpha_0$  be a number larger than  $\alpha_1$ . Then

$$\frac{\tilde{\sigma}_i(a \geq \alpha_k) - \tilde{\sigma}_i(a \geq \alpha_{k-1})}{\bar{\sigma}_i^*(a = \alpha_k)} \leq \frac{\tilde{\sigma}_i(a \geq \alpha_{k+1}) - \tilde{\sigma}_i(a \geq \alpha_k)}{\bar{\sigma}_i^*(a = \alpha_{k+1})}$$

for any  $k, k = 1, \dots, K-1$ .

*Proof.* From (i) to (ii). Consider act  $b$ , discussed in the proof of Lemma 2, which takes a value from a decreasing sequence of real numbers  $\beta_1, \dots, \beta_L$  such that  $b(\{\omega\}) = \beta_l$ , when the set of states,  $\Omega$ , is  $\{\omega_1, \dots, \omega_L\}$ . For such  $b$ ,

$$\begin{aligned} & \frac{\sigma_i(b \geq \beta_{l+1}) - \sigma_i(b \geq \beta_l)}{\bar{\sigma}_i^*(b = \beta_{l+1})} - \frac{\sigma_i(b \geq \beta_l) - \sigma_i(b \geq \beta_{l-1})}{\bar{\sigma}_i^*(b = \beta_l)} \\ &= \frac{\sum_{A \subseteq \{\omega_1, \dots, \omega_l\}} \varphi_{\sigma_i}(\{\omega_{l+1}\} \cup A) - \sum_{B \subseteq \{\omega_1, \dots, \omega_{l-1}\}} \varphi_{\sigma_i}(\{\omega_l\} \cup B)}{\frac{1}{L}} \\ &= \frac{\varphi_{\sigma_i}(\{\omega_1, \dots, \omega_l, \omega_{l+1}\})}{\frac{1}{L}} \geq 0 \end{aligned}$$

Now consider any act  $a$  that takes a value from a decreasing sequence of real

numbers  $\alpha_1, \dots, \alpha_K$ . Then for any  $k$ , there exist  $l_1, l_2, l_3$  such that

$$\begin{aligned} & \frac{\sigma_i(a \geq \alpha_k) - \sigma_i(a \geq \alpha_{k-1})}{\bar{\sigma}_i^*(a = \alpha_k)} \\ &= \sum_{l=l_1}^{l_2} \frac{\bar{\sigma}_i^*(b = \beta_k)}{\bar{\sigma}_i^*(\beta_{l_1} \leq b \leq \beta_{l_2})} \frac{\sigma_i(b \geq \beta_l) - \sigma_i(b \geq \beta_{k-1})}{\bar{\sigma}_i^*(b = \beta_k)} \end{aligned}$$

and

$$\begin{aligned} & \frac{\sigma_i(a \geq \alpha_{k+1}) - \sigma_i(a \geq \alpha_k)}{\bar{\sigma}_i^*(a = \alpha_{k+1})} \\ &= \sum_{l=l_2}^{l_3} \frac{\bar{\sigma}_i^*(b = \beta_k)}{\bar{\sigma}_i^*(\beta_{l_2} \leq b \leq \beta_{l_3})} \frac{\sigma_i(b \geq \beta_l) - \sigma_i(b \geq \beta_{k-1})}{\bar{\sigma}_i^*(b = \beta_k)} \end{aligned}$$

Note that we expressed the probability of act  $a$  as a weighted average of act  $b$ . From the property of the weighted average, we are guaranteed to have (44).

From (ii) to (i). Note that  $\sigma_i$  is convex if  $\sigma_i(A) + \sigma_i(B) \leq \sigma_i(A \cup B) - \sigma_i(A \cap B)$  for two events  $A$  and  $B$ . Without loss of generality, we may assume that  $A \cap B = \emptyset$ . (ii) implies

$$\frac{\sigma_i(A \cup B) - \sigma_i(A)}{\bar{\sigma}_i^*(B - A)} \geq \frac{\sigma_i(A) - \sigma_i(A \cap B)}{\bar{\sigma}_i^*(A - B)}$$

Rearranging the terms, we have

$$\bar{\sigma}_i^*(A - B)\sigma_i(A \cup B) \geq [\bar{\sigma}_i^*(A - B) + \bar{\sigma}_i^*(B - A)]\sigma_i(A) - \bar{\sigma}_i^*(B - A)\sigma_i(A \cap B)$$

By symmetry,

$$\bar{\sigma}_i^*(B - A)\sigma_i(A \cup B) \geq [\bar{\sigma}_i^*(B - A) + \bar{\sigma}_i^*(A - B)]\sigma_i(B) - \bar{\sigma}_i^*(A - B)\sigma_i(A \cap B)$$

Combining the last two inequalities, we prove that  $\sigma_i$  is convex.

The equivalence between (i) and (iii) can be shown easily. The proof of the equivalence between (iii) and (iv) is similar to the proof of the equivalence

between (i) and (ii). □

**Property 5.** (Convexity of  $g_i$ )  $g_i$  is convex if and only if

$$\frac{g_i(x_l) - g_i(x_{l-1})}{x_l - x_{l-1}} \leq \frac{g_i(x_{l+1}) - g_i(x_l)}{x_{l+1} - x_l} \quad (45)$$

for any  $l, l = 1, \dots, L-1$ . In the formula,  $x_l = \bar{\sigma}_i^*(A_l)$  for  $l \geq 1$ , and  $x_0 = 0$ .

*Proof.* The proof of “only if”: When  $g_i$  is convex,

$$g_i[\alpha x_j + (1 - \alpha)x_k] \leq \alpha g_i(x_j) + (1 - \alpha)g_i(x_k)$$

for any  $\alpha, 0 < \alpha < 1$ , such that  $\alpha x_j + (1 - \alpha)x_k$  lies in the domain of  $g_i$ . If  $x_j < x_k$ , then

$$\frac{g_i[\alpha x_j + (1 - \alpha)x_k] - g_i(x_j)}{(1 - \alpha)(x_k - x_j)} \leq \frac{g_i(x_k) - g_i[\alpha x_j + (1 - \alpha)x_k]}{\alpha(x_k - x_j)}$$

Choose  $x_j = x_{l-1}$ ,  $x_k = x_{l+1}$ , and  $\alpha = \frac{x_{l+1} - x_l}{x_{l+1} - x_{l-1}}$ . Then we obtain (45).

The proof of “if”: Without loss of generality, we may assume that  $x_j < x_k$ .

Choose  $l_1, l_2, l_3$  such that  $x_j = x_{l_1}$ ,  $\alpha x_j + (1 - \alpha)x_k = x_{l_2}$ , and  $x_k = x_{l_3}$ . Then

$$\frac{g_i[\alpha x_j + (1 - \alpha)x_k] - g_i(x_j)}{\alpha(x_k - x_j)} = \sum_{l=l_1+1}^{l_2} \frac{g_i(x_l) - g_i(x_{l-1})}{x_l - x_{l-1}}$$

and

$$\frac{g_i(x_k) - g_i[\alpha x_j + (1 - \alpha)x_k]}{\alpha(x_k - x_j)} = \sum_{l=l_2+1}^{l_3} \frac{g_i(x_l) - g_i(x_{l-1})}{x_l - x_{l-1}}$$

Note that the right hand side of the above two equations are weighted average.

From the property of the weighted average, these two equations together with (45) imply the convexity of  $g_i$ . □

**Lemma 4**

*Proof.* Recall that

$$U_i(a) = \bar{\sigma}_i(\Omega) \sum_{k=1}^K u_i(\alpha_k) \bar{\sigma}_i^*(a = \alpha_k) \\ + \sum_{k=1}^K u_i(\alpha_k) \bar{\sigma}_i^*(a = \alpha_k) \frac{[\tilde{\sigma}_i(a \geq \alpha_k) - \tilde{\sigma}_i(a \geq \alpha_{k-1})]}{\bar{\sigma}_i^*(a = \alpha_k)}$$

Let  $\frac{[\tilde{\sigma}_i(a \geq \alpha_k) - \tilde{\sigma}_i(a \geq \alpha_{k-1})]}{\bar{\sigma}_i^*(a = \alpha_k)}$  be  $w_k$ , and  $u_i(\alpha_k) \bar{\sigma}_i^*(a = \alpha_k)$  be  $x_k$ . Then Property 2 shows that (ii) implies (i). Property 3 shows that (i) implies (ii). The equivalence of (ii) to (iii) can be easily verified by examining Property 4 (ii) and Property 5.  $\square$

The proof of Lemma 5 uses the following concepts and properties:

*Total uncertainty* of nonadditive probability  $\sigma_i$  refers to the nonadditive component of  $\Omega$ , i.e.  $\tilde{\sigma}_i(\Omega)$ . We use the word “total” as  $\tilde{\sigma}_i(\Omega)$  provides an upper bound for uncertainty, as the following property shows.

**Property 6.**

$$\gamma_+(\sigma_i, a) \leq \frac{1}{2} \left(1 - \frac{1}{L}\right) \tilde{\sigma}_i(\Omega)$$

and

$$\gamma(\sigma_i, a) \leq \left(1 - \frac{1}{L}\right) \tilde{\sigma}_i(\Omega)$$

The equality is obtained if  $\tilde{\sigma}(A) = 0$  for  $A \neq \Omega$ , and act  $a$  takes a value from a decreasing sequence of real numbers  $\alpha_1, \dots, \alpha_L$  such that  $a(\{\omega_l\}) = \alpha_l$ , where  $\{\omega_1, \dots, \omega_L\}$  is the set of states,  $\Omega$ .

*Proof.* From the definition of downside uncertainty,

$$\begin{aligned}
\gamma_+(\sigma_i, a) &= \sum_{k=1}^{K-1} [\bar{\sigma}_i^*(a \geq \alpha_k) - \sigma_i(a \geq \alpha_k)] \bar{\sigma}_i^*(a = \alpha_{k+1}) \\
&\leq \sum_{k=1}^{K-1} [\bar{\sigma}_i^*(a \geq \alpha_k) - \bar{\sigma}_i(a \geq \alpha_k)] \bar{\sigma}_i^*(a = \alpha_{k+1}) \\
&= \sum_{k=1}^{K-1} \bar{\sigma}_i^*(a \geq \alpha_k) \bar{\sigma}_i^*(a = \alpha_{k+1}) [1 - \bar{\sigma}_i(\Omega)] \\
&= \sum_{k=1}^{K-1} \bar{\sigma}_i^*(a \geq \alpha_k) \bar{\sigma}_i^*(a = \alpha_{k+1}) \tilde{\sigma}_i(\Omega) \\
&\leq (1 - \frac{1}{L}) \tilde{\sigma}_i(\Omega)
\end{aligned}$$

The last inequality is derived in the proof of Lemma 2. The rest of the property can be verified easily.  $\square$

**Property 7.** *If two individuals  $i$  and  $j$  have identical additive extension expected utilities, then (i) their felicity functions,  $u_i$  and  $u_j$ , are identical, (ii) the additive components,  $\bar{\sigma}_i^*$  and  $\bar{\sigma}_j^*$ , are identical, and (iii) total uncertainties,  $\tilde{\sigma}_i(\Omega)$  and  $\tilde{\sigma}_j(\Omega)$ , are identical.*

*Proof.* To prove (i), consider constant acts with different payoffs. To prove (ii), consider acts that take the value of 1 at a particular element of  $\Omega$  and 0 elsewhere. (iii) is trivial.  $\square$

**Property 8.** *Let  $x_1, x_2, \dots, x_K$  and  $y_1, y_2, \dots, y_K$  be two sequences of numbers such that  $\sum_{k=1}^K x_k = \sum_{k=1}^K y_k$  and  $\sum_{k'=1}^k x_{k'} \geq \sum_{k'=1}^k y_{k'}$ . Let  $v_1, v_2, \dots, v_K$  and  $w_1, w_2, \dots, w_K$  be two increasing sequences of numbers such that  $v_{k+1} - v_k \leq w_{k+1} - w_k$  for all  $k$ ,  $1 \leq k \leq K - 1$ . Then  $\sum_{k=1}^K (y_k - x_k) v_k \leq \sum_{k=1}^K (y_k - x_k) w_k$ .*

*Proof.* This property is implied by Property 2.  $\square$

**Property 9.** Let  $P_K$  be a set of pairs of  $K$ -element sequences of numbers such that for any  $(x, y) \in P_K$ ,  $\sum_{k=1}^K x_k = \sum_{k=1}^K y_k$  and  $\sum_{k'=1}^k x_{k'} \geq \sum_{k'=1}^k y_{k'}$ . Let  $v_1, v_2, \dots, v_K$  and  $w_1, w_2, \dots, w_K$  be two sequences of numbers such that  $\sum_{k=1}^K (y_k - x_k)v_k \leq \sum_{k=1}^K (y_k - x_k)w_k$  for any  $(x, y) \in P_K$ . Then  $v_{k+1} - v_k \leq w_{k+1} - w_k$  for all  $k$ ,  $1 \leq k \leq K - 1$ .

*Proof.* This property is implied by Property 3.  $\square$

**Lemma 5**

*Proof.* Let act  $a$  be obtained by applying a mean preserving spread to act  $b$ . Then

$$\begin{aligned} & U_i(b) - U_i(a) \\ &= \bar{\sigma}_i(\Omega) \sum_{k=1}^K [u_i(\beta_k) - u_i(\alpha_k)] \bar{\sigma}_i^*(a = \alpha_k) \\ & \quad + \sum_{k=1}^K [u_i(\beta_k) - u_i(\alpha_k)] \bar{\sigma}_i^*(a = \alpha_k) \frac{[\tilde{\sigma}_i(a \geq \alpha_k) - \tilde{\sigma}_i(a \geq \alpha_{k-1})]}{\bar{\sigma}_i^*(a = \alpha_k)} \\ &= \sum_{k=1}^K [u_i(\beta_k) - u_i(\alpha_k)] \bar{\sigma}_i^*(a = \alpha_k) \frac{[\tilde{\sigma}_i(a \geq \alpha_k) - \tilde{\sigma}_i(a \geq \alpha_{k-1})]}{\bar{\sigma}_i^*(a = \alpha_k)} \end{aligned}$$

By Property 7,

$$\begin{aligned} & U_j(b) - U_j(a) \\ &= \sum_{k=1}^K [u_j(\beta_k) - u_j(\alpha_k)] \bar{\sigma}_j^*(a = \alpha_k) \frac{[\tilde{\sigma}_j(a \geq \alpha_k) - \tilde{\sigma}_j(a \geq \alpha_{k-1})]}{\bar{\sigma}_j^*(a = \alpha_k)} \end{aligned}$$

Note that  $c_j(a, b) \geq c_i(a, b)$  if and only if  $U_j(b) - U_j(a) \geq U_i(b) - U_i(a)$ . (This comes from Property 7.) Then by Property 8, (ii) implies (i), and by Property 9, (i) implies (ii). The equivalence between (ii) and (iii) is trivial.  $\square$

**Lemma 6**

*Proof.* Note that  $x_l - x_{l-1} = \frac{1}{L}$  for all  $l$ . Thus,

$$\begin{aligned}
\lambda_i &= \sum_{l=2}^L \frac{g_i''(x_l)}{g_i'(x_{l-1})} \leq \sum_{l=2}^L \frac{g_i''(x_l)}{g_i'(x_1)} \\
&= \frac{L [g_i'(x_L) - g_i'(x_1)]}{g_i'(x_1)} \\
&\leq \frac{L [g_i'(x_L) - \bar{\sigma}_i(\Omega)]}{\bar{\sigma}_i(\Omega)} \\
&\leq \frac{L \{L [1 - (1 - \frac{1}{L})\bar{\sigma}_i(\Omega)] - \bar{\sigma}_i(\Omega)\}}{\bar{\sigma}_i(\Omega)} \\
&= L^2 \frac{1 - \bar{\sigma}_i(\Omega)}{\bar{\sigma}_i(\Omega)}
\end{aligned}$$

It is easy to verify that the equality is obtained if and only if (30) is satisfied.  $\square$

## References

- Acar, Emmanuel, and Shane James. (1997). "Maximum Loss and Maximum Drawdown in Financial Markets." Unpublished manuscript.
- Alexander, Gordon J., and Alexandre M. Baptista. (2006). "Portfolio Selection with a Drawdown Constraint." *Journal of Banking and Finance* **30**, pp. 3171-3189.
- Arrow, Kenneth J. (1965). *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjo Hahnsson Foundation.
- Chamberlain, Gary. (2000). "Econometric Applications of Maximin Expected Utility." *Journal of Applied Econometrics* **15**, pp. 625-644.
- Chekhlov, Alexei, Stanislav Uryasev, and Michael Zabaranin. (2005). "Drawdown Measure in Portfolio Optimization." *International Journal of Theoretical and Applied Finance* **8**, pp. 13-58.
- Cvitanic, Jaksa, and Ioannis Karatzas. (1999). "On Dynamic Measures of Risk." *Finance and Stochastic* **3**, pp. 451-482.
- Dow, James, and Sergio Riberio de Costa Werlang. (1992). "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio." *Econometrica* **60**, pp. 197-204.
- Ellsberg, Daniel. (1961). "Risk, Ambiguity, and Savage Axioms." *Quarterly Journal of Economics* **75**, pp. 643-669.
- Ghirardato, Paolo. (1997). "On Independence for Non-Additive Measures, with a Fubini Theorem." *Journal of Economic Theory* **73**, pp. 261-291.
- Gilboa, Itzhak, and David Schmeidler. (1989). "Maximin Expected Utility with Non-unique Prior." *Journal of Mathematical Economics* **18**, pp. 141-153.
- Gilboa, Itzhak, and David Schmeidler. (1994). "Additive Representations of Non-additive Measures and the Choquet Integral." *Annals of Operations Research* **52**, pp. 43-65.

- Grossman, Sanford J., and Zhongquan Zhou. (1993). "Optimal Investment Strategies for Controlling Drawdowns." *Mathematical Finance* **3**, pp. 241-276.
- Harding, David, Georgia Nakou, and Ali Nejjar. (2003). "The Pros and Cons of Drawdown as a Statistical Measure of Risk for Investments." *AIMA Journal*, April, pp. 16-17.
- Hayes, Brian T. (2006). "Maximum Drawdowns of Hedge Funds with Serial Correlation." *Journal of Alternative Investments* **8**, pp. 26-38.
- Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. (2005). "A Smooth Model of Decision Making Under Uncertainty." *Econometrica* **73**, pp. 1849-1892.
- Leal, Ricardo Pereira Camara, and Beatriz Vaz de Melo Mendes. (2005). "Maximum Drawdown: Models and Applications." *Journal of Alternative Investments* **7**, pp. 83-91.
- Levy, Haim, and Harry M. Markowitz. (1979). "Approximating Expected Utility by a Function of Mean and Variance." *American Economic Review* **69**, pp. 308-317.
- Magdon-Ismail, Malik, Amir F. Atiya, Amrit Pratap, and Yaser S. Abu-Mostafa. (2004). "On the Maximum Drawdown of a Brownian Motion." *Journal of Applied Probability* **41**, pp. 147-161.
- Markowitz, Harry. (1952). "Portfolio Selection." *Journal of Finance* **7**, pp. 77-91.
- Von Neumann, John, and Oskar Morgenstern. (1947). *Theory of Games and Economic Behavior*, 2nd ed. Princeton: Princeton University Press.
- Pospisil, Libor, and Jan Vecer. (2008). "PDE Methods for the Maximum Drawdown." *The Journal of Computational Finance* **12**, pp. 59-76.
- Pratt, John W. (1964). "Risk Aversion in the Small and in the Large." *Econo-*

metrica **32**, pp. 122-136.

Rebonato, Ricardo, and Valerio Gaspari. (2006). "Analysis of Drawdowns and Drawups in the US\$ Interest-Rate Market." *Quantitative Finance* **6**, pp. 297-326.

Savage, Leonard J. (1972). *The Foundations of Statistics*, 2nd ed. New York: Dover Publications.

Schmeidler, David. (1986). "Integral Representation Without Additivity." *Proceedings of the American Mathematical Society* **97**, pp. 255-261.

Schmeidler, David. (1989). "Subjective Probability and Expected Utility Without Additivity." *Econometrica* **57**, pp. 571-587.