Maximum Drawdown and Asset Pricing

Daehwan Kim

Associate Professor, Department of Economics, Konkuk University, Seoul, 143-701, South Korea
+82 2 450 3937, dkim@konkuk.ac.kr

Maximum Drawdown and Asset Pricing

Abstract

Maximum drawdown refers to the largest cumulative loss of a portfolio within a given time interval. While it has been used by investment professionals as an important measure of portfolio risk for many years, its nature and its implications for asset pricing have not been well understood. The first part of the paper presents a rigorous argument for using maximum drawdown as a portfolio risk measure. The argument is based on liquidity preference of Keynes and rank dependent utility of Quiggin. We extend the Markowitz portfolio problem by including expected maximum drawdown as the third argument of the objective function. The second part of the paper investigates asset pricing implications. The marginal contribution of an individual asset to portfolio maximum drawdown—we call this "co-drawdown"—plays an important role. In an extended CAPM, there exists a linear relationship between expected return and co-drawdown.

Keywords

Maximum drawdown, Rank dependent utility, Markowitz portfolio problem, Co-drawdown, CAPM

JEL Classification

G11, G12, D81

1. Introduction

Maximum drawdown refers to the largest cumulative loss of a portfolio, i.e. the distance between the peak and the bottom within a given time interval. A survey of literature clearly indicates that investors consider maximum drawdown as a relevant and important measure of portfolio risk, sometimes as important as portfolio volatility.[1] However, the cause and the consequence of this phenomenon are not well understood. This paper aims to fill this gap in the literature. The goal of this paper is twofold: first, to provide a rigorous argument for maximum drawdown as a portfolio risk measure, and second, to investigate the asset pricing implications of incorporating maximum drawdown in the portfolio choice problem.

A theory of maximum drawdown as a portfolio risk measure can be built out of two important ideas in economics: liquidity preference as first proposed by Keynes (1964; originally published in 1936) and the rank dependent utility of Quiggin (1982). Keynes notes that "a need for liquid cash may conceivably arise" unexpectedly, and that investors want to protect themselves against such possibility. Given this possibility of random "liquidity shock," investors recognize that the actual holding period can be different from the original plan, and they want to ensure that the portfolio maintains a certain level of performance across different holding periods. Now, what if investors are averse to this "holding period risk" in the way suggested by the rank dependent utility theory of Quiggin (1982)? In this theory, utility is rank-dependent in the sense that decision makers give a higher "decision weight" to high-rank (i.e. low-payout) event. Thus, rank-dependent investors give a significant decision weight to the worst possibility, which in our case is represented by maximum drawdown.

Developing this idea further, we extend Markowitz (1952) portfolio problem and include "expected maximum drawdown" as the third argument of the objective function. In the Markowitz portfolio problem, investors' indirect utility is a function of portfolio mean return and variance, say, $V(\mu_p, \sigma_p^2)$.

---

[1] See, for example, Grossman and Zhou (1993), Magdon-Ismail, Atiya, Prtap, and Abu-Mostafa (2004), Chekhlov, Uryasev, and Zabarankin (2005), Alexander and Baptista (2006), Rebonato and Gaspari (2006), and Pospisil and Vecer (2008). Regulators take maximum drawdown seriously as well. The US Commodity Futures Trading Commision requires futures managers to report maximum drawdown in their performance reporting. See Harding, Nakou, and Nejjir (2003).

In our extension, expected maximum drawdown $\bar{\mathcal{M}}_p$ enters into the portfolio problem so that we may express the indirect utility as $V(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p)$.

The portfolio choice problem with expected maximum drawdown can be easily solved, just as the portfolio problem with skewness is solved by Kraus and Litzenberger (1976). A non-trivial aspect of the problem is how to represent the marginal contribution of an individual asset to portfolio maximum drawdown. Our analysis shows that there exists a simple formula for this, which we call "co-drawdown." Co-drawdown is measured by the return of an individual asset during "market maximum drawdown period," i.e. the period when the market experiences maximum drawdown. Co-drawdown has a number of desirable properties compared to other downside risk measures such as skewness, semi-variance, VaR, and expected shortfall.

With co-drawdown, we modify the two-fund separation result and also extend the CAPM. The relationship between an individual asset return and the market portfolio return can be extended so that the expected return of an individual asset is proportional to its expected co-drawdown as well as to its market beta. Also, market premium—the expected excess return of the market over the riskfree rate—is determined by the expected maximum drawdown as well as the volatility of the market portfolio. We interpret the expected maximum drawdown of the market as the premium demanded by investors for being exposed to liquidity shocks.

The rest of the paper is organized as follows. Section 2 presents an overview of related literature. Section 3 provides a theory of maximum drawdown as a portfolio risk measure and an extension of Markowitz portfolio problem. Section 4 solves the extended portfolio problem and obtains an extension of CAPM incorporating expected co-drawdown. Section 5 discusses additional properties of co-drawdown. Finally, Section 6 concludes the paper.

2. Literature Review

We start with a review of studies on maximum drawdown. Then we discuss Keynesian liquidity preference and the rank dependent utility theory, the two building blocks of our theory of maximum drawdown as a risk measure. We relate the rank dependent utility theory to a large literature on ambiguity aversion. Finally, we review the works that extend the Markowitz portfolio problem and the CAPM by considering downside risk measures.

Maximum Drawdown

Statistical properties of maximum drawdown are the topic of the studies by Magdon-Ismail, Atiya, Pratap, and Abu-Mostafa (2004), Rebonato and Gaspari (2006), and Pospisil and Vecer (2008). Magdon-Ismail et al. (2004) present a formula for the expected maximum drawdown of a Brownian motion, whereas Pospisil and Vecer (2008) examine the case of a geometric Brownian motion. Note that when price is a (geometric) Brownian motion, return is serially uncorrelated. This is a rather serious limitation: if return is serially uncorrelated, the distribution is completely characterized by mean and variance; there is no reason to consider maximum drawdown in an investor's decision. Needless to say, analyzing maximum drawdown under serial correlation is very complicated. A formula developed by Afonja (1972) is applicable, but carrying out further analysis from this formula is not easy either. Rebonato and Gaspari (2006) examine the properties of maximum drawdown empirically from daily interest rate data.

In this paper, we allow serial correlation in returns. So in that sense, our setup is more general than that of Magdon-Ismail et al. (2004) and Pospisil and Vecer (2008). In terms of mathematics, we do not report any new insights. Instead, we focus on the derivative of expected maximum drawdown, which we call co-drawdown, and its role in asset pricing.

Several studies examine portfolio choice problem with a constraint on maximum drawdown. Grossman and Zhou (1993) consider a portfolio of one risky asset and one riskfree asset in a continuous time setting. Chekhlov, Uryasev, and Zabarankin (2005) consider a multi-asset portfolio choice problem. In their study, however, expected maximum drawdown replaces, rather than complements, volatility. Alexander and Baptista (2006) consider both volatility and maximum drawdown, but in a single-period finite-state-space setting. Set in a single-period framework, maximum drawdown is identical to the minimum return. Having a finite state-space is also restrictive.

Our formulation of portfolio problem is more general in some respects. Unlike Grossman and Zhou (1993), we consider a portfolio of many assets, and, unlike Alexander and Baptista (2006), we consider many periods and an infinite state-space. Another difference from these studies is that we consider *expected* maximum drawdown, not maximum drawdown itself. For most continuous distributions including normal distribution, constraining maximum drawdown to be less than a fixed value leads to a cash-only position. So it does not make much sense. Grossman and Zhou (1993) allow continuous rebalancing so that it is possible to have a risky portfolio and guaranteed to be above a certain level. Alexander and Baptista (2006) consider discrete distributions, so it is possible to create a static portfolio whose value is bounded from below. We do not allow continuous rebalancing, nor do we deal with discrete distributions. So it is natural for us to work with expected maximum drawdown, not maximum drawdown itself.

Keynesian Liquidity Preference

In developing his theory of interest rates, Keynes (1964) proposes the concept of liquidity preference: A need for liquid cash may arise unexpectedly; when such a need arises, investors may incur a loss if the sale price of risky assets happens to be low at that particular moment; considering this possibility, investors give preference to liquid assets. Note that what generate liquidity preference are the existence of liquidity shock (i.e. sudden demand for liquidity) and the volatility of risky asset prices.

In this paper, we model liquidity shock by the variability of holding period. When there is sudden demand for liquidity, investors liquidate the risky portfolio; i.e., it is the end of holding period. Holding period begins when the opposite happens, i.e. when there is cash inflow. We work with a probability distribution over the set of holding periods. The volatility of risky asset prices is modeled in the standard way, i.e. by expressing returns as random variables.

Our interpretation of Keynesian liquidity preference is comparable to that of Hill (2009), who has termed "horizon uncertainty" to indicate holding period risk. Hill notes that this risk is "driven by uncertainty regarding investment horizon or holding period, the correlation across the cash flow demands of market participants, and short-run constraints on the market making capital." Thus, the overall liquidity condition of the market is relevant to holding period risk, and it is possible to relate our discussion to the literature on liquidity risk as well (for example, Amihud, 2002, Pastor and Stambaugh, 2003, and Acharya and Pedersen, 2005).[2]

Rank Dependence and Ambiguity Aversion

[2] Another liquidity literature is the one on the yield curve, where liquidity premium is implied by the slope of the yield curve. The relevance of Keynesian liquidity preference to such liquidity premium does not require elaboration.

The rank dependent utility (RDU) is often motivated by the Allais (1979) paradox, where decision makers appear to give ‘excessive’ preference for certainty. One way to solve the Allais paradox is by incorporating ‘how people feel about probabilities’ in the words of Wakker (2008). If a decision maker is not confident about the accuracy of the given probability, he may assign less than 100% to the probability, assigning extra significance to certainty. In the RDU of Quiggin (1982), a probability weighting function takes the place of probability in the expected utility theory. Its role is to “distort” the given probability so that it reflects the decision maker’s attitude toward the probability. When the probability weighting function is such that certainty receives more significance, we say that the decision maker is probabilistic-risk averse.

The RDU is a special case of the Choquet expected utility (CEU) of Schmeidler (1989).[3] This theory has been developed as a response to the Ellsberg (1961) paradox, where uncertainty cannot be adequately represented by probability. In the CEU theory, probability weighting function is replaced with a more general function called capacity. Capacity is not additive, its nonadditivity reflecting the ambiguity of the problem. As emphasized by Wakker (2008), the CEU and the RDU deal with two different situations. The CEU is applicable when there is no probability; the RUD is applicable when there is probability but it requires some “distortion.” Nonetheless, mathematically, the RDU is a special case of CEU. So we adopt certain insights from the CEU.

Since Schmeidler (1989), a large literature has emerged investigating various aspects of ambiguity and ambiguity aversion. Of special relevance to our study are those papers that examine the asset pricing implications of ambiguity aversion. Epstein and Wang (1994) present an intertemporal asset pricing model and Chen and Epstein (2002) study its continuous time version. Gollier (2011) focuses on the equity premium puzzle in the framework of the *smooth utility* of Klibanoff, Marinacci, and Mukerji (2005). Other authors focus on the cross-sectional implications. Kogan and Wang (2003) and Garlappi, Uppal, and Wang (2007) derive an equilibrium condition where expected returns are proportional to a measure of ambiguity. Boyle, Garlappi, Uppal, and Wang (2012) apply this model to the question of diversification benefit. Maccheroni, Marinacci, Rustichini, and Taboga (2009) present the monotone CAPM where CAPM beta is modified to reflect truncation to satisfy monotonicity. Their analysis is based on the *variational utility* of Maccheroni, Marinacci, and Rustichini (2006).

---

[3] The rank dependent utility is also a special case of probabilistically sophisticated utility of Machina and Schmeidler (1992), where utility depends on probability, but not in a linear form.

Markowitz, CAPM, and Downside Risk Measures

Many of the papers reviewed above—Kogan and Wang (2003), Garlappi et al. (2007), Boyle et al. (2012), and Maccheroni et al. (2009)—adopt a popular scheme that we also follow in the current paper. The scheme consists of three steps: (i) to modify the Markowitz problem, (ii) to examine the first-order condition, and then (iii) to arrive at asset pricing equation which can be interpreted as an extended CAPM. The way that the Markowitz problem is modified reflects the motivation of each research project. The papers mentioned above modify the Markowitz problem to incorporate ambiguity. Others have a goal of bringing downside risk measures into the discussion. We review several of such studies.

Kraus and Litzenberger (1976) introduce skewness into the Markowitz problem and extend the CAPM where there is another term, "co-skewness," reflecting assets' covariance with market volatility. Hogan and Warren (1972, 1974) replace portfolio variance with portfolio semi-variance in the Markowitz problem. The resulting CAPM-like formula has "co-semivariance" in place of usual covariance. Harlow and Rao (1989) generalize this model by replacing semi-variance with general "lower-partial moment." Extending the work by Tasche (2000), Bertsimas, Lauprete, and Samarov (2004) consider a portfolio problem with expected shortfall, and present a CAPM-like formula where "shortfall beta" measures the risk of individual assets.

Among popular downside risk measures, only value-at-risk (VaR) has not been integrated into the Markowitz-CAPM framework. This situation is due to the fact that VaR is not differentiable, making it difficult to come up with "co-VaR" or "VaR beta." Various properties of VaR and its comparison to expected shortfall have been studied by Artzner, Delbaen, Eber, and Heath (1999) and Rockafellar and Uryasev (2002).

3. Holding Period Risk, Rank Dependence, and Maximum Drawdown

In the classical portfolio problem of Markowitz (1952), an investor maximizes his indirect utility $V(\mu_p, \sigma_p^2)$ over all portfolios, where μ_p and σ_p^2 are the mean and variance of the investor's portfolio. We consider an extension of this problem which incorporates the expected maximum drawdown of the investor's portfolio, $\bar{\mathcal{M}}_{\mathrm{p}}$, into the indirect utility function; i.e., the indirect utility to be maximized is $V(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p)$. While several authors have studied the portfolio choice problem with maximum drawdown, no justification has been suggested for including maximum drawdown. The goal of our analysis is to do just that. We discuss a set of circumstances that lead to the indirect utility of the form $V\left(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p\right)$. We do not aim to identify the necessary and sufficient condition for maximum drawdown. We have a more modest goal: we present a situation which we believe is probable and in which maximum drawdown enters into the portfolio choice problem.

Our argument is based on two important ideas in economics: the Keynesian liquidity preference and Quiggin's (1982) rank dependent utility theory. As Keynes (1964) notes, a need for liquid cash may arise unexpectedly, and investors want to protect themselves against this possibility of "liquidity shock." To put it another way, investors recognize that the actual holding period can be different from the initial plan, and they want to ensure that the portfolio maintains a certain level of performance across all holding periods. Now, if investors are averse to this "holding period risk" in the way suggested by the rank dependent utility (RDU) theory, then they pay more attention to the worst possibility than probabilities warrant. In other words, investors give a higher "decision weight" to the worst possibility. The worst possibility is captured by maximum drawdown in our case.

Let us formalize the above idea in three steps. First, we introduce the holding period risk. Then we introduce the RDU theory. Finally, we combine these two.

Holding Period Risk

Suppose that the investor wants to construct a portfolio for the next T periods. We index time periods by t so that $t = 1, \cdots, T$. Let $r_{p,t}$ be the portfolio return for t. If there is no liquidity shock, then the investor holds the portfolio for all the periods, and the holding period return is $r_{p,1} + \cdots + r_{p,T}$.[4] If there are liquidity shocks, then the investor may hold the portfolio only for some periods. In general, the investor's holding period can be any subset of $\{1,2,\cdots,T\}$. Let us assume that the investor's holding period is one contiguous block of time periods,[5] i.e. $\{s, s+1, \cdots, t\}$ for some s and t,

[4] The formula is exact only if the returns are log-returns. If the returns are simple returns, the formula is ignoring compounding. Our discussion in the next section is on the no-compounding basis. The analysis can be extended to the case of compounding.

[5] If the holding period is not a contiguous block of time periods, it means that the investor creates and liquidates his portfolio at multiple times. If the investor's utility is time-separable, the optimal investment problem can be broken down into multiple *independent* problems, each of which involves creating one optimal portfolio for a contiguous block of time periods.

$1 \le s \le t \le \mathrm{T}$. Let us denote this holding period by $[\mathrm{s}, \mathrm{t}]$ and the holding period return by $r_{p,[s,t]}$. Then $r_{p,[s,t]} = r_{p,s} + \cdots + r_{p,t}$.

In words, our setup can be described in the following way: the investor has zero wealth initially. At the beginning of any time period, a positive liquidity shock may occur, in which case the investor receives one unit of cash to invest. The investor may face a negative liquidity shock any time, as long as he has positive wealth. The negative liquidity shock is always of the same size as the total wealth. There will be exactly one positive liquidity shock and one negative liquidity shock between period 1 and period T.[6]

Let us denote the set of possible holding periods by $\mathcal{H}$, i.e. $\mathcal{H} = \{[s,t]: 1 \le s \le t \le \mathrm{T}\}$. The investor facing liquidity shock recognizes that his holding period return can be any $r_{p,[s,t]}$ with $[\mathrm{s}, \mathrm{t}] \in \mathcal{H}$. Thus, there are two sources of risk for $r_{p,[s,t]}$. The first is the "return risk," i.e. the uncertainty regarding the realization of portfolio returns, $r_{p,1}, \cdots, r_{p,T}$. The second is the "holding period risk," i.e. the uncertainty regarding the realization of the holding period $[s, t]$.

The return risk can be represented by $\mathbb{R}^{\mathrm{T}}$ (the set of all possible values of $r_{p,1}, \cdots, r_{p,T}$) and a probability distribution P over $\mathbb{R}^{\mathrm{T}}$. Similarly, the holding period risk can be represented by $\mathcal{H}$ and a probability distribution Q over $\mathcal{H}$. The total risk can be represented by the product space $\mathbb{R}^{\mathrm{T}} \times \mathcal{H}$ and a probability distribution R over $\mathbb{R}^{\mathrm{T}} \times \mathcal{H}$. The investor's expected utility is $\mathrm{E}(r_{p,[s,t]}) = \int_{\mathbb{R}^T \times \mathcal{H}} r_{p,[s,t]} dR$. Note that P and Q are marginal distributions of R. In particular, if the return risk and the holding period risk are independent, R is the product of P and Q, i.e. $R(E) = P(E_1)Q(E_2)$. for $E \subseteq \mathbb{R}^{\mathrm{T}} \times \mathcal{H}$ with $E = E_1 \times E_2$, $E_1 \subseteq \mathbb{R}^{\mathrm{T}}$, and $\mathrm{E}_2 \subseteq \mathcal{H}$. Let *u* be the von-Neumann Morgenstern utility index so that the investor's instantaneous utility is $\mathrm{u}(r_{p,[s,t]})$. Then the investor's expected utility is $\mathrm{E}[\mathrm{u}(r_{p,[s,t]})] = \int_{\mathbb{R}^T} \int_{\mathcal{H}} u(r_{p,[s,t]}) dQdP$.

Literature suggests that the expected utility theory is not always applicable. Below we consider the possibility that the expected utility theory is applicable for the return risk, but not for the holding period risk. Our discussion is based on the RDU theory of Quiggin (1982). We relate our discussion to the ambiguity aversion of Schmeidler (1989) as well.

[6] See the previous footnote.

Rank Dependent Utility and Ambiguity Aversion

As seen in the Allais (1979) paradox, decision makers often give ‘excessive’ preference for certainty. This may be caused by the lack of confidence in the accuracy of the probability, which leads the decision makers to assign less than 100% to the probability. In the RDU theory of Quiggin (1982), this is done through a probability weighting function w.[7] One possibility is to assign a smaller weight for a small probability and a larger weight for a large probability (i.e. $w(x) \approx 0$ for x close to 0 and $w(x) \approx x$ for x close to 1). Then an event close to being certain ($x \approx 1$) has a larger effect than its probability suggests, generating the certainty effect in the Allais paradox. When the outcomes of a prospect x are rank ordered, i.e. $x_1 \geq \cdots \geq x_N$, the RDU of x is defined as

$$RDU(x) = \sum_{j=1}^{N} \pi_j u(x_j)$$

where $\pi_j = w(p_1 + \cdots + p_j) - w(p_1 + \cdots + p_{j-1})$, $\pi_1 = w(p_1)$, and p_i is the probability of x_i. When the probability weighting function is such that almost certain events have larger weights than other events, decision makers are said to have *probabilistic risk aversion*.

The RDU is a special case of the Choquet expected utility (CEU) of Schmeidler (1989). The CEU of prospect x with rank-ordered outcomes $x_1 \geq \cdots \geq x_N$ is defined by

$$CEU(x) = \int u(x)dv = \sum_{j=1}^{N} \mathrm{u}(\mathrm{x_j})\left[v(x \geq x_j) - v(x \geq x_{j-1})\right]$$

The probability-like function v is called a capacity. It shares many properties with the probability distribution, but it is not additive. Thus, $v(x \geq x_j) - v(x \geq x_{j-1}) \neq v(x = x_j)$. As follows from the formula, capacity v can be obtained from probability weighting function w and probability *p*: $v(x \geq x_j) = w[p(x \geq x_j)]$. The CEU is understood as reflecting decision makers’ aversion to ambiguity. Ambiguity arises when there is no adequate probability to describe uncertainty, as in the Ellsberg (1961) paradox, where decision makers show preference for the choice with a clearly-defined probability. As mentioned earlier, the situation where the CEU is applicable is distinctive from the situation where the RDU is applicable. For the former, ambiguity aversion is relevant, whereas for the latter probabilistic risk aversion is relevant. Nonetheless, the CEU formula “works” for the situation of probabilistic risk aversion as well. So we continue our discussion in terms of the CEU even though we are primarily interested in the situation of probabilistic risk aversion. This allows us to draw from extensive literature on ambiguity aversion which is phrased in the CEU framework.

[7] Three properties are required for w: (i) $w(0) = 0$, (ii) $w(0) = 0$, and (iii) *w* is continuous and increasing.

In the CEU framework, ambiguity arises as capacity v is not additive. For example, for two disjoint events, A and B, $v(A \cup B) - v(A) - v(B) \geq 0$. Extending this idea, Dow and Werlang (1992) define the ambiguity of event A as $v(S) - v(A) - v(A^c)$ where $S = A \cup A^c$ is the set of the state. Let us generalize this further, and define the ambiguity of a partition $\mathrm{E}_1, \cdots, E_k$ of S as $v(S) - v(E_1) - \cdots - v(E_k)$. Let us call the largest among the ambiguities of partitions the "total ambiguity" of capacity v, i.e. the total ambiguity of v is $\sup\{v(S) - v(E_1) - \cdots - v(E_k)$: $\mathrm{E}_1, \cdots, E_k$ are a partition of $S\}$.

When can we say that one decision maker is more ambiguity averse than another? Ghirardato and Marinacci (2002) show[8] that one decision maker is more ambiguity averse than another if and only if the capacity of the latter is greater than the capacity of the former at each event; i.e., decision maker 1 is more ambiguity averse than decision maker 2 if $v_1(E) \leq v_2(E)$ for all E.[9]

Based on the two ideas described above—"total ambiguity" and "being more ambiguity averse than someone else"—we would like to characterize "extreme aversion" in the following way: we call a decision maker "extremely ambiguity averse" if he is more ambiguity averse than anyone else given the total ambiguity. In terms of probability weighting function, we can define "total probabilistic risk," "being more probabilistic-risk averse than someone else," and "extremely probabilistic-risk averse" in a similar manner. The following lemma shows how it can be done.

Lemma 1. Fix probability p. Suppose that all probability weighting functions are convex. Then the following statements are true. (i) The total ambiguity of the capacity induced by a probability weighting function w is $\sup 1 - w(x)/x$ where $0 < x \leq 1$. (ii) Decision maker i whose capacity is induced by w_i is more ambiguity averse than decision maker j whose capacity is

[8] See Theorem 17 of the article.

[9] Schmeidler (1989) identifies the ambiguity aversion with the convexity of capacity v. Thus, it would seem natural to say that decision maker 1 is more ambiguity averse than decision maker 2 if v_1 is more convex than v_2. As clarified by Epstein (1999), ambiguity aversion can be identified with the convexity of capacity only in the Anscome-Aumann (1963) framework where each outcome of a prospect is a lottery rather than cash. In the Savage (1972; first edition in 1954) framework where each outcome of a prospect is cash, ambiguity aversion is identified with the monotonicity of capacity. Thus, Ghirardato and Marinacci's (2002) theorem is understandable.

induced by w_j if $w_i(x) \leq w_j(x)$ for all x with $0 \leq x < 1$. (iii) A decision maker whose capacity is induced by

$$w(x) = \begin{cases} ax & if \ \ 0 \leq x < 1 \\ 1 & if \ x = 1 \end{cases}$$

is extremely ambiguity averse.

Proof. Noting that if capacity v is induced by w and p, then $v(E) = w[p(E)]$, where E is an event, (i) can be easily verified geometrically. Consider the graph of $w: [0,1] \to [0,1]$. Draw a line connecting the origin and $(x, w(x))$. This line passes through $(1, w(x)/x)$, and the vertical distance between $(1, w(x)/x)$ and $(1,1)$ is $1 - w(x)/x$. Consider an event E_1 whose probability is x. Given the convexity of w, $v(S) - v(E_1) - \cdots - v(E_k)$ is greater than $1 - w(x)/x$ for any partition that includes E_1. Thus, the total ambiguity is greater than or equal to $\sup 1 - w(x)/x$. It is easy to show that, if the total ambiguity is strictly greater than $\sup 1 - w(x)/x$, we get to a contradiction. So (i) must be true. (ii) is obvious from the definition of "being more ambiguity averse than someone else." (iii) can be easily derived from (i) and (ii). □

Based on this lemma, we call $\sup 1 - w(x)/x$ the *total probabilistic risk* of w. Also, we say that decision maker i is *more probabilistic-risk averse than* decision maker j if $w_i(x) \leq w_j(x)$. Finally, we say that a decision maker is *extremely probabilistic-risk averse* when w is as shown in (iii) of Lemma 1.

Synthesis

We now combine the two ideas that we have developed above: (i) the existence of holding period risk in addition to return risk and (ii) the RDU theory and its specialization for extreme probabilistic risk-aversion.

We apply the RDU theory to holding period risk. In fact, in presenting his theory of liquidity preference, Keynes (1964) notes that the uncertainty underlying the liquidity preference is unlikely to be expressed by probability alone.[10] Given the uncertainty over holding period, it is probable that investors are more cautious and assign a larger weight to less favorable situation as the RDU theory suggests.

[10] See page 169 of Keynes (1964).

While we apply the RDU theory to holding period risk, we assume that the expected utility theory is still applicable to return risk. That is, investors are well informed of the return distribution and confident of its accuracy; there is no rank dependency as far as return risk is concerned. Rank dependency arises only for holding period risk, as investors are not confident of the accuracy of the probability on holding period.

We can model this situation in the following way. Recall that when return risk and holding period risk are independent, the expected utility is $\mathrm{E}\left[\mathrm{u}\left(r_{p,[s,t]}\right)\right] = \int_{\mathbb{R}^T} \int_{\mathcal{H}} u\left(r_{p,[s,t]}\right) dQdP$. As there is no rank dependence as far as $(\mathbb{R}^{\mathrm{T}}, \mathrm{P})$ is concerned, we leave the outer integral as it is. As for $(\mathcal{H}, \mathrm{Q})$, we allow rank dependence, thus the investor's utility becomes[11]

$$U\left(r_{p,[s,t]}\right) = \int_{\mathbb{R}^T} \sum_i \pi([s_i, t_i])\, u\left(r_{p,[s_i,t_i]}\right) dP$$

where $\{[s_i, t_i]\}$ is a rank-ordered sequence made of the elements of $\mathcal{H}$ given a particular realization of portfolio returns; i.e., $[s_i, t_i] \in \mathcal{H}$ for all i and, given $\mathrm{r}_{\mathrm{p},1}, \cdots, r_{p,T}$, for any i and j with $i < j$, $u\left(r_{p,[s_i,t_i]}\right) \geq u\left(r_{p,[s_j,t_j]}\right)$. π is the decision weight; i.e., $\pi([s_i, t_i]) = \mathrm{w}\left(\sum_{j=1}^{i} Q\left([s_j, t_j]\right)\right) - w\left(\sum_{j=1}^{i-1} Q\left([s_j, t_j]\right)\right)$ for some probability weighting function *w*.[12]

We now impose the extreme probabilistic risk aversion on *w* and π. So the form of *w* is as in Lemma 1. Then the investor's utility takes a familiar shape, as stated in the following theorem.

Theorem 1. Assume that (i) return risk and holding period risk are independent, (ii) when a holding period is fixed (i.e. when there is only return risk) the utility is not rank dependent, (iii) when returns are fixed (i.e. when there is only holding period risk), the utility is rank dependent and exhibits extreme probabilistic risk aversion given the total probabilistic risk of $1 - \alpha$. Then the investor's utility can be stated as follows:

$$U\left(r_{p,[s,t]}\right) = \alpha \int_{\mathbb{R}^T} \int_{\mathcal{H}} u\left(r_{p,[s,t]}\right) dQ\, dP + \int_{\mathbb{R}^T} u\left(\min_{[\mathrm{s,t}] \in \mathcal{H}} r_{p,[s,t]}\right) dP$$

---

[11] It is possible to obtain another expression by applying dP first and then dQ. See Ghirardato (1997).

[12] We can make this derivation more rigorous by starting from a capacity over product space $\mathbb{R}^{\mathrm{T}} \times \mathcal{H}$, assuming Moebius independence, and then assuming that there is no non-risk uncertainty in $\mathbb{R}^{\mathrm{T}}$ along the analysis of Ghirardato (1997).

$$= \alpha \int_{\mathcal{H}} \int_{\mathbb{R}^T} u(r_{p,[s,t]}) dP\, dQ + \int_{\mathbb{R}^T} u\left(\min_{[s,t]\in\mathcal{H}} r_{p,[s,t]}\right) dP$$

Proof. We only need to substitute the formula in Lemma 1 for w and π. □

The first part of the utility shown in Theorem 1 is the standard expected utility, and the second part is the expected utility of maximum drawdown. To make it more comparable to the objective function of the Markowitz problem, we propose the following approximation. We would like to express the investor's utility as a function of three parameters: T-period portfolio mean return μ_p, T-period portfolio variance σ_p^2, and the mean of maximum drawdown $\bar{\mathcal{M}}_p = \mathrm{E}^{\mathrm{P}}\left(\min_{[s,t]\in\mathcal{H}} r_{p,[s,t]}\right)$. Consider the first part of the utility: $A \equiv \alpha \int_{\mathcal{H}} \int_{\mathbb{R}^T} u(r_{p,[s,t]}) dP\, dQ$. The "inner" integral, $\int_{\mathbb{R}^T} u(r_{p,[s,t]}) dP$, is the standard expected utility, which Markowitz (1952) approximates as a function of mean return $\mathrm{E}^{\mathrm{P}}(r_{p,[s,t]})$ and variance $\mathrm{var}^{\mathrm{P}}(r_{p,[s,t]})$. Given the "outer" integral, we need further justification to approximate A as a function of μ_p and σ_p^2 (and also $\bar{\mathcal{M}}_p$). If the distribution of $r_{p,[s,t]}$ is "stationary" so that $\mathrm{E}^{\mathrm{P}}(r_{p,[s,t]})$ and $\mathrm{var}^{\mathrm{P}}(r_{p,[s,t]})$ depend only on the length of holding period $t-s$,[13] then A can be written as $\alpha \sum_{j=1}^{T} Q(t-s=j) f(\mu_j, \sigma_j^2)$, where $\mu_{p,j}$ and $\sigma_{p,j}^2$ are the mean and variance of $r_{p,[s,t]}$ given $t-s=j$, and $f(\,,)$ is a function comparable to Markowitz' indirect utility. Furthermore, under the stationarity, mean returns of different periods are identical so that $\mu_{p,j} = (j/T)\mu_p$. Unless serial correlation is zero, $\sigma_{p,j}^2$ cannot equal $(j/T)\sigma_p^2$. However, $\sigma_{p,j}^2$ can be well approximated by σ_p^2 and $\bar{\mathcal{M}}_p$ in some situations. If that is the case, A can be approximated by μ_p, σ_p^2, and $\bar{\mathcal{M}}_p$. Now consider the second part of the investor's utility, $B \equiv \int_{\mathbb{R}^T} u\left(\min_{[s,t]\in\mathcal{H}} r_{p,[s,t]}\right) dP$. One can think of approximating it with a function of mean maximum drawdown $\bar{\mathcal{M}}_p$ and variance of maximum drawdown $\mathrm{var}^{\mathrm{P}}\left(\min_{[s,t]\in\mathcal{H}} r_{p,[s,t]}\right)$. In general, this approximation can be rather inaccurate: even if returns are normal, maximum drawdown is not normal, making such approximation imprecise. However, in certain situations, $\bar{\mathcal{M}}_p$ and $\mathrm{var}^{\mathrm{P}}\left(\min_{[s,t]\in\mathcal{H}} r_{p,[s,t]}\right)$ together with σ_p^2 may approximate $\int_{\mathbb{R}^T} u\left(\min_{[s,t]\in\mathcal{H}} r_{p,[s,t]}\right) dP$. It is also possible to obtain good approximation only with $\bar{\mathcal{M}}_p$ and σ_p^2, i.e. without $\mathrm{var}^{\mathrm{P}}\left(\min_{[s,t]\in\mathcal{H}} r_{p,[s,t]}\right)$.

---

[13] If we interpret $\mathrm{E}^{\mathrm{P}}(r_{p,[s,t]})$ and $\mathrm{var}^{\mathrm{P}}(r_{p,[s,t]})$ in the conditional sense, then the "stationarity" would not lead to $\mathrm{E}^{\mathrm{P}}(r_{p,[s,t]})$ and $\mathrm{var}^{\mathrm{P}}(r_{p,[s,t]})$ being dependent only on $t-s$. What is necessary is the property that, conditional on the information available at the beginning of period 1, $\mathrm{E}^{\mathrm{P}}(r_{p,[s,t]})$ and $\mathrm{var}^{\mathrm{P}}(r_{p,[s,t]})$ depend only on $t-s$. Example 1 below discusses such an example. Example 2 discusses an example where such property fails but our approximation is nonetheless exact.

Examples 1 and 2 show two cases where our proposed approximation is exact.

Example 1. Suppose that, conditional on the information available at the beginning of period 1, portfolio returns $r_p = (r_{p,1}, \cdots, r_{p,T})'$ have a multivariate normal distribution with mean vector μ and variance-covariance matrix Σ where

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_1 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho & \cdots & \rho^{T-1} \\ \rho & \sigma_1^2 & & \\ & & \ddots & \\ & & & \sigma_1^2 \end{pmatrix}$$

Suppose also that we restrict the values of μ_p, σ_p^2 and ρ to the range where there is one-to-one correspondence between $(\mu_p, \sigma_p^2, \rho)$ and $(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p)$. Then the investor's utility can be expressed as $V\left(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p\right)$ for some function V.

Example 2. Suppose that portfolio returns follow AR(1) process, i.e. $r_{p,t+1} = \alpha + \beta r_{p,t} + \varepsilon_{t+1}$, where ε_{t+1} has a normal distribution with zero mean and the variance of σ^2. Suppose also that we restrict the values of α, β and σ^2 to the range where there is one-to-one correspondence between $(\alpha, \beta, \sigma^2)$ and $(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p)$. Then the investor's utility can be expressed as $V\left(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p\right)$ for some function V. All the moments are to be calculated conditional on $r_{p,0}$.

In Example 1, as the return distribution has three parameters μ_1, σ_1^2, and ρ, the utility can be written in terms of these three parameters. Thus, if we restrict the range of μ_p, σ_p^2 and ρ so that there is one-to-one correspondence between $(\mu_p, \sigma_p^2, \rho)$ and $(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p)$, it follows that the utility can be expressed as a function of μ_p, σ_p^2, and $, \bar{\mathcal{M}}_p$. In Example 2, conditional on $r_{p,0}$, means and variances of different time periods are not identical. So we are in a different situation from Example 1. Nonetheless, the return distribution, and thus the utility, depends only on three parameters α, $, \beta$, σ^2. So, when the range of these parameters is restricted, we can again express the utility as a function of μ_p, σ_p^2, and $, \bar{\mathcal{M}}_p$.

In these examples, it would certainly make more sense to express the utility in terms of the underlying parameters, i.e. μ_1, σ_1^2, and ρ in case of Fact 1 and α, β, and σ^2 in case of Fact 2. Also, any other re-parameterization of these underlying parameters should work as well. However, our re-parameterization in terms of μ_p, σ_p^2, and $, \bar{\mathcal{M}}_p$ is superior in that it reflects the structure of Theorem 1 well.

4. Portfolio Choice and its Implications on Asset Pricing

In the previous section, we have argued for the inclusion of expected maximum drawdown in the portfolio choice problem. In this section, we examine its consequences. We consider the portfolio choice problem where expected maximum drawdown enters into the objective function, and study its implications on asset pricing, i.e. how expected return of each asset is altered by the consideration of expected maximum drawdown.

We extend the standard single-period portfolio choice problem of Markowitz (1952) into a multi-period setup. The extension is not trivial because of one technical difficulty. In the single period problem, the portfolio return can be expressed as a linear combination of individual asset returns, say, $w_1 r_1 + \cdots + w_N r_N = w' r$. In the multi-period framework, it is not so simple. Individual asset returns should be cumulated first, to account for compounding effect, and weights can be applied only after that. The portfolio return might be expressed as $w_1 \prod(1 + r_{1,t}) + \cdots + w_n \prod\left(1 + r_{1,t}\right) - 1$, which cannot be represented as a linear combination of $w = \{w_i\}$ and $r = \{r_{i,t}\}$.[14] Given this difficulty, we face with two choices. The first choice is to make the discussion as rigorous as possible at the cost of giving up familiar matrix algebra and thus making the discussion rather abstract. The second choice is to make the discussion more intuitive at the cost of ignoring the compounding and thus making the discussion less complete. We choose the latter. We note, however, that the consideration of compounding does not alter the main result of this paper.[15]

We need to extend our notation to account for the multiplicity of assets. For clarity, we describe the portfolio choice problem again. The investor wants to construct a portfolio for the next *T* periods. We index time periods by *t* so that $t = 1, \cdots, T$. Suppose that there are *N* risky assets and one riskfree asset. We index the risky assets by *i* so that $i = 1, \cdots, N$. Let us denote the *excess* returns of risky asset *i* (over the riskfree rate) by *T*-vector r_i, and the excess returns of all risky assets by T-by-N matrix R, i.e., $R = [r_1, \cdots, r_N]$. Let us denote the risky portfolio weights by *N*-vector *w* so that $r_p = Rw$ is the return vector of the risky portfolio. If $1 is invested at the beginning of period 1, then

---

[14] Nor can it be represented as a linear combination of $\{w_i\}$ and log-returns $\{\log\left(1 + r_{i,t}\right)\}$.

[15] Let R be the T-by-N matrix of returns. Let c be the T-vector of 0 and 1, representing the holding period. Let w be the N-vector of portfolio weights. The holding period return of the portfolio can be written as $\mathrm{g}(R, c)' w$ where $\mathrm{g}(R, c)$ is an N-vector representing the holding period returns of individual assets. We can restate Lemmas and Theorems of this paper using this notation.

the cumulative excess profit at the end of period T is $\iota' r_p$. Let μ_p and σ_p^2 be the mean and variance of $\iota' r_p$, where ι is a vector of 1's. Let us denote expected maximum drawdown of the portfolio by $\bar{\mathcal{M}}_{\mathrm{p}}$. Then the portfolio choice problem can be expressed as $\max_{\mathrm{w}} V(\mu_p, \sigma_p^2, \bar{\mathcal{M}}_p)$, where V is the indirect utility function of the investor.

Three comments on the above setup are in order. First, other than allowing multiple periods, the above setup is very similar to that of Kraus and Litzenberger (1976), who include skewness in the place of $\bar{\mathcal{M}}_{\mathrm{p}}$. Some authors (e.g. Garlappi, Uppal, and Wang (2007)) prefer a more explicit parameterization of the indirect utility, typically a linear function. For example, we may express the indirect utility as $\mathrm{V} = \mu_q - \gamma\sigma_p^2 - \delta\bar{\mathcal{M}}_p$, where γ represents risk aversion and δ represents aversion to expected maximum drawdown.[16] In our solution we linearize the indirect utility; therefore, our setup is not any more general than those studies adopting a linear indirect utility function.

Second, considering our multi-period setup, a more relevant measure of variance could be the expected sample variance $\mathrm{E}\left[\sum(r_{p,t} - \bar{r}_p)^2\right]/T$, rather than the variance of the end-of-period-T value $\sigma_p^2 = \mathrm{var}\left[\sum r_{p,t}\right]$. The difference between these two is whether the serial correlation among $r_{\mathrm{p},1}, \cdots, r_{p,T}$ is added or subtracted. If we assume that the means of $r_{\mathrm{p},1}, \cdots, r_{p,T}$ are identical and denote the variance-covariance matrix of r_p by S, then $\mathrm{E}\left[\sum(r_{p,t} - \bar{r}_p)^2\right]/T = [\mathrm{tr}(S) - \iota' S\iota/T]/T$,[17] whereas $\sigma_p^2 = \mathrm{var}\left[\sum r_{p,t}\right] = \iota' S\iota$. Thus, for the expected sample variance, the off-diagonals of S are subtracted, whereas for the variance of the end-of-period-T value, they are added. For our analysis, it is unimportant which of these two are used as a measure of variance.

The last comment concerns the treatment of the riskfree rate. The above problem is expressed in terms of excess returns. We have subtracted the riskfree rate from risky asset returns. The portfolio choice problem determines the weight w^* of risky portfolio, and the weight on the riskfree asset is determined by $1 - \iota' w^*$. Obviously, the sum of w^* does not have to be 1. By choosing to express the portfolio choice problem in terms of excess returns, we are assuming that the investor defines a drawdown event relative to the riskfree rate. That is, to the investor, a drawdown event occurs when

---

[16] δ also represents thus probabilistic risk aversion and the aversion to liquidity shocks, given the discussion of the previous section.

[17] To see this, note that $\sum(r_{p,t} - \bar{r}_p)^2 = r_p'(I - \iota\iota'/T)r_p = \mathrm{tr}[(I - \iota\iota'/T)r_p r_p']$ and that $\mathrm{E}(r_p r_p') = S + E(r_p)E(r_p')$.

the risky portfolio underperforms the riskfree asset. This assumption turns out to be critical to the development of our theory. Without this assumption, in order to derive the equilibrium implications, we will have to assume instead that every investor owns identical portfolio of risky assets *and* the riskfree asset.[18] Fortunately, defining a drawdown event relative to the riskfree rate does not seem counter-intuitive.

We now present the solution to the portfolio problem. The first order condition is obtained by differentiating the objective function with respect to $w_1, \cdots, w_N$, and setting the results equal to zero, i.e.

$$\frac{\partial V}{\partial w_i} = \frac{\partial \mathrm{V}}{\partial \mu_p}\frac{\partial \mu_\mathrm{p}}{\partial \mathrm{w_i}} + \frac{\partial \mathrm{V}}{\partial \sigma_p^2}\frac{\partial \sigma_\mathrm{p}^2}{\partial \mathrm{w_i}} + \frac{\partial \mathrm{V}}{\partial \mathcal{M}(\mathrm{r_p})}\frac{\partial \bar{\mathcal{M}}_\mathrm{p}}{\partial \mathrm{w_i}} = 0$$

for $\mathrm{i} = 1, \cdots, \mathrm{N}$. Note that $\partial \mu_p / \partial w_i = \mathrm{E}(\iota' r_i)$, and $\partial \sigma_p^2 / \partial w_i = 2\mathrm{cov}(\iota' r_p, \iota' r_i)$. Assuming stationarity so that the means of $r_{i,1}, \cdots, r_{i,T}$ are identical and rearranging terms, the first order condition above can be written as

$$\mathrm{E}(r_{it}) = \gamma \frac{2}{T} \mathrm{cov}\big(\iota' r_p, \iota' r_i\big) + \delta \frac{1}{T} \frac{\partial \bar{\mathcal{M}}_\mathrm{p}}{\partial w_i}$$

where $\gamma = -(\partial \mathrm{V}/\, \partial \sigma_p^2)/(\partial V / \partial \mu_p)$ is the marginal rate of substitution between portfolio variance and portfolio mean, and $\delta = -(\partial \mathrm{V}/\, \partial \bar{\mathcal{M}}_\mathrm{p})/(\partial V / \partial \mu_p)$ is the marginal rate of substitution between expected maximum drawdown and portfolio mean.

The equation above has an intuitive interpretation. The expected excess return of an asset depends on its marginal contribution to portfolio variance and also to the expected maximum drawdown of the portfolio. The marginal contribution to portfolio variance is measured by the familiar "market beta"

[18] We have two-fund separation when different investors may invest different amounts of money in their risky portfolios, but their risky portfolios are identical. We show that the risky portfolios of different investors are identical by showing that when risky portfolio weights w^* satisfies the first order condition of one investor, kw^* satisfies the first order condition for some other investor. This requires that the terms in the first order conditions are homogeneous of degree one in w. Covariance between portfolio return and individual asset returns is homogeneous of degree one in w. Co-drawdown (to be defined later) is homogeneous of degree one in w only if the riskfree rate does not appear in the expected maximum drawdown term. Otherwise, co-drawdown would not be homogeneous of degree one, and we would not have two-fund separation result. In that case, we would have to assume that every investor owns identical portfolios of risk assets and the riskfree asset to discuss equilibrium implications. Such assumption would be very restrictive.

and its coefficient is the utility parameter indicating the sensitivity of utility to portfolio variance. The term involving the expected maximum drawdown of portfolio has the identical structure.

To discuss the implications of the above equation further, let us introduce some additional notation. Let C be a T(T+1)/2-by-T matrix that transforms a return vector into a cumulative return vector. Matrix C is a collection of rows $\{c(s,s')\}_{1\le s\le T, s'\ge s}$, where $c(s,s') = \{1(t \ge s, t \le s')\}_{1\le t\le T}$. That is,

$$\mathrm{C} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots \\ 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

With this matrix, the maximum drawdown of r_p can be expressed as $\min(Cr_p)$. Let $c^*(\ \ )$ be a T-vector such that $\min(Cr_p) = c^*(r_p)' r_p$, i.e. it transforms a return vector into the maximum drawdown. Then the expected maximum drawdown can be written as $\bar{\mathcal{M}}_p = E[\min(Cr_p)] = E[c^*(r_p)'r_p]$.[19] Based on this notation, Lemma 2 provides a simple representation of $\partial\bar{\mathcal{M}}_p/\partial w_i$. Note that differentiating $E[c^*(r_p)'r_p]$ with respect to w_i is not trivial since $c^*(r_p)$ is not a continuous function, and we may not change the order of derivative $\partial/\partial \mathrm{w}$ and expectation $\mathrm{E}[\ \ ]$.

Lemma 2. (Representation of Co-drawdown) Suppose that $\mathrm{E}(|r_{it}|) < \infty$ for all $\mathrm{i} = 1, \cdots, \mathrm{N}$ and $\mathrm{t} = 1, \cdots, \mathrm{T}$. Then

$$\frac{\partial}{\partial w_i} E[c^*(r_p)'r_p] = E\left[c^*(r_p)' \frac{\partial}{\partial w_i} r_p\right] = E\left[c^*(r_p)' \frac{\partial}{\partial w_i} r_i\right]$$

Proof. From the definition of partial derivative,

$$\frac{\partial}{\partial w_i} E[c^*(r_p)'r_p] = \lim_{\Delta\to 0} E\left[\frac{c^*(r_p + \Delta r_i)'(r_p + \Delta r_i) - c^*(r_p)'r_p}{\Delta}\right]$$

Since $c^*(\cdot)$ selects the worst drawdown periods, $c^*(r_p + \Delta r_i)'(r_p + \Delta r_i) - c^*(r_p)'(r_p + \Delta r_i) \le 0$. Thus,

[19] One could attempt to solve $E[\min(Cr_p)]$ directly. Afonja (1972) provides an explicit formula for $E[\min(x)]$ in the case that x is multivariate normal. Afonja's formula, however, is quite complicated and does not yield a simple interpretation.

$$c^*(r_p+\Delta r_i)'(r_p+\Delta r_i)-c^*(r_p)'r_p$$
$$=c^*(r_p+\Delta r_i)'(r_p+\Delta r_i)-c^*(r_p)'(r_p+\Delta r_i)+c^*(r_p)'(r_p+\Delta r_i)-c^*(r_p)'r_p$$
$$\le c^*(r_p)'\Delta r_i$$

Also, $c^*(r_p+\Delta r_i)'r_p-c^*(r_p)'r_p\ge 0$. Thus,

$$c^*(r_p+\Delta r_i)'(r_p+\Delta r_i)-c^*(r_p)'r_p$$
$$=c^*(r_p+\Delta r_i)'(r_p+\Delta r_i)-c^*(r_p+\Delta r_i)'r_p+c^*(r_p+\Delta r_i)'r_p-c^*(r_p)'r_p$$
$$\ge c^*(r_p+\Delta r_i)'\Delta r_i$$

Thus,

$$\lim_{\Delta\to 0}E\left[c^*(r_p+\Delta r_i)'r_i\right]\le \lim_{\Delta\to 0}E\left[\frac{c^*(r_p+\Delta r_i)'(r_p+\Delta r_i)-c^*(r_p)'r_p}{\Delta}\right]\le E\left[c^*(r_p)'r_i\right]$$

Note that $\left|c^*(r_p+\Delta r_i)'r_i\right|\le |r_{i1}|+\cdots+|r_{iT}|$, and that, from the condition in the given statement, $\mathrm{E}(|r_{i1}|+\cdots+|r_{iT}|)$ is finite, i.e. $(|r_{i1}|+\cdots+|r_{iT}|)f(r_{i1},\cdots,r_{iT})$ is integrable where $f(r_{i1},\cdots,r_{iT})$ is the probability density function. So we can apply the Lebesque's dominated convergence theorem[20], and obtain

$$\lim_{\Delta\to 0}E\left[c^*(r_p+\Delta r_i)'r_i\right]=E\left[\lim_{\Delta\to 0}c^*(r_p+\Delta r_i)'r_i\right]=E\left[c^*(r_p)'r_i\right]$$

Thus, the lower bound and the upper bound of $\lim_{\Delta\to 0}E\left[c^*(r_p+\Delta r_i)'(r_p+\Delta r_i)-c^*(r_p)'r_p\right]/\Delta$ are identical, and we have

$$\lim_{\Delta\to 0}E\left[\frac{c^*(r_p+\Delta r_i)'(r_p+\Delta r_i)-c^*(r_p)'r_p}{\Delta}\right]=E\left[c^*(r_p)'r_i\right]$$

This completes the proof. □

The remainder of this paper centers on $E\left[c^*(r_p)'r_i\right]$, which we call expected "co-drawdown." Expected co-drawdown indicates the contribution of the given asset to the maximum drawdown of the market portfolio. Clearly, expected co-drawdown is a "downside risk measure" in that it indicates the risk that the given asset may contribute to the downside risk of the market portfolio. We study the detailed properties of expected co-drawdown in the next section. Lemma 3 and Lemma 4 are the basic properties that follow immediately from Lemma 2.

---

[20] Lebesque's dominated convergence theorem: Let $\{f_k\}$ be a sequence of measurable functions such that $f_k\to f$ almost everywhere. If there exists an integrable function ϕ (i.e. $\int_{\mathrm{E}}\phi<\infty$) such that $|f_k|\le\phi$ almost everywhere for all *k*, then $\int_{\mathrm{E}}f_k\to\int_E f$. See Chapter 5 of Wheeden and Zygmund (1977).

Lemma 3. (Homogeneity of Co-drawdown) Expected co-drawdown is homogeneous of degree one in w. That is,

$$E[c^*(Rkw)'r_i] = kE[c^*(Rw)'r_i]$$

Proof. Immediate from the definition of $c^*(\quad)$. □

Lemma 4. (Linearity of Co-drawdown) Co-drawdown is linear in asset returns, i.e.

$$E\left[c^*(r_p)'(ar_1 + br_2)\right] = aE\left[c^*(r_p)'r_1\right] + aE\left[c^*(r_p)'r_1\right]$$

for any two numbers a, b, and two return vectors r_1, r_2.

Proof. It is obvious from the formula of co-drawdown. □

Lemma 3 together with the homogeneity of covariance function insures that the right hand side of the first order condition is homogeneous of degree one in w. That is, if w^* is the solution to the portfolio problem, then kw^* is the solution to the portfolio problem for some other values of γ and δ. More specifically, if w^* is the solution for γ and δ, then kw^* is the solution for γ/k and δ/k. This homogeneity ensures that different investors with different values of γ and δ choose the identical portfolios as long as the ratio of γ to δ is identical. So, assuming a constant ratio of γ to δ, we have the two-fund separation result. Then we may interpret r_p in the first order condition as the "market portfolio" of risk assets, and the first order condition itself as the equilibrium condition. Lemma 4 implies that the co-drawdown of the market portfolio is its maximum drawdown. The following theorem summarizes these ideas.

Theorem 2. (Equilibrium Condition) Suppose that the ratio of γ to δ is constant across investors. Then the following statements are true. (i) The two-fund separation holds, i.e. every investor holds the same portfolio of risky assets. Let us call this portfolio the market portfolio. (ii) In equilibrium, the expected return of individual asset is proportional to its market beta and its expected co-drawdown, i.e.

$$\mathrm{E}(r_{it}) = \gamma\frac{2}{T}\mathrm{cov}(\iota' r_M, \iota' r_i) + \delta\frac{1}{T}\,E[c^*(r_M)'r_i]$$

(iii) In equilibrium, the expected return of the market portfolio is proportional to its variance and its expected maximum drawdown, i.e.

$$\mathrm{E}(r_{M,t}) = \gamma \frac{2}{T} \mathrm{var}(\iota' r_M) + \delta \frac{1}{T}\ \bar{\mathcal{M}}_M$$

Proof. (i) follows from Lemma 2 and the homogeneity of covariance function. (ii) follows from Lemma 1 and statement (i) in this Lemma. (iii) follows from linearity. □

5. Properties of Co-drawdown and Comparison to Other Downside Risk Measures

In the previous section, we have discussed some basic properties of expected co-drawdown, i.e. it is homogeneous of degree one in *w* and linear in r_i. We discuss additional properties of co-drawdown in this section. As co-drawdown belongs to the class of downside risk measures, we compare co-drawdown to other downside risk measures as well.

Co-drawdown represents the behavior of asset returns *conditional on* the market portfolio returns. Thus, the non-systematic components of asset returns do not influence co-drawdown, while the systematic components, i.e. "beta," influence it. The following statement formalizes this idea.

Fact 1. (Beta Dependence) Suppose that the distribution of asset returns $r_i = \left(r_{i,1}, \cdots, r_{i,T}\right)$ (i.e. the returns over time of asset *i*) conditional on market returns $r_M = \left(r_{M,1}, \cdots, r_{M,T}\right)$ is such that the following linear regression model holds: $r_i = \alpha + \beta r_M + \varepsilon$, where $\mathrm{E}(\varepsilon | r_M) = 0$ and $\mathrm{var}(\varepsilon | r_M) = \Sigma_\varepsilon$. Then the following statements are true. (i) The co-drawdown of asset *i* does not depend on the non-systematic variances and serial-correlations; i.e., $\mathrm{E}[c^*(r_M)' r_i]$ does not depend on the diagonal and off-diagonal elements of Σ_ε. (ii) The co-drawdown of asset *i* depends on the correlations and cross-serial correlations between asset *i* and the market portfolio; i.e. $\mathrm{E}[c^*(r_M)' r_i]$ depends on the diagonal and off-diagonal elements of β.

Proof. By the law of iterated expectation, $\mathrm{E}[c^*(r_M)' r_i] = \mathrm{E}\{c^*(r_M)' \mathrm{E}[r_i | r_M]\} = \mathrm{E}\{c^*(r_M)'(\alpha + B r_M)\}$, from which the given statements become obvious. □

Thus, if an asset is highly correlated with the current or the previous returns of the market portfolio, its co-drawdown is going to be high. On the other hand, strong serial-correlation in asset returns is irrelevant if it is not related to the market portfolio. Note that Fact 1 reveals one difference between co-drawdown and beta. While the contemporaneous correlation with the market portfolio matters for both, the cross-serial correlation affects co-drawdown only. This is, however, not a fundamental difference between co-drawdown and beta. It is possible to modify the definition of beta to reflect the cross-serial correlation. A more fundamental difference between beta and co-drawdown is that co-drawdown is a downside risk measure. The following statement illustrates this point.

Fact 2. (Downside Beta Sensitivity) Let us partition the space of the market portfolio returns into the following three sets, $\mathrm{A} = \{r_M : r_M \ll 0\}$, $\mathrm{B} = \{r_M : r_M \gg 2\mu_M\}$, and $\mathrm{C} = \{r_M : r_M \notin \mathrm{A}, r_M \notin \mathrm{B}\}$, where $\gg$ indicates the element-by-element comparison, $r_M = (r_{M,1}, \cdots, r_{M,T})$, and μ_M is the mean of r_M. Suppose that the distribution of asset returns $r_i = (r_{i,1}, \cdots, r_{i,T})$ conditional on market returns r_M is such that (i) $r_i = \alpha_- + \beta_- r_M + \varepsilon_-$ on A, (ii) $r_i = \alpha_+ + \beta_+ r_M + \varepsilon_+$ on B, and (iii) $r_i = \alpha + \beta r_M + \varepsilon$ on C where ε_-, ε_+, and ε satisfy the assumption of the linear regression model; i.e. they have zero mean conditional on r_M. Suppose also that the distribution of r_M is symmetric around its mean μ_p. Then the co-drawdown of asset a is more sensitive to "downside beta" β_- than to "upside beta" β_+.

Proof. By the property of conditional expectation,

$$\begin{aligned}\mathrm{E}[c^*(r_M)' r_i] &= \Pr(A)\,\mathrm{E}\{c^*(r_M)' r_i | A\} + \Pr(\mathrm{B})\,\mathrm{E}\{c^*(r_M)' r_i | B\} + \Pr(\mathrm{C})\,\mathrm{E}\{c^*(r_M)' r_i | C\} \\ &= \Pr(A)\,\mathrm{E}\{c^*(r_M)'(\alpha_- + \beta_- r_M) | A\} + \Pr(B)\,\mathrm{E}\{c^*(r_M)'(\alpha_+ + \beta_+ r_M) | B\} \\ &\quad + \Pr(C)\,\mathrm{E}\{c^*(r_M)'(\alpha + \beta r_M) | C\}\end{aligned}$$

The derivative of $\mathrm{E}[c^*(r_M)' r_i]$ with respect to an element in β_- is $\Pr(A)\, E[c_s^*(r_M) r_{M,t} | A]$, whereas the derivative of $\mathrm{E}[c^*(r_M)' r_i]$ with respect to an element in β_+ is $\Pr(B)\, E[c_s^*(r_M) r_{M,t} | B]$, where $c_s^*(r_M)$ is the s-th element of $c^*(r_M)$. Note that on A, $\mathrm{c}_s^*(r_p) = 1$ for all s. By the symmetry of the distribution of r_p, $\Pr(A) = \Pr(\mathrm{B})$. Thus, $\Pr(A)\, E[c_s^*(r_M) r_{M,t} | A]$ is greater than $\Pr(B)\, E[c_s^*(r_M) r_{M,t} | B]$. □

Just as we have derived (expected) co-drawdown from (expected) maximum drawdown, we can derive a "co-measure" from any other portfolio risk measure as well.[21] While the original measure indicates the riskiness of a portfolio, the co-measure indicates the contribution of an asset to the riskiness of the portfolio. Formally, let $\mathrm{X}(r_p)$ be a risk measure of the portfolio return vector

[21] While the term "co-measure" itself is not widely used, adding prefix "co-" to the name of risk measure is not new. For example, Hogan and Warren (1974) talk about co-semivariance, which is derived from semivariance.

$r_p = Rw$. Then a "co-measure" is obtained by differentiating $\mathrm{X}(r_p)$ with respect to the weight of a particular asset, i.e. $\partial\mathrm{X}(r_p)/\partial w_i$. When $\mathrm{X}(r_p)$ is the sample variance, $\partial\mathrm{X}(r_p)/\partial w_i$ is the sample covariance. When $\mathrm{X}(r_p)$ is the (expected) maximum drawdown, $\partial\mathrm{X}(r_p)/\partial w_i$ is the (expected) co-drawdown, assuming that r_p is an optimal portfolio.

Given that co-drawdown is a downside risk measure, it is worth comparing it to other downside co-measures, i.e. the co-measures associated with downside risk measures. We consider four popular downside risk measures and co-measures associated with them: skewness (Kraus and Litzenberger, 1976), semi-variance (Hogan and Warren, 1972, 1974), value-at-risk (Artzner et al., 1999), and expected shortfall (Artzner et al., 1999; Tasche, 2000; Bertsimas et al., 2004). These measures are typically defined in terms of the final value of the portfolio $\iota' r_p$.[22] Thus, skewness can be represented by $\mathrm{E}\left\{\left(\iota' r_p - \iota'\mu_p\right)^3\right\}$ where $\mu_p = E(r_p)$, and semi-variance is represented by $\mathrm{E}\left\{\left(\iota' r_p - \iota'\mu_p\right)^2 1\left(\iota' r_p < x\right)\right\}$ for some x (typically $\mathrm{x} = 0$). Value-at-risk (VaR) is $\Phi_{\iota' r_p}^{-1}(q)$ for some q, $0 < q < 1$ (typically $q = 0.05$), where $\Phi_{\iota' r_p}$ is the cumulative distribution function of $\iota' r_p$. Expected shortfall is $\mathrm{E}\left[\iota' r_p 1\left(\iota' r_p < x\right)\right]$ for some x (typically $\mathrm{x} = 0$.).

It turns out that the co-measures associated with semi-variance and expected shortfall (let us call them "co-semivariance"[23] and "expected co-shortfall") share many properties with expected co-drawdown. Below we collect the properties of these two co-measures. Fact 3 is a consequence of the theorem of Hong and Warren (1972). Fact 4 is comparable to Lemma 5.6 of Tasche (2000) and Proposition 3 of Bertsimas et al. (2004).

Fact 3. (Properties of Co-semivariance) Let $\mathrm{X}(r_p)$ be the semivariance of portfolio, i.e. $\mathrm{X}(r_p) = \mathrm{E}\left[\left(\iota' r_p\right)^2 1\left(\iota' r_p < 0\right)\right]$. Let us call $\partial\mathrm{X}(r_p)/\partial w_i$ co-semivariance. Suppose that $\mathrm{E}(|r_{it}|) < \infty$ for all $\mathrm{i} = 1, \cdots, \mathrm{N}$ and $\mathrm{t} = 1, \cdots, \mathrm{T}$. Then the following statements are true. (i) Co-semivariance is homogeneous of degree one in weights, (ii) co-semivariance is linear in individual

[22] It is possible to define skewness and semi-variance in terms of each elements of r_p, i.e. the expected value of sample semi-variance and the expected value of sample skewness. Whether we define these measures one way or another does not influence our discussion in this paper.

[23] As has been noted before, Hong and Warren (1974) have introduced the term "co-seminvariance."

asset returns. (iii) Co-semivariance is beta-dependent in the sense of Fact 1. (iv) Co-semivariance is downside beta-sensitive in the sense of Fact 2.[24]

Proof. Hong and Warren (1972) show that co-semivariance is $2\mathrm{E}\left[\iota' r_p \iota' r_i 1\left(\iota' r_p < 0\right)\right]$. The homogeneity and the linearity of co-shortfall (i) and (ii) are immediate. The beta dependence (iii) and the downside beta dependence (iv) can be checked in the same way as in Fact 1 and Fact 2. □

Fact 4. (Properties of Expected Co-Shortfall) Let $\mathrm{X}\left(r_p\right)$ be the expected shortfall of portfolio, i.e. $\mathrm{X}\left(r_p\right) = \mathrm{E}\left[\iota' r_p 1\left(\iota' r_p < 0\right)\right]$. We call $\partial \mathrm{X}(r_p)/\partial w_i$ expected co-shortfall. Suppose that $\mathrm{E}(|r_{it}|) < \infty$ for all $\mathrm{i} = 1, \cdots, \mathrm{N}$ and $\mathrm{t} = 1, \cdots, \mathrm{T}$. Then the following statements are true. (i) Expected co-shortfall is homogeneous of degree one in weights, (ii) Expected co-shortfall is linear in individual asset returns. (iii) Expected co-shortfall is beta-dependent in the sense of Fact 1. (iv) Expected co-shortfall is downside beta-sensitive in the sense of Fact 2.[25]

Proof. We first show that $\partial \mathrm{X}(r_p)/\partial w_i = \mathrm{E}\left[\iota' r_i 1\left(\iota' r_p < 0\right)\right]$, from which all the statements follow. By the definition of partial derivative,

$$\frac{\partial \mathrm{X}\left(r_p\right)}{\partial w_i} = \lim_{\Delta \to 0} \frac{1}{\Delta} \mathrm{E}\left[\left(\iota' r_p + \Delta \iota' r_i\right) 1\left(\iota' r_p + \Delta \iota' r_i < 0\right) - \iota' r_p 1\left(\iota' r_p < 0\right)\right]$$

We add and subtract $\iota' r_p 1\left(\iota' r_p + \Delta \iota' r_i < 0\right)$ so that, inside the expectation operator, we have the sum of $\left(\iota' r_p + \Delta \iota' r_i\right) 1\left(\iota' r_p + \Delta \iota' r_i < 0\right) - \iota' r_p 1\left(\iota' r_p + \Delta \iota' r_i < 0\right)$ and $\iota' r_p 1\left(\iota' r_p + \Delta \iota' r_i < 0\right) - \iota' r_p 1\left(\iota' r_p < 0\right)$. The first of these simplifies to $\Delta \iota' r_i \left(\iota' r_p + \Delta \iota' r_i < 0\right)$. The second of these can be written as $\iota' r_p 1(A) - \iota' r_p 1(B)$, where $\mathrm{A} = \{r_p : \iota' r_p + \Delta \iota' r_i < 0 \text{ and} \iota' r_p \geq 0\}$ and $B = \{r_p : \iota' r_p + \Delta \iota' r_i \geq 0 \text{ and } \iota' r_p < 0\}$. Thus,

$$\frac{\partial \mathrm{X}\left(r_p\right)}{\partial w_i} = \lim_{\Delta \to 0} E\left[\iota' r_i \left(\iota' r_p + \Delta \iota' r_i < 0\right) + \frac{1}{\Delta} \iota' r_p 1(A) - \frac{1}{\Delta} \iota' r_p 1(B)\right]$$

On A, $0 \leq \iota' r_p < -\Delta \iota' r_i$, and $0 \leq \lim \mathrm{E}\left[\iota' r_p 1(A)/\Delta\right] \leq \lim \mathrm{E}[-\iota' r_i 1(A)] = 0$. (The last equality is guaranteed by Lebesque's dominated convergence theorem. Note that $\lim 1(A) = 0$ at each point.) On B, $-\Delta \iota' r_i \leq \iota' r_p < 0$, and $0 \geq \lim \mathrm{E}\left[\iota' r_p 1(B)/\Delta\right] \geq \lim \mathrm{E}[-\iota' r_i 1(B)] = 0$. (Again, the

[24] We can work with a more general formula for semivariance, i.e. $\mathrm{E}\left[\left(\iota' r_p - x\right)^2 1\left(\iota' r_p < x\right)\right]$ for some *x*. In that case, however, we have to give up the homogeneity of degree one in weights.

[25] We can work with a more general formula for expected shortfall, i.e. $\mathrm{E}\left[\iota' r_p 1\left(\iota' r_p < x\right)\right]$ for some *x*. In that case, however, we have to give up the homogeneity of degree one in weights.

last equality follows from Lebesque's dominated convergence theorem.) Finally, $\lim \mathrm{E}[\iota' r_i(\iota' r_p + \Delta\iota' r_i < 0)] = \mathrm{E}[\iota' r_i(\iota' r_p < 0)]$. Thus, we have shown that $\partial \mathrm{X}(r_p)/\partial w_i = \mathrm{E}[\iota' r_i 1(\iota' r_p < 0)]$. The homogeneity and the linearity of co-shortfall (i) and (ii) are immediate. The beta dependence (iii) and the downside beta dependence (iv) can be checked in the same way as in Fact 1 and Fact 2. □

Expected co-drawdown and expected co-shortfall share all these properties because both of them are the expectations of asset returns conditional on poor market performance. Their difference lies in how "poor market performance" is defined. For expected co-drawdown, poor market performance means that the market portfolio is experiencing the worst drawdown. For expected co-shortfall, poor market performance means that the cumulative return is negative. In this sense, these two measures are complementary to each other.

Lemma 2, Fact 3, and Fact 4 are based on the idea that certain conditioning arguments do not interfere with the application of the "chain rule." In Lemma 2, the conditioning argument is that the portfolio is experiencing maximum drawdown. In Fact 3 and Fact 4, the conditioning argument is that the portfolio has a negative return. Fact 3 shows that the chain rule works even for the second moment.

The co-measures associated with other downside risk measures do not share as many properties. Co-skewness is linear in individual asset returns, but not homogeneous of degree one in weights. Nor is it downside beta-dependent. Co-VaR is linear only when the distribution is normal. We state these properties more formally below.

Fact 5. (Properties of Co-skewness) Let $\mathrm{X}(r_p)$ be the skewness of portfolio, i.e. $\mathrm{X}(r_p) = \mathrm{E}\left[(\iota' r_p - \iota' \mu_p)^3\right]$. We call $\partial \mathrm{X}(r_p)/\partial w_i$ co-skewness. Then the following statements are true. (i) Co-skewness is linear in individual asset returns. (ii) Co-skewness is not homogeneous of degree one in weights. (iii) Co-skewness is beta-dependent. (iv) Co-skewness is not downside beta sensitive.

Proof. Note that $\partial \mathrm{X}(r_p)/\partial w_i = 3\mathrm{E}\left[(\iota' r_p - \iota' \mu_p)^2\right](\iota' r_i - \iota' \mu_i)$, from which the linearity (i) is immediate. It is also clear that co-skewness is not homogeneous of degree one in weights. Beta-dependence (iii) can be proven easily noting that ε has zero conditional expectation. Since co-skewness is symmetric, downside beta sensitivity (iv) does not obtain. □

Fact 6. (Properties of Co-VaR) Let $\mathrm{X}(r_p)$ be the VaR of portfolio, i.e. $\mathrm{X}(r_p) = \Phi^{-1}_{\iota' r_p}(q)$ for $q = 0.05$. We call $\partial \mathrm{X}(r_p)/\partial w_i$ co-VaR. If returns have a multivariate normal distribution, then co-VaR is linear in individual asset returns and also homogeneous of degree one in weights.

Proof. If returns have multivariate normal distribution, then VaR is $\iota'\mu_p - k\sqrt{w'\Sigma w}$, where k is approximately 1.96. Then co-VaR is $\iota'\mu_i - k\mathrm{cov}(r_p, r_i)/\sqrt{w'\Sigma w}$, from which linearity in returns and homogeneity in weights follow immediately. □

6. Concluding Remarks

We have shown that holding period risk and rank dependence may lead investors to consider expected maximum drawdown as a measure of portfolio risk. We have also shown that the systematic risk of an individual asset is proportional to its co-drawdown—the contribution of the asset to market maximum drawdown—as well as its market beta. As a downside risk measure, co-drawdown shares many properties with other downside risk measures such as co-semivariance and expected co-shortfall.

We have not carried out an empirical analysis of co-drawdown. One challenge is whether co-drawdown estimates from historical data would be reliable. Certainly, a proper estimation of co-drawdown requires a larger amount of data than the estimation of other downside risk measures. This requirement can be probably met by using a relatively higher frequency data set, say daily or weekly rather than monthly.

Ccareful empirical analysis allows to pursue a number of interesting questions. Given the link between liquidity preference and co-drawdown, one may examine whether the liquidity premium in the equity market is related to co-drawdown. Also, given the role of rank dependence in deriving co-drawdown, one may consider a possible link to ambiguity premium. There is also a possibility that equity value premium is related to co-drawdown, either directly or indirectly through liquidity or ambiguity.

Reference

Afonja, Biyi (1972), “The Moments of the Maximum of Correlated Normal and t-Variates,” *Journal of the Royal Statistical Society Series B*, 34(2), 251-262.

Alexander, Gordon J. and Alexandre M. Baptista (2006), “Portfolio selection with a drawdown constraint,” *Journal of Banking and Finance*, 30, 3171-3189.

Allais, Maurice (1979), “The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School,” M. Allais and O. Hagen eds. *Expected Utility and the Allais Paradox*, 27-145 [Originally published in French in 1953].

Amihud, Yakov (2002), “Illiquidity and Stock Returns: Cross-section and Time-series Effects,” *Journal of Financial Markets*, 5, 31-56.

Anscombe, F. J., and R. J. Aumann (1963) “A Definition of Subjective Probability,” *Annals of Mathematical Statistics*, 34(1), 199-205.

Acharya, Viral V., and Lasse Heje Pedersen (2005), “Asset Pricing with Liquidity Risk,” *Journal of Financial Economics*, 77, 375-410.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath (1999), “Coherent Measures of Risk,” *Mathematical Finance*, 9(3), 203-228.

Bertsimas, Dimitris, Geoffrey J. Lauprete, and Alexander Samarov (2004), “Shortfall as a Risk Measure: Properties, Optimization and Applications,” *Journal of Economic Dynamics and Control*, 28, 1353-1381.

Boyle, Phelim, Lorenzo Garlappi, Raman Uppal, and Tan Wang (2012), “Keynes meets Markowitz: the Trade-off Between Familiarity and Diversification,” *Management Science*, 58(2), 253-272.

Chekhlov, Alexei, Stanislav Uryasev, and Michael Zabarankin (2005), “Drawdown Measure in Portfolio Optimization,” *International Journal of Theoretical and Applied Finance*, 8(1), 13-58.

Chen, Zengjing, and Larry Epstein (2002), “Ambiguity, Risk, and Asset Returns in Continuous Time,” *Econometrica*, 70(4), 1403-1443.

Dow, James, and Sergio Ribeiro da Costa Werlang (1992), “Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio,” *Econometrica*, 60(1), 197-204.

Ellsberg, Daniel (1961), “Risk, Ambiguity, and the Savage Axioms,” *Quarterly Journal of Economics*, 75, 643-669.

Epstein, Larry G. (1999), “A Definition of Uncertainty Aversion,” *Review of Economic Studies*, 66, 579-608

Epstein, Larry G., and Tan Wang (1994), “ Intertemporal Asset Pricing under Knightian Uncertainty,” *Econometrica*, 62(2), 283-322.

Garlappi, Lorenzo, Raman Uppal, and Tan Wang (2007)), "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach," *Review of Financial Studies*, 20(1), 41-81.

Ghirardato, Paolo (1997), "On Independence for Non-Additive Measures, with a Fubini Theorem," *Journal of Economic Theory*, 73, 261-291.

Ghirardao, Paolo, and Massimo Marinacci (2002), "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251-289.

Gollier, Christian (2011), "Portfolio Choices and Asset Prices: The Comparative Statics of Ambiguity Aversion," *Review of Economic Studies*, 78, 1329-1344.

Grossman, Sanford J. and Zhongquan Zhou (1993), "Optimal Investment Strategies for Controlling Drawdowns," *Mathematical Finance* 3(3), 241-276.

Harlow, W. V., and Ramesh H. S. Rao (1989), "Asset Pricing in a Generalized Mean-Lower Martial Moment Framework: Theory and Evidence," *Journal of Financial and Quantitative Analysis*, 24(3), 285-311.

Harding, David, Georgia Nakou, and Ali Nejjir (2003), "The Pros and Cons of Drawdown as a Statistical measure of Risk for Investments," *AIMA Journal*, April, 16-17.

Hogan, William W., and James M. Warren (1972), "Computation of the Efficient Boundary in the E-S Portfolio Selection Model," *Journal of Financial and Quantitative Analysis*, 7(4), 1881-1896.

---------- (1974), "Toward the Development of an Equilibrium Capital-Market Model Based on Semivariance," *Journal of Financial and Quantitative Analysis*, 9(1), 1-11.

Hill, Joanne M. (2009), "A Perspective on Liquidity Risk and Horizon Uncertainty," *Journal of Portfolio Management*, Summer, 60-68.

Keynes, John Maynard (1964), *The General Theory of Employment, Interest, and Money*, New York: Harcourt Brace Jovanovich. [Originally published in 1936.]

Kogan, Leonid, and Tan Wang (2003) , "A Simple Theory of Asset Pricing Under Model Uncertainty," Working Paper.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji (2005), "A Smooth Model of Decision Making under Uncertainty," *Econometrica*, 73(6), 1849-1892

Kraus, Alan, and Robert H. Litzenberger (1976), "Skewness preference and the Valuation of Risk Assets," *Journal of Finance*, 31(4), 1085-1100

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini (2006), "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74(6), 2006, 1447-1498.

Maccheroni, Fabio, Massimo Marinacci, Aldo Rustichini, and Marco Taboga (2009), "Portfolio Selection with Monotone Mean-Variance Preferences," *Mathematical Finance*, 19(3), 487-521.

Machina, Mark J., and David Schmeidler (1992), "A More Robust Definition of Subjective Probability," *Econometrica*, 60(4), 745-780.

Magdon-Ismail, M., A. F. Atiya, A. Pratap, and Y. S. Abu-Mostafa (2004), "On the Maximum Drawdown of a Browninan Motion," *Journal of Applied Probability*, 41, 147-161.

Markowitz, Harry (1952), "Portfolio Selection," *Journal of Finance*, 7(1), 77-91.

Pastor, Lubos, and Robert F. Stambaugh (2003), "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 111(3), 642-685.

Pospisil, Libor, and Jan Vecer (2008), "PDE Methods for the Maximum Drawdown," *Journal of Computational Finance*, 12, 59-76.

Quiggin, John (1982), "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323-343.

Rebonato, Riccardo, and Valerio Gaspari (2006), "Analysis of Drawdowns and Drawups in the US$ Interest-rate Market," *Quantitative Finance*, 6(4), 297-326.

Rockafellar, R. Tyrrell, and Stanislav Uryasev (2002), "Conditional Value-at-Risk for General Loss Distributions," *Journal of Banking and Finance*, 26, 1443-1471.

Savage, Leonard J. (1972), *The Foundations of Statistics*, 2nd ed., New York: Dover Publications, Inc. [The first edition was published in 1954.]

Schmeidler, David (1989), "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57(3), 571-587

Tasche, Dirk (2000), "Risk Contributions and Performance Measurement," Working Paper.

Wakker, Peter P. (2008), "Uncertainty," *New Palgrave Dictionary of Economics*, New York: Palgrave Macmillan.

Wheeden, Richard L., and Antoni Zygmund (1977), *Measure and Integral*, New York: Marcel Dekker, Inc.