Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance
Author(s): Moshe Leshno and Haim Levy
Source: *Management Science*, Vol. 48, No. 8 (Aug., 2002), pp. 1074-1085
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/822676
Accessed: 08-06-2015 08:14 UTC

Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance

Moshe Leshno • Haim Levy
School of Business Administration, The Hebrew University of Jerusalem, Jerusalem 91905, Israel
leshnom@post.tau.ac.il • mshlevy@mscc.huji.ac.il

While "most" decision makers may prefer one uncertain prospect over another, stochastic dominance rules as well as other investment criteria, will not reveal this preference due to some extreme utility functions in the case of even a very small violation of these rules. Such strict rules relate to "all" utility functions in a given class including extreme ones which presumably rarely represents investors' preference. In this paper we establish almost stochastic dominance (ASD) rules which formally reveal a preference for "most" decision makers, but not for "all" of them. The ASD rules reveal that choices which probably conform with "most" decision makers also solve some debates, e.g., showing, as practitioners claim, an ASD preference for a higher proportion of stocks in the portfolio as the investment horizon increases, a conclusion which is not implied by the well-known stochastic dominance rules.
(*Stochastic Dominance; Almost Stochastic Dominance; Mean-Variance; Risk Aversion*)

1. Introduction

Employing investment decision-making rules under uncertainty, e.g., the Markovitz (1952) mean-variance (MV) rule or stochastic dominance (SD) rules[1] may fail to show dominance in cases where almost everyone would prefer one gamble to another. For example, consider a case where a decision maker (DM) is requested to choose from two prospects X and Y where $E(X) = 10^6 > E(Y) = 1$, and $\sigma_X = 10.1 > \sigma_Y = 10$. Hence, neither X nor Y dominates the other by MV rule, while we doubt that even one person out of a very large population of investors would select Y. Thus, the MV rule cannot distinguish between the two prospects although it seems that "most" DM would prefer X over Y. Technically, there is a utility function revealing a preference of Y. However, we assume in this paper that any utility function showing a preference for Y would be considered by most, if not all, DM as an extreme one.

Let us now illustrate the same issue with SD rules. Suppose that you have to choose between prospect X and prospect Y, where X yields 0 with a probability of 0.01 and a million dollars with a probability of 0.99, and Y yields one dollar with certainty. It would not be odd if in any sample of DM one takes, 100% of the subjects would choose X. Yet, SD rules (and the MV rule alike), which relate to "all" utility functions in a given class, reveal that neither X nor Y dominates the other. The reason for this SD result is that there is a utility function revealing a preference for Y even in this case. However, it seems to us that in practice *all* subjects would prefer X to Y (would you?), but SD and MV rules which relate to all subjects are unable to reveal this preference. The inability to reveal a preference in this case is induced by the fact that stochastic dominance rules (and the mean-variance rule) relate to "all" utility functions in a given class, including "extreme" utility functions which formally

[1] The detailed development of SD paradigm is given in the survey of Levy (1992, 1998).

Management Science © 2002 INFORMS
Vol. 48, No. 8, August 2002 pp. 1074–1085

0025-1909/02/4808/1074$5.00
1526-5501 electronic ISSN

are included in a given class of utility functions, but probably do not characterize the preference of any investor. For example, take the utility function such that for $x \leq 1$, $u(x) = x$, and for $x > 1$, $u(x) = 1$. This utility function belongs to the risk-averse class, and it is easy to show that investors with this preference would prefer Y over X. As we believe any sample of subjects we take would reveal 100% preference for X over Y, we claim that the above utility function is "extreme" and does not represent most, if any, investors, hence should be ruled out. It is irrelevant because it assumes that the marginal utility beyond $x = 1$ is zero; i.e., the investor is indifferent whether the higher outcome results of X are \$1, \$$10^6$, or for that matter a billion dollars. Whether an investor with such a utility exists or not, we will never know. If such an investor exists, he is not included in our group of investors, which we call "most" DM.

The inability of the MV rule to distinguish between two options, when it seems that such a preference exists, is not new. Baumol (1963), who realizes such possible scenarios, suggests another criterion called "Expected Gain-Confidence Limit Criterion" as a substitute to the MV decision rules. Baumol argued that an investment with relatively high standard deviation σ will be relatively safe if its expected value μ is sufficiently high. He therefore proposes the following risk index: $\mathrm{RI} = \mu - k\sigma$, where k is some constant selected by the investor representing his/her safety requirement such that the return is unlikely to fall below it.

In this paper we will establish the modified stochastic dominance rules to show how to obtain decisions which reveal a preference for X over Y in the above two examples. Our new rules relate to a subclass of the class of utilities to which the existing stochastic dominance rules relate. We develop *Almost* Stochastic Dominance rules (ASD). With these rules it is possible that the distributions of the return on two prospects under comparison do not obey any SD preference, but with "a small change" SD rules reveal a preference. Thus with ASD, X dominates Y (in the example above). A small change is made in the cumulative distribution such that it eliminates the possibility of a preference by extreme utility functions for what we believe is considered by most investors to be the inferior prospect. These rules are called ASD because they are appropriate for "most" investors in a given class of utility functions. Thus, these rules are appropriate for all utility functions after deleting "extreme" preferences like the utility function given above with $u'(x) = 0$ for $x > 1$. The advantages of ASD over SD and over the mean-variance rule are:

(a) ASD are able to rank otherwise unrankable alternatives.

(b) ASD eliminate from the SD efficient set alternatives which seem to be inferior by "most" investors.

(c) ASD sheds light on the debate related to optimal portfolio composition and the planned investor horizon. It is possible to establish a functional relationship between the percentage of equity in the portfolio and the planned investor's horizon. Namely, ASD may be employed by financial advisors in choosing portfolios for "young" versus "old" investors.

The paper is organized as follows. Section 2 presents the background and motivation for the paper. Section 3 contains the main results, followed by §4 with the concluding remarks.

2. Motivation and Background

We first define SD rules and demonstrate intuitively the concept of ASD with two examples. We show that in these two examples there is no stochastic dominance relationship. Yet, it seems that "most" DM prefer one prospect over the other. We say "it seems," as we leave it to the reader to judge our assertion. In the examples below, we relate to First-Degree Stochastic Dominance (FSD) and Second-Degree Stochastic Dominance (SSD). Therefore, let us first define these decision rules. Let X and Y be two random variables, and F and G denote the cumulative distribution functions of X and Y, respectively.

1. *FSD.* F dominates G by FSD ($F \succeq_1 G$) if $F(t) \leq G(t)$ for all $t \in R$ and a strict inequality holds for at least some t. $F \succeq_1 G$ iff $E_F u(X) \geq E_G u(X)$ for all $u \in \mathcal{U}_1$, where $\mathcal{U}_1$ is the set of all nondecreasing differentiable real-valued functions.

2. *SSD.* F dominates G by SSD ($F \succeq_2 G$) if $\int_{-\infty}^{x} [G(t) - F(t)]\, dt \geq 0$, for all $x \in R$ and a strict inequality holds for at least some x. $F \succeq_2 G$ iff $E_F u(X) \geq E_G u(X)$ for all $u \in \mathcal{U}_2$, where $\mathcal{U}_2$ is the set of all nondecreasing real-valued functions such that $u'' \leq 0$.

For proofs of SD rules see Fishburn (1964), Hanoch and Levy (1969), Hadar and Russell (1969), Rothschild and Stiglitz (1970). For a survey of SD rules and further analysis see Levy (1992, 1998).

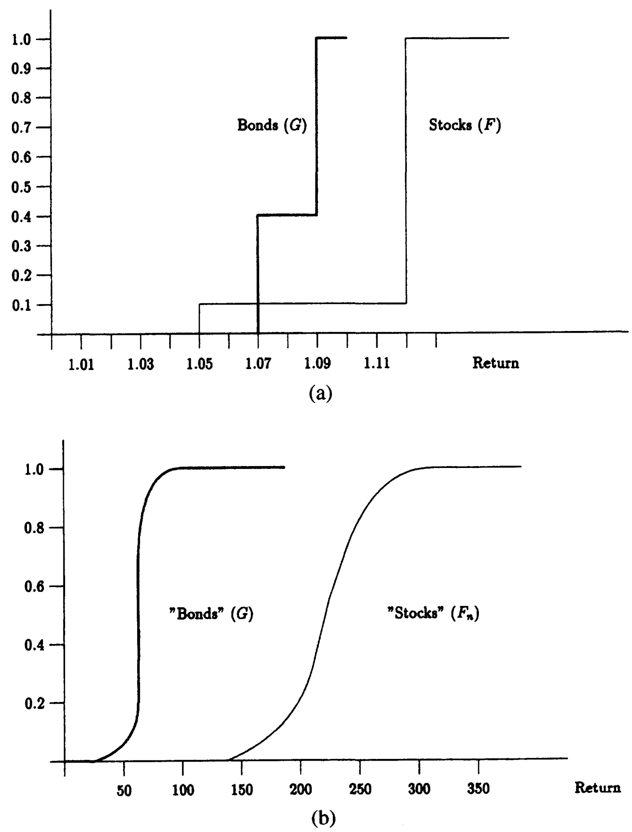
EXAMPLE 1: ALMOST FSD. In this example we focus on the relation between one-period and multiperiod rates of return on two investments. One of the critical issues in finance is the ongoing debate between practitioners and academics regarding the optimal mix of bonds and stocks as a function of the planned investment horizon. Practitioners claim that for long-run investors, equity dominates bonds (or that the proportion of equity in the portfolio should be increased) because the risk of stocks tends to cancel out over many investment periods (cross-time diversification, see, for example, Bernstein 1986). However, there is a contrasting view, argued mainly by academics, which asserts that no matter how long the investment horizon, equity will not dominate bonds (Samuelson 1994). Relying on strict mathematical arguments, Samuelson is correct in this debate. We demonstrate this with the following example and then suggest the ASD rule to solve this debate, showing that while academics are correct—that equity will never dominate bonds for *"all"* investors (no matter how long the horizon)—practitioners are correct that such dominance exists for "most" investors. Moreover, as the investment horizon increases, the group of "most" investors revealing the preference for equity as practitioners claim, increases.

We compare below two portfolios X and Y where X has a higher proportion of stocks than Y, and the rest is invested in a less risky asset called bonds. However, for simplicity of the presentation below we refer to stocks by X and to bonds by Y, recalling that both X and Y are portfolios of bonds and stocks, but X has a higher component of stocks. Thus, we consider two assets X (stocks) and Y (bonds), with cumulative distribution functions of F and G, respectively. The one-period return of X and Y is given by $X^{(1)} = 1+X$ and $Y^{(1)} = 1+Y$, respectively, when X and Y stand for *rates of return*. Take the following example:

Rate of Return: z	5%	7%	9%	12%
$Pr(X = z)$ (Stocks)	0.1	0	0	0.9
$Pr(Y = z)$ (Bonds)	0	0.4	0.6	0

The return after n periods is $X^{(n)} = \prod_{i=1}^{n}(1+X_i)$ and $Y^{(n)} = \prod_{i=1}^{n}(1+Y_i)$, respectively. Assume that the returns are independent and identical over time. For any n, the cumulative distribution of $X^{(n)}$ starts to the left of the cumulative distribution of $Y^{(n)}$, hence $X^{(n)}$ never dominates $Y^{(n)}$ by FSD. Indeed, there are some utility functions revealing a preference for $Y^{(n)}$ regardless of n, hence $X^{(n)}$ will not dominate $Y^{(n)}$ by FSD no matter how large n is. However, although $X^{(n)}$ does not dominate $Y^{(n)}$ by FSD, it can be easily shown that in the above specific example, as n increases the cumulative distribution of the stocks is shifted to the right much faster than the cumulative distribution of the bonds, hence the attractiveness of X relative to Y increases with an increase in n, which conforms with the practitioners claim. Consider Figures 1(a) and (b), which demonstrate the cumulative distributions for $n = 1$ and $n = 50$. Note that for $n = 50$, the

Figure 1 **(a) The Cumulative Distributions of F ("Stocks") and G ("Bonds") for One Period ($n = 1$)**
(b) F_n ("Stocks") and G_n ("Bonds") After $n = 50$ Periods



two distributions still intersect but the range where F_n is above G_n approaches zero, hence not shown in Figure 1(b). Thus, it seems that the attractiveness of X increases as n increases because the "negative" area enclosed between the two cumulative distributions decreases relative to the positive area. Moreover, $Pr(X^{(n)} > Y^{(n)}) \to 1$ as $n \to \infty$, yet we will never have FSD of F over G. However, it is worth mentioning that despite the fact that $Pr(X^{(n)} > Y^{(n)}) \to 1$ as $n \to \infty$, for a preference with a constant relative risk aversion (e.g., $u(x) = x^\alpha/\alpha$, $\alpha < 1$), if X is preferred to Y in the above example for $n = 1$, such a preference exists for all $n > 1$ (see Samuelson 1994). We believe that this utility function is "extreme." In our specific example, if Y is preferred for $n = 1$ it assigns a very low marginal utility for relatively high outcomes, hence the superiority of X over Y for large n is not revealed, despite the fact that the cumulative distribution of X is located almost entirely to the right of the cumulative distribution of Y. Whether a utility function revealing a preference of $Y^{(n)}$ over $X^{(n)}$ (a constant relative risk aversion), when $Pr(X^{(n)} > Y^{(n)}) \to 1$ as $n \to \infty$, should be considered a reasonable or an "extreme" function is rather subjective.[2]

In this paper we suggest a decision criterion such that for most nondecreasing utility functions u one can show that there exists an integer n such that $E_{X^{(n)}}(u) > E_{Y^{(n)}}(u)$ for "most" investors, i.e., "almost" FSD (AFSD) exists for the long-run investors (n periods, where n is large enough). Thus, we will show that for long investment horizons, "most" investors with a nondecreasing utility function would prefer to invest in X (more in equity) rather than in Y (more in bonds), as practitioners claim.[3] We shall show that the larger n is, the larger the group of "most" investors and it becomes closer to the group of "all" investors, which explains the practitioners' assertion that, in the long run, stocks dominate bonds.[4] Yet, for some "extreme" utility functions no matter how large n is, no dominance is revealed because this extreme utility function assigns an extremely high utility value to an area which is close to zero, or an extremely low utility value to high outcomes. Therefore, our AFSD relates to a class of nondecreasing utility functions, excluding the extreme ones. This will be shown later in the paper. Thus, AFSD is useful as it can be used for portfolio recommendation where the larger the planned investment horizon, the larger the equity component in the portfolio.

EXAMPLE 2: ALMOST SSD. Denote by F and G the cumulative distributions of X and Y, respectively. Suppose that X gets the values $9, 15, 30, 10^6$ with an equal probability and Y gets the values $5, 20, 25, 30$ with an equal probability. Figure 2 demonstrates F and G. The two cumulative distributions intersect, hence there is no FSD. It is easy to see also that there is no SSD because $\int_{-\infty}^{20}[G(t) - F(t)]\,dt = 1 - 1\frac{1}{4} = -\frac{1}{4} < 0$ (and of course also G does dominate F because F has a greater mean). Note that F does not dominate G by SSD because the second negative area is larger than the first positive area. Thus the $-\frac{1}{4}$ difference rules out a possible SSD no matter what the "positive" sizes of the following areas are. If we change the outcome 10^6 to 10^{10}, F still does not dominate G by SSD because U_2 includes an extreme utility function which assigns a relatively little weight to large outcomes, i.e., a utility function $u(x) \in U_2$ where for $x > 20$, $u' = 0$ is possible.

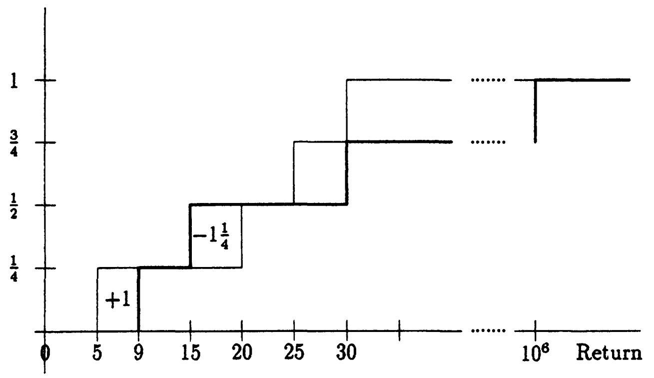
It seems to us, facing the above two prospects, that most investors would prefer X over Y. We do not claim that an extreme utility showing a preference for

[2] Consider $u(x) = 2x^{1/2}$, with $u'(x) = 1/x^{1/2}$; then for $x = 10{,}000$ the marginal utility is $1/100 = 0.01$, hence the right tail of the distribution of X is assigned only relative small additional utility relative to the right tail of Y.

[3] Benartzi and Thaler (1999) study experimentally the one-period ("one play") and multiperiod ("multiple plays") choices. They find that when the subjects observe the multiple play distributions they tend to accept the multiple (risky) play than they do the one-play case. Hence, for multiple play there is more preference to play than in the one play, i.e., in the specific case of our study to invest in stocks rather than bonds where the multiple play and one play stand for the multiperiod and one-period choices in our case.

[4] Levy (1996) showed that when riskless borrowing and lending are allowed, and the investors choose their portfolios by the mean-variance rule with no portfolio revisions, the opposite holds; i.e. as n increases, specialization in the asset with the lowest mean is obtained. However, here we do not allow the riskless asset, but simply compare between X and Y. Moreover, we do not rely here on the mean-variance rule but employ the maximum expected utility criterion.

Figure 2 **The Cumulative Distribution of F and G in Example 2**



Y does not exist. But we call such a preference an "extreme," hence, we rule it out. Thus, we want to conduct a modification in the "$-\frac{1}{4}$ area" such that a preference of F over G will be revealed. If the "$-\frac{1}{4}$ area" follows only by, say, a "$+\frac{1}{2}$ area," there is no dominance of F over G and we would like to keep it this way, because for some nonextreme utility function, indeed G is preferred over F. However, if F is like the one given in Figure 2, we would like to conduct a modification in the "$-\frac{1}{4}$ area" to reveal the superiority of F over G. Later on we see that the modification in the area is given by $\varepsilon\,(\varepsilon > 0)$ times the total area between F and G. Note, however, that ε is determined subjectively. It would be of interest to analyze it experimentally. It is possible to present a group of subjects with the distributions of F and G where there is no dominance of F over G due to a slight nagging negative area, as in Example 2. One can then increase the highest outcome of X gradually and check for various values the percentage of subjects selecting X. Such an experiment may indicate what is the relevant size of ε. More specifically, we can check what ε should be such that, say, $p = 95\%$ of choices are X. While for a relatively low $X_{\max}$, e.g., $X_{\max} = 50$, p may be relatively low, an increase in $X_{\max}$ gradually will be done until $p = 95\%$ will be obtained. For this specific $X_{\max}\varepsilon$ is calculated, i.e., only a minimum modification in F is required to reveal 95% of subjects choosing X rather than Y.

Discussion. The size of the efficient set is generally reduced by imposing restrictions on the utility functions, on the distributions of returns, or on both, because with the additional restrictions unrankable prospects may be ranked. Thus, the size of the ASD-efficient sets are generally smaller than the corresponding SD-efficient sets. Referring to SSD, Meyer adds to the risk-aversion assumption a restriction on the Arrow-Pratt absolute risk-aversion measure given by $r(x) = -(u''(x)/u'(x))$, hence his rule, generally, reduces the SSD-efficient set. Meyer suggests imposing upper and lower bounds on $r(x)$ (see Meyer 1977a), a case where the results are not expressed in a closed form. In a subsequent paper, Meyer (1977b) suggests SSD with respect to a function which generalizes SSD, but here the solution is provided only for a lower bound on $r(x)$.

In developing the ASD, like Meyer we also impose restrictions on the utility functions, hence reducing the efficient sets. However, two distinctions between the two approaches should be mentioned: First, we impose restrictions on the marginal utility function, hence reducing the FSD-efficient set and therefore eliminating prospects which are considered inferior by "most" investors with nondecreasing utility functions, while Meyer deals only with SSD. Secondly, when we analyze FSD the restriction is on $u'(x)$, and when we analyze SSD the restriction is on $u''(x)$, where $u'(x) > 0$ and $u''(x) < 0$ are the basic assumptions underlying the derivation FSD and SSD, respectively. In contrast, in the generalization of SSD Meyer imposes a constraint on the ratio $-(u''(x)/u'(x))$.

Each approach has its pros and cons. The Meyer approach seems to be more intuitive because he eliminates utility functions with an extreme risk aversion. However, in his approach there is no clue of what the upper and lower bounds should be on $r(x)$ to have "reasonable" decisions. Moreover, he does not allow a utility function with a linear segment (or close to linear) because in this case $-(u''(x)/u'(x)) = 0$ will be below the lower bound. In contrast, in our approach when there is a violation of FSD, we require that $u'(x)$ will not be given an extreme weight to the area of violation, and similarly, when there is a violation of SSD, we make sure that $u''(x)$ will not be given extremely high weight to the area of violation. Moreover, the restrictions on $u'(x)$ or $u''(x)$

are a function of the violation area, relative to the total area enclosed between F and G. In brief, if the SD violation area is small relative to the total area enclosed between F and G, we make sure that this small violation area will not get an extremely high utility weight, hence we can rank the two prospects even though FSD or SSD can not rank the prospects under consideration.

It is interesting to note that the myopic utility function is eliminated by both approaches. With a myopic utility function $u(x) = (x^\alpha/\alpha)$, with $0 < \alpha < 1$, we have $-(u''(x)/u'(x)) = ((1-\alpha)/x)$. Thus, for $x \to 0, -(u''(x)/u'(x)) \to \infty$, and for $x \to \infty$, $-(u''(x)/u'(x)) \to 0$; hence, for any relevant lower and upper bounds on $r(x)$ the myopic function risk aversion measure does not fall within the bounds. Similarly in our approach, the myopic utility function is not included within U_1^* or in U_2^* (see below).

3. Stochastic Dominance and Almost Stochastic Dominance

Having shown the above two examples and the definitions of FSD and SSD, we now turn to the formal definition of ASD rules, which reveal a preference, though not for all investors in a given set of preferences (e.g., risk averters), but for "most" of them.

For F and G, two cumulative distribution functions under consideration, we assume that the distributions have a finite support, say $[a, b]\,(-\infty < a < b < \infty)$. Without loss of generality we choose $a = 0$ and $b = 1$, namely the unit interval, i.e., $F(t) = 0$ for $t < 0$ and $F(t) = 1$ for $t \geq 1$. Denote by U_1 the set of all nondecreasing differentiable utility functions ($u \in U_1$ if $u' \geq 0$) and by U_2 the set of all concave utility functions twice differentiable ($u \in U_2$ if $u' \geq 0$ and $u'' \leq 0$). We start with defining the following two families of real-valued functions, which are subsets of U_1 and U_2, presumably not including the extreme utility functions as illustrated in the above examples. For every $0 < \varepsilon < 0.5$ define:[5]

$$U_1^*(\varepsilon) = \left\{ u \in U_1\colon\ u'(x) \leq \inf\{u'(x)\}\left[\frac{1}{\varepsilon} - 1\right], \forall x \in [0, 1] \right\} \tag{1}$$

[5] Originally we had a different definition for U_1^*. We thank the referee for suggesting this equivalent and more intuitive definition.

and respectively,

$$U_2^*(\varepsilon) = \left\{ u \in U_2\colon\ -u''(x) \leq \inf\{-u''(x)\}\left[\frac{1}{\varepsilon} - 1\right], \forall x \in [0, 1] \right\}. \tag{2}$$

These types of utility functions do not assign a relatively high marginal utility to very low values or a relatively low marginal utility to large values of x. To illustrate, the function $u(x)$ given by $u(x) = x$, for $x \leq \frac{1}{2}$ and $u(x) = \frac{1}{2}$ for $x > \frac{1}{2}$ is not included in U_1^* because $\inf\{u'(x)\} = 0$. Similarly, $u(x) = \log(x)$ defined on (0,1], is not in U_1^* (and hence not in U_2^*) because for $x \to 0$, $u'(x) \to \infty$; hence $u'(x)$ can not be smaller than $\inf\{u'(x)\}$, which is achieved at $x = 1$.

Obviously, for all ε, $U_1^*(\varepsilon) \subseteq U_1$ and $U_2^*(\varepsilon) \subseteq U_2$. We establish in this paper conditions for dominance of F over G for all $u \in U_1^*(\varepsilon)$ and then for all $u \in U_2^*(\varepsilon)$. As ε gets smaller, the sets U_i^* grow larger, but remain a strict subset of U_i. We also will see that if almost-dominance holds for *all* $0 < \varepsilon < 0.5$, then the standard FSD or SSD holds. Note that when ε approaches 0.5 the set $U_1(\varepsilon)$ contains risk-neutral utilities only, and $U_2(\varepsilon)$ contains only the linear utility functions and the quadratic utility functions. This implies that when $\varepsilon \to 0.5$, almost-FSD approaches expected value maximization criteria, meaning F ε-almost-FSD dominates G if and only if the expected value is larger under F than under G. For the ASSD, $\varepsilon \to 0.5$ implies that $U_2^*(\varepsilon)$ contains linear and quadratic utilities only. In this case, mean-variance dominance is sufficient to establish ε-almost-SSD, but not necessary. (See Hanoch and Levy 1970.)

Let us first define some notation that will be used later on:

$$S_1(F, G) = \{t \in [0, 1]\colon G(t) < F(t)\}, \tag{3}$$

$$S_2(F, G) = \left\{ t \in S_1(F, G)\colon \int_0^t G(x)\,dx < \int_0^t F(x)\,dx \right\}, \tag{4}$$

where F and G are two cumulative distribution functions of X and Y, respectively, and X and Y represent returns on two investments under consideration. As we shall see below, Equations (3) and (4) describe the regions where corrections are needed to get stochastic dominance. For two cumulative distribution functions F and G we define $\| F - G \| = \int_0^1 |F(t) - G(t)|\,dt$.

Definition: ASD. Let X and Y be two random variables, and F and G denote the cumulative distribution functions of X and Y, respectively. For $0 < \varepsilon < 0.5$ we define:

(1) AFSD. F dominates G by ε-Almost FSD ($F \succeq_1^{\text{almost}(\varepsilon)} G$) if and only if,

$$\int_{S_1} [F(t) - G(t)]\, dt \le \varepsilon \| F - G \|. \quad (5)$$

(2) ASSD. F dominates G by ε-Almost SSD ($F \succeq_2^{\text{almost}(\varepsilon)} G$) if and only if

$$\int_{S_2} [F(t) - G(t)]\, dt \le \varepsilon \| F - G \|, \quad (6)$$

and

$$E_F(X) \ge E_G(Y). \quad (7)$$

The left sides of (5) and (6) describe the "amount of correction" needed (for the definition of S_1 and S_2 see Equations (3) and (4)). We shall show that AFSD implies that $E_F(X) \ge E_G(Y)$, but in ASSD (6) does not imply $E_F(X) \ge E_G(Y)$, and (7) is an explicit required condition for ASSD.[6]

Note that ASD requires that the nagging negative area where G is above F has to be a small fraction (ε) of the total absolute area difference between F and g.

An alternative characterization of ASD that requires a distribution to be "close to" another distribution that dominates in the traditional sense of FSD or SSD is presented in the following proposition. Thus, ASD can be equivalently defined as follows (the proofs are given in Appendix A):

Proposition 1. *Let X and Y be two random variables, and F and G denote the cumulative distribution functions of X and Y, respectively.*

(1) AFSD. F dominates G by ε-AFSD ($F \succeq_1^{\text{almost}(\varepsilon)} G$) if and only if there exists a c.d.f. $\tilde{F}$ such that $\tilde{F}$ dominates G by the FSD rule and the following holds:

$$\| F - \tilde{F} \| = \int_0^1 | F(x) - \tilde{F}(x) |\, dx \le \varepsilon \| F - G \|. \quad (8)$$

(2) ASSD. F dominates G by ε-ASSD if there exists a c.d.f. $\tilde{F}$ such that $\tilde{F}$ dominates G by SSD rule and the following holds

$$\| F - \tilde{F} \| = \int_0^1 | F(x) - \tilde{F}(x) |\, dx \le \varepsilon \| F - G \|. \quad (9)$$

Thus, the difference between F and $\tilde{F}$ should be relatively small ($0 < \varepsilon < 0.5$). By having $\varepsilon < 0.5$ we assure that it is impossible to have both F dominates G and G dominates F by AFSD, because if F dominates G by AFSD, then $E_F(X) > E_G(Y)$ (see Proposition 2 below). While $E_F(X) > E_G(Y)$ is true for AFSD, for ASSD only a weaker inequality $E_F(X) \ge E_G(Y)$ is required (see Footnote 6). Formally, we state the following proposition regarding AFSD (the proofs of these proposition is presented in Appendix B):

Proposition 2.[7] *Let X and Y be two random variables, and F and G denote the cumulative distribution functions of X and Y, respectively.*

If F dominates G by ε-AFSD and F and G are not identical, then $E_F(X) > E_G(Y)$. Therefore it is impossible that F dominates G by ε-AFSD and G dominates F by ε-AFSD.

Theorem 1 given below contains the criteria for the almost FSD and almost SSD.

We will show below that F dominates G by ε-AFSD, if and only if $E_F(u) \ge E_G(u)$ for all $u \in U_1^*(\varepsilon)$. And F dominates G by ε-ASSD if and only if $E_F(u) \ge E_G(u)$ for all $u \in U_2^*(\varepsilon)$.

Theorem 1.

(1) AFSD. Let X and Y be two random variables, and F and G denote the cumulative distribution functions of X and Y, respectively. F dominates G by ε-almost FSD ($F \succeq_1^{\text{almost}(\varepsilon)} G$) if and only if for all u in $U_1^*(\varepsilon)$, $E_F(u) \ge E_G(u)$.

(2) ASSD. Let X and Y be two random variables, and let F and G denote the cumulative distribution functions of X and Y, respectively. F dominates G by ε-ASSD ($F \succeq_2^{\text{almost}(\varepsilon)} G$), if and only if for all u in $U_2^*(\varepsilon)$, $E_F(u) \ge E_G(u)$.

The proofs of Theorem 1 are given in Appendix C.

We show next that the definition of ASD is a generalization of the definition of SD. This is stated in the following proposition.

[6] This is similar to FSD, SSD, and Third Stochastic Dominance (TSD). In FSD and SSD the integral condition for dominance implies $E_F(X) \ge E_G(Y)$. In TSD one has to explicitly add the condition $E_F(X) \ge E_G(Y)$ because it does not follow from the integral condition. Note that in AFSD, $E_F(X) > E_G(Y)$, and therefore it is impossible that F dominates G by ε-almost-FSD and G dominates F by ε-almost-FSD. However, in ASSD in the specific case where $E_F(X) = E_G(Y)$, it is possible that F dominates G by ε_1-ASSD and G dominates F by ε_2-ASSD. In such a case we select the dominance corresponding to the smaller ε_i because it relates to a wider class of utility functions U_2^*.

[7] We thank Department Editor J. Smith for his suggestions to apply ASD only to $\varepsilon < 0.5$ and for providing the proof for this proposition.

PROPOSITION 3. *Let X and Y be two random variables, and F and G denote the cumulative distribution functions of X and Y, respectively.*

1. *FSD. F dominates G by FSD ($F \succeq_1 G$) if and only if for all $0 < \varepsilon < 0.5$, F dominates G by ε-AFSD ($F \succeq_1^{\text{almost}(\varepsilon)} G$).*

2. *SSD. F dominates G by SSD ($F \succeq_2 G$) if and only if for all $0 < \varepsilon < 0.5$, F dominates G by ε-ASSD ($F \succeq_2^{\text{almost}(\varepsilon)} G$).*

As the above two claims are for all $0 < \varepsilon < 0.5$, the two rules coincide. For a formal simple proof see Appendix D.

To illustrate the ASD concept and the required "area correction" let us turn back to Example 1 given in §2. We calculated $\varepsilon^{(n)}$ as follows,

$$\varepsilon^{(n)} = \frac{\int_{S_1}[F^{(n)}(t) - G^{(n)}(t)]\,dt}{\| G^{(n)} - F^{(n)} \|},$$

where $F^{(n)}$ and $G^{(n)}$ are the distribution of $\prod_{i=1}^{n}(1+X_i)$ and $\prod_{i=1}^{n}(1+Y_i)$, respectively. The following table summarizes the results:[8]

n	$\varepsilon^{(n)}$
1	0.095
2	0.063
5	0.020
10	0.005
50	5.4×10^{-7}

We can see from this table that $\varepsilon^{(n)} \to 0$ as $n \to \infty$, hence the set $U_1^*(\varepsilon)$ expands to include "most" utility functions. We believe that this is what practitioners claim; as $n \to \infty$ "most" investors would prefer equity to bonds. The fact that $\varepsilon^{(n)}$ decreases as n increases has a practical implication. Based on historical data, one can estimate the one-period probability distributions of stocks and bonds (for the source of such data, see Ibbotson Associates 2001). The financial analyst can construct a portfolio F_α which it consists of α percent of stocks and $1-\alpha$ percent of bonds (say, for $\alpha = 0, 0.1, 0.2, \ldots, 1$). Then, one can construct the n period distribution of each of the above distributions. Then for a given ε, e.g., $\varepsilon = 0.005$, one can find the portfolios corresponding to n (horizon) which are not AFSD inferior. For example, an investment recipe may be: For $n \le 5$, invest 40% or less in bonds, for $5 < n \le 10$ invest 20% or less in bonds, etc. Such a recipe conforms with the common recommendation to increase the proportion of equity in the portfolio as the investment horizon increases.

[8] The calculations were done using *MATLAB* Software.

4. Concluding Remarks

In some cases it seems that "most" decision makers (DM) would prefer one prospect over another, yet due to a very small violation of these rules, stochastic dominance (SD) rules (as well as the mean-variance rule) are unable to distinguish between the two prospects under consideration. One can increase the mean of one distribution as far as one wishes, yet a very small violation of SD criteria induces no dominance because there is some "extreme" utility function which assigns a very large weight to the violation area or a very small weight to very large outcomes, regardless of the big difference in the means of the two distributions. Thus, the "negative" area enclosed between the two cumulative distributions may be negligible relative to the "positive" area, yet SD is not revealed. Therefore, SD rules may lead to results, asserting no dominance in spite of the fact that it seems that "most" DM would claim that dominance exists. We suggest in this paper "almost stochastic dominance rules" (ASD) which are able to reveal that "most" investors would prefer one investment over the other, in spite of the fact that such preference is not revealed by the SD rules. Thus, the ASD rules can be considered as modified SD rules after eliminating the extreme utility functions, e.g., a function which assigns the same utility to, say, \$1 and to $\$10^6$. We do not claim that such extreme preference will never exist. But because we believe they are not typical, in constructing ASD rules we rule them out. Thus, ASD rules may reveal dominance in some cases where no-dominance relationship with SD rules prevails.

The ASD rules have several advantages over the SD (or MV) rules. First, ASD rules can rank alternatives which otherwise are unrankable. Second, ASD rules reveal preference which conforms with one's intuition, while SD may reveal a counterintuitive no-dominance decision. Finally, ASD rules are useful in portfolio construction for investors with various planned investment horizons. To be more specific, ASD rules shed light on the debate among practitioners and academics regarding the dominance of stocks over bonds for long-run investors. Practitioners claim that as the investment horizon increases, a portfolio with a relatively large component of stocks will dominate a portfolio with a relatively large component of bonds, but SD rules do not conform with this assertion. Indeed, we show that though "the stocks portfolio" generally does not dominate "the bonds portfolio" by SD, for long-run investors such dominance exists by ASD. Hence, indeed "most" long horizon investors will prefer stocks over bonds. Moreover, as the investment horizon increases, the group of "most" investors becomes closer to the group of "all" investors. However, academics are correct that for myopic utility functions with all assets invested (i.e., no constant exists in the argument of the utility function), the decision whether to invest in stocks or in bonds (or in some mix of the two) is invariant to the assumed investment horizon. Thus, investors with a myopic utility function may remain in the group of "all" investors but not in the group of "most" investors no matter how long the investment horizon is. Nevertheless, we demonstrate that some of the utility functions belonging to U_1 and not to U_1^* (or to U_2 and not to U_2^*) are indeed "extreme" where extreme is defined in the spirit of the above explanation. Thus, extreme utility functions can be the main explanation for the possible debate between practitioners and academics regarding the optimal portfolio mix as a function of the investment horizon. It is interesting to note that the myopic utility function is eliminated by Meyer (1997a), who imposes conditions on the risk-aversion measure. Having resolved the above issue, the ASD can be employed in practice by portfolio managers. For a given ε one can employ ASD to find the proportion of equity to be included in the portfolio corresponding to a given planned investment horizon. A recommendation of, say, no less than 40% bonds if $n < 5$, no less than 20% bonds if $5 < n \leq 10$, is possible, which may be very useful.

To sum up, the ASD conforms with the intuition asserting that in the long run stocks are more attractive than bonds, even though not for *all* investors. There are those stubborn myopic investors who completely ignore the investment horizon, and this is the reason why we have ASD, but not SD.

Acknowledgments

The revised version of the paper has greatly benefited from the comments of three anonymous referees, an associate editor, and J. Smith, the departmental editor of the journal. The authors gratefully acknowledge the financial support provided by the Krueger Foundation at the School of Business Administration, The Hebrew University, Jerusalem.

Appendix A

We show in this appendix the equivalence of the two definitions of ASD; namely, we prove Proposition 1.

Proof of Proposition 1. The proof for the SSD follows the same arguments as the proof for the FSD case; therefore, we omit the proof for the SSD case and focus on the FSD proof.

Part (a). Assume that there exists a c.d.f. $\widetilde{F}$ such that $\widetilde{F}$ dominates G by the FSD rule and the following holds: $\| F - \widetilde{F} \| = \int_0^1 |F(x) - \widetilde{F}(x)|\, dx \leq \varepsilon \| F(x) - G(x) \|$, and $\int_0^1 F(x)\, dx = \int_0^1 \widetilde{F}(x)\, dx$. The part that prevents F from dominating G is the set $S_1(F, G)$, namely $\{t \in R : G(t) < F(t)\}$ (see Equation (1)). As $\widetilde{F}$ dominates G, we have that for all $t \in R$, $G(t) \geq \widetilde{F}(t)$, and particularly that which holds for $t \in S_1$. Thus, for $t \in S_1$, $F(t) - \widetilde{F}(t) \geq F(t) - G(t)$. Thus we have that

$$\int_{S_1} |F(x) - G(x)|\, dx \leq \| F - \widetilde{F} \| \leq \varepsilon \| F - G \|. \tag{A1}$$

Therefore, $\int_{S_1} [F(t) - G(t)]\, dt \leq \varepsilon \| F - G \|$.

Part (b). We first prove the second part of the Proposition in the case where F and G have a finite number of points of increase and then generalize the proof to the case of infinite points of increase. Thus, we prove the following claim:

Claim 1. *Assume that X and Y are discrete random variables whose c.d.f.s F and G have a finite number of points of increase. Denote $H = G - F$. If $\int_{S_1} |H(x)|\, dx < \varepsilon \| F - G \|$, then there exists a c.d.f. $\widetilde{F}$ such that for all $x \in [0, 1]$, $G(x) > \widetilde{F}(x)$, and $\| \widetilde{F} - F \| < 2\varepsilon \| F - G \|$.*

Proof of Claim 1. $H = G - F$ is a step function with a finite number of steps. Let $I = [a_1, b_1)$ be the first negative step of H. If I does not exist, then F dominates G by the FSD rule. Let $J = [a_2, b_2)$ be the next first positive step of H to I from the left or from the right of I.

Assume that $b_2 \leq a_1$; i.e., I is to the left of J. Let γ_2 be the value of H on J and let $-\gamma_1$ be the value of H on I. Either,

$$\gamma_2(b_2 - a_2) \geq \gamma_1(b_1 - a_1), \tag{A2}$$

or

$$\gamma_2(b_2 - a_2) < \gamma_1(b_1 - a_1). \tag{A3}$$

If (A2) holds, let $\hat{b}_1 = b_1$; then there is a $\hat{b}_2$, satisfying $a_2 < \hat{b}_2 \le b_2$ such that,

$$\gamma_2(\hat{b}_2 - a_1) = \gamma_1(\hat{b}_1 - a_1). \tag{A4}$$

If (A3) holds, let $\hat{b}_2 = b_2$; then there is a $\hat{b}_1$, satisfying $a_1 < \hat{b}_1 < b_1$, such that (5) holds.

Define H_1 by,

$$H_1 = \begin{cases} +\gamma_2 & \text{for } \hat{b}_2 \le z < b_2 \\ -\gamma_1 & \text{for } a_1 \le z < \hat{b}_1 \\ 0 & \text{otherwise} \end{cases}.$$

In a similar way, one constructs H_1 if I is to the left of J. Define $F_1 = F + H_1$. It is easy to see that F_1 is a c.d.f. We use this technique to construct H_2 from $G - F_1$ and define F_2 by $F_2 = F_1 + H_2$. Because H is a step function with a finite number of steps, the process terminates after a finite number of iterations, say after n iterations. Define $\widetilde{H} = \sum_{i=1}^{n} H_i$, and $\widetilde{F}(x) = F(x) + \widetilde{H}(x)$. From the construction of $\widetilde{H}$ we have that $\int_{S_1} |\widetilde{H}(x)|\,dx = \int_{S_1}(F(x) - G(x))\,dx$, and by the assumption of the claim we have $\int_{S_1} |\widetilde{H}(x)|\,dx \le \varepsilon \| F - G \|$. For all $x \in [0,1]$ we have that $G(x) \ge \widetilde{F}(x) = F(x) + \widetilde{H}(x)$ and

$$\| F - \widetilde{F} \| = \int_0^1 |F(x) - \widetilde{F}(x)|\,dx = \sum_{i=1}^{n} \int_0^1 |H_i(x)|\,dx$$
$$= 2\int_{S_1} [F(x) - G(x)]\,dx \le 2\varepsilon \| F - G \|.$$

This completes the proof of the Theorem for the case where F and G have a finite number of points of increase. For the general case we approximate F and G by discrete F_n and G_n. Specifically, let X and Y be two random variables with c.d.f.s F and G, respectively, then define

$$G_n = \begin{cases} 0 & \text{for } x \le 0 \\ G(\frac{i}{n}) & \text{for } \frac{i}{n} \le x < \frac{i+1}{n} \quad i = 0, \dots, n-1 \\ 1 & x \ge 1 \end{cases},$$

and

$$F_n = \begin{cases} 0 & \text{for } x \le 0 \\ F(\frac{i+1}{n}) & \text{for } \frac{i}{n} \le x < \frac{i+1}{n} \quad i = 0, \dots, n-1 \\ 1 & x \ge 1 \end{cases}.$$

From the way we construct $\widetilde{H}_n$ it is easy to see that $\lim_{n,m\to\infty} \| \widetilde{H}_n - \widetilde{H}_m \| = 0$; i.e., $\{\widetilde{F}_n\}$ is a Cauchy sequence and thus the sequence $\{\widetilde{F}_n\}$ has a limit, say $\lim_{n\to\infty} \widetilde{F}_n = \widetilde{F}$. We have for all n

$$\| F - \widetilde{F} \| \le \| F - F_n \| + \| F_n - \widetilde{F}_n \| + \| \widetilde{F}_n - \widetilde{F} \|.$$

As $n \to \infty$ we have $\| F - F_n \| \to 0$, and $\| \widetilde{F}_n - \widetilde{F} \| \to 0$; we also have that for n large enough $\| F_n - \widetilde{F}_n \| \le \varepsilon$, thus we get that $\| F - \widetilde{F} \| \le 2\varepsilon$. Because $\widetilde{F}_n \to \widetilde{F}$, $G_n \to G$ and for all $x \in [0,1]$, $G_n(x) \ge \widetilde{F}_n(x)$ we have that $G(x) \ge \widetilde{F}(x)$ for all $x \in [0,1]$. □

This completes the proof of the Proposition. Note that by construction of $\widetilde{F}(x)$ we also have $\int_0^1 \widetilde{F}(x)\,dx = \int_0^1 F(x)\,dx$.

Appendix B

Proof of Proposition 2. We assume that F dominates G by ε-almost FSD but does not dominate G by FSD and that F and G are not identical. Then we have

$$0 < \int_{S_1} [F(t) - G(t)]\,dt \le \varepsilon \| F - G \|$$
$$= \varepsilon \left[\int_{\bar{S}_1} [G(t) - F(t)]\,dt - \int_{S_1} [G(t) - F(t)]\,dt \right],$$

where $0 < \varepsilon < 0.5$, or

$$0 < \varepsilon \int_{\bar{S}_1} [G(t) - F(t)]\,dt + (1 - \varepsilon) \int_{S_1} [G(t) - F(t)]\,dt,$$

or

$$0 < \int_{\bar{S}_1} [G(t) - F(t)]\,dt + \frac{(1-\varepsilon)}{\varepsilon} \int_{S_1} [G(t) - F(t)]\,dt.$$

As $\varepsilon < 0.5$, we have that $(1-\varepsilon)/\varepsilon > 1$, and because $\int_{S_1} [G(t) - F(t)]\,dt < 0$, we have that

$$0 < \int_{\bar{S}_1} [G(t) - F(t)]\,dt + \frac{(1-\varepsilon)}{\varepsilon} \int_{S_1} [G(t) - F(t)]\,dt$$
$$< \int_{S_1} [G(t) - F(t)]\,dt + \int_{\bar{S}_1} [G(t) - F(t)]\,dt$$
$$= E_F(X) - E_G(Y).$$

This completes the proof of the proposition. □

Appendix C

In this Appendix we provide a detailed proof of the first part of Theorem 1 and outline the proof of the second part of Theorem 1, which is very similar in its structure to the proof of the first part.

Proof of Theorem 1. AFSD Sufficiency. Generally for any a differentiable real-valued function u and a distribution F, we have that:

$$E_F(u) = \int_0^1 u(x)\,dF(x) = [u(x)F(x)]_0^1 - \int_0^1 u'(x)F(x)\,dx$$
$$= u(1) - \int_0^1 u'(x)F(x)\,dx.$$

Therefore,

$$E_F(u) - E_G(u) = \int_0^1 u'(x)[G(x) - F(x)]\,dx$$
$$= \int_{S_1} u'(x)[G(x) - F(x)]\,dx + \int_{\bar{S}_1} u'(x)[G(x) - F(x)]\,dx,$$

where $\bar{S}_1$ denotes the complement of S_1 in $[0,1]$. Let u be in $U_1^*(\varepsilon)$, and denote $\inf_{x\in[0,1]}\{u'(x)\} = \underline{\gamma}$ and $\sup_{x\in[0,1]}\{u'(x)\} = \bar{\gamma}$. We have

$$E_F(u) - E_G(u) = \int_{S_1} u'(x)[G(x) - F(x)]\,dx + \int_{\bar{S}_1} u'(x)[G(x) - F(x)]\,dx$$
$$\ge \bar{\gamma} \int_{S_1} [G(x) - F(x)]\,dx + \underline{\gamma} \int_{\bar{S}_1} [G(x) - F(x)]\,dx$$
$$= (\bar{\gamma} + \underline{\gamma}) \int_{S_1} [G(x) - F(x)]\,dx + \underline{\gamma} \| F - G \|.$$

Note that if $u'(x)=0$ for all x $E_F(X)=E_G(Y)$. In the rest of the proof we eliminate this trivial utility function. By the assumption that F dominates G by the ε-almost FSD we have that

$$\int_{S_1}[F(t)-G(t)]\,dt \le \varepsilon \,\|F-G\|.$$

Let $u\in U_1^*(\varepsilon)$, then by the definition of U_1^* (see Equation (1)) $\bar{\gamma}\le \underline{\gamma}[(1/\varepsilon)-1]$. Dividing by $\bar{\gamma}+\underline{\gamma}$, which is positive, yields $\varepsilon \le (\underline{\gamma}/\bar{\gamma}+\underline{\gamma})$. Therefore, we have

$$0\le \int_{S_1}[F(t)-G(t)]\,dt \le \varepsilon\,\|F-G\| \le \frac{\underline{\gamma}}{\bar{\gamma}+\underline{\gamma}}\|F-G\|.$$

Thus,

$$(\bar{\gamma}+\underline{\gamma})\int_{S_1}[G(x)-F(x)]\,dx+\underline{\gamma}\,\|F-G\|\ge 0,$$

or

$$E_F(u)-E_G(u)\ge 0.$$

Proof of Theorem 1—Necessity. Now assume that for all u in $U_1^*(\varepsilon)$, $E_F(u)\ge E_G(u)$. We next prove that F dominates G by the ε-almost FSD. It is sufficient to show that

$$\int_{S_1}[F(t)-G(t)]\,dt \le \varepsilon\,\|F-G\|.$$

Assume that $\int_{S_1}[F(t)-G(t)]\,dt > \varepsilon\,\|F-G\|$. We construct a function u in $U_1^*(\varepsilon)$ such that $E_F(u)<E_G(u)$, which contradicts the assumption that for all $u\in U_1^*$, $E_F(u)\ge E_G(u)$. Let $\bar{\gamma},\underline{\gamma}$ be positive real numbers such that $\varepsilon=(\underline{\gamma}/\bar{\gamma}+\underline{\gamma})$. With no loss of generality, we assume that S_1 is an interval and denote $S_1=[a,b]$, $C=\overline{[a,b]}$ (the complement of $[a,b]$ in $[0,1]$). Define

$$u(x)=\begin{cases}\underline{\gamma}x & \text{if } 0\le x\le a\\ \bar{\gamma}(x-a)+\underline{\gamma}a & \text{if } a\le x\le b\,.\\ \underline{\gamma}(x-b)+\bar{\gamma}b+\underline{\gamma}a & \text{if } b\le x\le 1\end{cases}$$

We have

$$\begin{aligned}E_F(u)-E_G(u) &= \int_a^b u'(x)[G(x)-F(x)]\,dx+\int_C u'(x)[G(x)-F(x)]\,dx\\ &= \bar{\gamma}\int_a^b[G(x)-F(x)]\,dx+\underline{\gamma}\int_C[G(x)-F(x)]\,dx.\end{aligned}$$

We assume that $\int_a^b[F(t)-G(t)]\,dt > (\underline{\gamma}/(\bar{\gamma}+\underline{\gamma})\,\|F-G\|$, thus we have that

$$\begin{aligned}E_F(u)-E_G(u) &= \bar{\gamma}\int_a^b[G(x)-F(x)]\,dx+\underline{\gamma}\int_C[G(x)-F(x)]\,dx\\ &= (\bar{\gamma}+\underline{\gamma})\int_a^b[G(x)-F(x)]\,dx+\underline{\gamma}\,\|F-G\|<0.\end{aligned}$$

This completes the proof of the first part of Theorem 1. □

Proof of ASSD. The proof follows from the same arguments of the proof of Theorem 1 and the following:

$$\begin{aligned}&E_F(u)-E_G(u)\\ &\quad=\int_0^1 u'(x)[G(x)-F(x)]\,dx\\ &\quad=u'(1)\int_0^1[G(t)-F(t)]\,dt+\int_0^1(-u''(x))\int_0^x[G(t)-F(t)]\,dt\,dx\\ &\quad=u'(1)\int_0^1[G(t)-F(t)]\,dt+\int_{S_2}(-u''(x))\int_0^x[G(t)-F(t)]\,dt\,dx\\ &\quad\quad+\int_{\bar{S}_2}(-u''(x))\int_0^x[G(t)-F(t)]\,dt\,dx.\end{aligned}$$

Note that to make sure that $E_F(u)\ge E_G(u)$, we have to assume that $E_F(X)\ge E_G(Y)$, which is equivalent to the condition $u'(1)\int_0^1[G(t)-F(t)]\,dt\ge 0$. □

Appendix D

Proof of Proposition 3. Let F dominate G by FSD ($F\succeq_1 G$), then for all t, $G(t)\ge F(t)$, therefore $S_1(F,G)=\{t\colon G(t)>F(t)\}=\varnothing$. Thus, we have for every $0<\varepsilon<0.5$

$$\int_{S_1}[F(t)-G(t)]\,dt=0\le\varepsilon\,\|F-G\|,$$

and F dominates G by ε-almost FSD ($F\succeq_1^{\text{almost}(\varepsilon)} G$). On the other hand, assume that for every $0<\varepsilon<0.5$, F dominates G by ε-almost FSD ($F\succeq_1^{\text{almost}(\varepsilon)} G$). If $\mu(S_1)=0$, where μ denotes the Lebesgue measure on R, then because F and G are nondecreasing and continuous from the right functions, for all t, $F(t)\le G(t)$; i.e., F dominates G by the FSD rule. If $\mu(S_1)>0$ and there is no FSD, we will show that there is also no AFSD for some $\varepsilon>0$. Denote $\int_{S_1}[F(t)-G(t)]\,dt=\varepsilon_0>0$. For $\varepsilon=(\varepsilon_0/2\,\|G-F\|)$, we have that $\int_{S_1}[F(t)-G(t)]\,dt=\varepsilon_0=2\varepsilon\,\|G-F\|>\varepsilon\|G-F\|$. In other words, F does not dominate G for any $\varepsilon>0$.

This completes the first part of Proposition 3. The proof of the second part is similar and is therefore omitted. □

References

Baumol, W. J. 1963. An expected gain in confidence limit criterion for portfolio selection. *Management Sci.* **10** 174–182.
Benartzi, S., R. H. Thaler. 1999. Risk aversion or myopia? Choices in repeated gambles and retirement investments. *Management Sci.* **45**(3) 364–381.
Bernstein, P. L. 1976. The time of your life. *J. Portfolio Management* **2**(4) 4–7.
Fishburn, P. C. 1964. *Decision and Values Theory.* Wiley, New York.
Hadar, J., W. R. Russell. 1969. Rules for ordering uncertain prospects. *Amer. Econom. Rev.* **59** 25–34.
Hanoch, G., H. Levy. 1969. The efficiency analysis of choices involving risk. *Rev. Econom. Stud.* **36** 335–346.
———, ———. 1970. Efficient portfolio selection with quadratic and cubic utility. *J. Bus.* **43** 181–189.

Ibbotson Associates. 2001. *2001 Stocks, Bonds, Bills, and Inflation Classic Edition Yearbook*. Ibbotson Associates, Chicago, IL.
Levy, H. 1992. Stochastic dominance and expected utility: Survey and analysis. *Management Sci.* **38**(4) 555–593.
———. 1996. Investment diversification and investment specialization of the assumed holding period. *Appl. Math. Finance* **3** 117–134.
———. 1998. *Stochastic Dominance: Investment Decision Making Under Uncertainty*. Kluwer Academic Publishers, Boston, MA.
Markowitz, H. M. 1952. Portfolio selection. *J. Finance* **7** 77–91.
Meyer, J. 1977a. Choice among distributions. *J. Econom. Theory* **14** 326–336.
———. 1977b. Second degree stochastic dominance with respect to a function. *Internat. Econom. Rev.* **18** 477–487.
Rothschild, M., J. E. Stiglitz. 1970. Increasing risk: I. definition. *J. Econom. Theory* **2** 225–243.
Samuelson, P. A. 1994. The long-term case for equities. *J. Portfolio Management* **21**(1) 15–24.

Accepted by James E. Smith; received June 15, 2001. This paper was with the authors 4 months for 4 revisions.