Stochastic Superiority

May 5, 2021

Liqun Liu
Private Enterprise Research Center
Texas A&M University
College Station, TX 77843
U.S.A.
lliu@tamu.edu

Jack Meyer
Department of Economics
Michigan State University
East Lansing, MI 48824
U.S.A.
jmeyer@msu.edu

Abstract

This paper introduces a definition of stochastic superiority. One random variable is stochastically superior to another whenever it stochastically dominates the other after the risk in each random variable has been optimally reduced. Stochastic superiority is implied by stochastic dominance, but the reverse is not true. Stochastic superiority allows more pairs of random alternatives to be ranked, and efficient sets to be smaller. A very strong sufficient condition for stochastic superiority is demonstrated to also be necessary when preferences are risk averse. This condition provides a relatively easy way to conduct stochastic superiority tests. As an alternative to "almost stochastic dominance", stochastic superiority also provides a natural solution to the "left tail problem" that arises often when comparing random alternatives.

1. Introduction

When choosing between risks a decision maker is often at the same time selecting an overall risk level. Choosing between two risky mutual funds is combined with deciding how much wealth to invest in the selected mutual fund and how much to hold in an insured savings account instead. Buying a car or house is combined with choosing how much auto or homeowner's insurance to purchase. When selecting a television or washing machine, extending the manufacturer's warranty is often an option. These secondary decisions are made simultaneously with the primary decision and are ways that decision makers can use to reduce the riskiness of their initial selection. The point being made with these examples is that when a decision maker is choosing between two risky alternatives (the primary decision), the decision of how much of the chosen risk to bear and how much to transfer to others is often made simultaneously. This secondary decision is relevant and important precisely because it is made concurrently with the primary decision. The decision maker's attitude toward risk affects both decisions, and this interaction can alter the choice between risky alternatives.[1]

This paper explicitly models a simultaneous secondary decision in the stochastic dominance context. Sets of decision makers are asked to rank or choose between two random variables $\tilde{x}$ and $\tilde{y}$ knowing that a portion of the risk they choose can be transferred to others. The analysis extends the literature on stochastic dominance that was started more than fifty years ago by Hadar and Russell (1969) and Hanoch and Levy (1969). The basic stochastic dominance question they pose is: when is random variable $\tilde{x}$ preferred or indifferent to random variable $\tilde{y}$

---

[1] Whereas the primary and secondary decisions are simultaneously made in real-world situations, the primary decision is typically made in isolation in a laboratory environment. Therefore, it is not surprising that the risk attitudes elicited based on choices made in the lab consistently fail to predict choices in the field (e.g., Charness et al. 2020).

by all decision makers with a utility function in some set U? This same question is addressed here with the additional feature that both x̃ and ỹ can be altered by choosing an optimal risk reducing change or transformation. This additional decision is made simultaneously with choosing between x̃ and ỹ, and can lead to a different selection. Because of this extra decision, individual decision makers may rank x̃ and ỹ differently than without it, and therefore the answer to the basic stochastic dominance question may also be different.[2]

Because the answer to the basic stochastic dominance question can be different when x̃ and ỹ can be transformed, a new term is required to discuss the partial order that is generated. The term used is "stochastically superior" or "stochastic superiority". Random variable x̃ is said to be stochastically superior to ỹ for a set of decision makers U if every decision maker in U weakly prefers the optimally transformed x̃ to the optimally transformed ỹ. The transformation applied to ỹ can be different from that applied to x̃, and these transformations can and likely do differ across decision makers. Each decision maker in U optimally selects a risk reducing transformation for x̃ and for ỹ at the same time he or she is choosing between x̃ and ỹ. All decisions are assumed to be made to maximize the decision maker's expected utility.

When decision makers can each optimally choose how to transform both x̃ and ỹ, the question of which of the two random variables is unanimously preferred to the other by sets of decision makers could well have fewer rather than more clear-cut answers; that is, there could be fewer rather than more unanimous rankings of x̃ and ỹ. The additional complication of a simultaneous risk level decision could make it more difficult for sets of decision makers to agree

---

[2] The concept of stochastic superiority developed here is related to a very similar portfolio stochastic dominance concept defined by Levy and Kroll (1976 and 1978), but the analysis and findings here differ significantly from theirs. This is discussed more fully in the body of the paper.

as to whether x̃ or ỹ is best. A main finding here, however, is that for the simultaneous risk level decision considered in this analysis, x̃ stochastically dominating ỹ implies, but is not implied by, x̃ being stochastically superior to ỹ. That is, for sets of decision makers, the inclusion of this simultaneous risk level decision makes unanimity for the primary decision of ranking x̃ and ỹ more rather than less likely. That stochastic superiority is implied by stochastic dominance is presented as Corollary 1 to Theorem 1 in the paper.

Stochastic superiority allows random variables to be ranked when stochastic dominance does not, and therefore leads to smaller efficient sets. As the title of the Hanoch and Levy (1969) paper indicates, reducing sets of random variables to efficient sets, ones containing only elements that at least one decision maker would choose, is one of the main goals of stochastic dominance analysis. This is true not only for first- and second-degree stochastic dominance as defined by Hadar and Russell (1969) and Hanoch and Levy (1969), but also for general nth-degree stochastic dominance and many other forms of stochastic dominance presented in the literature since then. Each new definition refines a previously established form of stochastic dominance in a way that leads to additional rankings and smaller efficient sets.

Leshno and Levy (2002), when defining almost stochastic dominance, are also concerned with obtaining additional rankings of random variables. They define a procedure that can be applied to any of the established forms of stochastic dominance such as first-, second- or nth-degree stochastic dominance. Almost stochastic dominance modifies an existing form of stochastic dominance by removing from the associated set of utility functions those utility functions that represent preferences that Leshno and Levy consider to be "extreme, pathological

or simply unrealistic".[3] By removing these extreme preferences and utility functions, unanimous rankings are obtained that these extreme preferences would otherwise prevent.[4]

The almost stochastic dominance procedure follows a long-established pattern. All stochastic dominance definitions, including those generated using almost stochastic dominance, propose a new form of stochastic dominance that allows more random variables to be ranked and accomplishes this by removing from consideration some of the utility functions previously allowed. This pattern was established at the very beginning by Hadar and Russell and Hanoch and Levy when they defined second-degree stochastic dominance by removing from consideration all increasing utility functions that are not concave. By removing these utility functions from consideration, second-degree stochastic dominance can rank more random variables and generate smaller efficient sets than first-degree stochastic dominance.

Like almost stochastic dominance, stochastic superiority is a modification procedure that can be applied to any established form of stochastic dominance and the goal is additional rankings and smaller efficient sets. Stochastic superiority obtains these additional rankings, however, not by removing utility functions, but by allowing decision makers to modify the random variables that are being compared. Stochastic superiority facilitates seemingly reasonable rankings not by discarding "extreme, pathological or simply unrealistic" preferences, but instead by allowing all decision makers, including those whose preferences seem to be extreme, to optimally reduce the risk associated with any random variable. Allowing this

---

[3] Tzeng et al. (2013), Guo et al. (2013), Denuit et al. (2014), Tsetlin et al. (2015), Chang et al. (2019), Luo and Tan (2020), and Liu and Meyer (2021) further refine and extend the analysis of almost stochastic dominance.

[4] Meyer (1977a and 1977b) also does this for first- and second-degree stochastic dominance by placing an upper bound and/or a lower bound on decision makers' Arrow-Pratt absolute risk aversion measure.

additional simultaneous risk reduction decision is often enough to cause the ranking of x̃ and ỹ

by those with extreme preferences to be aligned with the ranking by decision makers whose

preferences are not so extreme. Stochastic superiority and almost stochastic dominance have the

same goal, to rank more random variables and generate smaller efficient sets, but the way in

which this goal is accomplished is very different.

Consider the following example provided by Leshno and Levy (2002). Assume that two

random alternatives are available, x̃ which yields either $0 or $1,000,000 with probabilities 0.01

and 0.99 respectively, and ỹ which is $1 with certainty. Leshno and Levy point out that although

neither x̃ nor ỹ dominates the other in the first degree,[5] except for very extreme preferences, x̃

would be chosen over ỹ by those who prefer more to less. They provide an example of such

extreme risk preferences represented by the utility function $u(x) = x$ for $x \leq 1$ and $u(x) = 1$ for $x >$

1. Leshno and Levy argue that to ensure that x̃ is preferred to ỹ by all decision makers, utility

functions as extreme as this $u(x)$ must be excluded. The almost stochastic dominance

modification procedure excludes these utility functions and thereby ensures that all decision

makers unanimously prefer x̃ to ỹ.

To see how stochastic superiority deals with this same example, notice that random

variable x̃ can be interpreted as asset whose value is $1,000,000 but subject to total loss with

probability 0.01. Expected loss is then $10,000. Suppose coinsurance is available at a very high

price of $100,000, which equals a loading factor of 9. When the chosen share of loss that is

---

[5] Since x̃ has a positive probability mass to the left of the entire distribution of ỹ , x̃ does not stochastically dominate
ỹ in any degree. Of course, since $E(x̃) > E(ỹ)$, ỹ does not stochastically dominate x̃ in any degree either.

reimbursed is θ in [0, 1], the required insurance premium would be $100,000 (θ), and random

wealth x̃ becomes

$$\$1,000,000 \, (\theta) - \$100,000 \, (\theta), \qquad \text{with probability } 0.01$$

$$\$1,000,000 - \$100,000 \, (\theta), \qquad \text{with probability } 0.99$$

This transformed x̃ stochastically dominates $1 in the first degree whenever

($1,000,000 – $100,000) (θ) ≥ $1 or θ ≥ 1/900,000.

This example shows that every individual who prefers more to less, including the one

whose utility function is u(x) given above, would prefer x̃ to ỹ when a coinsurance arrangement

with a loading factor even as high as 9 is available. After x̃ has been chosen, individual decision

makers will choose a coinsurance rate θ in [0, 1] to maximize their own expected utility. Many

would choose not to insure at all, but those with extreme preferences including the example u(x),

would choose to coinsure at a rate greater than 1/900,000.

The paper is organized as follows. In the next section, the question of how to compare

one set of random variables with another set for groups of decision makers, the set dominance

issue, is discussed, and the notion of stochastic superiority is formally defined. Set dominance is

relevant for the definition of stochastic superiority because allowing a risk reducing secondary

decision associates a set of random variables with each original one, and as a result ranking a

pair of random variables is transformed into ranking two sets of random variables. In Section 3,

a simple and seemingly overly strong sufficient condition for stochastic superiority is presented.

One important implication of this sufficient condition is that stochastic dominance implies

stochastic superiority. In the same section, this same sufficient condition is demonstrated to be

necessary for convex sets of concave utility functions. Corollaries are provided that state these

general results for the specific cases of first, second and nth-degree stochastic superiority. The set of utility functions corresponding to each of these is a convex set of utility functions, and concavity of utility is assumed for all but first degree stochastic superiority. An important extension to situations where the cost of risk reduction differs for different random variables is discussed in Section 4. This material is particularly relevant when insurance is used to reduce risk since the cost of insurance is typically dependent on the expected loss associated with the indemnification contract. Section 5 is the application section with three example applications. The first two show precisely how stochastic superiority can be used to obtain a ranking of $\tilde{x}$ over $\tilde{y}$ without removing extreme preferences from the set of utility functions under consideration. The final application is an extensive analysis of the decision to self-protect using stochastic superiority as an analysis tool. A result is presented indicating when all risk averse decision makers choose a higher level of self-protection over a lower level. As those who have examined the self-protection literature know, such a result is impossible to obtain in the standard model without a secondary risk reduction decision. This last example adds significantly to the self-protection literature. Finally, the paper concludes and discusses extensions of this work in Section 6.

2. Set Dominance and Stochastic Superiority

How decision makers might rank sets of objects has been extensively studied.[6] The discussion of set dominance presented here differs from this general literature in two important ways. First, the objects in the sets to be ranked are random variables that are ranked using expected utility. Second, in the tradition of stochastic dominance, the relative desirability of one

---

[6] For a survey of this literature, see Barbera et al. (2004).

set of random variables over another is based on the unanimous preference for the former over the latter by groups rather than individual decision makers.

Let $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$ denote two sets of random variables. The supports of these random variables are assumed to lie in an interval [a, b]. The question under consideration is: when would all decision makers with utility functions u(x) in a set U rather be faced with the set $\{\tilde{x}_i\}$ than the set $\{\tilde{y}_j\}$? Throughout the paper, the utility functions in U are assumed to be continuously differentiable as many times as necessary. Most often the set U will be one of the many sets of utility functions associated with the various degrees and forms of stochastic dominance, but sets containing only a small finite number of elements are also possible. A general specification of U is used in the majority of the analysis. The set dominance definition used in this analysis is given below.

<u>Definition 1</u>: The set $\{\tilde{x}_i\}$ is said to dominate or be preferred or indifferent to the set $\{\tilde{y}_j\}$ on U if for every u(x) in U the best element from $\{\tilde{x}_i\}$ is preferred or indifferent to the best element from $\{\tilde{y}_j\}$.

In Definition 1, both the best element from a set and the relative desirability of the respective best elements are determined using expected utility. There is no requirement in this definition that the decision makers agree as to which element of either set is best. Using this definition to rank sets of random variables is consistent with assuming that after a set is selected, the decision maker then chooses and obtains the most preferred element from that set. There is no requirement on U imposed by this definition. U can contain a single element, a small finite

number of elements, or U can be any one of the various sets of utility functions employed in stochastic dominance analysis.

Definition 1 is not the only set dominance definition that is possible or reasonable. For game models it may be more accurate to assume that after a set is selected, someone else chooses the particular element from either $\{\tilde{x}_i\}$ or $\{\tilde{y}_j\}$, possibly in a malevolent way, or more neutrally, to assume that nature chooses from these sets in a random way. For our purpose of including a simultaneous risk reduction decision with a primary ranking decision, Definition 1 provides the desired concept of set dominance.

Since the work here is similar in other ways to that presented by Levy and Kroll (1976 and 1978) and Levy (2006, Chapter 4), it is important to point out that they do not use Definition 1 as their definition of set dominance.[7] Levy and Kroll indicate that the set $\{\tilde{x}_i\}$ LK-dominates the set $\{\tilde{y}_j\}$ on U if for every element of $\{\tilde{y}_j\}$ there exists an element of $\{\tilde{x}_i\}$ such that the latter element is preferred or indifferent to the former by all u(x) in U. When set $\{\tilde{x}_i\}$ dominates set $\{\tilde{y}_j\}$ by the LK definition it also does so by Definition 1, but the reverse is not true. LK-dominance is more difficult to satisfy than is Definition 1.

A restaurant analogy can be used to illustrate the differences between these two concepts of set dominance. Consider two restaurants, each offering a menu of food items. The definition of set dominance used here indicates that one restaurant is preferred or indifferent to the other by a group of diners if every diner in the group (weakly) prefers her or his favorite item on the one restaurant's menu to her or his favorite item at the other restaurant. This definition of set

---

[7] In Levy and Kroll's analysis, the sets $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$ are the returns from portfolios that are composed of a risky asset $\tilde{x}$ or $\tilde{y}$, and a riskless asset. In this regard, their decision model is very similar to that presented here.

dominance is consistent with assuming that after choosing which restaurant to visit, each person in the group can order their own most preferred item on that restaurant's menu. In contrast, one restaurant LK-dominates the other if for every item on the latter restaurant's menu, there exists an item on the former restaurant's menu that is (weakly) preferred by all diners in the group. For situations where all diners are required to order the same menu item, the LK definition can be useful. For the purposes here, the LK-dominance relation is stronger than needed or desired. It is because of this different set dominance definition that the work presented here differs from that of Levy and Kroll.

As indicated in the introduction, the fundamental stochastic dominance question was answered by Hadar and Russell (1969) and Hanoch and Levy (1969) when the set U consists of all nondecreasing utility functions, denoted U_1, or when the set U consists of all nondecreasing and weakly concave utility functions, denoted U_2. These are, of course, the well-known first-degree stochastic dominance (FSD) and second-degree stochastic dominance (SSD) definitions, respectively. In most of the literature since 1969, extending FSD and SSD involves specifying a different set of utility functions U which is a subset of U_2. For example, general nth-degree stochastic dominance considers sets of utility functions U_n whose kth derivatives (k = 1, 2, …, n) alternate in sign, the first derivative being nonnegative. The almost stochastic dominance (ASD) literature also follows this pattern of altering the set of utility functions and does so by excluding from U_1, U_2 or U_n, decision makers whose preferences are considered to be extreme.

The stochastic superiority definition presented below extends the FSD and SSD findings but does so in a different manner. Rather than considering subsets of U_2, instead assumptions are made that allow each decision maker to choose to transform the pair of alternatives, $\tilde{x}$ and $\tilde{y}$, that

are being compared. The possible transformations lead to a pair of sets $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$, that are

derived from $\tilde{x}$ and $\tilde{y}$, respectively, which the decision makers must then rank or choose between.

The ranking of $\tilde{x}$ and $\tilde{y}$ is determined by the ranking of sets $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$, and Definition 1 is

used when determining this ranking. This way of extending stochastic dominance is not the

usual one and requires a new term. The term used is stochastic superiority, and a general

definition of stochastic superiority is provided next.

Definition 2: Random variable $\tilde{x}$ is stochastically superior to random variable $\tilde{y}$ on U if for every

u(x) in U the most preferred element in set $\{\tilde{x}_i\}$ is preferred or indifferent to the most preferred

element in set $\{\tilde{y}_j\}$, where $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$ are derived from $\tilde{x}$ and $\tilde{y}$, respectively, by transforming $\tilde{x}$

and $\tilde{y}$ in a simultaneous secondary decision.

Whether $\tilde{x}$ is stochastically superior to $\tilde{y}$ or not depends critically on two things, the set of

utility functions under consideration, and the nature of the simultaneous secondary decision that

gives rise to the sets $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$. The sets of utility functions U considered in this analysis all

are well known and have been extensively discussed in the literature. Included in the analysis

are U_1, U_2 and U_n, the sets of utility functions that are associated with FSD, SSD and nth-degree

stochastic dominance, respectively. The focus here is not on new sets of utility functions.

In contrast, the sets $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$, and the simultaneous secondary decision that generates

these sets, are less well known in the stochastic dominance literature. While many simultaneous

secondary decisions can be considered, the focus here is on a particularly simple one which is

often available to the decision maker. The secondary decision is that of choosing an overall risk

level. That is, in addition to choosing between $\tilde{x}$ and $\tilde{y}$, the decision maker also decides how

much of the risk that results from choosing $\tilde{x}$ or $\tilde{y}$ to assume and how much of that risk to avoid or transfer to others. In this analysis, when choosing between risky alternatives, decision makers are simultaneously deciding how much to reduce the risk they have selected using a risk reduction tool such as holding a riskless asset or purchasing insurance. These risk reducing secondary decisions are represented by the set of linear transformations which are described next.

When $\tilde{x}$ is transformed linearly, and the α in the linear transformation $\alpha \cdot \tilde{x} + (1 - \alpha)\rho$ is restricted to be in [0, 1], the new random variable has smaller variance and less risk using a variety of measures of riskiness.[8] The transformed $\tilde{x}$ is also in the same location-scale family as $\tilde{x}$, and thus maintains many of its other general shape properties such as skewness or kurtosis.[9] Assuming changes to $\tilde{x}$ and $\tilde{y}$ are accomplished using such linear transformations is a simple assumption and one that maintains a focus on ranking $\tilde{x}$ relative to $\tilde{y}$. The assumption does not intentionally bias the choice between $\tilde{x}$ and $\tilde{y}$ in one direction or the other since each of these random variables can be transformed in the same way.

When ρ in the linear transformation is chosen to be equal to the mean value of $\tilde{x}$, the transformation is a Rothschild and Stiglitz (1970) mean preserving reduction in risk. While the analysis considers a range of values for ρ, the focus is not on mean preserving risk reductions. Instead, ρ is always assumed to be less than the mean of $\tilde{x}$. This is done to incorporate costly risk reduction into the simultaneous secondary risk reduction decision. The decision maker has

---

[8] Sandmo (1971) uses these linear transformations in his analysis of the firm under price uncertainty. If $\alpha > 1$, the transformation would not be risk reducing; if $\alpha < 0$, the risk-reducing transformation would go too far to preserve any risk structure of the original random prospect.

[9] For discussion of decision making over location-scale families of random alternatives, see, for example, Meyer (1987) and Wong and Ma (2008).

the opportunity to transform $\tilde{x}$ in a risk reducing manner but must give up a higher expected return in order to obtain this risk reduction. When diversifying using a riskless asset, this means the riskless return is assumed to be smaller than the expected returns to risky assets. When insuring, this implies that the price of insurance is assumed to be actuarially unfair; that is, the loading factor is positive. It is important to point out that the smaller the value for ρ, the more costly is risk reduction.

In summary, the sets $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$ are formed from $\tilde{x}$ and $\tilde{y}$ by selecting a value for α_i or β_j in [0, 1] so that $\tilde{x}_i = \alpha_i \cdot \tilde{x} + (1 - \alpha_i) \cdot \rho$ and $\tilde{y}_j = \beta_j \cdot \tilde{y} + (1 - \beta_j) \cdot \rho$, respectively. These chosen values for α and β represent how much of the risk of $\tilde{x}$ and $\tilde{y}$ is retained, and can range from no risk at all, $\alpha = 0$ or $\beta = 0$, to accepting all of the risk associated with $\tilde{x}$ and $\tilde{y}$, $\alpha = 1$ or $\beta = 1$. The general stochastic superiority definition can now be restated in a more specific way using these linear risk reducing transformations to generate the sets $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$.

Definition 3: Random variable $\tilde{x}$ is stochastically superior to random variable $\tilde{y}$ on U for a given ρ if for every u(x) in U, the most preferred element in set $\{\alpha \cdot \tilde{x} + (1 - \alpha)\rho\}_{\alpha \in [0, 1]}$ is preferred or indifferent to the most preferred element in set $\{\beta \cdot \tilde{y} + (1 - \beta)\rho\}_{\beta \in [0, 1]}$.

To avoid trivial comparisons and repeated qualifications, the value of ρ is assumed to be such that at least one of $\tilde{x}$ and $\tilde{y}$ does not stochastically dominate ρ on U and $\rho < \max \{E(\tilde{x}), E(\tilde{y})\}$. This is because when both $\tilde{x}$ and $\tilde{y}$ dominate ρ, risk reduction in either $\tilde{x}$ or $\tilde{y}$ is never expected utility increasing, and if ρ is greater than or equal to the mean of both $\tilde{x}$ and $\tilde{y}$, risk reduction is free and maximum risk reduction is chosen by all risk averse decision makers.

3. A Sufficient and A Necessary Condition for Stochastic Superiority

At first glance, finding a necessary and sufficient condition for stochastic superiority appears to be a difficult task. The most preferred element in $\{\tilde{x}_i\}$ must be compared with the most preferred element in $\{\tilde{y}_j\}$ for each utility function in set U. A sufficient and a necessary condition for stochastic superiority are provided in this section. An important first step in proving sufficiency is the following lemma. The lemma holds for an arbitrary set of utility functions U, and for more general linear transformations than the risk reducing ones employed in Definition 3.

<u>Lemma 1</u>: $\tilde{x}$ stochastically dominates $\tilde{y}$ for all u(x) in U implies $\lambda \cdot \tilde{x} + \gamma$ stochastically dominates $\lambda \cdot \tilde{y} + \gamma$ for all u(x) in U, where λ and γ are any constants with $\lambda \geq 0$.

Lemma 1 is quite obvious, easy to demonstrate, and is stated here because it is used when proving Theorem 1.[10] The set U in Lemma 1 can contain a single element, a small finite number of elements, or any one of the various sets of utility functions employed in standard stochastic dominance analysis.

<u>Theorem 1</u>: If for some α in [0, 1] and a fixed ρ, $\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho$ stochastically dominates $\tilde{y}$ for all u(x) in U, then $\tilde{x}$ is stochastically superior to $\tilde{y}$ for all u(x) in U for the given ρ.
<u>Proof:</u> See the appendix.

Theorem 1 provides a sufficient test condition for stochastic superiority. One only need find any element of the set $\{\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho\}$ that stochastically dominates the untransformed $\tilde{y}$

---

[10] For a comprehensive collection of the results on stochastic dominance including Lemma 1, see Denuit et al. (2005) and Shaked and Shanthikumar (2007).

and stochastic superiority of $\tilde{x}$ over $\tilde{y}$ holds. It is not necessary to find the most preferred element in $\{\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho\}$ nor to compare elements from this set with any element other than the untransformed element in the set $\{\beta \cdot \tilde{y} + (1 - \beta) \cdot \rho\}$.

To relate stochastic superiority to stochastic dominance, observe that the special case where $\alpha = 1$ indicates that stochastic dominance of $\tilde{x}$ over $\tilde{y}$ implies stochastic superiority of $\tilde{x}$ over $\tilde{y}$. This is formally stated as the following corollary.

Corollary 1: Stochastic dominance of $\tilde{x}$ over $\tilde{y}$ on U implies stochastic superiority of $\tilde{x}$ over $\tilde{y}$ on U for all ρ.

The sufficient condition in Theorem 1 appears to be overly strong. As with Lemma 1, Theorem 1 holds for an arbitrary set U which can contain only one or a small number of elements as well as any one of the sets of utility functions that are associated with the various conventional forms of stochastic dominance. In the search for weaker sufficient conditions, to our surprise, it was determined that the condition in Theorem 1 is not only sufficient, but also necessary for virtually all forms of stochastic dominance. The necessity result, Theorem 2, requires more of the set U and as a result it is stated and demonstrated as a separate theorem.

Theorem 2: Suppose $\tilde{x}$ is stochastically superior to $\tilde{y}$ for a given ρ on some convex set of concave utility functions U. Then there exists an α_0 in [0, 1] such that $\alpha_0 \cdot \tilde{x} + (1 - \alpha_0) \cdot \rho$ stochastically dominates $\tilde{y}$ on U.

Proof: See the appendix.

For the sets of $\{\tilde{x}_i\}$ and $\{\tilde{y}_j\}$ generated using linear transformations, it is straightforward to see that the condition in Theorem 1 is both sufficient and necessary for set $\{\tilde{x}_i\}$ to LK-dominate set $\{\tilde{y}_j\}$. So it is quite natural that this condition is also sufficient for stochastic superiority which is based on the weaker notion of set dominance given in Definition 1. It is important to mention again that Theorem 1 holds for arbitrary sets of utility functions.

The sufficient condition in Theorem 1 is demonstrated to be necessary in Theorem 2 when the set of utility functions is a convex set of concave utility functions. Recall that the sets of utility functions for all conventional forms of stochastic dominance other than first degree satisfy this requirement. The proof of Theorem 2 relies on the convexity of the set U and also on the concavity of the individual u(x). It is these restrictions on U that imply that the weak set dominance condition provided in Definition 1 is equivalent to the seemingly much stronger set dominance notion provided by Levy and Kroll (1976, 1978).[11] In summary, from the very beginning stochastic dominance has most often required unanimous preference by utility functions that come from a convex set of concave utility functions. Theorem 2 demonstrates the powerful nature of this assumption and uses it to provide a relatively simple and useful test condition for stochastic superiority.

To relate these sufficient and necessary conditions, which are stated quite generally, to more familiar sets of utility functions U_n, where $n \geq 1$, the following two corollaries are provided. Note that U_n is a convex set of concave utility functions for all $n \geq 2$.

---

[11] Fishburn (1974) exploits this same thing when demonstrating convex stochastic dominance theorems.

Corollary 2: If for some α in [0, 1] and a fixed ρ, $\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho$ stochastically dominates $\tilde{y}$ on U_n (or in the nth degree), where $n \geq 1$, then $\tilde{x}$ is stochastically superior to $\tilde{y}$ on U_n (or in the nth degree) for the given ρ.

Corollary 3: Suppose that $n \geq 2$. If $\tilde{x}$ is stochastically superior to $\tilde{y}$ in the nth degree for a given ρ, then there exists an α_0 in [0, 1] such that $\alpha_0 \cdot \tilde{x} + (1 - \alpha_0) \cdot \rho$ stochastically dominates $\tilde{y}$ in the nth degree.

Two additional properties of stochastic superiority on U_n are worth mentioning and follow directly from Definition 3. First, if $\tilde{x}$ is nth-degree stochastically superior to $\tilde{y}$ for a given ρ, then for any $n' > n$, $\tilde{x}$ is also n'th-degree stochastically superior to $\tilde{y}$ for that ρ. Second, if $\tilde{x}$ is nth-degree stochastically superior to $\tilde{y}$ for a given ρ, then $E(\tilde{x}) \geq E(\tilde{y})$. This is because that the risk neutral linear utility function is in U_n for all $n \geq 1$.

The parameter ρ in the linear risk reducing transformation is inversely related to the cost of risk reduction to the decision maker. Lower values for ρ imply a higher cost of reducing risk. While it is not obvious from Definition 3 whether a larger ρ would make stochastic superiority more likely, the following theorem gives an affirmative answer.

Theorem 3: If $\tilde{x}$ is stochastically superior to $\tilde{y}$ on some convex set of concave utility functions U for a given ρ, then $\tilde{x}$ is stochastically superior to $\tilde{y}$ on U for all ρ' such that $\rho' \geq \rho$.
Proof: See the appendix.

Theorem 3 indicates that when stochastic superiority of $\tilde{x}$ over $\tilde{y}$ holds for a given cost of risk reduction, then it also holds for all lower costs of risk reduction.

4. An Extension: Stochastic Superiority of ($\tilde{x}$, ρ_x) over ($\tilde{y}$, ρ_y)

All analysis to this point assumes that the same value for ρ is used when transforming both $\tilde{x}$ and $\tilde{y}$. Recall that ρ indicates the cost of risk reduction, and that smaller values for ρ indicate a larger cost of risk reduction. This section generalizes the definition of stochastic superiority to the situation where the cost of risk reduction is different for $\tilde{x}$ than for $\tilde{y}$. First, Definition 3 is simply restated for different costs of risk reduction for $\tilde{x}$ and $\tilde{y}$.

Definition 4: ($\tilde{x}$, ρ_x) is stochastically superior to ($\tilde{y}$, ρ_y) on U if for every u(x) in U, the most preferred element in set $\{\alpha \cdot \tilde{x} + (1 - \alpha)\rho_x\}_{\alpha \in [0, 1]}$ is preferred or indifferent to the most preferred element in set $\{\beta \cdot \tilde{y} + (1 - \beta)\rho_y\}_{\beta \in [0, 1]}$.

By definition, if ($\tilde{x}$, ρ_x) is stochastically superior to ($\tilde{y}$, ρ_y) on U, then for all $\rho_x' \geq \rho_x$ and $\rho_y' \leq \rho_y$, ($\tilde{x}$, ρ_x') is stochastically superior to ($\tilde{y}$, ρ_y') on U. The following property is less obvious and is derived from the sufficient and necessary conditions established in Theorems 1 and 2.

Theorem 4. Suppose that U is a convex set of concave utility functions. ($\tilde{x}$, ρ_x) is stochastically superior to ($\tilde{y}$, ρ_y) on U for any ρ_y such that $\rho_x \geq \rho_y$ if and only if ($\tilde{x}$, ρ_x) is stochastically superior to ($\tilde{y}$, ρ_x) on U.

Proof: See the appendix.

To illustrate situations where the cost of risk reduction is likely to differ across risky alternatives and at the same time to demonstrate that insurance is a method of risk reduction that is included in this analysis, consider the following general example. Suppose there are two

different risky assets: $\tilde{x} = w_1 - \tilde{L}_1$ and $\tilde{y} = w_2 - \tilde{L}_2$, where w_i, $i = 1$ or 2, is the nonrandom initial

wealth and $\tilde{L}_1 \geq 0$ and $\tilde{L}_2 \geq 0$ are two different random losses.[12] $\tilde{x}$ and $\tilde{y}$ could represent the

uninsured value of two houses. In this example, each house has a different value, and each house

is subject to a different random loss. These differences can occur because of different locations

and the likelihood of earthquakes or hurricanes at each, or various other house characteristics.

Traditional stochastic dominance analysis compares the two random variables representing the

uninsured houses. Stochastic superiority adds to this comparison the assumption that when

choosing between the two houses, the decision maker knows that he or she will purchase

insurance to partially indemnify losses and may choose to insure each house differently.

With co-insurance, $\tilde{x}$ can be transformed into any random final wealth in the set

$\{w_1 - \tilde{L}_1 + \theta(\tilde{L}_1 - m_1)\}_{\theta \in [0, 1]} = \{(1 - \theta) \cdot \tilde{x} + \theta(w_1 - m_1)\}_{\theta \in [0, 1]}$, where θ in [0, 1] is the share of

loss that is reimbursed, and $m_1 = (1 + \pi)E(\tilde{L}_1)$ is the insurance premium for full coverage when

the loading factor is $\pi > 0$. Similarly, $\tilde{y}$ can be transformed into any random final wealth in the

set $\{w_2 - \tilde{L}_2 + \theta(\tilde{L}_2 - m_2)\}_{\theta \in [0, 1]} = \{(1 - \theta) \cdot \tilde{y} + \theta(w_2 - m_2)\}_{\theta \in [0, 1]}$, where $m_2 = (1 + \pi)E(\tilde{L}_2)$.

Each of these sets are exactly in the form of the general linear transformation used when

stochastic superiority is defined (Definitions 3 or 4).

It is important to recognize, however, that the price of risk reduction is likely to differ

across these two assets due to differences in the loss variables. For these $\tilde{x}$ and $\tilde{y}$, the cost of risk

reduction for $\tilde{x}$ is $\rho_x = w_1 - m_1$ and for $\tilde{y}$ is $\rho_y = w_2 - m_2$ and these values may be different. In

the specific situation of coinsurance, the cost of risk reduction is partially determined by the

---

[12] Note that with the additional requirement that the minimum value of $\tilde{L}_i$ ($i = 1$ or 2) with a positive probability (density) be zero, the decomposition of $\tilde{x} = w_1 - \tilde{L}_1$ and $\tilde{y} = w_2 - \tilde{L}_2$ must each be unique.

loading factor. Theorems 3 and 4 can be used together to establish the following theorem indicating how the likelihood of stochastic superiority changes with changes in the loading factor.

Corollary 4: Suppose that U is a convex set of concave utility functions. If $\tilde{x}$ is stochastically superior to $\tilde{y}$ on U under a common loading factor $\pi_1 > 0$, then $\tilde{x}$ is stochastically superior to $\tilde{y}$ on U under any common loading factor π_2 such that $\pi_1 > \pi_2 \geq 0$.

5. Applications

Three examples are used to illustrate how stochastic superiority provides additional rankings of random variables. All three examples involve the so-called "left tail problem" where $\tilde{x}$ is better than $\tilde{y}$ for many or most decision makers, but stochastic dominance (of any degree) of $\tilde{x}$ over $\tilde{y}$ does not hold because $\tilde{x}$ has more probability mass than $\tilde{y}$ in the extreme left tail.

The first example is provided in Levy (2006).[13] Two alternatives are available, $\tilde{x}$ which yields either \$1 or \$1,000,000 with probabilities 0.1 and 0.9 respectively, and $\tilde{y}$ which is either \$2 or \$3 with those same probabilities. As Levy points out, neither $\tilde{x}$ nor $\tilde{y}$ dominates the other in the first degree, yet $\tilde{x}$ appears to be clearly better than $\tilde{y}$. Almost stochastic dominance remedies this and allows $\tilde{x}$ to be unanimously ranked higher than $\tilde{y}$ by eliminating those decision makers exhibiting extreme risk aversion from consideration. Stochastic superiority, on the other hand, indicates that $\tilde{x}$ is unanimously chosen over $\tilde{y}$ by all decision makers, including those with extreme risk aversion, when all decision makers are permitted to transfer some of the risk in $\tilde{x}$ and $\tilde{y}$ to others.

---

[13] Chapter 13, pages 331-332.

Specifically, suppose $\rho = 2.5$, a value such that it is larger than the smallest values of both $\tilde{x}$ and $\tilde{y}$ (so neither $\tilde{x}$ nor $\tilde{y}$ stochastically dominates ρ), and is also smaller than the means of both $\tilde{x}$ and $\tilde{y}$ (implying that risk reduction is costly for both $\tilde{x}$ and $\tilde{y}$). It can be readily seen that the transformation $(0.1) \cdot \tilde{x} + (0.9)(2.5)$ yields a random variable which stochastically dominates $\tilde{y}$ in the first degree. Thus, by Theorem 1, $\tilde{x}$ is stochastically superior to $\tilde{y}$ in the first degree for $\rho = 2.5$. This risk reducing transformation of $\tilde{x}$ eliminates the left tail issue that would lead some with extreme preferences to choose $\tilde{y}$ over $\tilde{x}$. This is an extreme transformation since it gives up 90% of the gain associated with the outcome $1,000,000$, but those with extreme enough preferences would choose to do this. Those with extreme preferences are willing to transfers risk to others even when the cost of such a transfer is very high. Of course, many decision makers would forego this risk reducing action because their preferences are less extreme. When this risk transfer is available, the ranking of $\tilde{x}$ and $\tilde{y}$ by those with extreme preferences are aligned with the majority.

The next example does not come from the almost stochastic dominance literature. In fact, the almost stochastic dominance modification procedure is unlikely to lead to a ranking of the two alternatives. Suppose one must choose between two goods that can either fail immediately or not fail at all. The first good fails with probability 0.3, and the second good fails with a higher probability 0.5. Assume that the value provided by each good is $1000 in the case on no failure and $0 in the case of failure. For risk averse buyers, these goods must sell for less than expected value. Assume the selling price of the first good is $500, and that of the second is $400. Which good does a risk averse decision maker choose?

When there is no refund and no warranty against failure, the first good generates a random outcome, denoted x̃, that yields a net loss of \$500 with probability 0.3 and a net gain of \$500 with probability 0.7. Similarly, the second good generates a random outcome, denoted ỹ, that yields a net loss of \$400 or a net gain of \$600 but with equal probability. Many risk averse decision makers would choose the first good over the second both because of the smaller probability of failure and the larger expected value, but many others would choose the second over the first because of the smaller net loss when failure occurs. It does not require extreme risk preferences to choose the second good over the first, making it difficult to know which preferences to eliminate in the almost stochastic dominance procedure.

It can be demonstrated that the first good is stochastically superior to the second in the second degree even though it does not stochastically dominate in the second degree. Specifically, suppose that the cost of risk transfer is high, $\rho = 0$. Even with this high cost of transferring risk to others, it can be readily seen that for $\rho = 0$ and $\alpha = 0.8$, the transformed x̃ (i.e., 0.8x̃) stochastically dominates ỹ in the second degree and therefore by Theorem 1, x̃ is stochastically superior to ỹ in the second degree for $\rho = 0$. This risk reducing transfer occurs by finding someone to take a 20% share in both outcomes associated with the purchase of the good. Many risk averse decision makers prefer the original x̃ to this transformed one, but this particular transformation of x̃ would be preferred to both the original x̃ and to ỹ by those with preferences putting significant weight on outcomes in the left tail. Even though x̃ does not stochastically dominant ỹ in the second degree, x̃ is preferred or indifferent to ỹ by all risk averse decision makers when the decision makers are allowed to transform x̃ (and ỹ) in this risk reducing way.

These two examples illustrate the way in which stochastic superiority can help remedy the left tail problem and allow the ranking of more pairs of random variables. In general, when random variable $\tilde{x}$ has a lowest outcome that is smaller than the lowest outcome for $\tilde{y}$, denoted $a_x < a_y$, one can choose an α in [0, 1] so that $\alpha \cdot a_x + (1 - \alpha) \cdot \rho \geq a_y$ as long as $\rho > a_y$. This latter condition is one often met since the riskless return is above the lowest outcome for all risky choices.

The above two examples have been constructed to illustrate a particular weakness in traditional stochastic dominance analysis referred to as the left tail problem. The final application is not a constructed example but instead an application to a model that is prominently discussed in the literature. In this application, stochastic superiority is shown to add a significant finding to the very large literature on self-protection.

The left tail problem arises in the model of self-protection introduced by Ehrlich and Becker (1972) and is the most prominent characteristic of that decision model. As Ehrlich and Becker point out, it is the case that choosing more self-protection always leads to a smaller lowest outcome than when less self-protection is selected. This fact implies that those decision makers who are extremely risk averse would always prefer lower levels of self-protection to higher levels no matter how effective self-protection is in reducing the probability of loss and increasing the expected value of the random outcome. As a result, without modifying the basic model of Ehrlich and Becker, it is impossible to establish any theorems which indicate that increased self-protection would make all risk averse decision makers better off.

This seemingly perverse feature of the self-protection model can be eliminated if the decision maker can both self-protect and also transfer some of the risk that is assumed to others.

Including such a risk transfer possibility is enough to present a positive result in which conditions are provided indicating when all risk averse decision makers would choose a higher level of self-protection over a lower level. These conditions depend on the parameter values in the self-protection model but are reasonable and are satisfied in many instances. The model of the decision to self-protect used here is that presented by Ehrlich and Becker (1972) and the notation used is that found in Eeckhoudt and Gollier (2005).

Eeckhoudt and Gollier assume that a decision maker begins with certain wealth w and that this wealth is subject to loss $L > 0$. This loss occurs with probability $p(e)$, where e represents the level of self-protection investment/cost chosen by the decision maker and $p(e)$ is decreasing in e. Let W denote final wealth, which can take on one of two values, either $(w - L - e)$ or $(w - e)$, with probabilities $p(e)$ and $1 - p(e)$, respectively.

Assume that there are two different levels of self-protection $e_1 \geq 0$ and $e_2 > e_1$. Also, to simplify notation, let $p(e_1) = p_1$ and $p(e_2) = p_2 < p_1$. Let $\tilde{x}$ denote the random wealth with the larger amount of self-protection, and $\tilde{y}$ the random wealth with the smaller amount of self-protection. Therefore, $\tilde{x}$ yields either $(w - L - e_2)$ or $(w - e_2)$, with probabilities p_2 and $(1 - p_2)$ respectively, and $\tilde{y}$ yields either $(w - L - e_1)$ or $(w - e_1)$, with probabilities p_1 and $(1 - p_1)$ respectively. The mean of $\tilde{x}$ is larger than the mean of $\tilde{y}$ if and only if

$$(p_1 - p_2)L > e_2 - e_1. \tag{1}$$

The term "effective" is used to describe the added level of self-protection when (1) holds. Later on a cost-benefit ratio proves useful, so the following definition is suggested.

<u>Definition 5</u>: Self-protection level e_2 is said to be an effective increase over e_1 whenever (1)

holds. In addition, $\dfrac{(e_2 - e_1)}{L(p_1 - p_2)}$ is said to be the cost-benefit ratio for the added level of self-

protection.

The cost-benefit ratio indicates how much additional effort must be expended per unit

reduction in the mean loss. Obviously, an increase in self-protection effort from e_1 to e_2 is

effective if and only if the associated cost-benefit ratio is less than one. When this ratio is less

than one, neither $\tilde{x}$ nor $\tilde{y}$ stochastically dominates the other in the second (or any) degree.[14]

Suppose that coinsurance is available to insure the self-protected asset at a loading factor

$\pi \geq 0$, and the individual can decide on a share of loss, denoted $(1 - \alpha)$, that is reimbursed.[15]

Then $\tilde{x} = w - e_2 - \tilde{L}_x$, where $\tilde{L}_x$ yields either L or 0 with probabilities p_2 and $(1 - p_2)$ respectively.

With coinsurance, $\tilde{x}$ can be transformed into any element of the form:

$$\{\alpha\tilde{x} + (1 - \alpha)\rho_x\}_{\alpha \in [0, 1]}$$

where $\rho_x = w - e_2 - (1 + \pi)E(\tilde{L}_x) = w - e_2 - (1 + \pi) p_2 L$.

Similarly, $\tilde{y} = w - e_1 - \tilde{L}_y$, where $\tilde{L}_y$ yields either L or 0 with probabilities p_1 and $(1 - p_1)$

respectively, can be transformed into any element of the form:

$$\{\alpha\tilde{y} + (1 - \alpha)\rho_y\}_{\alpha \in [0, 1]}$$

where $\rho_y = w - e_1 - (1 + \pi)E(\tilde{L}_y) = w - e_1 - (1 + \pi) p_1 L$.

---

[14] $\tilde{x}$ does not stochastically dominate $\tilde{y}$ in any degree because $\tilde{x}$ has a positive probability mass to the left of the entire distribution of $\tilde{y}$. $\tilde{y}$ does not stochastically dominate $\tilde{x}$ in any degree because $E(\tilde{x}) > E(\tilde{y})$.

[15] This means that an α-portion of the loss is retained by the individual and the remaining $(1 - \alpha)$-portion of the loss is transferred to the insurer.

The following theorem states that an effective increase in self-protection is stochastically superior to the status quo in the second degree, as long as the loading factor is below a certain critical level that is determined by the parameters of the self-protection model.

Theorem 5: Suppose that an increase in the level of self-protection from e_1 to e_2 is effective, and that coinsurance is available to insure the self-protected asset at a loading factor $\pi \geq 0$. Then all risk averse decision makers choose the larger level of self-protection over the smaller level if and only if the loading factor π satisfies

$$\pi \leq \frac{(1 - p_2)}{p_2}\left[1 - \frac{(e_2 - e_1)}{L(p_1 - p_2)}\right]. \tag{2}$$

Proof: See the appendix.

Note that the ratio in the brackets is the cost-benefit ratio for the increase in self-protection. It can be readily seen that (2) can be satisfied with $\pi > 0$ whenever the additional self-protection is effective (i.e., the cost-benefit ratio is less than one).

Theorem 5 indicates that when self-protection is effective, and the loading factor is small enough, all risk averse persons choose more self-protection. To see how this condition can be used, consider the following simple but reasonable parameter values. Let $p_1 = 0.2$, $p_2 = 0.1$, $L = 100$, $e_1 = 5$ and $e_2 = M$. If $M < 15$, self-protection is effective since the probability change decreases mean loss by 10. When $M = 10$, the cost-benefit ratio is 0.5, and loading factors $\pi \leq$ 4.5 are small enough that all risk averse decision makers choose the larger level of self-protection. When the added self-protection costs more, say $e_2 = 14$, the cost-benefit ratio increases to 9/10 and the right hand side of (2) becomes 0.9. Even so, loading factors less than

0.9 satisfy Theorem 5. Finally, as the cost of the added self-protection further increases so that $e_2 = 15$, the cost-benefit ratio increases to 1, and no positive loading factor would make every risk averse individual better off at the higher level of self-protection.

6. Conclusion

Including the opportunity to simultaneously reduce risk when making the primary decision of choosing between a pair of random alternatives can have an impact on that primary decision. This paper extends the notion of stochastic dominance to the notion of stochastic superiority by introducing such a risk reducing secondary decision. Stochastic superiority allows more pairs of random alternatives to be ranked than stochastic dominance. Both a sufficient and a necessary condition for stochastic superiority are established. These two conditions are identical when the relevant set of utility functions is a convex set of concave utility functions. One very important application of stochastic superiority is to provide an alternative to almost stochastic dominance when seeking to resolve the "left tail problem".

The analysis, theorems and proofs in this paper all assume decisions are made on the basis of expected utility (EU). Examples where violations of the EU axioms occur abound, including the well-known Allais paradox. As a result, many non-EU models of decision making under risk have been suggested.[16] This fact raises the question of the usefulness of stochastic superiority when the axioms of EU do not hold. Much like first or second degree stochastic dominance, first or second degree stochastic superiority can play a role in non-EU decision models. For instance, it is now common to define first degree stochastic dominance of F(x) over

---

[16] A prominent example is the prospect theory (e.g., Tversky and Kahneman 1992, Viscusi 1989, Neilson and Stowe 2002). See Quiggin (2014) for a survey of the non-EU models, and Bayrak and Hey (2020) for a comparison of alternative models.

G(x) using the necessary and sufficient condition derived in an EU decision model; that is, when

$F(x) \leq G(x)$ for all x. Even though this condition on random variables was interpreted within an

EU decision model, it has since been extensively used as the definition of first degree stochastic

dominance in non-EU decision models. In fact, proposed non-EU decision models are often

judged in part by how the decision model performs when such a first degree stochastic dominant

change occurs.

Similar uses for stochastic superiority in non-EU decision models are possible. First

degree stochastic superiority of $\tilde{x}$ over $\tilde{y}$ for fixed ρ can be defined as occurring when for some α

in [0, 1] and a fixed ρ, $\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho$ stochastically dominates $\tilde{y}$ in the first degree. Such a

definition does not assume EU even though it was formulated and interpreted assuming EU.

Non-EU decision models can then be judged by how the model performs when such a first

degree stochastically superior change occurs. Thus, even though all results in the paper are

derived assuming an EU decision model, applications to non-EU decision models are possible.

Exploring stochastic superiority in the various non-EU decision models would be too large an

addition to this paper, but can be a valuable avenue for future research.

There are also at least two other major directions in which this work can be extended or

used. First, the assumption of linear risk reducing transformations adopted in this paper could be

changed. A more lengthy title for this paper could involve the words "stochastic superiority with

respect to linear transformations" to make this obvious. Of course a random prospect's risk can

also be reduced through nonlinear transformations reflecting such things as deductible insurance

or options rather than coinsurance or diversification using a riskless asset. The framework here

allows such an extension, and using a concave or convex transformation would result in a different form of stochastic superiority.

The second direction of extension involves an application to a different literature. The work here is presented as part of decision making under risk. If the distributions describing random variables are instead assumed to represent distributions of income or wealth, then the findings in this paper can be applied to the literature that ranks income or wealth distributions and specifically the inequality literature. In that literature, the transformations of $\tilde{x}$ and $\tilde{y}$ have the very natural interpretation of using taxes, subsidies or other policies to redistribute income or wealth. Since linear or non-linear transformation alter both tails of a distribution, the effect on both the right and left tail can be discussed. For decision making under risk, the effect on left tail is the primary concern, but for income inequality the effect on each tail is important.

References

Barbera, S., Bossert, W. and Pattanaik, P.K. (2004). 'Ranking sets of objects', S. Barbera et al. (eds.), *Handbook of Utility Theory*, Springer Science+Business Media, Inc., New York.

Bayrak, O.K. and Hey, J.D. (2020). 'Decisions under risk: Dispersion and Skewness', *Journal of Risk and Uncertainty* 61: 1-24.

Chang J.-R., Liu, W.-H. and Hung, M.-W. (2019). 'Revisiting generalized almost stochastic dominance', *Annals of Operations Research* 281: 175-192.

Charness, G., Garcia, T., Offerman T. and Villeval, M.C. (2020). 'Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory?' *Journal of Risk and Uncertainty* 60: 99-123.

Denuit, M., Dhaene, J., Goovaerts, M. and Kaas, R. (2005). *Actuarial Theory for Dependent Risks: Measures, Orders and Models.* John Wiley & Sons: New York, NY.

Denuit, M., Huang R.J. and Tzeng L.Y. (2014). 'Bivariate almost stochastic dominance', *Economic Theory* 57: 377-405.

Eeckhoudt, L. and Gollier, C. (2005). 'The impact of prudence on optimal prevention', *Economic Theory* 26: 989-994.

Ehrlich, I. and Becker, G. (1972). 'Market insurance and self-insurance', *Journal of Political Economy* 80: 623-648.

Fishburn, P. (1974). 'Convex Stochastic Dominance with Continuous Distribution Functions', *Journal of Economic Theory* 7: 143-158.

Guo, X., Zhu, X., Wong, W.-K. and Zhu, L. (2013). 'A note on almost stochastic dominance', *Economics Letters* 121: 252-256.

Hadar, J. and Russell, W. (1969). 'Rules for ordering uncertain prospects', *American Economic Review* 59: 25–34.

Hanoch, G. and Levy, H. (1969). 'The efficiency analysis of choices involving risk', *Review of Economic Studies* 36: 335–346.

Leshno, M. and Levy, H. (2002). 'Preferred by "all" and preferred by "most" decision makers: Almost stochastic dominance', *Management Science* 48: 1074-1085.

Levy, H. (2006). *Stochastic Dominance: Investment Decision Making under Uncertainty* (2nd edition), Springer Science+Business Media, Inc., New York.

Levy, H. and Kroll, Y. (1976). 'Stochastic dominance with riskless assets', *Journal of Financial and Quantitative Analysis* 11: 743-777.

Levy, H. and Kroll, Y. (1978). 'Ordering uncertain options with borrowing and lending', *Journal of Finance* 33: 553-574.

Liu, L. and Meyer, J. (2021). 'Almost stochastic dominance: Magnitude constraints on risk aversion', manuscript.

Luo, C. and Tan, C.H. (2020). 'Almost stochastic dominance for most risk-averse decision makers', *Decision Analysis* 17: 169-184.

Meyer, J. (1977a). 'Second degree stochastic dominance with respect to a function', *International Economic Review* 18: 477-487.

Meyer, J. (1977b). 'Choice among distributions', *Journal of Economic Theory* 14: 326-336.

Meyer, J. (1987). 'Two-moments decision models and expected utility maximization', *American Economic Review* 77: 421-430.

Neilson, W. and Stowe, J. (2002). 'A further examination of cumulative prospect theory parameterizations', *Journal of Risk and Uncertainty* 24: 31-46.

Quiggin, J. (2014). 'Non-expected utility models under objective uncertainty', in M. Machina and W.K. Viscusi (eds.) *Handbook of the Economics of Risk and Uncertainty* (pp. 701-727), North Holland Press, London.

Rothschild, M. and Stiglitz, J. (1970). 'Increasing risk I: a definition', *Journal of Economic Theory* 2: 225-243.

Sandmo, A. (1971). 'On the Theory of the Competitive Firm Under Price Uncertainty', *American Economic Review* 61: 65-73.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders.* Springer: New York, NY.

Tsetlin, I., Winkler R.L, Huang R.J. and Tzeng L.Y. (2015) 'Generalized almost stochastic dominance', *Operations Research* 63: 363-377.

Tversky, A. and Kahneman, D. (1992). 'Advances in prospect theory: cumulative representation of uncertainty', *Journal of Risk and Uncertainty* 5: 297–323

Tzeng, L.Y., Huang, R.J. and Shih, P.-T. (2013). 'Revisiting almost second-degree stochastic dominance', *Management Science* 59: 1250-1254.

Viscusi, W.K. (1989). 'Prospective reference theory: toward an explanation of the paradoxes', *Journal of Risk and Uncertainty* 2: 235–263.

Wong, W.-K. and Ma, C. (2008). 'Preferences over location-scale family', *Economic Theory* 37: 119-146.

Appendix

Proof of Theorem 1

Assume $\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho$ stochastically dominates $\tilde{y}$ for all u(x) in U. Then from Lemma

1, for any λ in [0, 1] and $\gamma = (1 - \lambda)\rho$, $\lambda[\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho] + (1 - \lambda)\rho = \lambda \cdot \alpha \cdot \tilde{x} + (1 - \lambda \cdot \alpha) \cdot \rho$

stochastically dominates $\lambda \cdot \tilde{y} + (1 - \lambda)\rho$ for all u(x) in U. Put differently, for every element in the

set $\{\lambda \cdot \tilde{y} + (1 - \lambda) \cdot \rho\}$, there is an element in the set $\{\lambda \cdot \alpha \cdot \tilde{x} + (1 - \lambda \cdot \alpha) \cdot \rho\}$ that dominates the

former for all u(x) in U. This is of course sufficient for the optimal choice from the set

$\{\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho\}$ to be preferred or indifferent to the optimal choice from the set $\{\beta \cdot \tilde{y} + (1 - \beta) \cdot \rho\}$

for every utility function in the set U. Therefore, $\tilde{x}$ is stochastically superior to $\tilde{y}$ for all u(x) in U

for the given ρ.   Q.E.D.

Proof of Theorem 2

Let $U(\alpha) = Eu(\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho)$ for α in [0, 1] and a given ρ. What must be shown is that

if for any u(x) in U there exists an α such that $U(\alpha) \geq Eu(\tilde{y})$, then there exists an α_0 in [0, 1] such

that $U(\alpha_0) \geq Eu(\tilde{y})$ for all u(x) in U. Define A_u to be the set of all α in [0, 1] such that

$U(\alpha) \geq Eu(\tilde{y})$. For concave utility functions, A_u is a closed line segment in [0, 1]. Let A^* denote

the intersection of these line segments. If A^* is not empty, then any α_0 in A^* is such that $U(\alpha_0) \geq$

$Eu(\tilde{y})$ for all u(x) in U and we are done.

Assume instead that A^* is empty. We must show that this is not possible when $\tilde{x}$ is

stochastically superior to $\tilde{y}$. Let α_u denote the maximum value for α such that $U(\alpha) \geq Eu(\tilde{y})$ for

utility function u(x). That is, α_u is the right end point of the line segment A_u. Denote this set of

real numbers, one for each utility function, as $\{\alpha_u\}$. Let α_L denote the infimum or greatest lower

bound of this set $\{\alpha_u\}$. α_L exists because $\{\alpha_u\}$ is a bounded set of real numbers. Similarly, let

α_u^* denote the minimum value for α such that $U(\alpha) \geq Eu(\tilde{y})$. α_u^* is the left end point of the line

segment which is A_u. Denote this set of real numbers $\{\alpha_u^*\}$. Let α_R denote the supremum or

least upper bound of this set $\{\alpha_u^*\}$.

If $\alpha_L > \alpha_R$, then for all α in $[\alpha_R, \alpha_L]$, $U(\alpha) \geq Eu(\tilde{y})$ for all $u(x)$ in U and hence A^* is not

empty. When $\alpha_L = \alpha_R$, and α_L is in $\{\alpha_u\}$ and α_R is in $\{\alpha_u^*\}$, the same thing is true; that is

$U(\alpha_L) \geq Eu(\tilde{y})$ for all $u(x)$ in U. Thus, the hypothesis A^* empty rules this out.

If $\alpha_L < \alpha_R$, for any α_0 in $[\alpha_L, \alpha_R]$, there exists a utility function $u_1(x)$ such that $U_1(\alpha_0) <$

$Eu(\tilde{y})$ and the α^* which maximizes $U_1(\alpha)$ occurs at a point $\alpha^* < \alpha_0$ (because $\alpha_L < \alpha_0$). Similarly,

there also exists a utility function $u_2(x)$ such that $U_2(\alpha_0) < Eu(\tilde{y})$ and the α^* which maximizes

$U_2(\alpha)$ occurs at a point $\alpha^* > \alpha_0$ (because $\alpha_0 < \alpha_R$). This same thing is true when

$\alpha_L = \alpha_R$, and α_L is not in $\{\alpha_u\}$ and α_R is not in $\{\alpha_u^*\}$. Either of these two cases allows

construction of a utility function of the form $u_\lambda(x) = \lambda \cdot u_1(x) + (1 - \lambda) \cdot u_2(x)$ which satisfies

$U_\lambda(\alpha_0) < Eu(\tilde{y})$ and the maximum α^* for this utility function occurs at α_0. Hence stochastic

superiority is violated so this case is also not possible.

The final two cases to discuss are when $\alpha_L = \alpha_R$, and α_L is in $\{\alpha_u\}$ and α_R is not in $\{\alpha_u^*\}$,

or vice versa. The discussion is for the first of these cases and the second is resolved in a similar

fashion. When $\alpha_L = \alpha_R$, and α_L is in $\{\alpha_u\}$ and α_R is not in $\{\alpha_u^*\}$, there exists a utility function

$u_1(x)$ such that $U_1(\alpha_L) = Eu(\tilde{y})$ and the α^* which maximizes $U_1(\alpha)$ occurs at a point $\alpha^* < \alpha_L$.

There also exists a utility function $u_2(x)$ such that $U_2(\alpha_L) < Eu(\tilde{y})$ and the α^* which maximizes

$U_2(\alpha)$ occurs at a point $\alpha^* > \alpha_L$. This allows construction of a utility function of the form $u_\lambda(x) =$

$\lambda \cdot u_1(x) + (1 - \lambda) \cdot u_2(x)$ which satisfies $U_\lambda(\alpha_L) < Eu(\tilde{y})$ and the maximum α^* for this utility function occurs at α_L. Thus, this case is not possible either and it can be concluded that A^* is not empty when stochastic superiority of $\tilde{x}$ over $\tilde{y}$ holds. Q.E.D.

Proof of Theorem 3

Suppose that $\tilde{x}$ is stochastically superior to $\tilde{y}$ on some convex set of concave utility functions U for a given ρ. Then, from Theorem 2, there exists an α_0 in [0, 1] such that $\alpha_0 \cdot \tilde{x} + (1 - \alpha_0) \cdot \rho$ stochastically dominates $\tilde{y}$ on U. Since for $\rho' \geq \rho$, $\alpha_0 \cdot \tilde{x} + (1 - \alpha_0) \cdot \rho'$ stochastically dominates $\alpha_0 \cdot \tilde{x} + (1 - \alpha_0) \cdot \rho$, by transitivity it also stochastically dominates $\tilde{y}$ on U. Therefore, according to Theorem 1, $\tilde{x}$ is stochastically superior to $\tilde{y}$ on U for ρ'. Q.E.D.

Proof of Theorem 4

Sufficiency is obvious. To show necessity, suppose that $(\tilde{x}, \rho_x)$ is stochastically superior to $(\tilde{y}, \rho_y)$ on U for some ρ_y such that $\rho_x \geq \rho_y$. Then for every u(x) in U there exists an $\alpha \in [0, 1]$ such that $\alpha \cdot \tilde{x} + (1 - \alpha) \cdot \rho_x$ is preferred or indifferent to $\tilde{y}$. According to Theorem 2, there exists an $\alpha_0 \in [0, 1]$ such that $\alpha_0 \cdot \tilde{x} + (1 - \alpha_0) \cdot \rho_x$ stochastically dominates $\tilde{y}$ on U. From Theorem 1, this implies that $(\tilde{x}, \rho_x)$ is stochastically superior to $(\tilde{y}, \rho_x)$ on U. Q.E.D.

Proof of Theorem 5

(1) implies that $\rho_x > \rho_y$, and therefore according to Theorems 1, 2 and 4, $(\tilde{x}, \rho_x)$ is stochastically *superior* to $(\tilde{y}, \rho_y)$ in the second degree if and only if there is an α in [0, 1] such that $\alpha\tilde{x} + (1-\alpha)\rho_x$ stochastically *dominates* $\tilde{y}$ in the second degree, where $\rho_x = w - e_2 - (1 + \pi)p_2 L$.

To establish the condition under which there exists an α in [0,1] such that $\alpha\tilde{x} + (1-\alpha)\rho_x$ stochastically *dominates* $\tilde{y}$ in the second degree, note that $\alpha\tilde{x} + (1-\alpha)\rho_x$ is a binary distribution yielding a low outcome $\alpha(w - e_2 - L) + (1 - \alpha)(w - e_2 - (1 + \pi)p_2L)$ and a high outcome $\alpha(w - e_2) + (1 - \alpha)(w - e_2 - (1 + \pi)p_2L)$ with probabilities p_2 and $(1 - p_2)$ respectively. Similarly, $\tilde{y}$ is a binary distribution yielding a low outcome $w - e_1 - L$ and a high outcome $w - e_1$ with probabilities p_1 and $(1 - p_1)$ respectively. Because $p_2 < p_1$, $\alpha\tilde{x} + (1 - \alpha)\rho_x$ stochastically dominates $\tilde{y}$ in the second degree if and only if the low outcome of the former is no less than the low outcome of the latter, and the mean of the former is no less than the mean of the latter. These conditions are expressed as (3) and (4) given below:

$$\alpha(w - e_2 - L) + (1 - \alpha)(w - e_2 - (1 + \pi)p_2L) \geq w - e_1 - L \qquad (3)$$

and

$$\alpha(w - e_2) - \alpha\, p_2L + (1-\alpha)(w - e_2 - (1 + \pi)p_2L) \geq w - e_1 - p_1L. \qquad (4)$$

Now (3) and (4) can be equivalently expressed as:

$$\alpha \leq 1 - \frac{e_2 - e_1}{[1 - (1 + \pi)p_2]L} \qquad (5)$$

and

$$\alpha \geq 1 - \frac{(p_1 - p_2)L - (e_2 - e_1)}{\pi p_2 L} \qquad (6)$$

respectively.

There exists such an α that satisfies both (5) and (6) if and only if

$$\frac{(p_1 - p_2)L - (e_2 - e_1)}{\pi p_2 L} \geq \frac{e_2 - e_1}{[1 - (1 + \pi)p_2]L}$$

or equivalently

$$\pi \leq \frac{(1 - p_2)}{p_2}\left[1 - \frac{(e_2 - e_1)}{L(p_1 - p_2)}\right].$$

Q.E.D.