Who's afraid of selection bias? Robust inference in the presence of competitive selection

October 2013

Thomas Noe
Saïd Business School
University of Oxford

Working paper series | 2013-14

SAID BUSINESS SCHOOL

UNIVERSITY OF OXFORD

Who's afraid of selection bias? Robust inference in the presence of competitive selection

Thomas Noe
Balliol College/SBS

11 October, 2013

Abstract

This paper considers competitive selection dominance: what conditions on the unconditional distribution of of a random prospect will ensure that the prospect stochastically dominates a rival random prospect *conditioned* on competitive selection, i.e., conditioned on the prospect's realized value exceeding its rivals? Because standard distributional orders, such as stochastic dominance and the monotone likelihood ratio property (MLRP), do not provide either necessary or sufficient restrictions on the unconditional distributions to ensure selection dominance, new distribution orders are required. We provide the requisite orders, which we term *supermultiplicativity on average* and *geometric dominance*. These orderings generate conditions, satisfied by many, but not all, scale shifts of standard textbook distributions, under which the selection-conditioned distribution is stochastically dominant if and only if the unconditional distribution is stochastically dominant. When these conditions are satisfied, robust qualitative inferences concerning the unconditional population distribution can be drawn from the selection-conditioned subsample distribution and vice versa.

1 Introduction

This paper considers the question of when the dominance relations between distributions are preserved under endogenous competitive selection. This question is relevant in many economic contexts. In some cases, agents have information about an unconditional distribution and need to make inferences about a selection-conditioned distribution. For example, consider a firm which is going to be sold off either through an initial public offering (IPO) or a private equity buyout. Its current price reflects the unconditional value of the firm. The owners will choose the option that maximizes firm value. Since the current price reflects unconditional value, if one of the choices increases value, the other must lower it. An arbitrage trader has private information indicating that the owners will sell to a private equity firm. The trader, unlike the firm's owners, is not able to observe the value of the specific deal offered the owners. However, based on her knowledge about typical private equity and IPO deals, she can estimate the expected value of private equity buyouts and IPOs. Should the trader go long or short in the stock? Even if the arbitrage trader knows that private equity transactions on average generate more value, she must keep in mind that the owners will only select a particular IPO deal when they know it produces more value than the specific private equity deal they were offered. Can the trader be confident of earning gains on average from buying the stock?

In other cases a selection-conditioned distribution is known, and the agent aims to use this distribution to make inferences about an unconditional distribution. Consider an empirical researcher trying to infer the quality of Oxford and Cambridge graduates based on salary data for Oxbridge graduates hired by Goldman Sachs. If the researcher observes that Oxford graduates earn more, he cannot reach a conclusion that Oxford graduates are better without considering the fact that Goldman only hires the best graduates. Suppose that Goldman uses *fixed-criteria selection*, hiring only those candidates who pass a fixed threshold independent of rival candidate quality. Under this assumption, the selection inference problem would reduce to the following question: when does a distribution's being unconditionally better than another imply that it is better conditioned on sampling over any subset of realizations? When the chosen subset is the same for both distributions and independent of the realized sample, the answer to this question is provided by the monotone likelihood ratio property (MLRP) ordering. MLRP ordering implies that the MLRP dominant (and thus the stochastically dominant) distribution will remain stochastically dominant conditioned on sampling from any fixed subset of value levels.[1] Thus, if our researcher knows that the student-quality distribution is drawn from a family of MLRP ordered distributions, e.g., a family of normal distributions having the same variance, he can infer from significantly higher salaries for Goldman's Oxford hires that Oxford graduates are, in fact, better.

We address a different question from the one answered by the MLRP ordering. What if Goldman does not use fixed-criteria selection but instead uses *competitive selection*? For example, suppose that, for each opening,

[1]Technically, we must require the subset to be measurable. See Theorem 1.C.2 in Shaked and Shanthikumar (1994) for a derivation of this assertion. See Milgrom (1981) for a comprehensive discussion of the MLRP ordering in the context of economics and finance.

one Oxford and one Cambridge graduate are interviewed and Goldman picks the best candidate in the two-person pool. In this case, the threshold for hiring is not determined by a fixed cutoff independent of the quality of the other candidates but rather endogenously by the quality of the pool. Thus, the MLRP ordering criteria is not applicable. The aim of this paper is to develop ordering conditions, analogous to stochastic dominance and MLRP, that will permit ranking of distributions in such cases. We will develop necessary and sufficient conditions under which the unconditional stochastic dominance of a distribution implies that, conditioned on competitive selection, the distribution remains stochastically dominant. Perhaps more surprising, we will also develop conditions under which the unconditional stochastic dominance of a distribution implies that, conditioned on selection, the distribution is stochastically dominated.

Selection dominance implies that the expected value of draws selected from a given distribution by a value-maximizing decision maker is higher than the expected value produced when draws are selected from the rival distribution. We derive a necessary and sufficient condition for selection dominance, which we term *supermultiplicativity on average*. We find, using this criterion, that selection dominance does not generate a nice ordering over distributions. The problem with selection dominance is that it is not transitive. In other words, it is possible that $\tilde{X}$ selection dominates $\tilde{Y}$ and $\tilde{Y}$ selection dominates $\tilde{Z}$ but $\tilde{X}$ does not selection dominate $\tilde{Z}$. However, there exists a sufficient condition for selection dominance which does generate an order relation over random variables and their associated distribution functions—ordering by *geometric convexity* or simply geometric dominance. Moreover, many families of distributions generated by scale shifts of common textbook statistical distributions, e.g., exponential, Weibull, lognormal, are ordered in the same fashion by both stochastic dominance and geometric dominance. This implies that, for many important decision problems, unconditional rankings of population distributions can be used to infer selection-conditioned rankings and selection-conditioned rankings can be used to infer unconditional rankings.

The geometric convexity ordering can be informally described as the monotone likelihood ratio ordering (MLRP) on log-log graph paper. Recall that a distribution F dominates distribution G in the MLRP ordering if and only if the ratio of the distributions' derivatives, i.e., their probability densities, $F'/G' = f/g$, is increasing. When we specialize our order relations to the case where the distributions being compared have densities, we find that distribution F dominates distribution G under the geometric convexity ordering if and only if the ratio of the distributions' log derivatives $\log(F)'/\log(G)'$ is increasing. We show that when F dominates G in both the geometric convexity and stochastic dominance ordering, F also dominates G in the MLRP ordering. The converse implication does not hold, i.e., there exist cases in which one distribution dominates another in the MLRP ordering (and thus a fortiori with respect to stochastic dominance) but not in the geometric convexity ordering.

In fact, distribution F may dominate distribution G in the geometric convexity ordering and, at the same time, *G dominates F* under the MLRP ordering. In this case, selection is *certain* to reverse dominance, i.e., the stochastically dominant distribution surely yields a lower expected value conditioned on selection than the dominated distribution. Roughly speaking this "perverse" result occurs when the geometrically dominant distribution has a much fatter left tail than the geometrically dominated distribution. The fat left tail ensures both that the geometrically dominant distribution is rarely selected and that, when it is selected, its realized value is fairly high. These effects permit the geometrically dominant distribution to dominate the stochastically dominant distribution under competitive selection.

The attentive reader at this point might have noted that we have discussed a sufficient condition for selection dominance—geometric convexity—a great deal, without discussing selection dominance itself beyond noting that it is not an order relation. The reader might have also noted that, in our discussion of geometric convexity, distributions were always either stochastically dominated or stochastically dominant. This leads to the question of the role of selection dominance absent geometric convexity and the role of geometric dominance when random variables are not ordered by stochastic dominance. We have some characterizations that link these two questions. First, we show that a geometrically dominant distribution need not be either stochastically dominant or stochastically dominated. When the geometrically dominant distribution is neither stochastically dominant or dominated, the geometrically dominant distribution always crosses the geometrically dominated distribution once from above, implying that the geometrically dominant distribution has higher dispersion. Next, we show that a dispersion-increasing transformation, a transformation which leaves the probability of selection fixed while increasing tail weight, always leads to the transformed variable to be selection dominant even when the transformation does not produce geometric dominance.

The paper is organized as follows. Section 2 formalizes our research question. Section 3 develops necessary and sufficient conditions for selection and geometric dominance. Section 4 considers the order relations induced by selection dominance and geometric dominance. Section 5 characterizes the restrictions the geometric convexity order places on distributional properties. Section 6 analyzes geometric dominance in the case where it is consistent with stochastic dominance. Section 7 considers the case where the geometric dominance order reverses the stochastic dominance order. Section 8 considers geometric dominance in the absence of stochastic dominance. Section 9 considers selection dominance in the absence of geometric dominance.

2 Question

Suppose that $\tilde{X}$ and $\tilde{Y}$ are two independent random variables with distribution functions F and G respectively. One of these random variables will be selected by a decision maker. The selection will determine the value the decision maker receives. Let v be the decision maker's valuation function. Assume that v is strictly increasing the realization of the selected random variable. A value-maximizing decision maker will select $\tilde{X}$ whenever the the value it produces, $v(\tilde{X})$, exceeds the value produced by the alternative choice, $\tilde{Y}$. Because v is an increasing function, $v(\tilde{X}) > v(\tilde{Y})$ if and only if $\tilde{X} > \tilde{Y}$. The expected value of $\tilde{X}$ conditioned on $\tilde{X}$ being selected is thus

$\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y}]$. A natural question to pose in this context is when will value conditioned on $\tilde{X}$ being selected exceed value conditioned on $\tilde{Y}$ being selected? This motivates the following definition:

Definition. We say that $\tilde{X}$ (or its distribution function F) *selection dominates* $\tilde{Y}$ (or its distribution function G) if, for all increasing functions v,

$$\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y}] \geq \mathbb{E}[v(\tilde{Y})|\tilde{Y} > \tilde{X}]. \tag{1}$$

If the weak inequality between the conditional expectations in equation (1) is replaced by a strong inequality, we will say that F strictly selection dominates G. This paper considers the question of the restrictions that must be imposed on the unconditional distribution functions of $\tilde{X}$ and $\tilde{Y}$, F and G, in order to ensure that $\tilde{X}$ *selection dominates* $\tilde{Y}$. We also aim to determine when the selection dominant distribution is, in fact, the better distribution unconditionally, i.e. in the absence of selection. The criterion for dominance in the absence of selection is well known—stochastic dominance:

Definition. We say that $\tilde{X}$ *stochastically dominates* $\tilde{Y}$ if for all increasing functions v

$$\mathbb{E}[v(\tilde{X})] \geq \mathbb{E}[v(\tilde{Y})]. \tag{2}$$

A well-known result in economics and statistics is that $\tilde{X}$ *stochastically dominates* $\tilde{Y}$ if and only if $F(x) \leq G(x)$. Thus, stochastic dominance defines a partial order over distribution functions which determines when a decision maker with an increasing value function will prefer one distribution to another. What about selection dominance?

3 Basic results

We now turn to formalizing our problem. To avoid the problem of ties and indeterminacies, we impose the following restrictions on the distribution functions we consider:

Definition. Distribution functions F and G are an *admissible pair of distributions* if

1. F and G have common support $[\underline{x}, \bar{x}]$, $0 \leq \underline{x} < \bar{x} \leq \infty$.

2. F and G are continuous and mutually absolutely continuous.

3. $\int_0^\infty x\,dF(x) < \infty$ and $\int_0^\infty x\,dG(x) < \infty$.

A collection of distribution functions, $\{F_\alpha\}_{\alpha \in A}$ is admissible if, for all pairs $\alpha_1 \in A$, $\alpha_2 \in A$ in the collection, the pair F_{α_1} and F_{α_2} is admissible.

The condition that F and G have interval support is not necessary to derive our results. However, absent this assumption, stating some of our results would become more cumbersome. If we allowed for gaps in the support, we would need to identify points $x' \neq x''$ in $[\underline{x}, \bar{x}]$, at which $F(x') = G(x') = G(x'') = F(x'')$, and then state our results in terms of the resulting equivalence classes. The assumption that F and G are continuous and are mutually absolutely continuous eliminates the problem of ties and ensures that their inverse functions are continuous and increasing. Note that we do not assume that F are G are absolutely continuous with respect to Lebesgue measure and thus have associated probability density functions. The restriction of the support of the distributions to the non-negative real line is made simply for convenience. The assumption that F and G have finite expectations is made to ensure the expected value of the simple valuation function $v(x) = x$ is finite. Note that we allow for $\bar{x} = \infty$, and thus we do not assume a compact support for the distributions being compared.

For admissible pairs of distributions, we can express conditioning on selection as follows:

$$\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y}] = \frac{\mathbb{E}[v(\tilde{X})\, I_{\tilde{X}>\tilde{Y}}]}{\mathbb{E}[I_{\tilde{X}>\tilde{Y}}]} \quad \text{and} \quad \mathbb{E}[v(\tilde{Y})|\tilde{Y} > \tilde{X}] = \frac{\mathbb{E}[v(\tilde{Y})\, I_{\tilde{Y}>\tilde{X}}]}{\mathbb{E}[I_{\tilde{Y}>\tilde{X}}]}, \tag{3}$$

where, in the above expressions, I_S represents the indicator function for set S. By the independence of $\tilde{X}$ and $\tilde{Y}$ and Tonelli's Theorem, we have that

$$\mathbb{E}[v(\tilde{X})\, I_{\tilde{X}>\tilde{Y}}] = \int_{\underline{x}}^{\bar{x}} \int_{\underline{x}}^{\bar{x}} v(x)\, I_{y<x}\, dF(x)\, dG(y) = \int_{\underline{x}}^{\bar{x}} v(x) \left(\int_{\underline{x}}^{\bar{x}} I_{y<x}\, dG(y) \right) dF(x) = \int_{\underline{x}}^{\bar{x}} v(x)\, G(x)\, dF(x), \tag{4}$$

Using the same reformulation, we can express the other components of the conditional expectations as follows:

$$\mathbb{E}[v(\tilde{Y})\, I_{\tilde{Y}>\tilde{X}}] = \int_{\underline{x}}^{\bar{x}} v(x)\, F(x)\, dG(x); \quad \mathbb{E}[I_{\tilde{X}>\tilde{Y}}] = \int_{\underline{x}}^{\bar{x}} G(x)\, dF(x); \quad \mathbb{E}[I_{\tilde{Y}>\tilde{Y}}] = \int_{\underline{x}}^{\bar{x}} F(x)\, dG(x). \tag{5}$$

Using the expressions in (4) and (5) we can express the conditional expectations as follows:

$$\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y}] = \frac{\int_{\underline{x}}^{\bar{x}} v(x) G(x)\, dF(x)}{\int_{\underline{x}}^{\bar{x}} G(x)\, dF(x)} \quad \text{and} \quad \mathbb{E}[v(\tilde{Y})|\tilde{Y} > \tilde{X}] = \frac{\int_{\underline{x}}^{\bar{x}} v(x) F(x)\, dG(x)}{\int_{\underline{x}}^{\bar{x}} G(x)\, dF(x)}. \tag{6}$$

Define the probability distribution functions H and J by

$$H(x) = \frac{\int_{\underline{x}}^{x} G(s)\, dF(s)}{\int_{\underline{x}}^{\bar{x}} G(s)\, dF(s)} \quad \text{and} \quad J(x) = \frac{\int_{\underline{x}}^{x} F(s)\, dG(s)}{\int_{\underline{x}}^{\bar{x}} G(s)\, dF(s)}. \tag{7}$$

Using H and J, we can express the conditioning relation as a simple expectation with respect to the distribution functions as follows:

$$\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y}] = \int_{\underline{x}}^{\bar{x}} v(x)\, dH(x) \quad \text{and} \quad \mathbb{E}[v(\tilde{Y})|\tilde{Y} > \tilde{X}] = \int_{\underline{x}}^{\bar{x}} v(x)\, dJ(x). \tag{8}$$

Thus, for inequality (1) to hold, it is necessary and sufficient that H stochastically dominates J, i.e., $H(x) \leq J(x)$. From equation (7), we see that stochastic dominance is equivalent to the condition that for all $x > \underline{x}$,

$$\frac{\int_{\underline{x}}^{x} G(s)\, dF(s)}{\int_{\underline{x}}^{x} F(s)\, dG(s)} \leq \frac{\int_{\underline{x}}^{\bar{x}} G(s)\, dF(s)}{\int_{\underline{x}}^{\bar{x}} G(s)\, dF(s)}. \tag{9}$$

Integration by parts shows that

$$\begin{aligned} \int_{\underline{x}}^{x} G(s)\, dF(s) &= F(x)\, G(x) - \int_{\underline{x}}^{x} F(s)\, dG(s), \\ \int_{\underline{x}}^{\bar{x}} G(s)\, dF(s) &= 1 - \int_{\underline{x}}^{\bar{x}} F(s)\, dG(s). \end{aligned} \tag{10}$$

Substitution of equation (10) into equation (9) yields the following equivalent condition for (9)

$$\frac{1}{F(x)\, G(x)} \int_{\underline{x}}^{x} F(s)\, dG(s) \geq \int_{\underline{x}}^{\bar{x}} F(s)\, dG(s). \tag{11}$$

3.1 The quantile transformation function u

The key to deriving distributional restrictions inequality (11) is noting that only the behavior of the distributions relative to each other matters. Transforming both distributions by the same continuous increasing function will not affect the validity of (11). This permits us to reduce the dimensionality of the problem by using the quantile transform function. For any distribution function, G, define the inverse or quantile function of G, G^{-1} by

$$G^{-1}(t) = \inf\{x \in [0,\infty]\colon G(x) \geq t\},\ t \in [0,1].$$

Note that, because regular pairs of distributions increase over their support, G and F have well-defined increasing inverse functions. We represent the inverse function of F with F^{-1} and the inverse function of G with G^{-1}. Thus, the function $u = F \circ G^{-1}$ is well defined and $F = u \circ G$, where $\circ$ represents functional composition. Note that $u = F \circ G^{-1}$ maps quantiles of G into quantiles of F, e.g., $u(0.50) = 0.25$ implies that the median of distribution G equals the first quartile of distribution F. Thus, if for some $x^o \in [\underline{x}, \bar{x}]$, $F(x^o) \geq (\leq) G(x^o)$, then letting $t^o = G(x^o)$, we see that $u(t^o) \geq (\leq) t^o$. Similarly, a point at which the distribution functions meet maps into a unique point in the unit interval at which $u(t) = t$ and vice versa. We will refer to u simply as the *transform function*. The fact that G and F are continuous and increasing on their common support implies that $u(0) = 0$, $u(1) = 1$ and that u is increasing and continuous. We term the set of functions that have these properties admissible.

Definition. If a function $u\colon [0,1] \to [0,1]$ is increasing and continuous, with $u(0) = 0$ and $u(1) = 1$, we will call u an *admissible function*.

3.2 Conditions for selection dominance

Using the u transform, we can express equation (11) as

$$\frac{1}{u \circ G(x)\, G(x)} \int_{\underline{x}}^{x} u \circ G(s)\, dG(s) \geq \int_{\underline{x}}^{\bar{x}} u \circ G(s)\, dG(s). \tag{12}$$

Using the change of variables formula on on equation (12) shows that $\tilde{X}$ selection dominates $\tilde{Y}$ if and only if the following condition holds:

$$\text{For all } t \in [0,1], \quad \frac{1}{t}\int_0^t \frac{u(s)}{u(t)}\,ds \geq \int_0^1 u(s)\,ds. \tag{13}$$

For any admissible u, define

$$\Pi[u](t) = \frac{1}{t}\int_0^t \frac{u(s)}{u(t)}\,ds, \quad t \in (0,1]. \tag{14}$$

The continuity of u implies that $\Pi[u]$ is a continuous function defined on $(0,1]$. Using the Π representation, a necessary and sufficient condition for condition (14) to hold is that

$$\forall t \in (0,1), \quad \Pi[u](t) \geq \int_0^1 u(s)\,ds \equiv \Pi[u](1). \tag{15}$$

Note that we can express expression (15) in the following alternative forms by a change of variables in the integral:

$$\text{For all } t \in (0,1) \int_0^1 \frac{u(t\,s)}{u(t)}\,ds \geq \int_0^1 u(s)\,ds. \tag{16}$$

Or, equivalently

$$\text{For all } t \in (0,1), \int_0^1 \Big(u(t\,s) - u(s)\,u(t)\Big)\,ds \geq 0. \tag{17}$$

Thus, the question of selection dominance reduces to the question of what restrictions that must be imposed on the transformation $u = F \circ G^{-1}$ in order to ensure that (17) holds. The following result, which is essentially specialized and simplified statement of Theorem 1 in Finol and Wójtowicz (2000), will prove to be of great assistance in answering this question.

Lemma 1. *If u is admissible, then the following statements are equivalent:*

(i) For all $s \in (0,1]$ and $t \in (0,1]$, the function $t \to u(t\,s)/u(t)$ is nonincreasing

(ii) u is geometrically convex, *i.e., for all $(s,t) \in (0,1] \times (0,1]$ and $\alpha \in (0,1)$, $u(s^{\alpha}\,t^{1-\alpha}) \leq u(s)^{\alpha}\,u(t)^{1-\alpha}$.*

(iii) The conjugate function $\hat{u} : (-\infty,0) \to (-\infty,0)$ defined by $\hat{u}(y) = \log \circ u \circ \exp(y)$ is continuous, convex and increasing with $\lim_{y\to 0} \hat{u}(y) = 0$ and $\lim_{y\to -\infty} \hat{u}(y) = -\infty$

The three equivalent conditions, (i), (i), *and* (iii) *imply that u is supermultiplicative, i.e.,*

$$\forall (s,t) \in (0,1] \times (0,1], \;\; u(st) \geq u(s)\,u(t). \tag{18}$$

Proof. Theorem 1 in Finol and Wójtowicz (2000) establishes all of the results except the assertions that $\hat{u}$ is increasing with $\lim_{y\to 0} \hat{u}(y) = 0$ and $\lim_{t\to -\infty} u(t) = -\infty$. Because $\hat{u} = log \circ u \circ \exp$, $\hat{u}$ being increasing follows because u is regular and thus increasing and log and exp are increasing functions. $\lim_{y\to 0} \hat{u}(y) = 0$ follows because u, exp, and log are continuous and increasing and $u(0) = 0$; thus,

$$\lim_{y\to 0} \hat{u}(y) = \log(1) = 0. \tag{19}$$

In the sequel, we extend the definition of $\hat{u}$ to 0 by setting its value at 0 to its limiting value when approaching 0, which is 0. That is, we define $\hat{u}(0) = 0$. Next note that $\lim_{y\to-\infty} \hat{u}(y) = -\infty$ because, u, exp, and log are continuous and increasing and $u(0) = 0$; thus,

$$\lim_{y\to-\infty} \hat{u}(y) = \log_{y\to-\infty} (u(\exp[y])) = \lim_{t\to 0} \log(t) = -\infty. \tag{20}$$

□

Thus, a sufficient condition for the two equivalent conditions for selection dominance, (16) and (17), to hold is that u is geometrically convex. The characterization of geometric convexity provided by condition (iii) of Lemma 1 provides some intuition for the properties of geometrically convex functions. The geometric convexity of u is equivalent to $\log u(t) = \hat{u}(\log t)$, with $\hat{u}$ convex. In other words, $\log u(t)$ is a convex function of $\log(t)$. Thus, geometric convexity is equivalent to u being convex when plotted on a graph in which both the ordinate and abscissa have been log scaled. Geometric convexity is a weaker condition than logarithmic convexity. Logarithmic convexity requires that the logarithm of the function is convex. Logarithmic convexity holds whenever for all $\alpha \in (0,1]$, $u(\alpha s + (1-\alpha)t) \le u(s)^{\alpha} u(t)^{1-\alpha}$.

Thus, by the geometric mean–arithmetic mean inequality, logarithmic convexity implies geometric convexity (and in fact, by a different argument, it implies convexity as well) but the converse implication does not hold. Logarithmic convexity is equivalent to the function being convex when plotted on a graph in which only the ordinate has been log scaled. As we we will see in many subsequent examples, in general, geometric convexity neither implies nor is implied by convexity.

These observations allow us to define conditions on u which ensure selection dominance. If expression (17) holds we will say that u is *supermultiplicative on average*. If expression (17) holds with the weak inequality replaced by a strong inequality we will say that u is *strictly supermultiplicative on average*. The motivation for describing the satisfaction of expression (17) as supermultiplicativity on average is apparent from the definition of supermultiplicativity given by expression (18). Based on Lemma (1), we will say that u is *geometrically convex* if condition (iii) of Lemma 1 is satisfied. If this condition is satisfied by a conjugate function which is strictly convex we will say that u is *strictly geometrically convex*.

Theorem 1. *Suppose that F and G are an admissible pair of distribution functions; Let $u = F \circ G^{-1}$, then*

(i) F (strictly) selection dominates G, if and only if u satisfies the (strict) supermultiplicative on average condition, expression (17).

(ii) If u is (strictly) geometrically convex, F (strictly) selection dominates G.

Proof. Part (i) follows from the derivation above. Part (ii) follows from (17) and part (i) of Lemma 1. □

4 Geometric and selection dominance as ordering

In the case of unconditional comparisons of distributions, the stochastic dominance partial order, $\underset{sd}{\succcurlyeq}$, defined by

$$F \underset{sd}{\succcurlyeq} G \text{ if } F(x) \leq G(x), x \in (\underline{x}, \bar{x}), \tag{21}$$

is a necessary and sufficient condition for F to dominate G. If the inequality in expression (21) is strict we will say that F strictly stochastically dominates G. Our aim is to find a analogous partial order in the presence of selection. Thus, we will define relations between distribution functions based on the selection dominance and geometric convexity and examine the extent to which we can order random variables in the presence of selection.

Definition. Suppose that F and G are an admissible pair of distributions and that $u = F \circ G^{-1}$, then

(i) If (17) holds we will say that F *selection dominates* G under the average supermultiplicative relation; If (17) holds with the weak inequality replaced by a strong inequality, we will say that F strictly selection dominates G.

(ii) If u is geometrically convex, we will say that F dominates G under the geometric convexity relation. If u is strictly geometrically convex, we will say that F strictly dominates G under the geometric convexity relation.

As the next lemma, Lemma 2, reports, the selection dominance relation is not transitive and thus not even a preorder over the set of distribution functions. Thus, in the presence of selection, there is no order relation over distribution functions which is necessary and sufficient to ensure that one distribution dominates another. However, geometric convexity is a preorder over distribution functions and as we showed in Proposition 1, geometric convexity is a sufficient condition for selection dominance. Thus, the Lemma shows that it is impossible to derive a distributional order that completely determines whether a random variable is selection dominant but an ordering does exist that will capture a subset of selection-dominance relations. Whether this result has any value, of course, depends on whether the subset of selection dominance relations captured by the ordering is large and interesting. This is the topic we will begin to address in the next section of the paper.

Lemma 2. *(i) The geometric convexity relation is a preorder over the set of distributions. (ii) The selection dominance relation is not a preorder over the set of distributions because selection dominance is not transitive.*

Proof. To prove (i) note that we need to show that the geometric convexity relation is is reflexive and transitive. If F, G, and K are three admissible distribution functions, then the relation is transitive (i) if F dominates F, and transitive (ii) if F dominates G and G dominates K implies that F dominates K. To show this, note that, by definition, F dominates itself if the the function $u = F \circ F^{-1}$ is geometrically convex. Because $u = F \circ F^{-1}$ is the identity, its geometric convexity is immediate. Now consider transitivity. Transitivity will hold whenever the functions, $u_1 = F \circ G^{-1}$ and $u_2 = G \circ K^{-1}$ being geometrically convex implies that the function $u_3 = G \circ K^{-1}$ is geometrically convex. Geometric convexity holds if and only if $\hat{u}_3 = \hat{u}(y) = \log \circ u_3 \circ \exp$ is convex. Because $u_3 = u_1 \circ u_2$,

$$\hat{u}_3 = \hat{u}(y) = \log \circ u_1 \circ u_2 \circ \exp = (\log \circ u_1 \circ \exp) \circ (\log \circ u_2 \circ \exp) = \hat{u}_1 \circ \hat{u}_2.$$

Because u_1 and u_2 are geometrically convex, $\hat{u}_1$ and $\hat{u}_2$ are convex. Because, u_1 and u_2 are nondecreasing, $\hat{u}_1$ and $\hat{u}_2$ are nondecreasing. The composition of nondecreasing convex functions is convex. Thus, $\hat{u}_1 \circ \hat{u}_2$ is convex. Thus, the geometric convexity relation is a preorder. However, it is not a partial order because the geometric dominance relation fails to satisfy antisymmetry, i.e., F dominating G and G dominating F in the order does not imply that $F = G$. This can be seen by taking $F(x) = x^2$ and $G(x) = x$, $x \in [0,1]$, then $\hat{u}(y) = 2y\, y < 0$. Because $\hat{u}$ is linear F dominates G and G dominates F in the geometric convexity ordering but $F \neq G$. Thus, the geometric dominance relation is not antisymmetric. Now consider(i). The selection dominance order is clearly reflexive; however, it is not transitive. This can be verified by a counterexample available in Appendix C. Thus, selection dominance is not a preorder. □

We represent the order relation defined by geometric convexity in the standard fashion. When F dominates G in the geometric convexity ordering, we will write $F \succcurlyeq_g G$. Also, to avoid unnecessary verbiage, when F dominates G in the geometric convex ordering we will simply say that F geometrically dominates G.

As pointed out in the proof of Lemma 2, geometric dominance is not a partial order relation because it is possible for distinct distributions, F and G, $F \succcurlyeq_g G$ and $G \succcurlyeq_g F$ When this occurs, we will say that F and G are *geometrically equivalent*. If, for an admissible pair of distributions, F and G, F selection dominates G and G selection dominates F, we will say that the pair is *selection equivalent*. The pair of distributions is selection equivalent if and only if, conditioned on selection, the expected value of any increasing value function is the same under both distributions. In this case the differences in the distributions are reflected entirely in their probability of being selected. In general, geometric dominance is a sufficient but not necessary condition for selection dominance. However, for selection equivalence, it is both a necessary and sufficient condition. As the next lemma shows, selection equivalence imposes a very strong condition on distribution functions: they must be related by a power transform.

Lemma 3. *For an admissible pair of distributions functions, F and G, the following statements are equivalent:*

(i) F and G are geometrically equivalent

(ii) F and G are selection equivalent

(iii) $F(x) = G(x)^p$ for some $p > 0$.

Proof. See Appendix A.

Examples of equivalences among textbook distributions are not hard to find. For example, clearly, all distributions of the form $F(x) = x^p$, $x \in [0,1]$ are geometrically equivalent to each other. A less obvious example is provided by Fréchet distributions. All Fréchet distributions that have the same shape parameter α are geometrically equivalent. To see this, note that if two Fréchet distributions, F_1 and F_2, differ only by the scale parameter, σ, then

$$F_1(x) = \exp\left[-\left(\frac{x}{\sigma_1}\right)^{-\alpha}\right] \text{ and } F_2(x) = \exp\left[-\left(\frac{x}{\sigma_2}\right)^{-\alpha}\right], \quad x > 0, \tag{22}$$

then, letting $p = (\sigma_2/\sigma_1)^{-\alpha}$, we see that $F_2(x) = (F_1(x))^p$.

5 Characterizing geometrically dominant distributions

We now turn to characterizing the restrictions that geometric dominance imposes on the relation between the dominant and dominated distribution. As the next theorem shows, geometric dominance imposes strong restrictions. In essence, the geometrically dominant distribution either lies below the the dominated distribution, it which case it is stochastically dominant, lies above the dominated distribution, in which case it is stochastically dominated, or crosses the dominated distribution once from below. Thus geometric dominance rules out crossings from above or multiple crossings of the distribution functions. This is a stronger restriction on distribution functions than that imposed by, for example, second-order stochastic dominance, which is consistent with multiple crossings.

Theorem 2. *Suppose that F and G are an admissible pair of distributions and let $u = F \circ G^{-1}$. Suppose that F geometrically dominates G, i.e, that u is geometrically convex.*

(i) If, on some neighborhood of $\underline{x}$, $F(x) < G(x)$, then for all $x \in (\underline{x}, \bar{x})$, $F(x) < G(x)$, and thus F strictly stochastically dominates G

(ii) If, on some neighborhood of $\underline{x}$, $F(x) > G(x)$, then either

(a) $F(x) > G(x)$ for all $x \in (\underline{x}, \bar{x})$ and thus G strictly stochastically dominates F, or

(b) There exists a point $x^o \in (\underline{x}, \bar{x})$ such that for all $x \in (\underline{x}, x^o)$, $F(x) \geq G(x)$ and for all $x \in (x^o, \bar{x})$, $F(x) \leq G(x)$. If F strictly geometrically dominates G, then for all $x \in (\underline{x}, x^o)$, $F(x) > G(x)$ and for all $x \in (x^o, \bar{x})$, $F(x) < G(x)$. In this case, if the mean payoff under F is the same as the mean payoff under G, then F is a mean-preserving risk-increasing shift of G as defined by Diamond and Stiglitz (1974).

Proof (Sketch). The formal proof of this result is presented in Appendix A. The intuition behind the proof of this result is that geometric convexity implies convexity of the transform function when this function is plotted using logarithmic scaling. This convexity imposes strong restrictions on the behavior of the underlying distributions. At quantiles where the two distribution functions, F and G, cross, i.e., points where $F(x) = G(x)$, the transform function, u meets the identity function at a corresponding point, i.e., $u(t) = t$. The conjugate function, $\hat{u}$, which is just the log scaled u function, also meets the its identity function at a corresponding point in the log-scaled space. The convexity of the conjugate function then places strong restrictions on how and how often it can meet the identity function. If it starts below the identity function, it can only meet it once. However, since the conjugate must meet the identity at its endpoint, which corresponds to the point where both distribution functions equal 1, it cannot meet the identity at any other point. If the conjugate function starts out above the identity, convexity implies that it can cross the identity at most twice. Again, because one of these crossings must occur at the endpoint, it crosses the identity at most once before reaching the endpoint. Translating these properties back from the log scaled space restricts the u transform function's crossings of its identity function, and translating the u functions behavior back to the underlying distributions yields the results.

Theorem 2 shows that the geometric convexity imposes strong conditions on how distributions ordered by geometric dominance cross. These conditions, in turn, permit us to characterize to some extent the shape of the transform function, u. These characterizations are provided by the next theorem, Theorem 3.

Theorem 3. *Suppose that F and G are an admissible pair of distributions and let $u = F \circ G^{-1}$; Assume that F geometrically dominates G, i.e., u is geometrically convex and that, on some neighborhood of $\bar{x}$, either $F(x) > G(x)$ or $F(x) < G(x)$. Then, one of the following three mutually exclusive characterizations of the distributions and transform function must hold.*

(i) $F(x) < G(x) x \in (\underline{x}, \bar{x})$*, and u is convex. If u is strictly geometrically convex, u is strictly convex.*

(ii) $F(x) > G(x) x \in (\underline{x}, \bar{x})$ *and $t \to u(t)/t$ is nonincreasing. In this case, if u is strictly geometrically convex, then $t \to u(t)/t$ is decreasing and $\lim_{t \to 0} u(t)/t = \infty$.*

(iii) On some neighborhood of $\underline{x}$, $F(x) > G(x)$ and on some neighborhood of $\bar{x}$, $F(x) \leq G(x)$. In this case, there exists $t^o \in (0,1)$ with $u(t^o) \leq t^o$ such that $t \to u(t)/t$ is nonincreasing for $t \leq t^o$ and u is convex for $t \geq t^o$. If u is strictly geometrically convex and, on some neighborhood of $\bar{x}$, $F(x) < G(x)$, then there exists $t^o \in (0,1)$ with $u(t^o) \leq t^o$, such that $t \to u(t)/t$ is decreasing for $t \leq t^o$ and $\lim_{t \to 0} u(t)/t = \infty$ and, for $t > t^o$, u is strictly convex.

Proof (Sketch of proof). For a formal proof, see Appendix A. The basic idea behind the proof is that the fact that F lies below G implies that $\hat{u}(y) - y \leq 0$, where $\hat{u}(y)$ is the conjugate function to the transform function, u. Because $\hat{u}(y)$ is convex, $\hat{u}(y) - y$ is convex. Thus, if $\hat{u}(y) - y$ were decreasing at some point, by convexity, its slope would have to be at least as small at all points less than that point. Because the domain of $\hat{u}(y)$ is the entire non-positive real line, this would imply that eventually $\hat{u}(y) - y > 0$, which is impossible. This argument shows that $\hat{u}(y) - y$ is nondecreasing. Translating the result back to the transform function, u, is then shown to imply that u is convex. This argument is used to prove part (i). The other parts of the theorem are proved in a similar fashion by showing that F lying above G and the convexity of the conjugate function imply that $\hat{u}(y) - y$ is decreasing and then translating this result back to the transform function, u.

First note that the condition that, on some neighborhood of $\bar{x}$, either $F(x) > G(x)$ or $F(x) < G(x)$, is harmless. If it were not satisfied, then because the distribution functions are continuous, there would be an interval around the lower endpoint where the distributions were equal. In that case, we could extend our proof by conditioning on both distributions exceeding this endpoint and then applying the Theorem. To avoid trying the reader's patience even more than we already have, we will not formalize this argument. Thus, in essence, Theorems 2 and 3 divide geometric dominance relations into three possible configurations, and within each configuration, place strong restrictions on the way the underlying distributions cross and on the shape of the transform function u. When the geometrically dominant distribution is also stochastically dominant, u is convex. Intuitively this implies that the transform takes quantiles of the geometrically dominated distribution and reduces them at an ever increasing rate. The higher the quantile, the bigger the shift. In Section 6 we will show that when the distributions have densities, the convexity of the shift implies MLRP ordering between the distributions. When the geometrically dominant distribution is stochastically dominated, we obtained a weaker but still significant restriction on the shape of the transform function–that $t \to u(t)/t$ is decreasing. When geometric dominance is strict, the limit of $u(t)/t$ as t approaches 0 is infinite. This implies that the transform shifts quantiles upward at a rate approaching infinity around the lowest quantiles of the stochastically dominant but geometrically dominated distribution. Thus, the geometrically dominant distribution has a much higher probability of producing realizations near the bottom of the common support of the two distributions. This is the basis for the selection dominance of the geometrically dominant distribution. The very low realizations near the bottom of the support are almost never selected when they are realized. Thus, conditioned on selection, the value of the geometrically dominant distribution can be higher despite its being stochastically dominated. When the geometrically dominant distribution is neither dominant nor dominated stochastically, the shape restrictions are a mixture of those obtained in the two other cases. At the low end of the support, the shift greatly increases the probability of low realizations. At the high end, the shift lowers the quantiles at an increasing rate. Thus, loosely, because we do not mean to impose an equality of means between the two distributions, we can say in this case, that the geometrically dominant distribution is dispersive, increasing weight on both high and low realizations. Finally, note that the point t^o, defined for this case in Theorem 3 part (iii), is not a cutoff between the regions where $t \to u(t)/t$ is decreasing and u is convex. It might well be the case that $u(t)/t$ continues to decrease after u becomes convex. $u(t)/t$ is a measure of the average slope of u starting from 0. Even when the slope starts to increase at t^o, its level may still be below the average of the slope between 0 and t^o and thus $u(t)/t$ may well continue to decrease. In fact, $t \to u(t)/t$ being decreasing over its entire domain is not inconsistent with u becoming convex after t reaches t^o.

It is not possible to strengthen the results in part (ii) of Theorem 3 from its assertion that $t \to u(t)/t$ is decreasing to the assertion that u is concave. A simple counterexample to concavity is provided by

$$u(t) = \exp\left[\max\left[-\sqrt{-\log(t)}, 2^{-1/2}\log(t)\right]\right],\ t \in (0,1],\ u(0) = 0.$$

u is admissible at it is continuous and increasing, $u(0) = 0$, and $u(1) = 1$. The conjugate function to u is given by $\hat{u}(y) = \max\left[-\sqrt{-y}, 2^{-1/2}y\right],\ y \leq 0$. $\hat{u}(y)$ is convex as it is the maximum of two convex functions. $u(t) > t,\ t \in (0,1)$. Thus, u satisfies the conditions of Theorem part (ii) of Theorem 3. However, the derivative from the right of u evaluated at $t = e^{-2}$ exceeds the derivative from the the left evaluated at the same point, which is inconsistent with concavity.

5.1 Geometric convexity in terms of probability densities

Theorems 2 and 3 establish a natural structure for relating geometric convexity to standard order relations and for investigating the distributional conditions implied by geometric convexity. We have three cases: one where the geometrically dominant distribution is stochastically dominant, in which case the transform function, u, must be convex, the second where the geometrically dominant distribution is stochastically dominated, and the third, where the geometrically dominant distribution is dispersion increasing. We investigate each of these cases in the subsequent sections. First, we require a few preliminary results that will provide a simple test for geometric dominance when the distribution functions have densities and characterize these density functions based on the behavior of the transform function u. The first of these results, Lemma 4, is a simple test for geometric convexity when u is differentiable.

Lemma 4. *Suppose u is differentiable, then u is (strictly) geometrically convex if and only if the function R defined by*

$$R(t) = \frac{u'(t)\,t}{u(t)}, \tag{23}$$

is (increasing) nondecreasing over $(0,1]$

Proof. The result then follows from observations offered after Theorem 1 in Finol and Wójtowicz (2000). □

We aim to characterize the relation between positive geometric dominance and the standard orderings of distributions used in economics and finance. These orderings are typically formulated under the assumption that the random variables under consideration have absolutely continuous distribution functions with respect to Lebesgue measure and thus are characterized by their probability density functions. We have not imposed this assumption thus far. In order to relate geometric dominance to these ordering relations, we will have to impose sufficient regularity conditions. To avoid tedious discussions and lengthy statements of proofs, we impose a single regularity condition that will be sufficient for all of the subsequent derivations.

Definition. An admissible pair of distribution functions F and G are *regularly related* if

(i) F and G are twice differentiable.

(ii) On $(\underline{x}, \bar{x})$, their probability density functions, f and g are positive.

Regularity implies that u is differentiable and thus Lemma 4 can be used to verify that u is geometrically convex. It also implies that F and G have probability density functions. Under regularity, Lemma 4 permits us to produce a simple mapping between the properties of u, such as convexity and geometric convexity, and the properties of the underlying distribution and density functions of the random variables generating u. This mapping is provided by by the next result, Lemma 5.

Lemma 5. *Suppose that F and G are regularly related and $u = F \circ G^{-1}$.*

(i) u is (strictly) convex if and only if $x \to f(x)/g(x)$ is (increasing) nondecreasing over $(\underline{x}, \bar{x})$, i.e. F dominates G in the MLRP order.

(ii) u is (strictly) geometrically convex if and only if $x \to \frac{f(x)}{g(x)} \frac{G(x)}{F(x)}$ is (increasing) nondecreasing over $(\underline{x}, \bar{x})$.

(iii) u is (strictly) logarithmically convex if and only if $x \to \frac{f(x)}{g(x)} \frac{1}{F(x)}$ is (increasing) nondecreasing over $(\underline{x}, \bar{x})$.

Proof. See Appendix A.

Regularity allows us to compare our ordering with the standard orderings used in finance and economics. First note that Lemma 5 implies that MLRP ordering is equivalent to u being convex. MLRP implies that F/G is increasing and thus G/F is decreasing. Geometric dominance requires ((f/g) (G/F)) to be increasing, because (G/F) is decreasing when f/g is increasing; whether the distributions are geometrically ordered as well will depend on the rate of increase of f/g relative to G/F. When F is dominated in the MLRP ordering f/g is decreasing. This implies that G/F is increasing. Thus, it appears, as will be verified below, that, if the rate of increase of G/F is sufficiently large relative to f/g, an MLRP dominated distribution can be geometrically dominant. It is also worth noting that the density condition for geometric dominance, part (ii) of Lemma 5, can be expressed as $(\log(F))'/(\log(G))'$ being increasing. This formulation of geometric dominance yields the following simple corollary to Lemma 5.

Corollary 1. *If F (strictly) geometrically dominates G, then the ratio $\log(F(x))/\log(G(x))$, is (increasing) nondecreasing in x.*

Proof. Because $\log(G) < 0$, over $(\underline{x}, \bar{x})$, it never changes sign over over $(\underline{x}, \bar{x})$. Next note that $\lim_{x \to \bar{x}} \log \circ G(x) = 0$ Part (ii) of Lemma 5 shows that (strict) geometric dominance implies that that $(\log \circ F)'/(\log \circ F)'$ is (increasing) nondecreasing. The assertion follows from these observations and the Monotone LHôpital Theorem (see Propositon 1.1 in (Pinelis, 2002)). □

Part (ii) of Lemma 5 and its corollary establish a rather tight analogy between MLRP and geometric dominance. The MLRP property can be expressed as F'/G' being increasing and MLRP implies that the ratio F/G is increasing. Geometric dominance can be expressed as the ratio $(\log \circ F)'/(\log \circ F)'$ being increasing and geometric dominance implies that the ratio $\log \circ F/\log \circ F$ is increasing. Thus, geometric dominance can be thought of a a log transformed version of MLRP. We also see from part (iii) of Lemma 5 that the logarithmic convexity of the transform function, u, implies both geometric dominance and MLRP dominance. However, the logarithmic convexity of the transform function is not a very useful notion of dominance because it is hardly ever satisfied when the two distributions are generated by textbook distributions. To see this, note that, by part iii of Lemma 5,

u being logarithmic convex is equivalent to $(f/F)(1/g)$ being increasing. If F is logarithmically concave and g is not always nondecreasing on it support, this condition cannot be satisfied. As most standard textbook distribution functions are logarithmically concave and most do not have monotonically increasing densities, the logarithmic convexity of u is not a very useful basis for ordering "typical" distribution functions.[2] In contrast, geometric convexity orders many textbook distributions.

6 Positive geometric convexity: Selection robust inference

We aim to analyze case (i) of Theorem 2. Since geometric dominance is a sufficient condition for selection dominance, if F dominates G under both geometric dominance and stochastic dominance, F's dominance over G is robust to selection effects. In this case, selection preserves the ordering of distributions. When one distribution is unconditionally better than another, it is also better conditioned on selection. This leads us to consider the effect of imposing both geometric and stochastic dominance and examining the implications of the dual ordering for the properties of the distribution functions which it orders.

Definition. If F geometrically dominates G ($F \underset{g}{\succcurlyeq} G$) and F stochastically dominates G ($F \underset{sd}{\succcurlyeq} G$), then we will say that F *positively geometrically dominates* G and write $F \underset{g+}{\succcurlyeq} G$. If in addition F strictly geometrically dominates G, we will say that *strictly positively geometrically dominates* G and write $F \underset{g+}{\succ} G$.

Proposition 1. *Let $\tilde{X}$ and $\tilde{Y}$ be two random variables whose distribution functions F and G are an admissible pair. Suppose that F positively geometrically dominates G, then*

(i) $\tilde{X}$ selection dominates $\tilde{Y}$, i.e., for any increasing valuation function, v, $\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y}] \geq \mathbb{E}[v(\tilde{Y})|\tilde{Y} > \tilde{X}]$.

(ii) $\tilde{X}$ dominates $\tilde{Y}$ in the MLRP ordering.

(iii) The probability that $\tilde{X}$ will be selected is higher, i.e.,$\mathbb{P}[\tilde{X} > \tilde{Y}] \geq 1/2$.

Proof. (i) follows from Theorem1. (ii) follows from Corollary 3, Lemma 5, and the definition of the MLRP ordering. (iii) follows by the following argument: The probability of F being selected is given by

$$\int_0^1 (1-u(s))\,ds \tag{24}$$

Because positive geometric dominance implies stochastic dominance and stochastic dominance implies that $u(t) \leq t$ we see that

$$\int_0^1 (1-u(s)\,ds \leq \int_0^1 (1-s)\,ds = \frac{1}{2}. \tag{25}$$

□

Proposition 1 shows that positive geometric convexity is a rather tame ordering of distributions. It simply

[2]See Bagnoli and Bergstrom (2005) for an exhaustive discussion of the logarithmic concavity of distribution functions and verification of the prevalence of logarithmic concavity

represents a strengthening of the MLRP ordering condition. Moreover, geometric dominance has an added implication not shared by geometric convexity; it implies not only that, regardless of selection, the expected value under the dominant distribution is higher under any increasing value function but also that the dominant distribution is more likely to be selected. Moreover, because the positive geometric dominance ordering inherits antisymmetry from stochastic dominance, it is also not only a preorder but also a partial order over admissible distribution pairs. This result is recorded in Lemma 6.

Lemma 6. *The positive geometric convexity order* $\underset{g+}{\succcurlyeq}$ *is a partial order.*

Proof. Note that because both the stochastic dominance and the geometric convexity orderings are reflexive and transitive, positive geometric dominance is reflexive and transitive. To see that it is antisymmetric, note that the stochastic dominance ordering is antisymmetric. Thus, $F \underset{g+}{\succcurlyeq} G$ and $G \underset{g+}{\succcurlyeq} F$ implies that $F \underset{sd}{\succcurlyeq} G$ and $G \underset{sd}{\succcurlyeq} F$, which implies that $F = G$. □

6.1 Selection robust families of distributions

Thus far, we have not verified that any specific distributions satisfy the geometric dominance condition. In this section, we will show that positive geometric dominance orders many standard families of distributions defined by scale shifts. These families include the Weibull distribution (of which the Exponential distribution is a special case), the log-logistic distribution, and the Kumaraswamy distribution when the shape parameter exceeds 1. In the next section, Section 6.2, we will derive a simple condition for families of distributions generated by multiplicative scale shifts to be ordered by geometric dominance. We will then verify that the half-normal and lognormal distributions satisfy this condition. Because changing the scale parameter for the half normal and lognormal distributions produces a multiplicative scale shift, this result verifies that scale families of lognormal and half normal distributions are ordered by positive geometric dominance. Within such families, one can infer that the expected value of any increasing function is higher conditioned on selection if and only if it is higher unconditionally. This observation motivates the following definition.

Definition. Consider a family of admissible distribution functions indexed by α, $\{K_\alpha\}_\alpha$ we will say that the family of distribution functions is *selection robust* if for any two members of the family $K_{\alpha'}$, $K_{\alpha''}$, $\alpha' \neq \alpha''$,$K_{\alpha'}$ stochastically dominates $K_{\alpha''}$ implies that $K_{\alpha'}$ selection dominates $K_{\alpha''}$. The family is strictly selection robust if $K_{\alpha'}$ strictly stochastically dominates $K_{\alpha''}$ implies that $K_{\alpha'}$ strictly selection dominates $K_{\alpha''}$.

Geometric dominance can establish selection robustness for a large class of distributions.

Weibull distribution The Weibull distribution is a probability distribution defined by,

$$K(x) = K_\alpha^W(x;\lambda) = 1 - e^{-\left(\frac{x}{\alpha}\right)^\lambda}, \quad x \geq 0,\ \alpha > 0,\ \lambda > 0 \tag{26}$$

λ is called the shape parameter and α is called the scale parameter. The distribution is used in labor economics to model worker quality (Dunn and Holtz-Eakin, 2000).

Result. Any family of Weibull distributions with a common shape parameter λ is selection robust.

Proof. Consider two distributions in the family $K_{\alpha_F} = K^W_{\alpha_F}(\cdot;\lambda)$ and $K_{\alpha_F} = K^W_{\alpha_G}(\cdot;\lambda)$, $\alpha_F > \alpha_G$. Let u represent $K_{\alpha_F} \circ K^{-1}_{\alpha_G}$. u is given by

$$u(t) = 1-(1-t)^r,\ t \in [0,1], \quad r = \left(\frac{\alpha_G}{\alpha_F}\right)^{\lambda} \tag{27}$$

It is easy to see that $K_{\alpha_F} \underset{sd}{\succ} K_{\alpha_G}$. Thus, in order for $K_{\alpha_F} \underset{g+}{\succcurlyeq} K_{\alpha_G}$, we need only show that $K_{\alpha_F} \underset{g}{\succ} K_{\alpha_G}$ By Lemma 1, a necessary and sufficient condition for strict geometric dominance is that the conjugate function to u, $\hat{u}$ to be strictly convex. The conjugate function is given by

$$\hat{u}(y,r) = \log\left(1-(1-e^y)^r\right),\ y \le 0, \quad r = \left(\frac{\alpha_G}{\alpha_F}\right)^{\lambda}. \tag{28}$$

$\hat{u}$ is strictly convex if and only if $r > 1$, i.e., if and only $\alpha_F > \alpha_G$. By Theorem 1 geometric dominance implies selection dominance. □

Thus, a scale family of Weibull distributed random variables with a common shape parameter is strictly selection robust. Because convexity of u implies MLRP ordering, it is also easy to see directly that the family is ordered by MLRP, as we expected given Theorem 3.

Log-logistic distribution A distribution function is log-logistic if it is given by the function

$$K^{\mathrm{LL}}_{\alpha}(x;\beta) = \frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}, \quad x>0,\ \alpha>0,\ \beta>1. \tag{29}$$

β is called the shape parameter of the distribution and α is called the scale parameter. In economics, the log-logistic distribution is used to model the the distribution of income ((Fisk, 1961)). The requirement that $\beta > 1$ is imposed to ensure that the distribution has a finite expectation, as required by admissibility.

Result. Any family of log-logistic distributions with a common shape parameter λ is strictly selection robust.

Proof. Consider any two members of the family, α_F and α_G, with $\alpha_F > \alpha_G$. Let $K_{\alpha_F} = K^{\mathrm{LL}}_{\alpha_F}(\cdot;\lambda)$ and $K_{\alpha_F} = K^{\mathrm{LL}}_{\alpha_G}(\cdot;\lambda)$. First note that, for any log-logistic distribution,

$$K^{-1}_{\alpha}(x) = \alpha\left(\frac{t}{1-t}\right)^{1/\beta}. \tag{30}$$

Thus,

$$u(t) = K_{\alpha_F} \circ K^{-1}_{\alpha_G}(t) = \frac{t}{r(1-t)+t}, \quad r = \left(\frac{\alpha_F}{\alpha_G}\right)^{\beta}. \tag{31}$$

In this case, rather than using the conjugate function to test for geometric convexity, we use Lemma 4. Note that the function R as defined in that lemma is given by

$$R(t) = \frac{r}{(r-1)(1-t)+1}, \tag{32}$$

which is increasing if and only if $r > 1$, i.e., if and only if $\alpha_F > \alpha_G$. Also, when $\alpha_F > \alpha_G$, F strictly stochastically dominates G. By Theorem 1 geometric dominance implies selection dominance. Thus, log-logistic scale families are strictly selection robust. □

Kumaraswamy distribution with $b > 1$ The Kumaraswamy distribution has been applied to model recovery rates in models of CDO defaults (Höcht and Zagst, 2010). This distribution is supported by the unit interval and is defined by

$$K_\alpha^{\mathrm{Ks}}(x;b) = 1-(1-x^\alpha)^b.\; x \in [0,1],\; \alpha > 0,\; b > 0. \tag{33}$$

Result. For any fixed $b > 1$, the family of Kumaraswamy distributions indexed by the sale parameter $\alpha > 0$, $\left\{K_\alpha^{\mathrm{Ks}}(\cdot;b)\right\}$, is strictly selection robust.

Proof. Is is apparent from inspection that if $\alpha_F > \alpha_B$ then $K_{\alpha_F}^{-1}$ strictly stochastically dominates $K_{\alpha_G}^{-1}$. The proof that $K_{\alpha_F}^{-1}$ strictly geometrically dominates $K_{\alpha_G}^{-1}$is both tedious and elementary and thus is deferred to Appendix B. Because geometric dominance implies selection dominance by Theorem 1, the result follows. In an example, which will be discussed extensively in Section 7, we will show that, when $b < 1$, families of Kumaraswamy distributions are not selection robust.

6.2 Conditions for multiplicative scale shifts to preserve dominance under selection

In many cases, the tests for geometric convexity developed thus far are difficult to apply in practice because closed-form expressions for the quantile functions of the distributions do not exist. In this section, we develop a simple test for selection robustness when the pair of distributions results from a multiplicative scaling of the underlying random variable.This test is useful because for many common distributions defined over the non-negative real line, the scale parameter shifts can be represented by multiplicative scaling. For example, if $\tilde{X}$ is a lognormal random variable with parameters (μ,σ), then $\tilde{X}$ is identically distributed to $e^\mu\,\tilde{Y}$, where $\tilde{Y}$is a lognormally distributed random variable with parameters $(0,\sigma)$. Thus, if scale shifts preserve dominance under selection, families of lognormal distributions with common log variance are selection robust. In this section, using this test, we will verify selection dominance for some common distributions. In the next section, we will show by means of a counterexample, that distributions exist which fail to satisfy the conditions of the test.

To initiate the analysis, suppose $\tilde{Y}$ is a random variable whose whose distribution function, G, is absolutely continuous and whose support equals $[0,\infty)$. Suppose that the distribution of $\tilde{X}$, F, is given by a scale shift of $\tilde{Y}$, that is $\tilde{X} \sim S\tilde{Y}$, $S > 1$. In this case, the distribution of $\tilde{Y}$, G, is given by $G(x) = F(Sx)$ and its density is given by $g(x) = Sf(Sx)$. Note that F strictly stochastically dominates G. In order to keep our expressions compact, we define the following operator. For any differentiable function, $\phi : \mathbb{R}^+ \to \mathbb{R}$, define the operator, $D_{\mathscr{L}}$ by $D_{\mathscr{L}}(\phi)(x) = (\log \circ \phi(x)$'. By induction, for $D^{n-1}_{\mathscr{L}}(\phi)$ define

$$D^n_{\mathscr{L}}(\phi)(x) = D_{\mathscr{L}}\left(D^{n-1}_{\mathscr{L}}(\phi)\right)(x). \tag{34}$$

$D_{\mathscr{L}}$ simply represents log differentiation. Applying these definitions we obtain

$$\frac{f(x)}{F(x)} = D_{\mathscr{L}}(F)(x) \quad \text{and} \quad \frac{g(x)}{G(x)} = \frac{Sf(Sx)}{F(Sx)} = SD_{\mathscr{L}}(F)(Sx). \tag{35}$$

From Lemma 5, F geometrically dominates G and thus F is selection dominant by Theorem 1, if

$$\frac{D_{\mathscr{L}}(F)(x)}{D_{\mathscr{L}}(G)(x)} = \frac{D_{\mathscr{L}}(F)(x)}{SD_{\mathscr{L}}(F)(Sx)} \tag{36}$$

is increasing. This function will be increasing precisely when the difference in the log derivatives of $D_{\mathscr{L}}(F)(x)$ and $D_{\mathscr{L}}(G)(x)$ is is decreasing, i.e., when

$$D^2_{\mathscr{L}}(F)(x) \geq SD^2_{\mathscr{L}}(F)(Sx). \tag{37}$$

We can multiply both sides of this equation by x without changing the direction of the inequality. This yields, the condition

$$xD^2_{\mathscr{L}}(F)(x) \geq (Sx)D^2_{\mathscr{L}}(F)(Sx). \tag{38}$$

This condition will be satisfied for all scale shifts $S > 1$ if and only if the function:

$$x \to xD^2_{\mathscr{L}}(F)(x) = x\left(\frac{f'(x)}{f(x)} - \frac{f(x)}{F(x)}\right) \text{ is decreasing.} \tag{39}$$

These observations yield the following result:

Proposition 2. *If the distribution of $\tilde{X}$, F, results from a simple scaling up of $\tilde{Y}$, with a scale factor $S > 1$, then X will strictly geometrically dominate Y and thus always produce a higher value conditioned on selection, whenever the the second log derivative of F, $D^2_{\mathscr{L}}(F)$, satisfies condition* (39).

6.2.1 Examples

As we demonstrate below, condition (39) is satisfied by many common textbook distributions over the positive real line.

Half Normal Consider scale shifts of the the standard half normal distribution, $F(x) = \text{erf}\left(\frac{x}{\sqrt{2}}\right)$, $x \geq 0$. In this case we can verify condition (39) by a simple differentiation:

$$(xD^2_{\mathscr{L}}(F)(x))' = \left(\frac{f'(x)}{f(x)} - \frac{f(x)}{F(x)}\right)' = -x - \left(x + \frac{f(x)}{F(x)}\right)\left(1 - \frac{xf(x)}{F(x)}\right), \quad x > 0. \tag{40}$$

Because f is decreasing for $x \geq 0$, $(xf(x))/F(x) < 1$ for $x > 0$. Thus, $(xD^2_{\mathscr{L}}(F)(x))' < 0$ and hence condition (39) is satisfied. This implies that upward scale shifts of a half normally distributed random variable are strictly selection dominant.

Log-normal Finally, consider the scaling of a standard log-normal distribution. Let, Φ equal

$$\Phi(x) = xD^2_{\mathscr{L}}(F)(x), \tag{41}$$

where F is the standard log-normal distribution. Define the function $\eta : \mathbb{R} \to \mathbb{R}$ by $\eta(y) = \Phi(\exp(y))$. Because $\exp(\cdot)$ is increasing and thus order preserving, if we can show that η is decreasing, then it must be the case that Φ is decreasing, and thus condition (39) is satisfied. Note first that

$$\eta(y) = -(y+1) - \frac{1}{m(-y)}, \tag{42}$$

where m is the Mills Ratio, $m = (1-F)/f$. Because the Mills ratio is positive, η is negative. To see that it is decreasing note that, by a result in Sampford (1953), the derivative of $1/m(x)$ is strictly less than 1 for all real x. Thus, η and hence Φ are strictly decreasing. This implies that upward scale shifts of log-normally distributed random variable are strictly selection dominant.

6.2.2 Counterexample

Not all upward scale shifts lead to geometric dominance. However, counterexamples are not easy to construct. One essentially needs a distribution that has a density with a very flat left tail, a very compact middle region, and a very thin but long right tail. In this case, one can think of the selection of a given distribution as signaling both the value of the random variable selected and the region the of density from which it was drawn. Upward scaling has almost no effect on the left tail because it has almost no probability mass. Thus, scaling only has a significant effect when the random variable is in the middle or upper region. The flat left tail serves to make the middle region fairly small. Thus, dispersion over the middle region is small. Therefore, scaling has a major effect on the probability of selection in this region. Conditioned on being in the upper tail, the superior, i.e., stochastically dominant, random variable is of course more valuable in expectation. However, because of the much higher dispersion of value over the upper tail, this does not translate into as large of a difference in the probability of being selected as it does in the middle region. For this reason, selection of the superior random variable is correlated with the middle region being realized. Hence, selection of the inferior random variable signals that the selected random variable is in the long right tail. Because the value difference between the right tail and the middle region dwarfs value differences within either region, and because selection of the inferior distribution signals that the right tail has been realized, the inferior distribution is selection dominant. To produce and example that verifies that such reversals are indeed possible, we take a distribution with a very long right tail, the Fréchet distribution, modify it a bit to produce a flat left tail, and then verify that selection reversal occurs. Note that, as pointed out in Section 4, scale shifts of the Fréchet distribution could not produce selection reversals because scaled Fréchet distributions are geometrically equivalent to each other. The distribution we select for the counterexample is

$$F(x) = \exp\left[-\frac{1}{(x+\log(1+x))^2}\right]. \tag{43}$$

The expression enclosed in square brackets converges to $-\infty$ as $x \to 0$ and to 0 as $x \to \infty$, and is increasing, so F defines a distribution function over $[0,\infty)$. This distribution has a finite expectation, as can be verified by noting that

$$1 - F(x) < 1 - F_{\text{Fréchet}}(x), \tag{44}$$

where $F_{\text{Fréchet}}$ is a Fréchet Distribution with shape parameter 2 and scale parameter 1. Because this Fréchet distribution has finite expectation, so so does F. Calculations show that

$$xD^2_{\mathscr{L}}(F)(x) = -\left(\frac{x+2}{x+1}\right)\left(\frac{x}{(x+2)^2} + \frac{3x}{x+\log(x+1)}\right), \tag{45}$$

$$F^{-1}(t) = W\left(\exp\left[\frac{1}{\sqrt{-\log(t)}}+1\right]\right) - 1, \tag{46}$$

$$u(t) = F \circ G^{-1} = F\left(\frac{1}{S}F^{-1}(t)\right), \tag{47}$$

where $W(\cdot)$ is Lambert's W-Function. In Figure 1 we plot scalings effects for the case of $s = 1.2$. We see that $xD^2_{\mathscr{L}}(F)(x)$ is not decreasing and in fact it is increasing until $x \approx 5.38$ and then it decreases. However, the probability of a realization in excess of 5.38, is less than 2%. Thus, condition (39) is violated over almost all the probability mass of F. This fact can also be verified by inspecting Panel C, which plots the geometric dominance test function defined in Proposition 6.2. The increase in $xD^2_{\mathscr{L}}(F)(x)$ in the extreme right tail is reflected in the sharp upturn in $R(t)$ around $t = 0.98$. In Table 1 we verify that in fact the scaled up distribution, F, has a lower expected value conditioned on being selected.

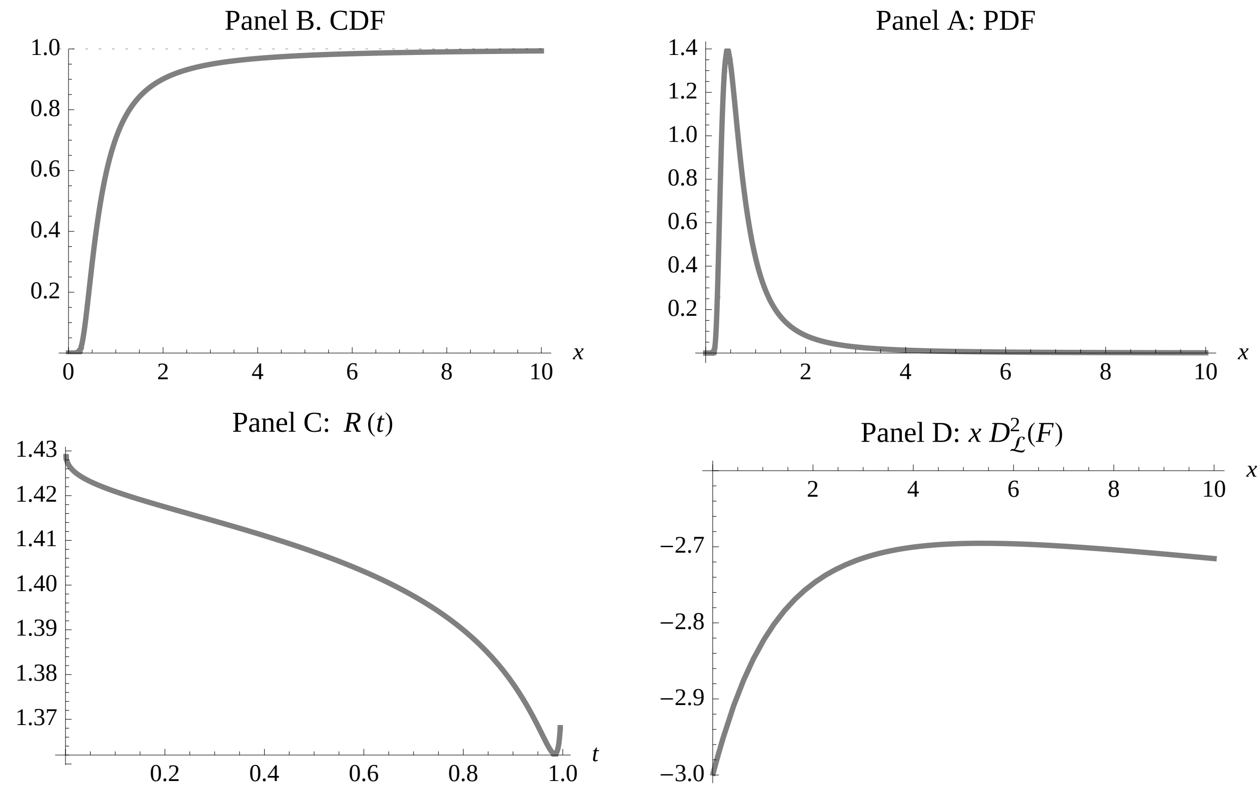

Figure 1: Counterexample to up scaling leading to geometric dominance. In the figure, F is given by (43).

$\mathbb{E}[\tilde{X}\vert\tilde{X} > \tilde{Y}] == 1.5339$	$\mathbb{E}[\tilde{Y}\vert\tilde{Y} > \tilde{X}] = 1.5452$
$\mathbb{P}[\tilde{X} > \tilde{Y}] = 0.58196$	$\mathbb{P}[\tilde{Y} > \tilde{X}] = 0.41804$
$\mathbb{E}[\tilde{X}] = 1.1354$	$\mathbb{E}[\tilde{Y}] = 0.94661$

Table 1: Counterexample to up scaling producing selection dominance. In the table, $\tilde{X} \sim F$ and $\tilde{Y} \sim (1/S)\tilde{X}$, with $S = 1.2$ and F given by (43).

7 Negative geometric dominance: Selection reversal

As shown in part (ii.a) of 2 a distribution can be geometrically dominant yet stochastically dominated. We now turn to this case. When distributions have this property, better is worse conditioned on selection. We begin by formalizing this relation.

Definition. If F geometrically dominates G ($F \underset{g}{\succcurlyeq} G$) and G stochastically dominates F ($F \underset{sd}{\preccurlyeq} G$), then we will say that F *negatively geometrically dominates* G and write $F \underset{g-}{\succcurlyeq} G$. If in addition, F strictly geometrically dominates G, we will say that F *strictly negatively geometrically dominates* G and write $F \underset{g-}{\succ} G$.

The following proposition, Proposition 3, uses Theorem 3 and Lemma 5 to characterize distribution functions that are ordered by negative geometric dominance. In the case of negative geometric dominance, we do not have the fine control over the ratios between densities that we had in the case of positive geometric dominance.

However, negative geometric dominance does restrict the ratio between the distribution functions and the behavior of the densities around the lower endpoint of their support.

Proposition 3. *Let $\tilde{X}$ and $\tilde{Y}$ be two random variables with regularly related distribution functions F and G. Suppose that F negatively geometrically dominates G, then*

(i) $\tilde{X}$ selection dominates $\tilde{Y}$, i.e., for any increasing valuation function, v, $\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y}] \geq \mathbb{E}[v(\tilde{Y})|\tilde{Y} > \tilde{X}]$

(ii) The ratio F/G is non-increasing.

(iii) If F strictly negatively geometrically dominates G, then F/G is decreasing and

$$\lim_{x \to \underline{x}} \frac{F(x)}{G(x)} = \lim_{x \to \underline{x}} \frac{f(x)}{g(x)} = \infty.$$

(iv) The probability that $\tilde{X}$ will be selected is lower, i.e.,$\mathbb{P}[\tilde{X} > \tilde{Y}] \leq 1/2$.

(v) The negative geometric convexity order $\underset{g-}{\succcurlyeq}$ is a partial order.

Proof. (i) follows from Theorem1. To prove (ii), note that part (ii) implies that $t \to u(t)/t$ is decreasing. Noting that $u = F \circ G^{-1}$ and making the substitution, $t = G^{-1}(x)$ shows that $t \to u(t)/t$ being nonincreasing implies that $x \to F(x)/G(x)$ is nonincreasing. Part (iii) follows from (ii) in like fashion after noting that the limits of F/G and f/g are the same by L'Hôpital's rule. (iv) follows from exactly the same argument that was used to prove part (iii) of Proposition 1. Finally, part (v), follows from the same argument used to prove Lemma 6. □

Just as the conditions in Proposition 1 were not sufficient for positive geometric dominance, the conditions in Proposition 3 are not sufficient for negative geometric dominance. However, in both cases, geometric dominance does impose fairly strong necessary conditions on the underlying distribution functions. In the case at hand, negative geometric dominance, the restriction that F/G be non-increasing established in part (ii), implies that G, the negative geometrically dominated distribution, dominates the negatively geometrically dominant distribution, F, in the reverse hazard rate order. This order is not nearly as important as the MLRP order. However, it has found some applications in financial economics (Kijima and Ohnishi, 1999). The key restriction imposed by strict negative geometric dominance, is that F/G converges to infinity as x approaches the left endpoint of the distribution functions' support. The probability weight on the left tail of the negatively geometrically dominant distribution grows explosively relative to the dominated distribution's weight. We term this behavior "left-tail explosion." Negative geometric dominance can only occur when left-tail explosion occurs. A left-tail explosion ensures that low realizations from the geometrically dominated distribution are almost never selected and this censoring increases the selection-conditioned value of the distribution. Of course, the left-tail explosion reduces the dominant distributions selection probability. This is the result recorded in part (iv) of the Proposition.

However, in contrast to positive geometric dominance, negative geometric dominance is not consistent with any unconditional ordering of distribution functions. In fact, not only are negatively dominant distributions never dominant under standard unconditional orderings, they can, as we show later by means of an example, be dominated in the strongest unconditional ordering, MLRP ordering. More importantly, because geometric dominance implies selection dominance by Theorem 1, when distributions are ordered by negative geometric dominance, the dominant distribution in the absence of selection is always dominated in the presence of selection. Thus, selection reverses qualitative inferences regarding the underlying random variables. This observation motivates the following definition:

Definition. Consider a family of admissible distribution functions indexed by α, $\{K_\alpha\}_\alpha$. We will say that the family of distribution functions is *selection reversing* if for any two members of the family $K_{\alpha'}$, $K_{\alpha''}$, $\alpha' \neq \alpha''$,$K_{\alpha'}$ stochastically dominates $K_{\alpha''}$ implies that $K_{\alpha''}$ selection dominates $K_{\alpha'}$. The family is strictly selection robust if $K_{\alpha'}$ strictly stochastically dominates $K_{\alpha''}$ implies that $K_{\alpha''}$ strictly selection dominates $K_{\alpha'}$.

Are there reasonable families of distributions that are selection reversing? Yes, but not many textbook distributions have this property. One that does is the Kumaraswamy distribution with shape parameter $b < 1$.

Example 1. For any fixed $b < 1$, the family of Kumaraswamy distributions indexed by the sale parameter $\alpha > 0$, $\{K_\alpha^{\mathrm{Ks}}(\cdot;b)\}$, is selection reversing.

Proof. As was noted in the discussion of the Kumaraswamy distribution in the previous section, $\alpha_F < \alpha_G$ if and only if $K_{\alpha_G}^{\mathrm{Ks}} \underset{sd}{\succ} K_{\alpha_G}^{\mathrm{Ks}}$. In Appendix B, we show that, when the shape parameter $b < 1$, $\alpha_F < \alpha_G$ if and only if $K_{\alpha_F}^{\mathrm{Ks}} \underset{g}{\succ} K_{\alpha_G}^{\mathrm{Ks}}$. Since strict geometric dominance implies strict selection dominance, the family of distributions,$\{K_\alpha^{\mathrm{Ks}}(\cdot;b)\}$, is strictly selection reversing for $b < 1$. □

Because the proof of this result is not terribly intuitive and involves tedious calculations, we will focus on a special case of this result, the case were $b = 1/2$, $\alpha_F = 1$, and $\alpha_G = 2$. In this case, the two distributions corresponding to α_F and α_B, F and G, are

$$F(x) = 1 - \sqrt{1-x}, \quad G(x) = 1 - \sqrt{1-x^2}. \tag{48}$$

Because of the simple form of the Kumaraswamy distribution, we can explicitly solve for the transform function, u, and its conjugate, $\hat{u}$ These functions are presented below:

$$u(t) = 1 - \sqrt{1 - \sqrt{(1-t)\,t + t}},\ t \in [0,1]; \quad \hat{u}(y) = \log\left(1 - \sqrt{1 - \sqrt{e^y\,(1-e^y) + e^y}}\right),\ y \in (-\infty, 0]. \tag{49}$$

The distribution functions, densities, geometric dominance condition, and the likelihood ratio, are plotted in Figure 2. As one can see from inspecting Panel A of Figure 2, F is strictly dominated by G in the stochastic dominance ordering. In fact, as Panel D of Figure 2 shows, f/g is decreasing and therefore F is even dominated by G in the MLRP ordering. The geometric convexity condition for F to strictly geometrically dominate G (given in Lemma 5) is that $(f/g)\,(G/F)$ is increasing. This condition is verified in Panel C of Figure 2. The fat-tail explosion is evident if we consider the behavior of the ratio F/G at selected quantiles of the G distribution. At the 1% quantile of G the ratio $F : G$ is more than 7:1. At the 0.1% quantile, the ratio is more than 20:1. This tail explosion is evidenced in Panel B of Figure 2 by the probability densities of F and G near the lower end point of their support.

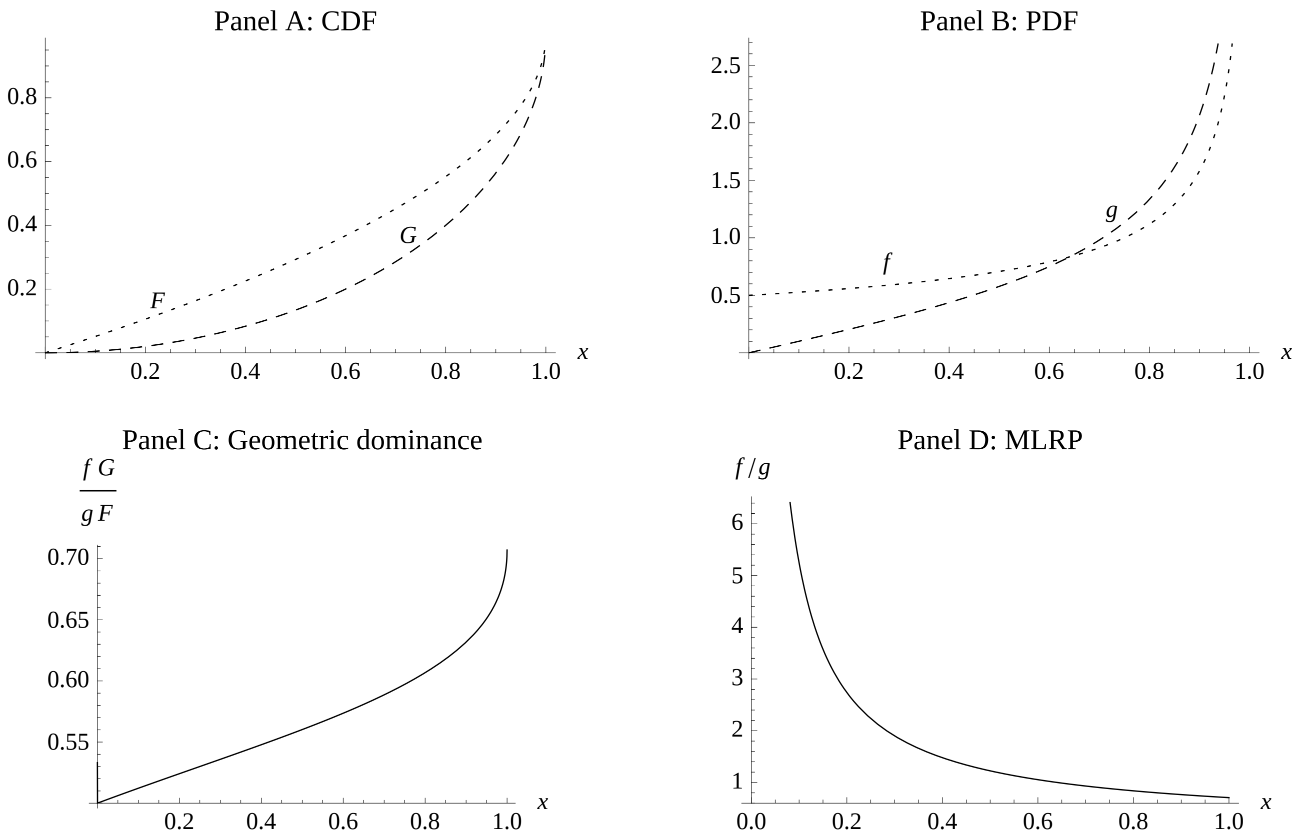

Figure 2: Kumaraswamy distribution with $b = 1/2$. In the figure,

Figure 3 plots the transform function, u, and its conjugate function, $\hat{u}$. As can be seen from Panels A and B of Figure 3, in this example, the transform distribution, u, is concave yet its conjugate function, $\hat{u}$, is convex. Thus, logarithmic scaling, although it preserves order relations, can completely alter the shape of the transform function. The convexity of $\hat{u}$ also again verifies the geometric dominance of F.

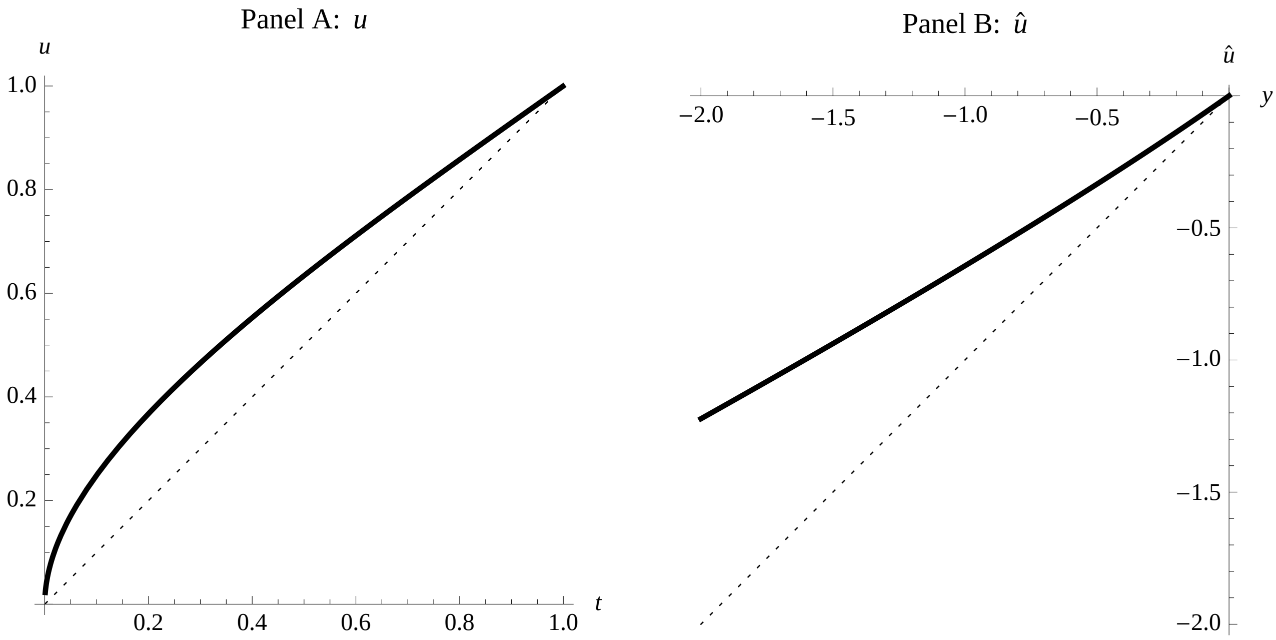

Figure 3: Transform and conjugate transform functions for the Kumaraswamy distribution. In the figure, u and $\hat{u}$ are given by (49)

The selection dominance of F is also illustrated in Table 2, which computes value under selection assuming that the valuation function, v, is simply $v(x) = x$. Note first that, as expected, value is higher conditioned on the random variable generated by F, which is $\tilde{X}$, being selected than it is when the random variable generated by G, $\tilde{Y}$, is selected. Note also that the that the stochastic superiority of $\tilde{Y}$ is absorbed by the probability of selection. $\tilde{Y}$'s probability of being selected is more than 60% higher than $\tilde{X}$'s even though $\tilde{Y}$'s expected value is only 17% higher.

$\mathbb{E}[\tilde{X}\|\tilde{X} > \tilde{Y}] = 0.88284$	$\mathbb{E}[\tilde{Y}\|\tilde{Y} > \tilde{X}] = 0.8731$
$\mathbb{P}[\tilde{X} > \tilde{Y}] = 0.39052$	$\mathbb{P}[\tilde{Y} > \tilde{X}] = 0.60948$
$\mathbb{E}[\tilde{X}] = 0.66667$	$\mathbb{E}[\tilde{Y}] = 0.7854$

Table 2: Valuations under selection: Kumaraswamy Distribution. In the table, $\tilde{X} \sim F$ and $\tilde{Y} \sim G$, where F and G are given by (48).

8 Geometric convexity without stochastic dominance: Dispersion induced dominance

Finally, we consider the the case first identified in part (ii.b) of Theorem 2—geometric dominance when distributions are not ordered by stochastic dominance. This case is perhaps the least interesting because neither distribution is unconditionally better than the other. Thus there is no question of selection robustness or selection reversal to resolve. However, the results in this section have value in that they illustrate that geometric dominance can result simply from from dispersion. We will illustrate this result with the following pair of distributions.

$$F(x) = \begin{cases} e^{-c\sqrt{\frac{1}{x}}} & x \in (0,1] \\ 0 & x = 0 \end{cases}, \quad c > 0, \quad G(x) = x, \quad x \in [0,1]. \tag{50}$$

A simple calculus exercise shows that F is increasing with $F(0) = 0$ and $F(1) = 1$. G is the uniform distribution over $[0,1]$. Thus both F and G are distribution functions. Since both F and G have bounded common support, $[0,1]$, they are both clearly integrable. Thus, the pair of distributions is admissible and, in fact, regular. Explicit computation of the transform function, u, and its conjugate, $\hat{u}$, yields

$$u(t) = \begin{cases} e^{-c\sqrt{\frac{1}{x}}} & x \in (0,1] \\ 0 & x = 0 \end{cases}; \quad \hat{u}(t) = -c\sqrt{-y},\ y \in (-\infty,0]. \tag{51}$$

These functions are plotted in Figure 4 assuming that the parameter $c = 0.70$. Either by inspection of the graphs or explicit calculation it is apparent that $\hat{u}$ is strictly convex and thus F strictly geometrically dominates G.

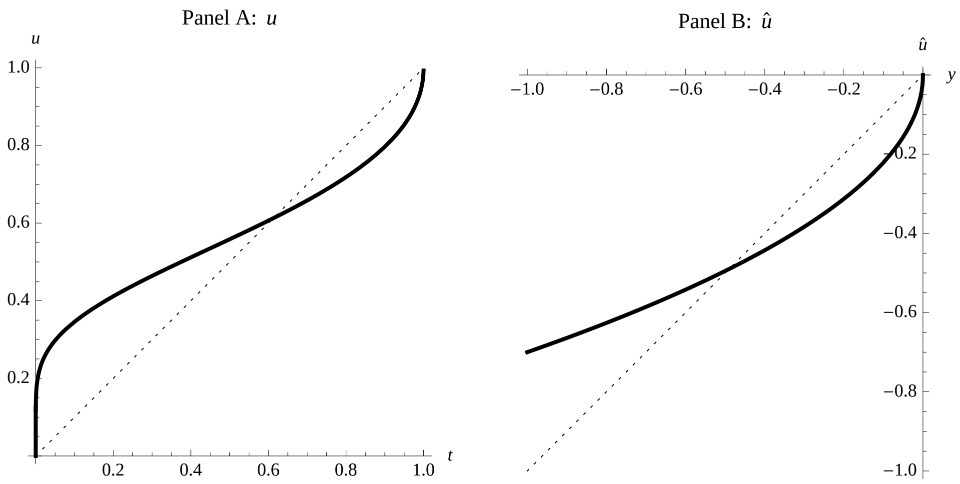

Figure 4: Geometrically dominant distribution increases dispersion. In the graph u and $\hat{u}$ are given by (51), with the parameter $c = 0.70$

Figure 4 illustrates the properties of dispersive geometrically dominant distributions developed in Theorem 2 and Theorem 3. u crosses the diagonal once from above. $u(t)/t$ explodes around 0 and, after crossing the diagonal, u becomes convex. Thus, using the results for both negative and positive geometric dominance provided in Propositions 1 and 3, it possible to deduce that the ratio F/G increases to infinity as x converges to the lower bound of the distributions' support, 0. In contrast, the upper end of the distribution of F dominates the distribution of G under the MLRP ordering. Thus, while negatively and positively geometrically dominant distributions garner their dominance from only one device, MLRP superiority in the case of positively dominant distributions, and left-tail explosions in the case of negatively dominant distributions, dispersive geometrically dominant distributions can use both devices.

Combining left-tail explosions with MLRP dominance in the right tail leads to strong selection superiority, which is illustrated by Table 3. In this table, the valuation function is simply $v(x) = x$. Selection dominance is not produced by the unconditional superiority of F, as is apparent from noting that F's expected value is lower than G's expected value.

$$\begin{array}{rr} \mathbb{E}[\tilde{X}|\tilde{X} > \tilde{Y}] = 0.7784 & \mathbb{E}[\tilde{Y}|\tilde{Y} > \tilde{X}] = 0.5854 \\ \mathbb{P}[\tilde{X} > \tilde{Y}] = 0.4352 & \mathbb{P}[\tilde{Y} > \tilde{X}] = 0.5648 \\ \mathbb{E}[\tilde{X}] = 0.4352 & \mathbb{E}[\tilde{Y}] = 0.5000 \end{array}$$

Table 3: Expected payoffs under selection when geometrically dominant distribution is dispersive. In the table, $\tilde{X} \sim F$ and $\tilde{Y} \sim G$, where F and G are given by (50) and $c = 0.70$.

9 Selection dominance without geometric dominance

In the last four sections of the paper, we focused on the geometric dominance order, a sufficient but not necessary condition for selection dominance. Given our objective, to define an ordering over distributions analogous to stochastic dominance, this focus is not surprising. As shown in Lemma 2, selection dominance does not even define a preorder over distribution functions. As shown by Theorem 1, selection dominance is equivalent to supermultiplicativity on average which roughly speaking is geometric convexity on average. Thus, the question of selection dominance, unlike the question of geometric dominance, cannot be resolved by tests comparing distributions and densities at specific points; average values matter and distributions that violate the distributional criteria for geometric convexity can be selection dominant as long as the violations "average out." In fact, this dependence on average values is the root cause for selection dominance being intransitive. Violations of geometric convexity that are small enough to average out in one comparison of distributions may not be small enough to average out in another.

Section 8 provides us with a clue for finding a case where selection dominance is easy to verify—dispersion increasing transformations. In fact, we will show that dispersion increasing transformations that render the transformed distribution MLRP dominant in its upper tail and MLRP dominated in its lower tail, while leaving the probability of selection fixed, always render the "riskier" transformed distribution selection dominant even if such transforms do not produce geometric dominance. The logic behind this result is fairly apparent. Both MLRP dominance in the upper tail and being MLRP dominated in the lower tail, favor selection dominance. Being dominant in the upper tail implies that values are higher when selected; being dominated in the lower tail means that low draws from the distribution are unlikely to be selected, thus increasing the expected value conditioned on selection. The problem of finding conditions for selection dominance is thus much easier when distributions are variability ordered rather than size ordered. In essence, variability leads to selection dominance provided the variability transformation is restricted using the now standard MLRP ordering. In contrast, when the unconditional distributions are ordered by size and not variability, MLRP dominant distributions can be geometrically dominated and thus selection dominated, as we showed in Section 7. As shown in Lemma 5 the MLRP ordering for regularly-related distributions is equivalent to the transform function, u, being convex. This observation motivates the definition of a selection-probability preserving, dispersion increasing transformation which is provided below.

Definition. Let u be a regular function the following properties:

(i) $\int_0^1 u(s)\,ds = 1/2$

(ii) There exist $t^* \in (0,1)$ such that u restricted to $[0,t^*]$ is strictly concave and u restricted to $[t^*,1]$ is strictly concave,

then u is *a selection-probability preserving, dispersion increasing transformation*.

The next result, Proposition 4, shows that such transformations always lead to selection dominance.

Proposition 4. *If F and G are regularly related distributions, $u = F \circ G^{-1}$, and u is a selection-probability preserving, dispersion increasing transformation, then Fstrictly selection dominates G.*

Proof. See Appendix A.

The very straight forward intuition behind this proposition is illustrated by the following example. In this example, the distribution functions are given by

$$F(x) = \frac{1}{2}\left((2x-1)^3+1\right),\ x \in [0,1] \quad G(x) = x,\ x \in [0,1]. \tag{52}$$

Because, $G(x) = x$, the transform function is simply given by $u(t) = F(t)$. Figure 5 plots the distribution functions, densities, and the R, and Π functions associated with these distributions. Panel A shows that, in fact, F and G are distribution functions over [0,1]. Panel B, plots the associated densities and shows the dispersive nature of the the u transform. Panel C plots the R function defined in Lemma 4. Because R being increasing is a necessary and sufficient condition for geometric convexity, it is clear that F does not geometrically dominate G. Panel D plots the $\Pi(u)$ function. By Theorem 1, supermultiplicativity on average of u is equivalent to the selection dominance of F over G. The supermultiplicativity on average condition is satisfied whenever, $\Pi(u)[t] \geq \Pi(u)[1], t \in [0,1)$. Panel D verifies this condition for selection dominance. Expected values under the two distribution functions are numerically evaluated in Table 4 for the case where $v(x) = x$. As the table shows, the expected value under selection is much higher for F, the selection probability preserving, dispersion increasing transform of G.

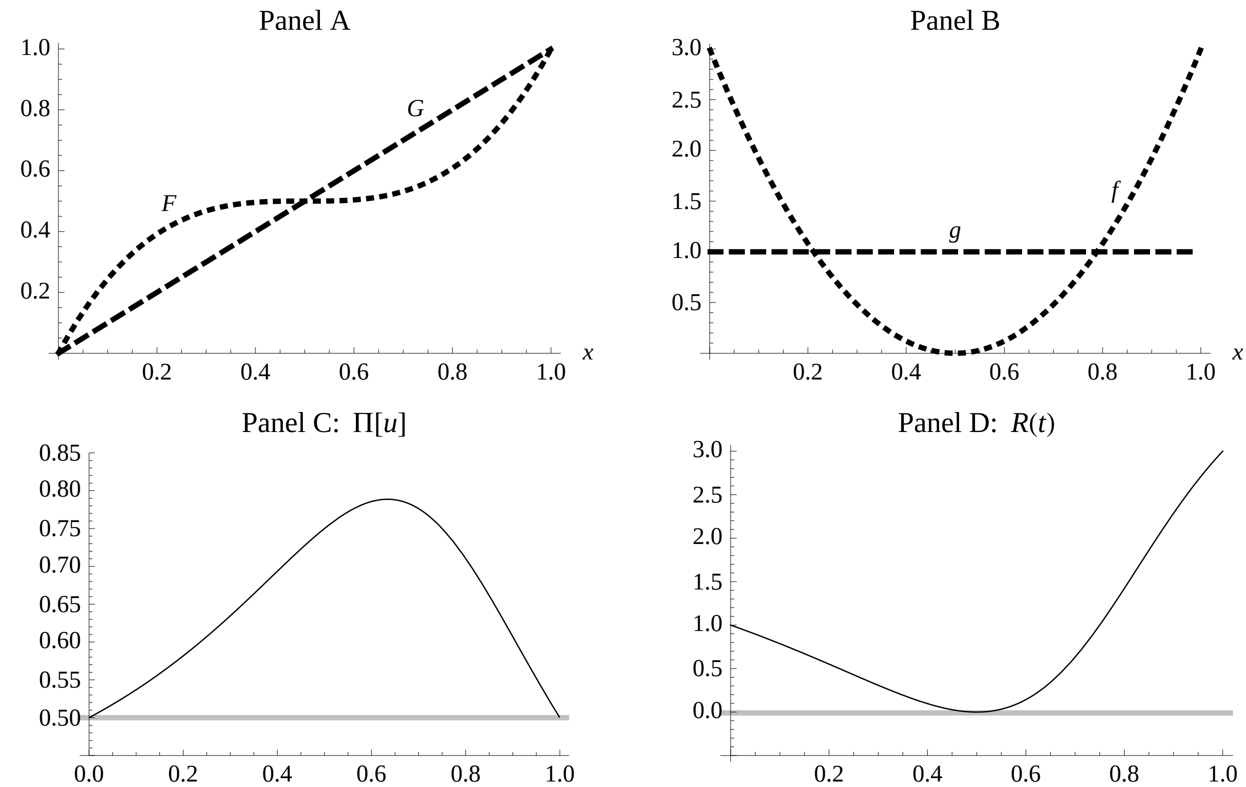

Figure 5: Dispersion increasing distributions and selection dominance. In the figure, F and G are given by (52).

$\mathbb{E}[\tilde{X}\mid\tilde{X} > \tilde{Y}] = 0.8000$	$\mathbb{E}[\tilde{Y}\mid\tilde{Y} > \tilde{X}] = 0.6000$
$\mathbb{P}[\tilde{X} > \tilde{Y}] = 0.5000$	$\mathbb{P}[\tilde{Y} > \tilde{X}] = 0.5000$
$\mathbb{E}[\tilde{X}] = 0.5000$	$\mathbb{E}[\tilde{Y}] = 0.5000$

Table 4: Selection dominant dispersion increasing transformation. In the table, $\tilde{X} \sim F$ and $\tilde{Y} \sim G$, where F and G are given by (52)

References

BAGNOLI, M., AND T. BERGSTROM (2005): "Log-Concave Probability and Its Applications," *Economic Theory*, 26(2), pp. 445–469.

CHAN, W., F. PROSCHAN, AND J. SETHURAMAN (1990): "Convex-Ordering among Functions, with Applications to Reliability and Mathematical Statistics," *Lecture Notes-Monograph Series*, 16, pp. 121–134.

DIAMOND, P. A., AND J. E. STIGLITZ (1974): "Increases in risk and in risk aversion," *Journal of Economic Theory*, 8(3), 337 – 360.

DUNN, T., AND D. HOLTZ-EAKIN (2000): "Financial Capital, Human Capital, and the Transition to SelfEmployment: Evidence from Intergenerational Links," *Journal of Labor Economics*, 18(2), pp. 282–305.

FINOL, C. E., AND M. WÓJTOWICZ (2000): "Multiplicative properties of real functions with applications to classical functions," *Aequationes Mathematicae*, 59(1-2), 134–149.

FISK, P. R. (1961): "The Graduation of Income Distributions," *Econometrica*, 29(2), pp. 171–185.

HARDY, G. H., J. E. LITTLEWOOD, AND G. POLYA (1952): *Inequalities*. Cambridge university press.

HÖCHT, S., AND R. ZAGST (2010): "Pricing distressed CDOs with stochastic recovery," *Review of Derivatives Research*, 13(3), 219–244.

KIJIMA, M., AND M. OHNISHI (1999): "Stochastic orders and their applications in financial optimization," *Mathematical Methods of Operations Research*, 50(2), 351–372.

MILGROM, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 12(2), pp. 380–391.

PINELIS, I. (2002): "L'Hospital Type rules for monotonicity, with applications," *J. Ineq. Pure & Appl. Math*, 3(1).

SAMPFORD, M. R. (1953): "Some Inequalities on Mill's Ratio and Related Functions," *The Annals of Mathematical Statistics*, 24(1), pp. 130–132.

SHAKED, M., AND J. G. SHANTHIKUMAR (1994): *Stochastic Orders and Their Applications*. Academic Press.

Appendix A Proofs of Selected Propositions

Proof of Lemma 3. We first show that (i) implies (iii). Let $u = F \circ G^{-1}$. Then $F \underset{g}{\succcurlyeq} G$ if and only if u is geometrically convex. $G \underset{g}{\succcurlyeq} F$ if and only if u^{-1} is geometrically convex. By part (iii) of Lemma 1, geometric convexity implies that the conjugate functions to u and u^{-1} are both convex. Thus $\hat{u}$ and $\widehat{u^{-1}}$ are both convex. Because $\widehat{u^{-1}} = \hat{u}^{-1}$, $\hat{u}$ and its inverse must both be increasing convex functions equal to 0 at $y = 0$. Thus, $\hat{u}$ must be a linear function of the form $\hat{u}(y) = p\,y,\ p > 0$. Thus,

$$u(t) = \exp(p\log(t)) = t^p,\ p > 0. \tag{A-1}$$

Next, noting that the conjugate function to $u(t) = t^p$ is linear shows that (iii) implies (i). By Theorem 1, (i) implies (ii). Thus, to complete the proof we need only show that (ii) implies (iii). To prove this note that, by Theorem 1 selection dominance is equivalent to the supermultiplicativity on average condition given by expression (17) being satisfied. Because u is strictly increasing and continuous, u^{-1} satisfying this condition is equivalent to $\Pi[u^{-1}](u(t)) \geq \Pi[u^{-1}](1)$. Thus, both F and G will be selection dominant if and only if

$$\Pi[u^{-1}](u(t)) \geq \Pi[u^{-1}](1) = \int_0^1 u^{-1}(s)\,ds,\ t \in [0,1], \tag{A-2}$$

$$\Pi[u](t) \geq \Pi[u](1) = \int_0^1 u(s)\,ds,\ t \in [0,1]. \tag{A-3}$$

Expanding the definition of $\Pi[u^{-1}](u(t))$, we see that

$$\Pi[u^{-1}](u(t)) = \frac{1}{u(t)\,t}\int_0^t u^{-1}(s).\,ds. \tag{A-4}$$

Young's Theorem (see for example Theorem 156 in Hardy, Littlewood, and Polya (1952)) implies that

$$\int_0^t u(s)\,ds + \int_0^{u(t)} u^{-1}(s)\,ds = t\,u(t). \tag{A-5}$$

Equations (A-5) and (A-4) and imply that

$$\Pi[u^{-1}](u(t)) = 1 - \Pi[u](t). \tag{A-6}$$

Letting $t = 1$ in (A-5) shows that

$$\int_0^1 u(s)\,ds + \int_0^1 u^{-1}(s)\,ds = 1. \tag{A-7}$$

Thus, if we let c equal the first integral in (A-7), we see that the supermultiplicativity on average condition being satisfied for both u and u^{-1} implies that

$$\Pi[u](t) \geq c \quad \text{and} \quad 1 - \Pi[u](t) \geq 1 - c. \tag{A-8}$$

Thus, $\Pi[u^{-1}](u(t)) = c$. This implies that for all $t \in (0,1]$,

$$\frac{1}{c}\frac{1}{t}\int_0^t u(s)\,ds = u(t). \tag{A-9}$$

Because u is identically equal to the left-hand-side of equation (A-9), and because u is continuous and thus its integral is differentiable, u must be differentiable. Differentiation of equation (A-9) shows that u must satisfy the differential equation,

$$(1-c)\,u(t) - c\,t\,u'(t) = 0, \quad u(1) = 1. \tag{A-10}$$

This differential equation has a unique solution, $u(t) = t^{(1-c)/c}$. □

Proof of Theorem 2. First consider (i). $F < G$ on some neighborhood of $\underline{x}$, implies that $u(t) < t$ on some neighborhood of 0. The geometric convexity of u implies by Lemma 1 that conjugate function to u, $\hat{u}(y) = \log \circ u \circ \exp(y)$, is an increasing convex function defined over $(-\infty, 0]$. The conjugate function to the the identify function $\mathrm{id}(t) = t$ is simply $\widehat{\mathrm{id}}(y) = y$, the identify function. Because conjugation preserves order relations, and because $u(t) > \mathrm{id}(t)$ in a neighborhood of 0, condition (ii) implies that there exists $\underline{y} < 0$, such that $\hat{u}(y) < \widehat{\mathrm{id}}(y)$ when $y < \underline{y}$. Because $\hat{u}(y)$ is convex and $\widehat{\mathrm{id}}(y)$ is linear and because the functions meet at 0, they cannot meet at any other point. Thus, $\hat{u}(y) < \mathrm{id}(y)$ for all $y < 0$. The order-preserving nature of conjugation then ensures that $u(t) < t$, for $t < 1$. The definition of u then implies that $F(x) < G(x),\ x \in (\underline{x}, \bar{x})$, Thus, F strictly stochastically dominates G.

Now consider (ii). $F > G$ on an open neighborhood of $\underline{x}$, implies that $u(t) > t$ on some open neighborhood of 0. Thus, for the same reasons as advanced in the proof of part (i), there exists $\underline{y} > 0$, such that $\hat{u}(y) > \widehat{\mathrm{id}}(y)$ when $y < \underline{y}$. Because $\hat{u}(y)$ is continous, either (case (a)) $\hat{u}(y) > \widehat{\mathrm{id}}(y),\ y < 0$ or (case (b)) there exists $y^o < y$ such that $\hat{u}(y^o) = \widehat{\mathrm{id}}(y^o)$. In case (a), $\hat{u}(y) > \widehat{\mathrm{id}}(y),\ y < 0$ implies that $u(t) > t,\ t \in (0,1)$. The definition of u then implies that $F(x) > G(x),\ x \in (\underline{x}, \bar{x})$. In case (b), because $\hat{u}(y)$ is convex, it must be the case that for all $y > y^o$, $\hat{u}(y) \geq \widehat{\mathrm{id}}(y)$. Let $t^o = \exp(y^o)$, then reversing the transformation we have that $u(t^o) = t^o$, and for all $t \in (t^o, 1)$, $u(t) < t$. Letting $x = F^{-1}(t^o) = G^{-1}(t^o)$ establishes the result. If F strictly geometrically dominates G then $\hat{u}$ is strictly convex. The fact that $\hat{u}$ and $\widehat{\mathrm{id}}$ meet at 0, the strict convexity of $\hat{u}$, and the fact that $\hat{u}(y) > \widehat{\mathrm{id}}(y)$ when $y < \underline{y}$, then imply that $\hat{u}$ and $\widehat{\mathrm{id}}$ meet at, at most, one other point. Case (b) assumes that they meet and thus they must meet at exactly one point; call this point y^o. Convexity implies that for $0 > y > y^o$, $\hat{u}(y) < \widehat{\mathrm{id}}(y)$. Translating these results back to the u function and then back to the underlying distributions, then yields the result. The assertion that F is is a mean-preserving risk-increasing shift of G when the mean of F equals the mean of G then follows simply by from the fact that a F crosses G once from above and their means are by assumption equal. □

Proof of Theorem 3. For any real valued function of a single variable, f, let $D_+ f(x)$ represent the right derivative derivative of f evaluated at x. Note that a convex function has a right derivative at all points on the interior of its domain and that a necessary and sufficient condition for convexity of a function is that its right derivative is nondecreasing.

Proof of part (i)*:* Let

$$\hat{v}(y) = \hat{u}(y) - y, \quad y \leq 0. \tag{A-11}$$

First note that, because u is strictly geometrically convex, $\hat{u}$ is convex and thus $\hat{v}$ is convex. Also note that $\hat{v}(0) = 0$. As shown in Theorem 2, condition (i) implies that $\hat{u}(y) < y$ for all $y < 0$. Thus, $\hat{v}(y) < 0$ for $y < 0$. Next, we show that that $\hat{v}$ is increasing over $(-\infty, 0]$. To see this, note that, because v is convex, and thus has a right derivative, if it were not nondecreasing, there would be at least one point $y^o < 0$, such that $D_+\hat{v}(y^o) < 0$. Convexity implies that the line through y^o with slope $D_+\hat{v}(y^o)$ lies below $\hat{v}$. Thus, were $\hat{v}$ not nondecreasing, it would be the case that for all $y \leq 0$,

$$\hat{v}(y) \geq \hat{v}(y^o) + D_+\hat{v}(y^o)(y - y^o), \quad D_+\hat{v}(y^o) < 0, \tag{A-12}$$

which implies, for y sufficiently small, $\hat{v}(y) > 0$, a contradiction that establishes that $\hat{v}$ is nondecreasing. Using the definition of the conjugate function given in Lemma 1 we can see that

$$\hat{v} \circ \log(t) = \hat{u} \circ \log(t) - \log(t). \tag{A-13}$$

The definition of the conjugate function implies that $u(t) = \exp \circ \hat{u} \circ \log(t)$. Thus,

$$u(t) = t \exp[\hat{v}(\log(t)] \tag{A-14}$$

Because $\hat{u}$ is convex, it has left and right derivatives. Thus, u has left and right derivatives. We differentiate equation (A-14) and obtain

$$\begin{aligned} D_+u(t) = \exp[\hat{v}(\log(t))] + t \exp[\hat{v}(\log(t))] D_+\hat{v}(t)\, \frac{1}{t} = \\ \exp[\hat{v}(\log(t))] + \exp[\hat{v}(\log(t))] D_+\hat{v}(t), \quad t \in (0,1). \end{aligned} \tag{A-15}$$

Because $\hat{v}$ is nondecreasing and and convex, $D_+\hat{v}$ is positive and nondecreasing. Thus, equation (A-15) implies that D_+u is nondecreasing over (0, 1), which implies that u is convex. By replacing the assumption that u is geometrically convex with the assumption that it is strictly geometrically convex, it is straightforward to modify the proof to establish that u is strictly convex when F strictly dominates G.

Proof of part (ii)*:* Again define $\hat{v}(t) = \hat{u}(t) - t$ as in the proof of part (i). Note that, for the analogous reasons to those given in the proof of part (i), Theorem 2 implies that

$$\hat{v}(y) \geq 0, y \leq 0 \text{ and } \hat{v}(0) = 0. \tag{A-16}$$

Because $\hat{v}$ is convex, (A-16) implies that $\hat{v}$ is nonincreasing. To see this, note that $\hat{v}(y)$ cannot be increasing for all $y < 0$, otherwise $\hat{v}(0) > 0$, contradicting (A-16). Because $\hat{v}$ is convex, were it to be increasing anywhere it would have to be on an interval of the form $(y', 0]$. Thus, because it would be nonincreasing on $(-\infty, y']$, $\hat{v}$ would have a global minimizing value, say y'' in the interior of $(-\infty, 0]$. By (A-16), $\hat{v}(y'') \geq 0$. $\hat{v}$ would be increasing on $(y', 0]$, $\hat{v}(0) > \hat{v}(y'') \geq 0$, again contradicting (A-16). Thus, $\hat{v}$ is nonincreasing. Using equation (A-14), we can write

$$\frac{u(t)}{t} = \exp[\hat{v}(\log(t))],\ t \in (0,1]. \tag{A-17}$$

Because, exp and log are increasing functions and $\hat{v}$ is non-increasing, thus $t \to u(t)/t$ is non-increasing. When u is strictly geometrically convex, $\hat{u}$ is strictly convex and thus $\hat{v}$ is strictly decreasing and strictly convex. Thus, $t \to u(t)/t$ is decreasing. Because $\hat{v}$ is strictly convex and decreasing it is bounded from below by support line with a negative slope. Thus $\lim_{y \to -\infty} \hat{v}(y) = \infty$. Hence, (A-17) implies that, $\lim_{t \to 0} u(t)/t = \infty$.
Proof of part (iii)*:* The proof of part (iii) simply amounts to combining the arguments from the proofs of parts (ii) and (i) and therefore will be omitted. □

Proof of Lemma 5. From Chan, Proschan, and Sethuraman (1990) we see that

$$u(t) = \int_0^t \phi \circ G^{-1}(s)\, ds, \tag{A-18}$$

where ϕ is the Radon-Nikodym derivative of G with respect to F. If G and F are a regular pair of distributions, ϕ is absolutely continuous with repect to Lebesgue measure and is given by $\phi = f/g$. Thus,

$$u'(t) = \phi \circ G^{-1}(t). \tag{A-19}$$

First consider (i). Because G is continuous, G increasing and thus G^{-1} is increasing over $[\underline{x}, \bar{x}]$. Hence u is nondecreasing if and only if ϕ is nondecreasing, i.e., f/g is nondecreasing. Now consider (ii), For regularly related distributions, geometric convexity requires that R, defined in Lemma 4, be nondecreasing. Substituting the definitions of u and u' from equations (A-18) and $u = F \circ G^{-1}$ into R shows that

$$R(t) = \frac{\phi \circ G^{-1}(t)\, t}{F \circ G^{-1}(t)}. \tag{A-20}$$

Now make the substitution $s = G(t)$. This yields

$$R \circ G(s) = \frac{\phi(s)\, G(s)}{F(s)}. \tag{A-21}$$

Because G is increasing, $R \circ G(s)$ is nondecreasing if and only if $s \to \phi(s)\, G(s)/F(s)$ is nondecreasing. A similar argument establishes (iii). □

Proof of Proposition 4. To prove Proposition 4 we require the following lemma:

Lemma A.1. (i) If u is (strictly) convex at over $[0, \varepsilon)$, then whenever $0 < t \leq \varepsilon$

$$\Pi[u](t)(<) \leq \frac{1}{2}. \tag{A-22}$$

(ii) If u is (strictly) concave at over $[0,\varepsilon)$, then whenever $0 < t \leq \varepsilon$

$$\Pi[u](t) \geq (>) \frac{1}{2}. \tag{A-23}$$

Proof. We prove only (i) in the case, where convexity is not strict, because the proof for (ii) and the strictly convex case of (i) are virtually identical. First note that $u(0) = 0$; this fact combined with the convexity of u over $[0,\varepsilon)$ implies that the function $t \to u(t)/t$ is increasing. Hence,

$$0 < s < t \leq \varepsilon \Rightarrow \frac{u(s)}{s} \leq \frac{u(t)}{t}. \tag{A-24}$$

Therefore,

$$\frac{u(s)\,t}{u(t)\,s} \leq 1, \tag{A-25}$$

which implies that

$$\frac{u(s)}{u(t)\,t} \leq \frac{s}{t^2}. \tag{A-26}$$

If we integrate both sides of inequality (A-26) over $(0,t]$, we obtain

$$\Pi[u](t) = \int_0^t \frac{u(s)}{u(t)\,t}\,ds \leq \int_0^t \frac{s}{t^2}\,ds = \frac{1}{2}. \tag{A-27}$$

□

We need to show that for $t \in (0,1)$, $\Pi[u](t) > \Pi[u](1)$. First note that Lemma A.1 implies that that for $t \in (0,t^*]$, $\Pi[u](t) > 1/2$. By assumption (i), $\Pi[u](1) = 1/2$. Thus, by the continuity of $\Pi[u]$ if for some $t \in [t^*,1]$, it is the case that, $\Pi[u](t) \leq \Pi[u](1)$, then there must be two points, t' and t'' in $[t^*,1]$ such that, $t'' > t'$ and $\Pi[u](t') = \Pi[u](t'') = 1/2$. Expanding the definition of Π shows that this implies that

$$2\int_0^{t''} u(s)\,ds = u(t'')t'', \tag{A-28}$$

$$2\int_0^{t'} u(s)\,ds = u(t')t'. \tag{A-29}$$

Let $t^o = \inf\{t\colon u'(t) \geq u(t)/t\}$ if $\{t\colon u'(t) \geq u(t)/t\}$ is non empty and define $t^o = 1$ otherwise. Concavity below t^* implies that $t^o > t^*$. The continuity of u' implies that when $t^o < 1$, $u'(t^o) = u(t^o)/t^o$. For a strictly convex function such as u when restricted to $[t^*,1]$, the function $t \to t\,u'(t) - u(t)$ is strictly increasing. Thus,

$$\forall\, t > t^o, \quad u'(t) > \frac{u(t)}{t}. \tag{A-30}$$

Next, we claim that $t' \geq t^o$. To see this note that, if $t' < t^o$, then for all $s \in (0,t']$, $u(s) > u'(s)\,s$. Thus,

$$\int_0^{t'} u(s)\,ds > \int_0^{t'} u'(s)\,s\,ds = u(t')\,t' - \int_0^{t'} u(s).\,ds \tag{A-31}$$

Thus,

$$2\int_0^{t'} u(s)\,ds > u(t')\,t', \tag{A-32}$$

which contradicts (A-29). This implies that

$$\forall s \in (t',1),\; u'(t) > \frac{u(t)}{t}. \tag{A-33}$$

To complete the proof subtract (A-28) from (A-29). This yields

$$2\int_{t'}^{t''} u(s)\,ds = u(t'')t'' - u(t')t' = \int_{t'}^{t''} d(u(s)s) = \int_{t'}^{t''} u(s)\,ds + \int_{t'}^{t''} s\,u'(s).\,ds \tag{A-34}$$

Thus,

$$\int_{t'}^{t''} (u(s) - su'(s))\,ds = \int_{t'}^{t''} s\left(\frac{u(s)}{s} - u'(s)\right) ds = 0. \tag{A-35}$$

Expression (A-35) contradicts (A-30) and this contradiction establishes the result. □

Appendix B Proof of selection robustness and selection reversal for the Kumaraswamy Distribution

$$\frac{(\log\circ F(x))'}{(\log\circ G(x))'} = \frac{\alpha_F}{\alpha_G}\,\frac{\gamma(x^{\alpha_G})}{\gamma(x^{\alpha_F})}, \quad \gamma(x) = \left(\frac{1}{x} - 1\right)\left(\frac{1}{(1-x)^b} - 1\right),\; x \in [0,1]. \tag{B-1}$$

Lemma 5 shows that the left-hand side being increasing is a necessary and sufficient condition for geometric dominance. If $\alpha_F < \alpha_G$, then $x^{\alpha_F} > x^{\alpha_G}$. This observation combined with (B-1) shows that if the function, $x \to \log\circ\gamma(x)$ is decreasing, the right-hand side of equation (B-1) and thus the left as well is increasing and hence $F \underset{g}{\succcurlyeq} G$. Similarly, if the function, $x \to \log\circ\gamma(x)$ is increasing, then $G \underset{g}{\succcurlyeq} F$. It is clear from inspection that $G \underset{sd}{\succcurlyeq} F$ whenever $\alpha_F < \alpha_G$. Thus, if $\log\circ\gamma$ is decreasing $F \underset{g-}{\succcurlyeq} G$ and if $\log\circ\gamma$ is increasing $G \underset{g+}{\succcurlyeq} F$. The derivative of $\log\circ\gamma$ is given by

$$(\log\circ\gamma(x))' = \left(\frac{1}{x\,(1-(1-x)^b)\,(1-x)}\right)\left(x b - \left(1-(1-x)^b\right)\right). \tag{B-2}$$

The denominator on the right-hand side of equation (B-2) is positive. Thus, the sign of the expression depends on the numerator. Examining the numerator we see that

$$x b - \left(1-(1-x)^b\right) = \int_0^x \left(b - b(1-y)^{b-1}\right) dy. \tag{B-3}$$

Thus, $b < 1$, $(\log\circ\gamma(x))' < 0$ and thus $\log\circ\gamma$ is decreasing. If $b > 1$, $(\log\circ\gamma(x))' > 0$ and thus $\log\circ\gamma$ is increasing. Thus, when $\alpha_F < \alpha_G$, if $b < 1$, $F \underset{g-}{\succcurlyeq} G$ and, if $b > 1$, $G \underset{g+}{\succcurlyeq} F$. Of course, when $b = 1$, F and G are selection equivalent.

Appendix C Example of the intransitivity of selection dominance

Define the functions: $u_1 : [0,1] \to [0,1]$, $u_1 : [0,1] \to [0,1]$ as follows. Let

$$u_o(t) = \begin{cases} \frac{1}{2}\frac{t}{\eta_o} & \text{if } t \in [0,\eta_o) \\ \frac{1}{2} & \text{if } t \in [\eta_o, 1-\eta_o) \\ \frac{1}{2} + \frac{1}{2}\frac{t-(1-\eta_0)}{\eta_o} & \text{if } t \in [\eta_o, 1] \end{cases}, \tag{C-1}$$

where $\eta_o = 3/50$.

$$u_1(t) = p_o\, t + (1-p_o)\, u_o(t) \tag{C-2}$$

$$u_2(t) = \frac{(t+1)\log(t+1) - c_o\, t}{2\log(2) - c_o}, \tag{C-3}$$

where $c_o = 9/10$ and $p_o = 1/10$. It is easy to verify that u_1 and u_2 are an admissible functions. Thus, these functions define an admissible collection of distributions distributions, F, G, and H over the unit interval:

$$H(x) = x, \quad G(x) = u_2 \circ H(x), \quad F(x) = u_1 \circ G(x). \tag{C-4}$$

These distributions, as well as their associated selection-dominance functions, Π, defined in equation (14), are graphed in Figure C. Panels B and C of Figure C verify that u_1 and u_2 satisfy the supermultiplicativity on average condition given by by expression (15). Theorem 1 shows that supermultiplicativity on average of the u function is is necessary and sufficient for selection dominance. Thus, F selection dominates G and G selection dominates H. Because $F(x) = u_1 \circ u_2 \circ H(x)$, for F to selection dominate H it is necessary for $u_1 \circ u_2$ to satisfy the supermultiplicativity on average condition given in expression (15). This condition requires that $\Pi[u_1 \circ u_2]$ have a minimum value at $t = 1$. As Panel D shows, this is not the case. Thus, F does not selection dominate H. Hence, the selection dominance relation is not transitive.

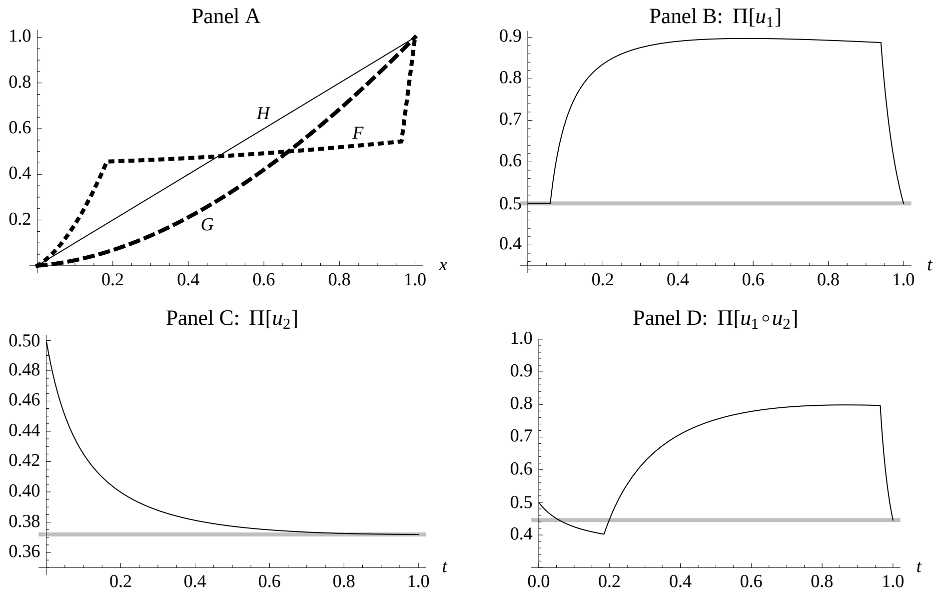

Figure 6: Counterexample to transitivity of selection dominance. Panel A plots the distribution functions, F G and H. Panel B plots the function, $\Pi(u_1)$, (defined in expression (14)) used to test the selection dominance of F over G, Panel C plots the function, $\Pi(u_2)$, used to test the selection dominance of G over H. Panel D plots the function, $\Pi(u_1 \circ u_2)$, used to test the selection dominance of F over H.