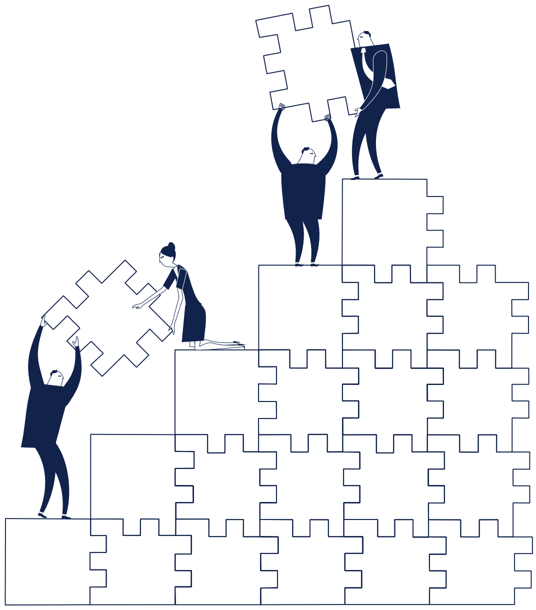
Saïd Business School
Research Papers

June 2015

Stochastic Orders and the Anatomy of Competitive Selection

Thomas Noe
Balliol College, University of Oxford; Saïd Business School, University of Oxford



Saïd Business School RP 2015-7

Stochastic orders and the anatomy of competitive selection

Thomas Noe
Balliol College/SBS*

1 June, 2015

Abstract

This paper determines the conditions under which stochastic orderings of random variables, e.g., stochastic dominance and the monotone likelihood ordering, are preserved or reversed by conditioning on competitive selection. A new stochastic order over unconditional distributions is introduced—*geometric dominance*—which is sufficient for a random variable, conditioned on competitive selection, to be stochastically dominant and both necessary and sufficient for it to be dominant under the monotone likelihood ratio ordering. Using geometric dominance, we provide an "anatomy of selection bias" by identifying the conditions under which competitive-selection conditioning preserves and reverses unconditional stochastic order relations. We show that, for all standard error distributions (Normal, Logistic, Laplace *inter alia*), competitive selection preserves stochastic order relations. To illustrate possible uses of these results, we develop simple conditions for signing treatment effects in self-selected samples and derive the selection-conditioned predictions of two auction-models (first-price and all-pay) from the models' unconditional predictions.

Key Words: stochastic orders, convexity, selection bias, stochastic dominance

The author would like to thank Alexei Boulatov and the participants in the Higher School of Economics research symposium for many helpful comments on an early draft of this paper. In addition, the author is grateful to Charles Manski, Todd Kaplan, and Thomas Hellmann, for many insightful comments on more recent drafts. The usual disclaimer applies.

*Park End Street, Oxford OX11HP UK, email:thomas.noe@sbs.ox.ac.uk

1 Introduction

Whether random prospects satisfy stochastic ordering relations has important consequences. For instance, a wealth maximizing investor who is attempting to construct an optimal portfolio can discard stochastically dominated assets from consideration (Hanoch and Levy, 1969). A statistician can construct a uniformly most powerful test if the parameter set is ordered by the monotone likelihood ratio property (MLRP) (Casella and Berger, 2002, Theorem 8.3.17). Theoretical comparative statics based on the single-crossing property and interval dominance order are robust to the introduction of uncertainty if compared probability density functions are ordered by the MLRP property.[1]

Whether or not random variables, or their associated distribution functions, satisfy these stochastic order relations, of course, will be affected by conditioning. One of the most ubiquitous forms of conditioning is competitive selection, or "self selection" in statistical parlance. Competitive selection occurs when the distributions of each of the compared random variables is conditioned on the random variable producing a higher payoff than its rival. Conditioning on competitive selection is a ubiquitous problem in statistics, econometrics, economic theory, and real-life decision making. Selection problems appear to come in two versions, which we term "posterior," and "prior." In both, the problem is to relate the unconditional and competitive selection-conditioned properties of the random variables considered. The posterior (or "statistical") selection problem has received by far the most attention. It involves reasoning from the expected valuation of the competitive-selection-conditioned distribution to the expected valuation of the underlying unconditional distribution. For example, suppose we find that one drug treatment, say treatment A, produces better outcomes on average another treatment, say treatment B. However, patients were not randomly assigned to treatments A or B, but rather picked the treatment that produced the best outcome in their specific case. The problem is inferring whether treatment A would have produced a better outcome than treatment B had patients been randomly assigned to treatments.

The prior selection problem arises when a theorist or decision maker has prior beliefs about expected valuations of unconditional distributions but must make a decision or develop an empirical prediction relating to the expected valuations under the competitive selection-conditioned distribution. For example, consider a theorist trying to spin an empirical prediction on the terms of procurement contracts out a comparative static result which predicts that financially constrained firms will bid less aggressively in first-price sealed-bid procurement auctions. The empirical prediction concerns the expected value of competitive-selection-

[1] See Milgrom and Shannon (1994) for a definition of the single crossing property or Topkis (1998) for a more extensive treatment of the closely related concept of supermodularity; See Quah and Strulovici (2009) for a definition of interval dominance order its relation to MLRP. See Athey (2002) for an analysis of the role of MLRP in applying monotone comparative statics to uncertain prospects; see Milgrom (1981) for a comprehensive discussion of the MLRP ordering in the context of economics and finance.

conditioned winning bids while the theoretical comparative static relates to unconditional bids. Developing competitive-selection-conditioned predictions from the theory substantially facilitates its testing because such predictions, being predictions about the selected sample, such predictions are robust to competitive-selection-biased data.

This paper will provide a fairly complete answer to these sorts of questions. We start by deriving the necessary and sufficient conditions on unconditional distributions which ensure that one distribution competitive selection stochastically dominates (CSSD) another, i.e., one distribution, conditioned on competitive selection, stochastically dominates another. We show that CSSD depends only on how the quantile functions of the two distributions are related and not on properties such as the range of their support or their absolute continuity with respect to Lebesgue measure. We develop a function, the quantile transform function, whose behavior completely determines whether the distributions are related by CSSD. These necessary and sufficient conditions for the CSSD relation do not define an order relations over distribution functions because the CSSD relation is not transitive. However, two sufficient conditions for CSSD based on the character of the transform function, do order distribution functions. These conditions relate to the "geometric convexity" of the quantile transform function. Geometric convexity, a recent topic of extensive mathematical research, especially with respect to its connection with Mulholland's inequality, has received little attention in economics and statistical research.[2] Yet it captures the exact property—supermultiplicativity of the quantile transform function—that generates competitive selection conditioned dominance. The weakest and most general of these two conditions is for the quantile transform function to be geometrically convex on average. The second, stronger, and less general condition for CSSD is that the quantile transform function be geometrically convex. The second condition is much more tractable and has the added advantage that it is also necessary and sufficient for one distribution to dominate another, conditioned on competitive selection, under the likelihood ratio ordering. Thus, geometric dominance imposes the same restrictions on competitive-selection-conditioned distributions as the monotone likelihood ratio property (MLRP) imposes on unconditioned distributions.

Translating these characterizations of the quantile transform into characterizations of the underlying distributions yields simple characterizations of the distributional requirements for competitive-selection conditioned dominance. In the case of geometric dominance, the condition is simply that the ratio between the reverse hazard rates of the dominating and dominated distributions be increasing.[3]

Using these results we show that geometric dominance implies that either (a) the geometri-

[2]See Jarczyk and Matkowski (2002) for applications of geometric convexity to Mulholland's inequality and see Baricz (2010) for an analysis of the geometric concavity of univariate distribution functions.

[3]The reverse hazard rate of a distribution is the ratio between its probability density function and cumulative distribution function.

cally dominant distribution is stochastically dominant, or (b) the geometrically dominant distribution is stochastically dominated or (c) the distribution function of the geometrically dominant distribution crosses the dominated distribution's function once from above, i.e., roughly the geometrically dominant distribution is more dispersive.

We call case (a), the case where the geometric dominance and stochastic dominance orders agree,"selection preservation" and term case (b) "selection reversal." We derive the restrictions on the distribution functions implied by both selection preservation and selection reversal and show that preservation and reversal impose non-analogous restrictions on the behavior of the underlying distribution functions. Preservation requires that the distributions be ordered by the monotone likelihood ratio property (MLRP). This condition can be interpreted as the dominance of the stochastically dominant distribution being fairly uniform over its support. Reversal is consistent with but does not require the geometrically dominant distribution to be MLRP dominated. However reversal imposes very strong conditions on the lower tail behavior of the compared distributions—the ratio between the lower tail weight of the geometrically dominant and dominated distributions must explode as the lower endpoint of the support of the distributions is approached. Roughly speaking, selection reversal only occurs when the stochastically dominant distribution is dominant because its worst draws are much better than the worst draws from the stochastically dominated distribution rather than because its best draws are better than the best draws from the stochastically dominated distribution. Very bad draws from the stochastically dominated distribution "admit" fairly bad draws from the stochastically dominant distribution into its selection-conditional sample. At the same time these bad draws cannot compete with the draws from the stochastically dominant distribution and thus are excluded from its selection conditioned sample, resulting in the selection-conditioned stochastic dominance of the unconditionally dominated distribution. This result, and others developed in the paper, provide an anatomy of competitive selection bias, a characterization of the distributional properties of random variables that render the effect of competitive selection on inference mild or severe.

We next turn to applying these characterizations to standard textbook statistical distributions. As a desideratum, we first develop simple characterizations for geometric dominance that apply when the compared distributions are related by a location or scale shift. These conditions relate to the logconcavity or geometric concavity of the reverse hazard rate of the standardized member of the distribution's location/scale family. Using these conditions, and the more general results developed earlier, we completely characterize selection preservation and reversal for a large number of distribution families, including the Normal, Lognormal, Gamma, Logistic and Log-logistic. We show that, within all of these families and some others, competitive selection preserves dominance.

Finally, using the anatomy of selection bias developed in the preceding sections of the paper, we apply our results to illustrative inference problems. We first consider a posterior inference problem—the classic self-selection problem first posed in Manski (1990) of signing average treatment effects in the presence of self selection bias. Our results show that under standard error law assumptions, the self-selection effect is fairly innocuous in that the sign of the average treatment effect will always be identified by standard linear regression models.

Next, we apply our results to two related prior inference problems: deriving the relation between the unconditional and winning-contingent predictions in first-price and all-pay auction models. In the first-price auction setting, we show the stochastic dominance of the high-valuation bidder's unconditional bid distribution is preserved by competitive selection, i.e., the high-valuation bidder's bid distribution conditioned on the high-valuation bidder winning the auction stochastically dominates the low-valuation bidder's bid conditioned on the low valuation bidder winning the auction. In contrast, we show that, in an all-pay auction setting, selection reverses inference, although the unconditional bid distribution of the high-valuation bidder stochastically dominates the bid distribution of the low-valuation bidder, the winning bid distribution of the high-valuation bidder is stochastically dominated by the winning bid distribution of the low-valuation bidder.

The analysis in this paper derives the relation between unconditional outcome distributions and competitive-selection-conditioned distributions. Its place in the literature depends on whether one views the results in terms of reasoning from the properties of the unconditional distribution to the competitive-selection conditioned distribution or vice versa. Viewed from the perspective of determining the effect of competitive selection conditioning on the unconditional distribution of outcomes, this paper is part of a vast research program in economics, mathematics, statistics, and operations research that has examined stochastic orders.[4] Like these research papers, this paper identifies a property of interest, in our case competitive selection dominance, and derives an ordering over distribution functions which implies that the property holds. Perhaps its closest relative is the Milgrom (1981) analysis of the MLRP ordering. Like Milgrom, we consider conditions under which conditioning preserves stochastic dominance. Under the MLRP ordering, the MLRP dominant (and thus the stochastically dominant) distribution will remain stochastically dominant conditioned on sampling from any fixed subset of value levels.[5] In our analysis, in contrast to Milgrom's, conditioning is not with respect to a fixed range but rather based on the realized values of the compared distribution.

If viewed from the perspective of determining the properties of unconditional distributions from selection-conditioned distributions, this paper is related to identifying the sign of treatment effects in the presence of selectivity bias. In fact, one application we consider maps exactly into the Manski (1990) formulation of the problem of identifying the sign of treatment effects in the presence of self-selection bias. Manski's approach to this problem is nonpara-

[4]See Shaked and Shanthikumar (1994) for an exhaustive discussion of these results.

[5]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.

metric and involves imposing restrictions on the supports of the distributions of the treatment effects and using ancillary information about the probability of selection. Manski's aim is to determine the sign of the expectation of the unconditional treatment effect. Our approach is "quasi-parametric" in that it does not impose a specific distributional form on treatment effects but does impose stochastic order relations on outcomes under the alternative treatments. Our identification of the sign does not use any ancillary information, such as the probability of selection or any instrumental variables. Moreover, our approach signs all monotonically increasing functions of the unconditional treatment effect rather than simply its expectation.

The results in this paper are also, albeit more loosely, related to papers that impose specific distributional conditions and aim to identify average treatment effects rather than simply the sign of treatment effects. For example, the Heckman (1979) correction for selection bias imposes a normality assumption and identifies average unconditional treatment effects using the cross-sectional variation in the probability of selection estimated by an ancillary Probit prediction model. Identification results from the fact that the unconditional mean of a Normal random variable is linearly related to its truncated expectation and its inverse Mills ratio evaluated at the truncation point. Perhaps the closest relation to our paper in this literature is Heckman and Honoré (1990), which examines the conditions under which imposing qualitative distributional restrictions (e.g., logconcavity of the probability density functions) permits identification of average treatment effects. Our stochastic order approach differs from this approach because it does not use ancillary information, imposes weaker distributional conditions, and aims to identify cases where one treatment's effect stochastically dominates another's rather than identify the magnitude of unconditional expected treatment effects. Finally, our technical development is indebted to recent mathematical research related to supermultiplicative functions and geometric convexity, for example, Niculescu and Persson (2004) and Kuczma (2000).

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 in the realization of the selected random variable.[6] A value-maximizing decision maker will be willing to select $\tilde{X}$ whenever the value it produces, $v(\tilde{X})$, at least equals the value produced by the alternative choice, $\tilde{Y}$. Because v is a strictly increasing function, $v(\tilde{X}) \geq v(\tilde{Y})$ if and only if $\tilde{X} \geq \tilde{Y}$. The expected value of

[6]We follow the standard convention in convex analysis and use the term "increasing" to signify that a function is weakly monotonically increasing (e.g., nondecreasing) and the term "strictly increasing" to signify that it is strongly monotonically increasing.

$\tilde{X}$ conditioned on $\tilde{X}$ being selected is thus $\mathbb{E}[v(\tilde{X})|\tilde{X} \geq \tilde{Y}]$. A natural question to pose in this context is when will the 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) dominates $\tilde{Y}$ (or its distribution function G) under *competitive selection stochastic dominance* (CSSD) if, for all integrable strictly increasing functions v,

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

If the inequality in (1) holds strictly, we will say that $\tilde{X}$ *strictly CSSD dominates* $\tilde{Y}$.

This paper considers 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 $\tilde{X}$ CSSD dominates $\tilde{Y}$. Our primary aim to determine how competitive selection affects dominance, i.e., to determine the conditions that must be imposed on the unconditional distributions, F and G, in order to ensure CSSD dominance and to determine how these conditions relate to unconditional dominance.

The criterion for dominance in the absence of competitive selection is well known—stochastic dominance:

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

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 $[\underline{x}, \bar{x}]$, $-\infty \leq \underline{x} < \bar{x} \leq \infty$.

2. G is continuous and absolutely continuous with respect to F and F is continuous except perhaps at $\underline{x}$.

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

A collection of distribution functions is admissible if all pairs in the collection are admissible.

Note that condition 1 does not rule out unbounded supports. It does rule out gaps in the support. This no-gaps condition is not necessary to derive our results. However, absent this condition, 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 of points.

The assumption that the supports of the two distributions are identical involves little loss of generality. First suppose that the upper bound of support of F exceeds the upper bound of the support of G. In this case, F could never be CSSD dominated by G. To see this, simply consider a sequence of valuation functions whose limit equaled a valuation function that is constant on the intersections of the supports and positive on the part of the support of F not contained in G. Next suppose that the lower bound of the support of F is less than the lower bound of the support of G, then draws from F below the lower bound of G's support would never be selected. Thus, the competitive selection-conditioned distribution of F would be the same as it would have been if F simply had a point mass at the lower bound of G's support. This case is analyzed in the paper.

The assumption that G is continuous and absolutely continuous with respect to F eliminates the problem of ties. This assumption combined with the assumption that F is continuous except perhaps at $\underline{x}$ ensures that on their common support, $[\underline{x}, \bar{x}]$ F and G are continuous and strictly increasing. Note that we do not assume that F or G are absolutely continuous with respect to Lebesgue measure and thus have associated probability density functions. The assumption that F and G have finite expectations is simply made to ensure the the expected value of the canonical valuation function, $v(x) = x$, is finite.

For admissible pairs of distributions, we can express conditioning on competitive 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 Fubini's Theorem, we have that

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

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

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

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

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

Define the probability distribution functions H and J by

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

Using H and J, we can express the conditioning relation as a simple expectations over distributions:

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

Thus, for competitive selection stochastic dominance (CSSD), formalized by inequality (1), to hold, it is necessary and sufficient that H stochastically dominate J, i.e., $H(x) \le J(x)$. From equation (7), we see that stochastic dominance of H over J is equivalent to the condition that for all $x > \underline{x}$,

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

Integration by parts shows that

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

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

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

As the subsequent analysis will show, the conditions on the unconditional distributions, F and G, that are necessary and sufficient for CSSD are neither very tractable nor do they yield an order relation over the unconditional densities. However, we can immediately show that the conditions for MLRP dominance of one competitive-selection-conditioned distribution over another are transparent when both unconditional distributions are absolutely continuous. Note that in this case, the densities of the competitive-selection-conditioned distributions, H and J, defined in equation (9), are

$$h(x) = f(x)G(x), \quad j(x) = g(x)F(x), \quad x \in (\underline{x}, \bar{x})$$

If the ratio h/j is increasing, the competitive-selection-conditioned distribution of F dominates the competitive-selection-conditioned distribution of G in the MLRP ordering. This motivates the following definition and Lemma.[7]

[7]Since we are only concerned in this paper with competitive selection, for the sake of avoiding long strings of

Definition. If $x \hookrightarrow (f(x)\,G(x))/(g(x)\,F(x))$ is increasing then F dominates G in the *competitive selection-conditioned MLRP* ordering and strictly dominates if $x \hookrightarrow (f(x)\,G(x))/(g(x)\,F(x))$ is strictly increasing.

Using the well-known properties of the MLRP ordering, we see that competitive-selection-conditioned MLRP imposes stronger restrictions on the expected valuations than CSSD.

Lemma 1. *Let v be an increasing integrable function. Let B be a measurable set, and assume that $\mathbb{P}[\tilde{X} \in B] > 0$ and $\mathbb{P}[\tilde{X} \in B] > 0$. Let $\tilde{X}$ and $\tilde{Y}$ be independent random variables with distribution functions F and G respectively. Then if F competitive selection conditioned MLRP dominates G then*

$$\mathbb{E}[v(\tilde{X})|\tilde{X} > \tilde{Y} \;\&\; \tilde{X} \in B] \geq \mathbb{E}[v(\tilde{Y})|\tilde{Y} > \tilde{X} \;\&\; \tilde{Y} \in B],$$

with the weak inequality being replaced by a strict inequality if the dominance is strict.

Proof. Follows directly from the MLRP dominance of the selection-conditioned distribution of F and Theorem 1.C.2 in Shaked and Shanthikumar (1994). □

Thus, if two distributions are ordered by competitive selection-conditioned MLRP, the dominant distribution not only produces higher expected value conditioned on selection but also when conditioning on both selection and any measurable subset of the range of possible realizations.

3.1 The quantile transform function, u

The key to deriving the distributional restrictions implied by inequality (11) is noting that only the behavior of the distributions relative to each other matters. Transforming both distributions by the same continuous strictly increasing function will not affect the validity of (11). This permits us to reduce the dimensionality of the problem by using a transform function. First note that, because regular pairs of distributions are continuous and strictly increase over their support, G and F have well-defined strictly 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. We will refer to u simply as the *quantile transform function* or simply the *transform function*.

The fact that G and F regularly related implies that u is continuous on $(0,1)$ and that $\lim_{t\uparrow 1} u(t) = 1$ and $\lim_{t\downarrow 0} u(t) = F(\underline{x})$ Thus, we can extend the definition of the quantile transform function, u, by defining $u(1) = 1$ and $u(0) = F(\underline{x})$. The resulting function, $u : [0,1] \to [0,1]$, will be strictly increasing and continuous.

compound adjectives, we will frequently simply use the term "selection" to signify competitive selection.

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

An intuitive representation for u is provided by noting that, restricted to $(0,1)$, the graph of u is described by the parametric equations:

$$t = G(x), \quad u = F(x), \quad \underline{x} \leq x \leq \bar{x}.$$

The values of u at the endpoints 0 and 1 are simply the limit points of the graph. Figure 1 sug-

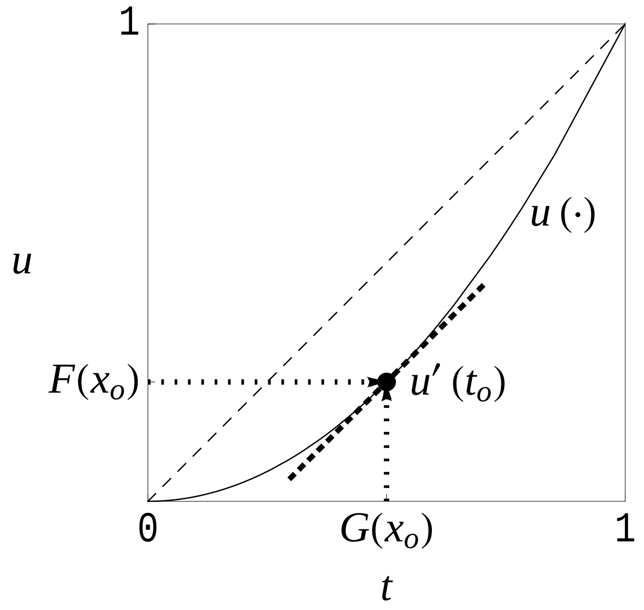


Figure 1: *Parametric representation of the transform function u.* Each point in the graph of the transform function (t,u) represents the ordered pair $(G(x), F(x))$. The graph is produced by varying x across the support of the distribution functions,$[\underline{x},\bar{x}]$

gests three important properties of the quantile transform function, u, which we will formalize and use in the sequel. First, it is apparent that $F[x] \leq G[x]$, $x \in [\underline{x},\bar{x}]$, if and only if $u(t) \leq t$, $t \in [0,1]$. Thus, stochastic dominance of F over G is equivalent to $u(t) \leq t$. Second, because the parametric representation of a curve is not unique, two pairs of distribution functions may generate exactly the same transform function even when the pairs are very different. In fact, this can occur even when the pairs selected are standard textbook distributions. This fact makes determining selection dominance using the u transform function sometimes much more efficient than determining selection dominance using the underlying distribution functions. Third, when F and G are differentiable, the slope of the graph of the transform function, u' at a given point equals the likelihood ratio, $F'/G' = f/g$ at the value of x which produces that point.

3.2 Conditions for CSSD dominance

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

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

Using the change of variables formula on equation (12) shows that $\tilde{X}$ CSSD 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

$$U(t) = \begin{cases} \frac{1}{t} \int_0^t u(s)\, ds & \text{if } t \in (0,1], \\ u(0) & \text{if } t = 0. \end{cases} . \tag{14}$$

$U(t)$ represents the average value of u over the interval $[0,t]$. Note that U is continuous over $[0,1]$ and that the continuity of u ensures that U is continuously differentiable over $(0,1]$. Next, note that $(U(t)\,t)' = u(t)$. Therefore,

$$t\,U'(t) + U(t) = u(t), \quad t \in (0,1]. \tag{15}$$

Finally, define the function, $\Pi[u]$ by

$$\Pi[u](t) = U(t)/u(t) = \frac{1}{t} \int_0^t \frac{u(s)}{u(t)}\, ds, \quad t \in (0,1]. \tag{16}$$

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 (16) to hold is that

$$\forall t \in (0,1), \quad \Pi[u](t) \geq \int_0^1 u(s)\, ds \equiv \Pi[u](1). \tag{17}$$

These observations provide the necessary and sufficient conditions on unconditional distribution of F and G which ensure that the competitive selection-conditioned distribution of F stochastically dominates the competitive selection distribution of G. This result is recorded below.

Theorem 1. *Suppose that F and G are an admissible pair of distribution functions; Let $u = F \circ G^{-1}$, then F (strictly) CSSD dominates G, if and only if u satisfies condition* (17).

Although Theorem 1 provides necessary and sufficient conditions for CSSD, it does have two limitations. First, the relation between the unconditional distributions and the Π function which provides the necessary and sufficient conditions is opaque. Second, as we will show later, the CSSD relation has a rather significant drawback—it does not define an order relation over the unconditional distributions.

For this reason we develop two sufficient conditions for CSSD. These conditions are much more transparently related to the underlying unconditional distributions and do define an order relation. The first and perhaps most obvious sufficient condition for (17) is that $t \hookrightarrow \Pi[u](t)$ be decreasing. Inspecting (17) shows that $t \hookrightarrow \Pi[u](t)$ being (strictly) decreasing is equivalent to the condition that

$$t \hookrightarrow \frac{u(t)}{U(t)} \text{ is (strictly) increasing.} \tag{18}$$

To develop the second sufficient condition, first note that we can rewrite expression (17) in the following alternative forms by a change of variables in the integral:

$$\text{For all } t \in (0,1), \int_0^1 \Big(u(t\,s) - u(s)\,u(t) \Big)\, ds \geq 0. \tag{19}$$

A sufficient condition for (19) to hold is that

$$\text{for all } s,t \in (0,1),\ u(st) \geq u(s)\,u(t). \tag{20}$$

Whether these two sufficient conditions hold depends on whether $u(st) > u(s)\,u(t)$, i.e., on whether u is "supermultiplicative." As we show in the next section, the answer to this question hinges of whether u is "geometrically convex."

3.3 Geometric convexity

The basic results from the theory of geometric convexity required to demonstrate this assertion are provided below. The definition of geometric convexity results from replacing the arithmetic mean used to define convexity with the geometric mean.

Definition (Geometric Convexity)**.** Let I and J be subintervals of $(0,\infty)$, then a continuous function $\phi : I \to J$ is *geometrically convex* if, for all $s,t,\in I$ and $\alpha \in (0,1)$,

$$\phi(s^{\alpha}\, t^{1-\alpha}) \leq \phi(s)^{\alpha}\, \phi(t)^{1-\alpha}. \tag{21}$$

If the weak inequalities are replaced with strict inequalities, we will say that ϕ is *strictly geometrically convex*. Geometric concavity is defined analogously by reversing the inequality in (21).

The following theorem, which is essentially specialized and simplified statement of Theorem 1 in Finol and Wójtowicz (2000) and Lemma 2.1.1 in Niculescu and Persson (2004), provides the basic characterization.

Lemma 2. *Let I and J be subintervals of $(0,\infty)$ let $\phi : I \to J$ be a continuous increasing function. Then the following statement are equivalent.*

(i) ϕ *is* geometrically convex.

(ii) The conjugate function $\hat{\phi} : \log(I) \to \Re$ defined by $\hat{\phi}(y) = \log \circ \phi \circ \exp(y)$ is continuous, convex and increasing.

(iii) If I is an interval of the form $(0,A]$*,* $A < \infty$*, The equivalent conditions,* (i) *and* (ii) *imply that* ϕ *is supermultiplicative, i.e.,*

$$\forall (s,t) \in (0,A] \times (0,A], \quad \phi(st) \geq \phi(s)\,\phi(t). \tag{22}$$

These statements remain valid if we replace "convex" with "concave," "super" with "sub" and reverse the inequalities.

Proof. Theorem 1 in Finol and Wójtowicz (2000) establishes all of the results except the assertions that $\hat{u}$ is increasing. Thus result simply follows from the fact that the we have added the hypothesis that ϕ is increasing and thus conjugate map, which involves the composition of ϕ with increasing functions is also increasing. □

If ϕ is differentiable, then the following Lemma provides a simple test for verifying geometric convexity.

Lemma 3. *Suppose* ϕ *is differentiable, then* ϕ *is (strictly) geometrically convex if and only if the function R defined by*

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

is (strictly) increasing over I

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

The geometric convexity of ϕ is equivalent to $\log \phi(t) = \hat{\phi}(\log t)$, with $\hat{\phi}$ convex. In fact, just as the transform function, u, has a parametric representation in terms of the distribution functions, so the conjugate transform function, $\hat{u}$, has a parametric representation in terms of the logs of the distribution functions, i.e.,

$$y = \log \circ G(x), \quad \hat{u} = \log \circ F(x), \quad \underline{x} < x \leq \bar{x}.$$

This parametric representation is depicted in Figure 2. The figure suggests, as we will demonstrate later, that, when the distributions are differentiable, the slope of $\hat{u}(\cdot)$ is determined by the ratio between $(\log \circ F)' = f/F$ and $(\log \circ G)' = g/G$. Geometric convexity neither implies or is implied by convexity.[8] Convexity is a weaker condition than logconvexity. For increasing functions, geometric convexity is a weaker condition than logconvexity. [9]

[8]E.g., consider the functions $\phi_1(t) = 1/2 + 1/2\sqrt{t}$ and $\phi_2(t) = e^{-(1-t)}t^3$, $t \in (0,1]$. Using the derivative test provided by Lemma 3, is easy to verify that ϕ_1 is both strictly concave and strictly geometrically convex while ϕ_2 is both strictly convex and strictly geometrically concave.

[9]logconvexity requires $\log \circ \phi$ to be convex. Given our assumption that ϕ is increasing, the geometric mean–arithmetic mean inequality ensures that logconvexity implies geometric convexity. By a different argument, logconvexity implies convexity as well.

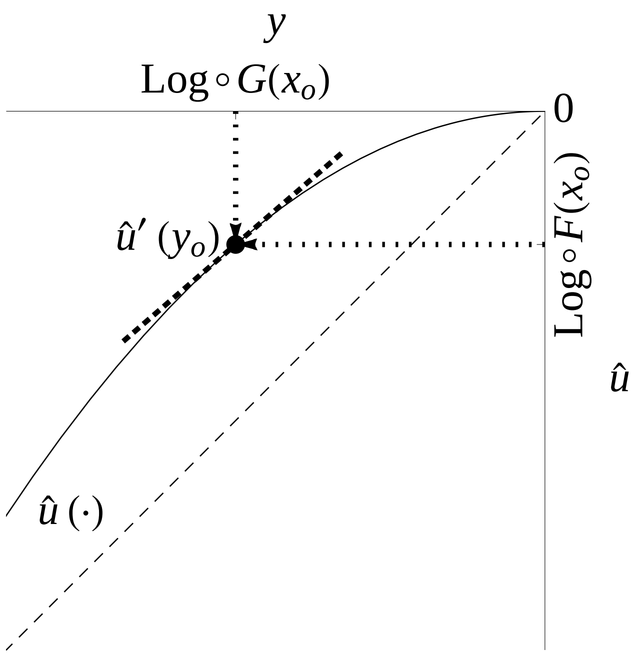


Figure 2: *Parametric representation of the conjugate transform function u.* Each point in the graph of the conjugate transform function $(y, \hat{u})$ represents an ordered pair $(\log \circ G(x), \log \circ F(x))$. The graph is produced by varying x across the support of these distributions. $(\underline{x}, \bar{x}]$

.

3.4 Geometric convexity and selection dominance

Using Lemma 2 the following result is immediate.

Theorem 2. *Suppose that F and G are an admissible pair of distribution functions; Let $u = F \circ G^{-1}$, then if u is (strictly) geometrically convex, F (strictly) CSSD dominates G.*

Proof. As discussed above, condition (20) is sufficient for CSSD. By Lemma 2, u satisfies this condition when u is geometrically convex. □

A slightly less obvious but weaker sufficient condition for CSSD is that the average function, U defined in equation (14) is geometrically convex. This result is stated and proved below.

Theorem 3. *Suppose that F and G are an admissible pair of distribution functions; Let $u = F \circ G^{-1}$, then if u is (strictly)* geometrically convex on average*, i.e., the average function U defined by equation* (14)*, is (strictly) geometrically convex, F (strictly) CSSD dominates G.*

Proof. By Lemma 3, U being (strictly) geometrically convex implies that $t \hookrightarrow (U'(t)\,t)/U(t)$

is (strictly) increasing. Equation (15), implies that

$$\frac{U'(t)\,t}{U(t)} = \frac{u(t)}{U(t)} - 1. \tag{24}$$

The right hand side of this equation is the sufficient condition for CSSD given in equation (18).

□

Next note that u being geometrically convex implies that U is geometrically convex on average. This result is recorded below.

Lemma 4. *If u is an admissible function, and u is (strictly) geometrically convex, then it is (strictly) geometrically convex on average.*

Proof. Let,

$$I(t) = \int_0^t u(s)\,ds.$$

By Montel's theorem (Niculescu and Persson, 2004, Theorem 2.4.1), u being (strictly) geometrically convex implies that I is (strictly) geometrically convex. Next note that $U(t) = I(t)/t$. Using the conjugate expression definition for geometric convexity given in Lemma 2.ii, we see that $\hat{U}(y) = \widehat{I/t)}(y) = \hat{I}(y) - y$. Because I is (strictly) geometrically convex, $\hat{I}$ is (strictly) convex. The sum of a linear function and a (strictly) convex function is (strictly) convex. Thus, $\hat{I}(y) - y = \hat{U}(y)$ is (strictly) convex. By definition, this implies that U is (strictly) geometrically convex. □

3.5 Interpreting the average transform function, U

The interpretation of the average transform function, U, is not quite as straightforward as the interpretation of the transform function. First, note that a simple change of variables shows that

$$U \circ G(x) = \frac{1}{G(x)} \int_{\underline{x}}^{x} F(z)\,dG(z). \tag{25}$$

Thus, the average transform function evaluated at $G(x)$ represents the probability that draws from F will be less than draws from G conditioned on the draws from G being less than x, i.e.,

$$U \circ G(x) = \mathbb{P}\Big[\tilde{X} \le \tilde{Y} | \tilde{Y} \le x\Big], \quad \tilde{X} \underset{d}{\sim} F,\ \tilde{Y} \underset{d}{\sim} G.$$

If $x = \bar{x}$ then $U \circ G(\bar{x})$ is simply the probability that draws from F will be less than draws from G, i.e.,

$$U \circ G(\bar{x}) = U(1) = \int_0^1 u(s)\,ds = \mathbb{P}\Big[\tilde{X} \le \tilde{Y}\Big]. \tag{26}$$

Thus, if F (strictly) dominates G by stochastic dominance, $U(1)(<) \le 1/2$, In general, because (strict) stochastic dominance implies that $u(t)(<) \le t$, integrating and averaging both sides over the interval $(0,t]$ and dividing by t shows that,

$$U(t)(<) \leq \frac{1}{2}t. \tag{27}$$

The reverse inequality hold when F is stochastically dominated by G.

4 Geometric and CSSD dominance as orderings

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

$$F \underset{sd}{\succcurlyeq} G \text{ if } F(x) \leq G(x), x \in (\underline{x}, \bar{x}), \tag{28}$$

is a necessary and sufficient condition for F to dominate G. If the inequality in expression (28) 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. Suppose that F and G are an admissible pair of distributions and that $u = F \circ G^{-1}$, then if u is (strictly) geometrically convex, we will say that F *(strictly) geometrically dominates* G. If U is (strictly) geometrically convex, we will say that F *(strictly) geometrically on average dominates* G. When F dominates G in the (average) geometric convexity ordering, we will write $(F \underset{ga}{\succcurlyeq} G) F \underset{g}{\succcurlyeq} G$.

As the next lemma, Lemma 5, 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 competitive selection, there is no order relation over distribution functions which is necessary and sufficient to ensure that one distribution dominates another. However, geometric dominance and geometric dominance on average are preorders over distribution functions.

Lemma 5. *(i) The Geometric dominance and average geometric dominance relations are preorders over the set of distributions. (ii) The CSSD dominance relation is not a preorder over the set of distributions because CSSD dominance is not transitive.*

Proof. See Appendix A.

We next show that, although geometric dominance and geometric dominance on average are pre-orders, neither is a partial order because it is possible for F and G to be distinct distributions yet $F \underset{g}{\succcurlyeq} G$ and $G \underset{g}{\succcurlyeq} F$ ($F \underset{ga}{\succcurlyeq} G$ and $G \underset{ga}{\succcurlyeq} F$) When this occurs, we will say that F and G are *geometrically equivalent (on average)*. If, for an admissible pair of distributions, F and G, F CSSD dominates G and G also CSSD dominates F, we will say that the pair is *CSSD equivalent*.

A pair of distributions is selection equivalent, the stochastic dominance is reflected in their probability of being selected rather than dominance conditioned on being selected. In general, neither geometric dominance nor geometric dominance on average are necessary conditions for CSSD dominance. However, as the next lemma demonstrates, geometric and average geometric dominance are both necessary and sufficient conditions for CSSD equivalence. Moreover, CSSD equivalence imposes very strong conditions on distribution functions: they must be related by a power transform.

Lemma 6. *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 geometrically equivalent on average.

(iii) F and G are CSSD equivalent.

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

Proof. See Appendix A.

This result is very intuitive, especially under the assumption that the p in part (iv) of the Lemma is a positive integer. In which case, if condition (iv) is satisfied, F is the distribution of the maximal order statistic resulting from p independent draws from G. Thus, one can think of G as competing against p independent copies of itself for selection. Each of the p copies of G is selected if its value exceeds the value of the one draw from G and the value realized by the other $p-1$ copies, i.e., when its value exceeds the value of p independent draws from the G distribution. This is exactly the same condition that a draw from G must satisfy for selection. Thus, the expected value conditioned on selection for each of the p copies is the same as G's expected value. Since the expected value of draws from F conditioned on selection equals the average of the selection-conditioned values of the copies, the selection-conditioned expectation of G is the same as the selection-conditioned expectation of F. Of course, if p is large, draws from F are much more likely to be selected than draws from G, but the selection-conditioned distributions of F and G will still be identical.

This result has two significant implications: First, it applies that our results, which compare selection-conditioned distributions of two random variables, extend to multiple random variables under the assumptions that the random variables are independent draws from two different distributions and conditioning is with respect to the draw being the maximum draw. The second implication is that, if we compare distributions that are different powers of the same underlying distribution, these distributions are selection equivalent under all three relations. For example, all distributions of the form $F(x) = x^p$, $x \in [0,1]$ are equivalent. A less obvious example is provided by the Extreme-value type distributions (e.g., Fréchet and Gumbel distributions) with the same shape but different location or scale parameters.

5 Geometric convexity in terms of probability densities

We aim to characterize the extent to which our order relations can be applied to the standard distributions used in economics and finance research. Typically, these distributions have absolutely continuous distribution functions with respect to Lebesgue measure and thus are characterized by their probability density functions. Restricting attention to distribution functions that are sufficiently regular so that they can be characterized by their density functions permits us to both develop more intuitive characterizations of geometric and average geometric dominance and to develop tests that are frequently easy to apply to verify dominance relations between distributions. These conditions and tests are applicable when the distributions' relationship satisfies the following regularity conditions.

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

(i) $F(0) = G(0) = 0$.

(ii) F and G are absolutely continuous with respect to Lebesgue measure on $(\underline{x}, \bar{x})$.

(iii) On $(\underline{x}, \bar{x})$, F and G have density functions, f and g that are strictly positive and continuous.

Regularity implies that u is differentiable and thus Lemma 3 can be used to verify that u is geometrically convex. It also implies that F and G have probability density functions. Under regularity, Lemma 3 permits us to produce a simple mapping between the properties of the transform function, u, and the properties of the underlying distribution and density functions of the random variables generating u. This mapping is provided by the next result, Lemma 7.

Lemma 7. *Suppose that F and G are regularly related and $u = F \circ G^{-1}$.*

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

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

(iii) U is (strictly) convex on average if and only if

$$x \hookrightarrow \frac{\int_{\underline{x}}^{x} f(z)\, G(z)\, dz}{\int_{\underline{x}}^{x} g(z)\, G(z)\, dz}$$

is (strictly) increasing over $(\underline{x}, \bar{x})$.

(iv) U is (strictly) geometrically convex on average if and only if

$$x \hookrightarrow \frac{\int_{\underline{x}}^{x} f(z)\, G(z)\, dz}{\int_{\underline{x}}^{x} g(z)\, F(z)\, dz}$$

is (strictly) increasing over $(\underline{x}, \bar{x})$.

(v) F CSSD dominates G if and only if

$$\sup_{x \in (\underline{x}, \bar{x})} \frac{\int_{\underline{x}}^{x} f(z)\, G(z)\, dz}{\int_{\underline{x}}^{x} g(z)\, F(z)\, dz} = \frac{\int_{\underline{x}}^{\bar{x}} f(z)\, G(z)\, dz}{\int_{\underline{x}}^{\bar{x}} g(z)\, F(z)\, dz}.$$

Proof. See Appendix A.

An immediate and interesting consequence of Lemma 7 is that a geometric dominance relation between the unconditional distributions implies that their competitive selection-conditioned distributions are ordered by MLRP.

Lemma 8. *If F and G are regularly related, F (strictly) geometrically dominates G if and only if F (strictly) dominates G under the competitive selection conditioned MLRP ordering.*

Thus, Lemmas 1 and 8 show that geometric dominance implies that not only is the selection-conditioned distribution F stochastically dominant conditioned on selection, it is also stochastically dominant if we condition not only on competitive selection but also the realizations of the competitively selected distribution lying in any interval. Thus, as well as being a sufficient condition for CSSD dominance, geometric dominance is a necessary and sufficient condition for the sort of dominance which, in the case of unconditional distributions, is implied by MLRP.

Note that $(\log \circ F)' = f/F$ represents the reverse hazard ratio, r_F, of the distribution F. Thus, the geometric dominance condition can be expressed as the ratio of reverse hazard ratios, r_F/r_G, being increasing. Most textbook statistical distributions are log concave and thus these distributions have decreasing reverse hazard ratios. In fact, it is not possible for a distribution with support equal to the non-negative real line to have reverse hazard ratio which is uniformly increasing.[10] Thus, for most distributions, geometric dominance can be framed as the condition that the rate of decrease of the geometrically dominant distribution's reverse hazard rate is slower than the rate of decrease of the geometrically dominated distribution's reverse hazard rate.

The analogy between MLRP ordering of unconditional distributions and the geometric dominance ordering of selection-conditioned distributions is very close. MLRP can be expressed as the ratio F'/G' being increasing while geometric dominance can be expressed as the ratio $(\log \circ F)'/(\log \circ F)'$ being increasing. MLRP dominance is equivalent to the transform function, u being convex. Geometric dominance is equivalent to u being being geometrically convex. The MLRP ordering implies dominance even if valuations are conditioned on a subset of possible realizations. Geometric dominance implies selection-conditioned dominance even if realizations of the selected variable are conditioned on a subset of possible realizations.[11]

[10]See Block, Savits, and Singh (1998) for derivation of these assertions and more details regarding reverse hazard ratios.

[11]The logconvexity of the transform function, u, implies both geometric dominance and MLRP dominance. However, the logconvexity of the transform function is not a very useful notion of dominance because it is hardly ever satisfied by textbook distributions. u being logconvex is equivalent to $(f/F)(1/g)$ being increasing. If F is logconcave and g is not always increasing on it support, this condition cannot be satisfied. As most standard textbook distributions are logconcave and most do not have monotonically increasing densities, the logconvexity

Geometric dominance, although a stronger condition than CSSD and average geometric dominance, in practice, is much easier to verify because verification only requires inspection of the graph of its defining ratio, $(f\,G)/(F\,g)$. In contrast, neither CSSD nor geometric dominance on average can be identified directly from the distribution functions being compared. However, it is possible to develop a simple, sufficient, and intuitive condition for CSSD which depends only on the behavior of the ratio $(f\,G)/(F\,g)$. This condition is presented below.

Lemma 9. *If (a)* $\lim_{x\to\underline{x}}(f(x)\,G(x))/(g(x)\,F(x)) = 1$ *(b) F strictly stochastically dominates G, and (c) The ratio* $(f(x)\,G(x))/(g(x)\,F(x))$ *is either increasing or U-shaped, first decreasing and then increasing, then F CSSD dominates G. Moreover, a sufficient condition for* (a) *is that* $\lim_{x\to\underline{x}} f(x)/g(x) > 0$.

Proof. See Appendix A.

Lemma 9 shows that CSSD dominance given stochastic dominance, is essentially a tale of the two tails of the compared distributions. If the left tails of the two distributions are "proportional" at some fixed ratio and the right-tail advantage of the stochastically dominant distribution, measured by the ratio $(f(x)\,G(x))/(g(x)\,F(x))$, is ultimately increasing, the dominance of the stochastically dominant distribution is preserved by selection. As we show in the sequel, not only are these two tail conditions sufficient for CSSD dominance of stochastically dominant distributions, their violation is necessary for competitive selection to completely reverse the dominance ordering and render the stochastically dominated distribution geometrically dominant.

The conditions developed in this section are aimed to provide intuition into the anatomy of conditional dominance viewed from the perspective of the compared distributions. However, it is worth noting, as is apparent from the u-transform approach taken in Section 3, selection-based notions of dominance depend only on the quantile transform function, u. Properties of the underlying distribution that are not captured by u are irrelevant for selection-conditioned dominance. Thus, it is not possible to ensure that things "work out right," i.e., the stochastic dominance of the unconditional distribution implies stochastic dominance of the conditional distribution, simply by imposing standard restrictions on the "absolute" properties of the unconditional distributions. Simple counterexamples to this approach can always be constructed as follows. Take any "nice" distribution (e.g., smooth, with logconcave density, and compact support) G. Then choose a transform function, u, with the properties that (a) $u(t) > t,\, t \in (0,1)$, (b) $u(t)$ is strictly concave, and (c) $u(t)$ is strictly geometrically convex. Such functions are not hard to construct. One example is $u(t) = 1-\sqrt{1-\sqrt{t+t\,(1-t)}}$. Now let $F = u\circ G$. (a)

of u is not a very useful basis for ordering "typical" distribution functions. See Bagnoli and Bergstrom (2005) for an exhaustive discussion of the logconcavity of distribution and density functions and verification of the prevalence of logconcavity.

implies that F is strictly stochastically dominated, (b) implies that F is strictly dominated in the MLRP order, and (c) implies that F is strictly geometrically dominant. As long as G is "nice enough" so that the distortion induced by u will not lead to the nice properties being destroyed, we have two very nice distributions under which selection completely reverses the direction of inference. For example, if we take G to be a Beta distribution with both shape parameters equal to 5, G has a very logconcave density, a compact support, and is infinitely differentiable, as nice as one might hope. Moreover, $F = u \circ G$ inherits all of these properties. Yet selection reverses dominance. Thus, the effect of selection on dominance cannot be pinned down by placing "absolute" restrictions on the compared distributions.

6 Origins of geometric dominance

Our primary aim is to uncover the relation between conditional and unconditional stochastic orderings. However, if only to be able to assess the limitations of our analysis, we need to develop some understanding of how conditional dominance between distributions can arise, in general, even in the absence of stochastic dominance. Because, CSSD is not even an order relation, the question of the distributional characteristics which lead to CSSD dominance in general is difficult to frame. Geometric dominance on average, as noted earlier cannot be directly connected to the shape of the underlying distributions. In contrast, as we show in the next theorem, geometric dominance, even absent any stochastic dominance relation between the compared distributions, places very strong and simple restrictions on the compared distributions.

Theorem 4. *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 $x \in (x^o, \bar{x})$, $F(x) < G(x)$, i.e., the geometrically dominated distribution crosses the dominated distribution once from below.*

Proof (Sketch). The formal proof of this result is presented in Appendix A. The intuition behind the proof of this result is that strict geometric convexity implies strict 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 its identity function at a corresponding point in the log-scaled space. The strict 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. □

In essence, Theorem 4 divides geometric dominance relations into three possible configurations: In the first case, developed in part (i) of the Theorem, the geometrically dominant distribution is stochastically dominant. In the second, developed in part (ii.a), the geometrically dominant distribution is stochastically dominated. In the third case, developed in part (ii.b) the geometrically dominant distribution is in some since dispersion increasing. Because our aim is to relate unconditional stochastic dominance to stochastic dominance conditioned on competitive selection, our focus in the subsequent analysis will be on the first two cases, i.e., the first case, where competitive selection preserves stochastic dominance, and the second case, were competitive selection reverses stochastic dominance. However, when considering the application of our results to economic and econometric models, we will need to consider the implications of Theorem 4 for the robustness of our results to variation in the dispersion of the compared distributions.

7 Competitive selection preservation and reversal

7.1 Selection preservation

We aim to first analyze case (i) of Theorem 4. In this case, selection preserves the ordering of distributions: when one distribution is unconditionally better than another, it is also better conditioned on selection. When competitive selection preservation takes the strong form of geometric dominance, preservation imposes sharp, standard, and fairly easy to verify restrictions on the properties of the unconditional distributions. Weaker forms of selection dominance naturally impose weaker restrictions.

Theorem 5. *Suppose that F and G are an admissible pair of distributions and let $u \circ G = F$. Suppose that F is strictly stochastically dominates G. Then,* *i. If F strictly geometrically dominates G, then*

a. The transform function, u, is convex and thus,

b. if F and G are regularly related, F strictly dominates G in the MLRP order.

ii. If F strictly geometrically dominates G on average, then the average transform function, U, is convex.

Proof. See Appendix A.

Theorem 5 shows that, when geometric dominance preserves the stochastic dominance ordering, it implies the MLRP ordering. The MLRP ordering controls the behavior of the distribution over its entire support. Thus, selection preservation requires that the dominant distribution's "superiority" over the dominated distribution be in some sense uniform over the support. As discussed in Section 5, the rate of increase in the likelihood ratio must be sufficiently great to compensate for the increase in the reciprocal ratio of the distributions. In other words, the relative superiority of the dominant distribution needs to increase sufficiently fast as values move from the lower to the upper end of the distributions' support.

7.2 Selection reversal

As shown in part (ii.a) of Theorem 4 a distribution can be geometrically dominant yet stochastically dominated. In this case, selection reverses dominance. Perhaps the "easiest" way to induce selection reversal is for the selection dominant distribution to be a mixture between a point mass at $\underline{x}$ and the rival distribution, i.e., $F(x) = p + (1-p)\,G(x)$, $x \in [\underline{x}, \bar{x}]$ and $p \in (0,1)$. In this case, the transform function is $u(t) = p + (1-p)\,t$, $t \in [0,1]$ and the conjugate transform function is $\hat{u}(y) = \log(p + (1-p)\,e^y)$, $y \le 0$. Note that

$$\hat{u}''(y) = \frac{e^y\,(1-p)\,p}{(e^y\,(1-p)+p)^2} > 0,$$

and, thus, $\hat{u}$ is strictly convex, and therefore F strictly geometrically dominates G; yet F is strictly stochastically dominated by G. Intuitively, the geometric dominance of F results because, whenever an $\underline{x}$ realization from F is pitted against a draw from G, the draw from G is selected regardless of its quality. This effect, which one might term the "admission effect" lowers the average value of selected draws from G. Moreover, because G places no point mass at $\underline{x}$, the $\underline{x}$ draws from F are never selected and thus have no effect on the quality of selected draws from F. This admission effect starkly illustrates the difference between competitive and fixed-threshold selection: under competitive selection the distribution of one alternative affects not only its selection-conditioned quality but also the selection-conditioned quality of its rival. In fact, as the next lemma shows, a point mass at $\underline{x}$ always strictly strengthens selection dominance.

Lemma 10. *Suppose let F_o be and admissible distribution function with $F_o(\underline{x}) = 0$ such that F_o (CSSD) (geometrically on average) (geometrically) dominates G, let $F = p + (1-p)F_o$, then F strictly (CSSD) (geometrically on average) (geometrically) dominates G.*

Proof. See Appendix A.

Lemma 10 shows that a point mass at $\underline{x}$ favors selection reversal. The next result shows that "something like" a point mass at $\underline{x}$ is the *only* way to induce selection reversal: a point mass per se is not required but for selection reversal to occur the ratio between the left tail weights assigned by the two distributions must explode as the upper bound on the left-tail shrinks to $\underline{x}$.

Theorem 6. *Suppose that F and G are an admissible pair of distributions and let $u \circ G = F$. Suppose that F is strictly stochastically dominated by G. Then,*

i. if F CSSD dominates G

$$\limsup_{x \to \underline{x}} \frac{F(x)}{G(x)} = \infty.$$

ii. If, in addition, F strictly geometrically dominates G on average, then there exists $x^o \in (\underline{x}, \bar{x})$ such that for all $x < x^o$, the function

$$a.\ x \hookrightarrow \frac{\int_{\underline{x}}^{x} \frac{F(z)}{G(z)} G(z)\, dG(z)}{\int_{\underline{x}}^{x} G(z)\, dG(z)} \text{ is strictly increasing and } b.\ \lim_{x \to \underline{x}} \frac{F(x)}{G(x)} = \infty.$$

iii. If in addition, F strictly geometrically dominates G, then for all $x \in (\underline{x}, \bar{x})$, the function

$$a.\ x \hookrightarrow \frac{F(x)}{G(x)} \text{ is strictly increasing and } b.\ \lim_{x \to \underline{x}} \frac{F(x)}{G(x)} = \infty.$$

In the case of reversals, we do not have the fine control over the ratios between densities that we had in the case of selection preservation. In this case, the likelihood ratios of the unconditional distributions need not be ordered. However, selection reversal does restrict the ratio between the distribution functions and their densities around the lower endpoint of their common support: The probability weight on the left tail of the selection dominant but stochastically dominated distribution must grow explosively relative to the selection dominated but stochastically dominant distribution's tail weight. We term this behavior a "left-tail explosion." In the language of asymptotic analysis, the relation between the geometrically dominant distribution, F and geometrically dominated distribution, G, must satisfy $G = o(F)$, $x \to \underline{x}$, i.e., the left tail of F has a higher order of magnitude than the left tail of G.

A left-tail explosion reproduces the effect of the point mass. As we shall see, selection reversal can result even when comparing two distributions that are both "size" parameter shifts of the same textbook distribution. However, we shall also see that, in contrast to selection preservation, in the world of textbook distributions, reversal is quite uncommon.

8 Distributions

Thus far, we have not verified that any specific distributions satisfy the geometric dominance condition. We now turn to this task. We consider families of standard textbook distributions. Typically, these distribution families can be represented by two parameters: a "shape" parameter and a "size" parameter. Typically, under standard parameterizations, for a fixed shape parameter, the map between the size parameter and the and the distribution is an order preserving with respect to the standard order of the real line and stochastic dominance, i.e., increasing the size parameter increases the distribution in the stochastic dominance ordering. In a few cases, noted in the table below, the textbook parameterization is order reversing. In these cases, we reparameterize to ensure that, under our parameterization, the map is order preserving. We then ask the question of whether, holding the shape parameter fixed, the map between the size parameter and the distribution preserves selection-conditioned dominance as defined by the three relations: CSSD, geometric dominance on average, and geometric dominance. If, under at least one of these relations, the map from size to the distribution functions is order preserving, we say that selection *preserves* stochastic dominance and report the strongest ordering condition under which the map is order preserving. If the map is order *reversing* for at least one of the selection-conditioned orders, we say that selection is reversed and report the strongest ordering condition under which order reversal occurs. If the distributions are selection equivalent, we say that selection is *neutralized*. Of course, it is possible that, conditioned on selection, the distributions are not ordered. However, for the textbook distributions we examine, this possibility is never actualized.

If, within one distribution family, for a fixed shape parameter, the size parameter of one distribution is larger than the size parameter of the other, we call the distribution with the larger size parameter the *upsized distribution* and we call the other distribution the *original distribution*. In most, but not all cases considered, the size parameter is either a location or scale parameter. In the case of location/scale parameters it is possible to develop very simple criterion for geometric dominance. These criteria are defined and developed below.

Suppose that G is an absolutely continuous distribution function with support $(-\infty,\infty)$ and that the density of g is continuous over $(-\infty,\infty)$. Then if $F(x) = G(x-c)$, $c > 0$ we will say that F is a *c-upshift* of G and call F the *upshifted distribution* and G the *original distribution*. Similarly, suppose that G is an absolutely continuous distribution function with support $[0,\infty)$ and that the density of g is continuous over $(0,\infty)$. If $F(x) = G(x/s)$, $s > 1$ we will say that F is an *s-upscaling* of G and call F the *upscaled distribution* and G the *original distribution*. As the next lemmas show, for upshifts and up-scalings, geometric dominance only depends on the generalized convexity properties of the reverse hazard rate of the original distribution.

Lemma 11. *All c-upshifts of G (strictly) geometrically dominate G if and only if the reverse* *hazard rate of G, $r = g/G$ is (strictly) logconcave.*

Proof. See Appendix A.

Lemma 12. *All s-up-scalings of G (strictly) geometrically dominate G if and only if the reverse hazard rate of G, $r = g/G$ is (strictly) geometrically concave.*

Proof. See Appendix A.

Because scaling and translation of both distributions by the same factor does not affect selection dominance, in the case of location and/or scale families, Lemmas 11 and 12 provide simple test for selection dominance. Simply take the "standard" member of the location/scale family and examine the behavior of its reverse hazard rate.

As pointed out in Section 5, Geometric dominance depends only on the relation between the compared distributions not their "absolute" properties. Thus, it is impossible to rule out selection reversing dominance simply by placing standard restrictions on the distributions being compared. However, if we restrict attention to location and scale shifts, we have restricted the u transform function to some extent. Thus, a natural question is whether the location/scale restriction rules out selection reversal. The answer to this question is provided by Lemma 13. The Lemma shows that restricting the relation between the compared unconditional distributions does provide some leverage in restricting selection reversal, but only for scale shifts and, even for scale shifts, only a limited restriction which also requires strong additional assumptions.

Lemma 13. *Suppose that G has a strictly logconcave continuously differentiable density g, then, (i) if F is any upscaling of G, G cannot geometrically dominate F. (ii) The condition that the density of G is strictly logconcave cannot be weakened to the condition that the G itself is strictly logconcave.(iii) the condition that F is an upscaling of G cannot be replaced by the condition that F is an upshift of G.*

Proof. See Appendix A.

The lemma shows that if we restrict attention to scale shifts of distributions with logconcave densities, the upshifted distribution cannot be conditionally dominated in the strongest sense, geometric dominance. Even in this case, as we illustrate later with one of our "textbook" distributions, this result does not imply that the upscaled distribution is geometrically dominant. When upsizing results from translation, even this restriction does not hold. As the example in the proof of Lemma 13 shows, it is possible for an upshifted distribution to be geometrically dominated by the original distribution even if the original distribution has a logconcave density. The recipe for producing this effect is to pick the original distribution so that, although it support is the entire real line, it has very little mass below a fixed point, and, above this point the density shoots up. In this case, the upshifts primary effect is to create a mass of realizations of the original distribution that have almost no chance of being selected. This mass is "uncompetitive." However, if matched with realizations from the upshifted distribution, this mass permits low realizations of the upshifted distribution to be selected, lowering its selection-conditioned quality. Thus, the upshift acts as a sort of pseudo-mass point at the pseudo-lower bound. With upscaling, this recipe fails to produce complete reversal because upscaling has a more significant effect on the upper tail of the distribution than upshifting.

However, the restrictions imposed by Lemma 13 are fairly weak and never actually verify selection preservation. Thus, it appears that verification requires studying specific distribution functions. A task to which we now turn. Using Lemmas 11 and 12 as well as the more general characterizations of selection dominance developed earlier, we derive the effects of size shifts under competitive selection for many standard families of distributions. The results of these derivations are reported in Table 8. The derivations themselves are presented in Appendix D (supplementary material).

Table 1: **Selection preservation and reversal**

	Parameters			**Effect of selection on dominance**
Distribution	*Size*	*Type*	*Shape (fixed)*	*Under selection dominance is*
Normal $F[x] = \frac{1}{2} - \frac{1}{2}\operatorname{Erf}\left(\frac{\mu - x}{\sqrt{2}\sigma}\right)$ $x \in (-\infty, \infty)$	μ	location	$\sigma > 0$	*Preserved* Strictly geo. Dominant
Logistic $F[x] = \frac{1}{1+e^{-\frac{x-\mu}{s}}}$ $x \in (-\infty, \infty)$	μ	location	$s > 0$	*Preserved* Strictly geo. Dominant
Laplace $F[x] = 1/2 \exp\left(\frac{x-\mu}{s}\right)$ if $x \geq 0$ $F[x] = 1 - 1/2 \exp\left(-\left(\frac{x-\mu}{s}\right)\right)$ if $x < 0$ $x \in (-\infty, \infty)$	μ	location	$s > 0$	*Preserved* Strictly Avg. geo. dominant but not geo. dominant
Gumbel $F[x] = \exp\left(-e^{-\frac{x-\mu}{s}}\right)$ $x \in (-\infty, \infty)$	μ	location	$s > 0$	*Neutralized* Selection equivalent

Table 1: (continued)

Distribution	Parameters: Size	Parameters: Type	Parameters: Shape (fixed)	Effect of selection on dominance: Under selection dominance is
Gamma $F[x]=(1/\Gamma(\alpha))\int_0^{x/s} z^{\alpha-1}e^{-z}dz$ $x\in[0,\infty)$	s	scale	$\alpha>0$	*Preserved* Strictly geo. Dominant
Generalized Exponential[a] $F[x]=\left(1-e^{-\frac{x}{s}}\right)^b$ $x\in[0,\infty)$	s	scale	$b>0$	*Preserved* Strictly geo. Dominant
Weibull[a] $F[x]=1-e^{-\left(\frac{x}{a}\right)^\lambda}$ $x\in[0,\infty)$	$a>0$	scale	$\lambda>0$	*Preserved* Strictly geo. dominant
Pareto[b] $F[x]=1-\left(\frac{x_m}{x}\right)^{\frac{\mu}{\mu-1}}$ $x\in[x_m,\infty)$	$\mu>1$	neither	$x_m>0$	*Preserved* Strictly geo. Dominant
Kumaraswamy $F[x]=1-(1-x^\alpha)^b$ $x\in[0,1]$	$\alpha>0$	neither	$b>0$	*Preserved* if $b>1$ Strictly geo. dominant *Neutralized* if $b=1$ selection equivalent *Reversed* for $b<1$ Strictly geo. dominated
Lognormal $F[x]=\frac{1}{2}-\frac{1}{2}\operatorname{Erf}\left(\frac{\mu-\log(x)}{\sqrt{2}\sigma}\right)$ $x\in[0,\infty)$	μ	scale	$\sigma>0$	*Preserved* Strictly geo. dominant
Fréchet $F[x]=\exp\left(-\left(\frac{x}{s}\right)^{-\alpha}\right)$ $x\in[0,\infty)$	s	scale	$\alpha>0$	*Neutralized* Selection equivalent
Log-logistic $F[x]=\frac{1}{1+\left(\frac{x}{\alpha}\right)^{-\beta}}$ $x\in[0,\infty)$	$\alpha>0$	scale	$\beta>1$	*Preserved* Strictly geo. dominant

Table 1: (continued)

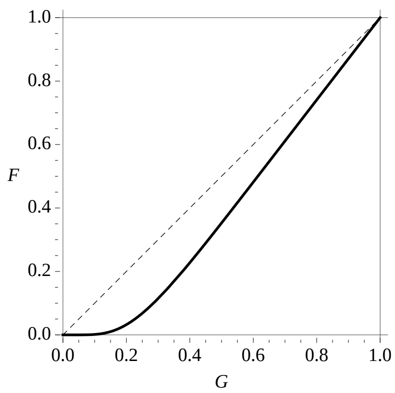
Distribution	Parameters: Size	Parameters: Type	Parameters: Shape (fixed)	Effect of selection on dominance: Under selection dominance is
Gompertz[c] $F[x] = 1 - \exp\left(\eta\ \left(1 - e^{x/s}\right)\right)$ $x \in [0, \infty)$	$s > 0$	scale	$\eta > 0$	*Preserved* Strictly geo. dominant if $\eta \geq 1$ Strictly CSSD dominant but not geo. dominant and not always geo. dominant on average if $\eta < 1$

[a] The exponential distribution is a special case of the Weibull distribution, the case where $\lambda = 1$ and of the Generalized Exponential where $b = 1$.
[b] Standard parameterization changed to ensure that increasing the size parameter produces stochastically larger distributions. Replacing μ with $\alpha/(1-\alpha)$ produces the standard parameterization.
[c] Standard parameterization of distribution modified to ensure that increases in the scale parameter up scales rather than down scales. Replacing s with $1/b$ produces the standard parameterization.

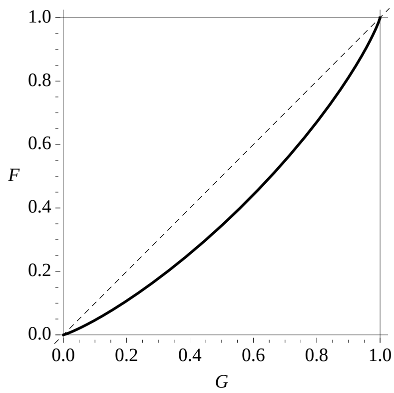
The first and most striking observation about Table 8 is that size parameter shifts lead to strict geometric dominance for the upsized distribution in the vast majority of cases. Because, geometric convexity implies selection conditioned MLRP dominance, and as shown in Theorem 5, MLRP dominance is a necessary condition for geometric dominance when the geometrically dominant distribution is stochastically dominant, these result show that, for the overwhelming majority of distributions surveyed, the monotone likelihood ordering is robust to competitive selection.

Given the results thus far, most of the exceptional cases where selection is not preserved or partially preserved are quite easy to rationalize. The Gumbel and Fréchet distributions are extreme value type distributions and as explained in Section 1, such distributions are always equivalent conditioned on selection. For two distributions on our list, the Laplace and Gompertz, upsizing does not result in geometric dominance but some other form of dominance is preserved. For the Gompertz distribution when $\eta < 1$, the upscaled distribution is CSSD dominant but not geometrically dominant and sometimes not even geometrically dominant on average. The failure of geometric dominance in this case is also easy to explain. When the Gompertz shape parameter, η is less than 1, the upscaled distribution is not MLRP dominant, thus, by Theorem 5 it cannot be geometrically dominant. In contrast, for the Laplace distribution, the upshifted distribution is MLRP dominant. As MLRP dominance of the upshifted distribution is a necessary but not sufficient condition for geometric dominance, the failure of geometric dominance is, of course, consistent with our earlier results. In the case of the Laplace distribution, the failure of geometric convexity results from non-convexity of the conjugate transform function around the quantile of the original distribution that maps into the mean of the upshifted distribution. At this point, the upshifted density has a kink which leads to nonconvexity. Averaging out the conjugate transform leads to convexity and thus the upshifted distribution is always geometrically dominant on average.

The only distribution in the table in which selection completely reverses dominance is the Kumaraswamy distribution for shape parameters $b < 1$. In this case, the upshifted distribution is strictly MLRP dominant but strictly geometrically dominated. What accounts for selection causing a complete reversal of dominance in this case? Perhaps the best way to understand this reversal case is to compare it with the "normal" case where selection is preserved. What better exemplar can one find of the normal case than the Normal distribution itself. Thus, below, in Figure 3, we plot the u transform functions of the Kumaraswamy and Normal distributions. In both cases $u = F \circ G^{-1}$ maps the relation between a distribution, G, and its upsizing, F.



A. $u_{\text{Kumaraswamy}}$



B. u_{Normal}

Figure 3: *Plots of the transform function, u, for the Kumaraswamy and Normal Distributions, in both cases F is an upsizing of G.* In the Kumaraswamy case, the common shape parameter is $b = 1/8$ and the size parameter for F is $\alpha = 8$ and for G is $\alpha = 1$. In the Normal case, the common shape parameter is $\sigma = 1$ and the size parameter for F is $\mu = 0.40$ and for G is $\mu = 0$.

As explained in Section 3.5, the graph of the transform function can be thought of as a parametric plot where each point on the graph is given by $(G(x), F(x))$, for some point x in the support of the distributions. Thus the abscissa in Figure 3 is labeled G and the ordinate is labeled F. From the plot we see that u, in both the Kumaraswamy and Normal cases, is convex and lies below the diagonal, showing that in both cases the upscaled distribution is stochastically and MLRP dominant. However, as the graph makes apparent, the nature of dominance in the two cases is very different. In the case of the Kumaraswamy, the convexity of u is concentrated in the the very lowest quantiles. In fact, as the graph shows, u is essentially flat below the 10% quantile, showing that 10% of the mass of the original distribution is essentially beneath the entire mass of the upsized distribution. Convexity rapidly falls as the quantile increases and, at the top end of the quantile range, u is essentially linear. Roughly, in the Kumaraswamy case, the worst draws of the worse (original) distribution are much worse than draws from the best (upsized) distribution but the best draws from the worse distribution are not that much worse than the best draws from the best. In contrast, for the Normal distribution, dominance is fairly uniform across the quantiles and, in fact, can be approximated, to the point of visual indistinguishability, by a quadratic polynomial. As shown in Theorems 5 and 6, uniform dominance in the case of selection preservation and dominance concentrated in the left-tail are always the signature characteristics of selection preservation and reversal. Thus, geometric convexity of the conjugate transform, $\hat{u}$, in the Normal case, indicating selection preservation, and geometric concavity of $\hat{u}$ in the Kumaraswamy case, indicating selection reversal, are expected. This expectation is verified by Figure 4 Given these observations it is not

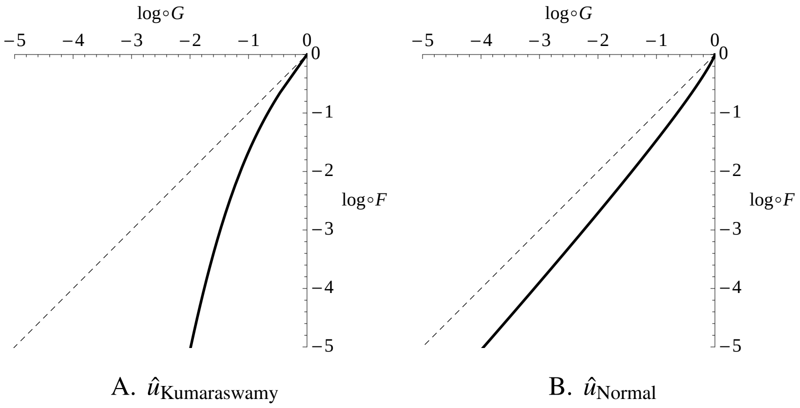


Figure 4: *Plots of the conjugate transform function, $\hat{u}$, for the Kumaraswamy and Normal Distributions, in both cases F is an upsizing of G.* In the Kumaraswamy case, the common shape parameter is $b = 1/8$ and the size parameter for F is $\alpha = 8$ and for G is $\alpha = 1$. In the Normal case, the common shape parameter is $\sigma = 1$ and the size parameter for F is $\mu = 0.40$ and for G is $\mu = 0$.

surprising that, for Kumaraswamy distributions, selection reversals can be fairly dramatic. For example, when the original distribution is Kumaraswamy with size parameter $\alpha_F = 1/20$ and shape parameter $b = 1/2$ and the upsized distribution is distributed Kumaraswamy with size parameter $\alpha_F = 1$ and shape parameter $b = 1/2$, the unconditional expectation of the original distribution is approximately 0.195, and the unconditional expectation of the upsized distribution is approximately 0.667. However, conditioned on selection, the expectation of the original distribution equals approximately 0.762 and the expectation of the upscaled distribution equals approximately 0.701.[12]

[12]The numbers result from numerical integration. The generating code is available upon request.

9 Applications

In this section, we illustrate, through concrete examples, how the order relations developed above determine the toxicity of competitive selection's effect on qualitative inference. First we consider the posterior problem of inferring the properties of the unconditional distributions from selection-conditioned observations. We show that in a classical linear regression model of treatment effects, the assumption of standard error law distributions, implies that the advantage of one treatment over another is fairly uniform over the range of possible responses. Thus, selection is innocuous in that it never reverses the stochastic ordering of the treatments, i.e., the sign of the treatment effect can be identified by simple, naive, statistical approaches that ignore the problem of selection bias. Next we illustrate the use of the ordering relations in prior problems. Here we derive selection conditioned predictions from unconditional predictions of two auction models.

9.1 Identifying the sign of treatment effects in standard regression models in the presence of self-selection.

In this section, we apply the results developed above to the problem of identifying the sign of treatment effects in the presence of self-selection bias. The development follows Manski (1990).[13]Manski analyzes the following problem: consider a population of subjects. Assume that a random sample of the population is drawn and for each subject in the sample, the econometrician observes x and z, where x represents a vector of individual covariates of the subjects and z is a indicator function which equals 1 if the subject received drug treatment A and equals 0 if the subject received drug treatment B. Let y_A be a scalar measure of the outcome of treatment A and let y_B represent the outcome of treatment B. The econometrician's difficulty is selectivity bias: the econometrician observes $\tilde{y}_A$ only when the subject receives treatment A and observes $\tilde{y}_B$ only when the subject receives treatment B. The econometrician's aim is to infer the difference in unconditional treatment effects, the difference in the average effects of the drugs on subjects with covariates x under random assignment, i.e., to infer $E[\tilde{y}_A|\tilde{x}] - E[\tilde{y}_B|\tilde{x}]$. One version of the problem considered by Manski assumes that selection is based on self selection, i.e., treatment A is selected by a subject if and only if $\tilde{y}_A > \tilde{y}_B$.

Under self-selection, the problem of determining the sign of the treatment effect from the observed data becomes a problem of inferring $E[\tilde{y}_A|\tilde{x}] - E[\tilde{y}_B|\tilde{x}] > 0$ from $E[\tilde{y}_A|\tilde{x}, \tilde{y}_A > \tilde{y}_B] - E[\tilde{y}_B|\tilde{x}, \tilde{y}_B > \tilde{y}_A]$. Manski considers the question of how non-parametric restrictions on the supports of the random variates can be used to identify the sign of the treatment effect. We consider a different question, in a standard parametric regression, what restrictions on the

[13]See also Manski (1989) and Manski (1997)

error-term distribution permit identification of the sign of the treatment effect? In a standard parametric regression (linear or non-linear), the distribution of outcomes under treatment A and B can be expressed as

$$\tilde{y} = t_B(\tilde{x}) + (t_A(\tilde{x}) - t_B(\tilde{x}))I_A + \Phi(\tilde{x}) + \tilde{\varepsilon}. \tag{29}$$

where Φ is a real-valued function of the covariates, x, and $\tilde{\varepsilon}$, is an independent and identically distributed error term.

For any fixed vector of characteristics, x, the distributions of $\tilde{y}_A$ and $\tilde{y}_B$ are location translations of the distribution of the error term $\tilde{\varepsilon}$, the translation being given by the average difference in treatment effects given x. If the distribution of the error term comes from one of the translation-shift families of distributions identified in Table 8, e.g., $\tilde{\varepsilon}$'s distribution is Normal, Logistic, or Laplace, then stochastic dominance is preserved by conditioning, i.e., $E[\tilde{y}_A|\tilde{x}, \tilde{y}_A > \tilde{y}_B] - E[\tilde{y}_B|\tilde{x}, \tilde{y}_B > \tilde{y}_A]$ implies that $E[\tilde{y}_A|\tilde{x}] - E[\tilde{y}_B|\tilde{x}] > 0$. Thus the sign (but not the magnitude) of the treatment effect can be identified by the standard regression model even in the presence of self-selection. Moreover, if the error term is Normally or Logistically distributed, translations of the error term are ordered by geometric dominance, and thus, even, if in addition to self-selection bias, the self-selected sample is truncated based on observed values of $\tilde{y}$, a simple regression would still correctly identify the sign of the treatment effect. Essentially, the same argument shows that the sign of the treatment effect in a standard Logit regression can be identified even in the presence of self selection bias.

Note also that identification of the sign of the treatment effect does not depend on information about the propensity of subjects to choose a given treatment. Thus, the average treatment effect could be determined from a study that sampled a predetermined number of patients receiving each treatment, provided that the sample for each treatment was random. This contrasts from both standard parametric and nonparametric identification approaches where the propensity to choose the treatment is a key element in the identification strategy.

Of course, these results rest on the assumption that, save for self-selection and perhaps truncation, the regression model is correctly specified. As we demonstrated in Section 6, dispersion alone can induce geometric dominance. Thus, if the error term is not homoscedastic and varies with the treatment selected, the misspecified model might not identify the sign of the treatment effect. Even absent dispersion effects, our previous results point to a very simple case where selection preservation would fail. Suppose that the two drugs produce identically distributed effects for most subjects but for a fraction, ρ, of patients, drug B produces an allergic reaction with extremely negative side effects. In this case, allergic patients would choose drug A even if its effects were mediocre, while drug B would only be selected if it produced better effects than drug A. Thus, becasue of the "admission effects" discussed in Section 7.2, selection reversal would occur and the inferior drug B, would produce better results on average when it was selected.

10 Selection-conditioned theoretical predictions

In the previous section, we illustrated one application of selection conditioned stochastic orders: inferring unconditional relations between treatment effects from selection-conditioned samples. In this section we consider another application: developing theoretical predictions about selection conditioned outcomes. Because these predictions predict the behavior of the self-selected sample, self selection presents no obstacle to identification. Thus, selection-conditioned predictions generate easily testable model implications. A model will be able to produce selection conditioned predictions whenever the distributions produced by the model are ordered by a selection dominance ordering. We illustrate this application of selection dominance orders with two examples from the auction literature, a first-price auction with asymmetric bidders, and an all-pay auction with asymmetric bidders. In both examples, one bidder is *ex ante* stronger than the other, e.g., one bidder might be receiving state subsidies or have observable structural characteristics (e.g., size) that render the bidder more liekly to be able to extract value from acquiring the auctioned good. In both cases we show that the model's predictions about unconditional differences in bidding behavior of the two types of bidders can be extended to generate predictions about the bidding behavior conditioned on the bidder winning the auctioned good. This extension is quite useful because, as Laffont, Ossard, and Vuong (1995) point out, frequently bid data is only available for the winning bid in an auction because only the winning bid is paid. In essence, under this approach, the problem of selection bias is resolved by the unconditional theoretical predictions of the model being adjusted to account for selection rather than the traditional approach of adjusting the econometrics to permit the testing of unconditional predictions with selection-conditioned data.

10.1 Example: First-price sealed bid auctions

In our first example, a first-price auction we follow a special case of the development in Kaplan and Zamir (2012). They consider a first-price auction with two risk-neutral bidders, H and L. The auction is a private value auction with each bidder's valuation independently distributed. The valuation of H is uniformly distributed over the interval $[0, v_H]$ and the valuation of L is uniformly distributed over the interval $[0, v_L]$. We assume that $v_H > v_L > 0$. The auctioneer may set a reserve price of m where $m \in [0, v_L]$. In the special case where $m = 0$, i.e., the reserve price is not binding, the model reduces to the auction model developed in Criesmer, Levitan, and Shubik (1967). We extend our results to the general case of a binding reserve price in Appendix C. In this section, we concentrate on the special case where $m = 0$. In this case, the equilibrium inverse bid functions for the two bidders, $B_i^{-1} : [0, \bar{b}] \to \Re$, $i = H, L$,

$\bar{b} = (v_H v_L)/(v_H + v_L)$ in this auction are given by

$$B_L^{-1}(b) = \frac{2b\,(v_H v_L)^2}{(v_H v_L)^2 + b^2\,\left(v_H^2 - v_L^2\right)}, \tag{30}$$

$$B_H^{-1}(b) = \frac{2b\,(v_H v_L)^2}{(v_H v_L)^2 - b^2\,\left(v_H^2 - v_L^2\right)}. \tag{31}$$

Given the uniform distribution assumption, the distribution of bids is given by

$$\mathbb{P}[\tilde{b}_i \leq b] = \mathbb{P}[B_i(\tilde{v}_i) \leq b] = \mathbb{P}[\tilde{v}_i \leq B_i^{-1}(b)] = \frac{1}{v_i} B_i^{-1}(b).$$

Thus, the equilibrium bid distributions for the two bidders are given by

$$F_L(b) = \frac{1}{v_L}\left(\frac{2b\,(v_H v_L)^2}{(v_H v_L)^2 + b^2\,\left(v_H^2 - v_L^2\right)}\right), \quad b \in [0, \bar{b}], \tag{32}$$

$$F_H(b) = \frac{1}{v_H}\left(\frac{2b\,(v_H v_L)^2}{(v_H v_L)^2 - b^2\,\left(v_H^2 - v_L^2\right)}\right), \quad b \in [0, \bar{b}]. \tag{33}$$

It is clear from inspection that F_H strictly stochastically dominates F_L. Thus, the "stronger bidder" H will bid more aggressively for the auctioned good, i.e., the Kaplan and Zamir (2012) model predicts that higher bids on average from the stronger bidder. Our earlier results easily generate selection conditioned predictions from this unconditional prediction. To see this, note that by Lemma 7, strict geometric dominance of F_H over F_L is verified if the ratio between the reverse hazard rates of the distributions, r_H/r_L is increasing. Simple calculations show that the ratio between the reverse hazard rates is given by

$$\frac{r_H(b)}{r_L(b)} = \left(\frac{v_H^2 v_L^2 + b^2\,\left(v_H^2 - v_L^2\right)}{v_H^2 v_L^2 - b^2\,\left(v_H^2 - v_L^2\right)}\right)^2.$$

Because this ratio is clearly increasing, F_H srictly geometrically dominates F_L and thus the selection conditioned distribution of H MLRP dominates the selection conditioned distribution of F_L. Hence, this model predicts that winning bids of the stronger bidder will stochastically dominate winning bids of the weaker bidder. In fact, even if in addition to selection conditioning, the sample of winning bids is truncated because, for example, only winning bids higher than some threshold are reported, MLRP dominance of the conditional distributions implies that the winning bid of the stronger bidder will still, on average, be higher. These selection-conditioned predictions can be tested even using naive econometrics that fail to account for selection bias because the predictions themselves are about selection-bias conditioned behavior of the bids.

10.2 All-pay auctions

In the previous example, using the selection dominance orderings, we were able to generate theoretical predictions about selection conditioned behavior from unconditional predictions. In that example, the selection-conditioned predictions were in the same direction as unconditional predictions. In this section, we develop a simple example were we are able to make equally strong selection-conditioned predictions but the predictions are the reverse of the unconditional predictions. The example considers an all-pay auction between two symmetrically informed bidders. Bidder H places a valuation of v_H on the auctioned good and bidder L places a valuation of v_L, where $v_H > v_L > 0$. In an all-pay auction bidders pay their bid regardless of whether they win or lose. The high bidder receives the good and, we assume, if the two bids are equal the good is allocated randomly, with each bidder having an equal probability of receiving the good. It is easy to see that the equilibrium bid distribution for H and L is given as follows: H randomizes uniformly over the interval $(0, v_L)$. Bidder L randomizes by bidding 0 with probability $\rho = (v_H - v_L)/v_H$ and, with probability $1 - \rho$, randomizes uniformly over the interval $(0, v_L)$. To see that these strategies constitute an equilibrium, simply note that the payoff to H from a bid of b_H conditioned on F_L is given by

$$\begin{cases} \frac{1}{2}(v_H - v_L) & \text{if } b_H = 0 \\ v_H - v_L & \text{if } b_H \in (0, v_L) \\ v_H - b_H & \text{if } b_H \geq v_L. \end{cases}$$

Thus, all bids b_H in $(0, v_L)$ are best responses to F_L, verifying that H's strategy of uniformly randomizing over $(0, v_L))$ is a best response for H. Similarly, the payoff to L from a bid of $b_L \geq 0$ conditioned on F_H is given by

$$\begin{cases} 0 & \text{if } b_L < v_L \\ v_L - b & \text{if } b_L \geq v_L. \end{cases}$$

Thus, all bids in $[0, v_L)$ are best responses for L, verifying that L's strategy of mixing between a point mass at 0 and uniform randomization over $(0, v_L$ is a best response for L. In fact, F_H and F_L essentially constitute the unique equilibrium for the game. However, this assertion is more difficult to prove and we refer the reader to Baye, Kovenock, and Vries (1993).

It is clear that F_H strictly stochastically dominates F_L. Thus this model predicts that H's bids are stochastically larger than L's. However, because $F_L = \rho + (1 - \rho) F_H$, Lemma 10, shows that H's bid, conditioned on H winning, is stochastically dominated by L's bid, conditioned on L winning. Thus, the all-pay auction model predicts that winning bids of the stronger bidder will be stochastically *smaller* than winning bids of the weaker bidder. Again, this selection conditioned prediction could be tested on a sample of winning bids without correcting for selection bias.

11 Conclusion

In this paper we characterized the effect of competitive conditioning on stochastic orders. In our analysis, the conditioning event was competitive selection, i.e., the realized value of one of the compared random variable exceeding the realized value of another. In this setting, we derived necessary and sufficient conditions for competitive selection conditioning preserving or reversing stochastic order relations between compared distributions. This analysis set determinant bounds on the validity of inferences from the unconditional quality of random variables to their selection-conditioned quality and vice versa.

The analysis initiated in this paper can be extended in a number of directions. The most obvious and natural extension is to consider conditioning on other events associated with competitive selection. For example, one might consider, “deselection conditioned dominance,” where deselection denotes the event that the realized value of a random variable is less than the realized value of the compared random variable. This extension has some obvious applications. For example, in two bidder English and second price auctions, the minimum reservation bid determines the price at which the auctioned good is sold. Thus, an analysis of deselection dominance could determine the distributional conditions under which a bidder, having a stochastically dominant valuation, would pay more or less in expectation for the auctioned good. The extension of these results to deselection dominance is straightforward: it simply involves replacing the geometric convexity condition for dominance with a geometric concavity condition and replacing the quantile transform of the distribution functions with the quantile transform of the complementary distribution function, $1-F$. Conditioning on other order statistics, e.g., the median, is much more difficult because dominance cannot be characterized by geometric convexity or concavity but rather depends on finer properties of the conjugate transform function.

Our analysis also assumed that random draws from the compared distributions were independent. Relaxing this condition is a tractable problem under the assumption that each variable is the sum of a common term and an independent error term. Extending the analysis further is problematic because, under such extensions, dominance would depend not only on the unconditional distributions alone but also on the nature of their structural dependence. Another interesting extension would be to consider noisy selection. Selection contingent, for example, on the random variable being selected by quantile response utility function.

In short, this paper answers a question of when stochastic orders are preserved by competitive selection. This question is of considerable relevance to many economic and statistical problems. At the same time, it initiates an extensible research program into the preservation of stochastic order under conditioning.

References

ATHEY, S. (2002): "Monotone comparative statics under uncertainty," *Quarterly Journal of Economics*, 117(1), 187–223.

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

BARICZ, Á. (2010): "Geometrically concave univariate distributions," *Journal of Mathematical Analysis and Applications*, 363(1), 182–196.

BAYE, M. R., D. KOVENOCK, AND C. G. D. VRIES (1993): "Rigging the Lobbying Process: An Application of the All-Pay Auction," *The American Economic Review*, 83(1), pp. 289–294.

BLOCK, H. W., T. H. SAVITS, AND H. SINGH (1998): "The reversed hazard rate function," *Probability in the Engineering and Informational Sciences*, 12(1), 69–90.

CASELLA, G., AND R. L. BERGER (2002): *Statistical inference*, vol. 2. Duxbury Pacific Grove, CA.

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.

CRIESMER, J. H., R. E. LEVITAN, AND M. SHUBIK (1967): "Toward a study of bidding processes part IV-games with unknown costs," *Naval Research Logistics Quarterly*, 14(4), 415–433.

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

HANOCH, G., AND H. LEVY (1969): "The efficiency analysis of choices involving risk," *The Review of Economic Studies*, 36(3), 335–346.

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

HECKMAN, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), pp. 153–161.

HECKMAN, J. J., AND B. E. HONORÉ (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58(5), pp. 1121–1149.

JARCZYK, W., AND J. MATKOWSKI (2002): "On Mulholland's inequality," *Proceedings of the American Mathematical Society*, 130(11), 3243–3247.

KAPLAN, T. R., AND S. ZAMIR (2012): "Asymmetric first-price auctions with uniform distributions: analytic solutions to the general case," *Economic Theory*, 50(2), 269–302.

KUCZMA, M. (2000): *An Introduction to the Theory of Functional Equations and Inequalities*. Birkhäuser Verlag.

LAFFONT, J.-J., H. OSSARD, AND Q. VUONG (1995): "Econometrics of first-price auctions," *Econometrica*, 63(4), 953–980.

MANSKI, C. F. (1989): "Anatomy of the selection problem," *Journal of Human Resources*, 24(3), 343–360.

MANSKI, C. F. (1990): "Nonparametric Bounds on Treatment Effects," *The American Economic Review*, 80(2), pp. 319–323.

MANSKI, C. F. (1997): "Monotone treatment response," *Econometrica*, 65(6), 1311–1334.

MILGROM, P., AND C. SHANNON (1994): "Monotone comparative statics," *Econometrica*, 62(1), 157–180.

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

NICULESCU, C. P., AND L.-E. PERSSON (2004): *Convex Functions and Their Applications: A Contemporary Approach*. Springer-Verlag Berlin Heidelberg.

PINELIS, I. (2002): "L'Hospital Type rules for monotonicity, with applications," *J. Ineq. Pure & Appl. Math*, 3(1), Art. 5, 5 pp.

——— (2006): "On L'Hospital-type rules for monotonicity," *J. Inequal. Pure Appl. Math*, 7(2), Art. 40, 19 pp.

QUAH, J. K.-H., AND B. STRULOVICI (2009): "Comparative statics, informativeness, and the interval dominance order," *Econometrica*, 77(6), 1949–1992.

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

TOPKIS, D. M. (1998): *Supermodularity and complementarity*. Princeton university press.

Appendix A Proofs of Selected Propositions

Proof of Lemma 5. To prove (i) note that we need to show that the geometric convexity relation 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 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 increasing, $\hat{u}_1$ and $\hat{u}_2$ are increasing. The composition of increasing convex functions is convex. Thus, $\hat{u}_1 \circ \hat{u}_2$ is convex. Thus, the geometric convexity relation is a preorder. Exactly the same proof, with u replaced by U, verifies that geometric convexity on average is a preorder.

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 B. □

Proof of Lemma 6. First note that, by Theorem 4, (i) implies (ii). By Theorem 2, (ii) implies (iii).

Next, we show that (i) implies (iv). To see this, 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 (ii) of Lemma 2, 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 $\hat{u}(y) = py,\ p > 0$. Thus,

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

Thus, we have shown that (i) implies (iv).

Next, we show (i) implies (iii). To see this note that the conjugate function to $u(t) = t^p$ is linear shows that (iv) implies (i). Thus, by Theorem 2, (i) implies (iii).

To complete the proof we need only show that (iii) implies (iv). To prove this note that, by Theorem 1, CSSD dominance is equivalent to

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

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

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

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

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

Equations (A-4) and (A-3) and imply that

$$\Pi[u^{-1}](u(t)) = 1 - \Pi[u](t). \tag{A-5}$$

Letting $t = 1$ in (A-4) shows that

$$\int_0^1 u(s)\,ds + \int_0^1 u^{-1}(s)\,ds = 1. \tag{A-6}$$

Thus, if we let c equal the first integral in (A-6), we see that both inequalites in expression (A-2) implies that

$$\Pi[u](t) \geq c \quad \text{and} \quad 1 - \Pi[u](t) \geq 1 - c. \tag{A-7}$$

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

$$\frac{1}{c}\frac{1}{t}\int_0^t u(s)\,ds = u(t). \tag{A-8}$$

Because u is identically equal to the left-hand-side of equation (A-8), and because u is continuous and thus its integral is differentiable, u must be differentiable. Differentiation of equation (A-8) shows that u must satisfy the differential equation,

$$(1-c)\,u(t) - c\,t\,u'(t) = 0, \quad u(1) = 1. \tag{A-9}$$

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

Proof of Lemma 7. From Chan, Proschan, and Sethuraman (1990) we see that

$$u(t) = \int_0^t \phi \circ G^{-1}(s)\,ds, \tag{A-10}$$

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 respect to Lebesgue measure and is given by $\phi = f/g$. Thus,

$$u'(t) = \phi \circ G^{-1}(t). \tag{A-11}$$

First consider (i). Because G is continuous and its support is $[\underline{x}, \bar{x}]$, G strictly increasing and thus G^{-1} is strictly increasing over $[\underline{x}, \bar{x}]$. Hence u increasing if and only if ϕ is increasing, i.e., f/g is increasing.

Now consider (ii), For regularly related distributions, geometric convexity requires that

R, defined in Lemma 3, be increasing. Substituting the definitions of u and u' from equations (A-10) and $u = F \circ G^{-1}$ into R shows that

$$R(t) = \frac{\phi \circ G^{-1}(t)\,t}{F \circ G^{-1}(t)}. \tag{A-12}$$

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

$$R \circ G(s) = \frac{\phi(s)\,G(s)}{F(s)}. \tag{A-13}$$

Because G is strictly increasing, $R \circ G(s)$ is increasing if and only if $s \to \phi(s)\,G(s)/F(s)$ is increasing.

Next, consider (iii). By equation (15), $U' = u(t)/t - U(t)/t$. This equation of can be rewritten as

$$U'(t) = \frac{\frac{1}{2}\int_0^t \Big(u(t) - u(s)\Big)\,ds}{\int_0^t s\,ds}. \tag{A-14}$$

Because the distributions are regularly related, u' exists and is absolutely continuous, thus

$$u(t) - u(s) = \int_s^t u'(r)\,dr.$$

Thus, expression (A-14) can be rewritten as

$$U'(t) = \frac{\frac{1}{2}\int_0^t \Big(\int_s^t u'(r)\,dr\Big)\,ds}{\int_0^t s\,ds}. \tag{A-15}$$

Using Fubini's Theorem to reverse the order of integration yields

$$U'(t) = \frac{\frac{1}{2}\int_0^t u'(r)\Big(\int_0^r ds\Big)\,dr}{\int_0^t s\,ds} = \frac{1/2\int_0^t u'(s)s\,ds}{\int_0^t s\,ds} = \frac{1/2\int_0^t u'(s)s\,ds}{1/2\,t^2} = \frac{\int_0^t u'(s)s\,ds}{t^2}.$$

Finally, note that as shown in the proof of part (i), $u' \circ G^{-1} = f/g$. Using this result and performing a change of variables shows that

$$U' \circ G(x) = \frac{\int_{\underline{x}}^x G(z)\,f(z)\,dz}{G(x)^2} = \frac{\int_{\underline{x}}^x G(z)\,f(z)\,dz}{\int_{\underline{x}}^x G(z)\,g(z)\,dz}. \tag{A-16}$$

Because G^{-1} is strictly increasing, this expression must be strictly increasing if U' is to be strictly increasing as required by the strict convexity of U.

Now consider part (iv). Note that equations (25) and (A-16) imply that

$$R_U \circ G(x) = \frac{G(x)\,U' \circ G(x)}{U \circ G(x)} = \frac{\int_{\underline{x}}^x G(z)\,f(z)\,dz}{\int_{\underline{x}}^x F(z)\,g(z)\,dz}, \text{ where } R_U(t) = \frac{U'(t)\,t}{U(t)}. \tag{A-17}$$

Because G is strictly increasing, R_U is (strictly) increasing if and only if $R \circ G$ is (strictly) increasing. R_U being (strictly) increasing necessary and sufficient for U being (strictly) geometrically convex, i.e., for F to (strictly) strictly geometrically on average dominate G.

Now consider part (v), Define

$$\mathscr{R}(x) = \frac{\int_{\underline{x}}^{x} G(z)\, f(z)\, dz}{\int_{\underline{x}}^{x} g(z)\, F(z)\, dz}, \quad x \in (\underline{x}, \bar{x}). \tag{A-18}$$

Equations (16) and (24) imply that

$$\Pi[u](t) = \frac{1}{1 + R_U(t)}, \tag{A-19}$$

where R_U is defined in equation (A-17). Using this observation we see that

$$\Pi[u] \circ G = \frac{1}{1 + \mathscr{R}(x)}. \tag{A-20}$$

Thus, because G is strictly increasing, The condition that CSSD condition given by equation (17) is equivalent to the condition that

$$\sup_{x \in (\underline{x}, \bar{x}]} \mathscr{R}(x) = \mathscr{R}(\bar{x}). \tag{A-21}$$

□

Proof of Lemma 9. Define,

$$\rho(x) = \frac{f(x)\, G(x)}{F(x)\, g(x)}, \quad x \in (\underline{x}, \bar{x}). \tag{A-22}$$

We first show under the hypotheses of the Lemma, $\mathscr{R}$ is quasiconvex: If ρ is increasing, the Monotone L'Hospital Theorem (Pinelis, 2002, Proposition 1) implies that R is increasing and thus quasiconvex. If ρ is U-Shaped, then the Monotone L'Hopital Rule for oscillatory functions (Pinelis, 2006, Proposition 4.3) implies that $\mathscr{R}$ is either decreasing or U-shaped, decreasing and then increasing, and thus in this case also it is quasiconvex.

Because $\mathscr{R}$ is quasiconvex, if $x \in (x', \bar{x})$,

$$\mathscr{R}(x) \leq \max[\mathscr{R}(x'), \mathscr{R}(\bar{x})] \tag{A-23}$$

The hypothesis of the is Lemma is that $\lim_{x \to \underline{x}} \mathscr{R}(x) = 1$. Moreover, $\mathscr{R}(\bar{x}) > 1$. This follows from equation (A-20), and the fact that $\Pi[u](1)$ represents the probability that draws from the strictly stochastically dominated distribution, G exceed draws from the strictly stochastically dominant distribution, F and thus $\Pi[u](1) < \frac{1}{2}$. Thus, for all $x \in (\underline{x}, \bar{x})$, if x' is sufficiently close to $\underline{x}$,

$$\mathscr{R}(x) \leq \max[\mathscr{R}(x'), \mathscr{R}(\bar{x})] \leq \max[1, \mathscr{R}(\bar{x})] = \mathscr{R}(\bar{x}),$$

which implies by Lemma 7.v, that F CSSD G.

Finally, we show that the conditions imposed on the densities in the Lemma are sufficient to ensure that $\lim_{x \to \underline{x}} \mathscr{R}(x) = 1$. First note that the stochastic dominance of F and the continuity of f imply that in a sufficiently small neighborhood of $\underline{x}$, $f(x) \leq g(x)$. Thus, convergence f/g to infinity is not possible. Hence, the sufficient condition given in the Lemma implies that

$f/g \to L$, were L is finite and positive. Thus by L'Hosptial's rule

$$\lim_{x \to \underline{x}} \frac{F(x)}{G(x)} = L.$$

Thus,

$$\lim_{x \to \underline{x}} \frac{f(x)\,G(x)}{g(x)\,F(x)} = \lim_{x \to \underline{x}} \frac{f(x)/g(x)}{F(x)/G(x)} = L/L = 1.$$

□

Proof of Theorem 4. 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 2 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 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 continuous, 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. □

Proof of Theorem 5. We start by proving assertion (a) of part (i). Assertion (b) follows from (a) and part (i) of Lemma 7.

For any real valued function of a single variable, f, let $D_+ f(x)$ represent the right 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.

Define the function $\hat{v}(y) : (-\infty, 0] \to \mathbb{R}$ by

$$\hat{v}(y) = \hat{u}(y) - y, \quad y \leq 0. \tag{A-24}$$

Where $\hat{u}$ is conjugate function to u defined in Lemma 2. First note that, because u is strictly geometrically convex, $\hat{u}$ is strictly convex and thus $\hat{v}$ is strictly convex. The hypothesis of strict stochastic dominance implies that $u(t) < t, t \in (0,1)$. Because the conjugate transform is order preserving, this implies that $\hat{u}(y) - y \leq 0$, Thus, $\hat{v}$, which is defined on the non-positive real line, is bounded from above by the x-axis. Thus $\hat{v}$ must be increasing. To see this note that if $\hat{v}$ were decreasing anywhere, it would have a support line with a negative slope. For y sufficiently small this support line would cross the x-axis. Because $\hat{v}$ is convex its graph lies above all of its support lines. Thus, $\hat{v}$ would cross the x-axis. Because, by hypothesis, such crossing is impossible, it must be the case that $\hat{v}$ is increasing.

Using the definition of the conjugate function given in Lemma 2 we can see that

$$\hat{v} \circ \log(t) = \hat{u} \circ \log(t) - \log(t). \tag{A-25}$$

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

$$u(t) = t \exp[\hat{v}(\log(t))]. \tag{A-26}$$

Because $\hat{u}$ is convex, it has left and right derivatives. Thus, u has left and right derivatives. We differentiate equation (A-26) and obtain

$$\begin{aligned} D_+u(t) = \exp[\hat{v}(\log(t))] + t \exp[\hat{v}(\log(t))] D_+\hat{v}(t) \frac{1}{t} = \\ \exp[\hat{v}(\log(t))] + \exp[\hat{v}(\log(t))] D_+\hat{v}(t), \quad t \in (0,1). \end{aligned} \tag{A-27}$$

Because $\hat{v}$ is increasing and convex, $D_+\hat{v}$ is positive and increasing. Thus, equation (A-27) implies that D_+u is increasing over (0, 1), which implies that u is convex.

The proof of part (ii) is almost identical. As discussed in Section 3.5, stochastic dominance implies that the average transform function, U satisfies, $U(t) \leq {}^1\!/\!{}_2 t$. This implies that is conjugate transform, $\hat{U}$ satisfies $\hat{U} \leq \log({}^1\!/\!{}_2) + y$. We use this fact to define the function $\hat{\gamma}(y) : (-\infty, 0] \to \mathbb{R}$ by

$$\hat{v}(y) = \hat{u}(y) - y - \log(1/2), \quad y \leq 0. \tag{A-28}$$

From this point on, the proof of part (ii) proceeds in exactly the same fashion as the proof of part (a). □

Proof of Lemma 10. To prove the assertion for geometric convexity and geometric convexity on average define the map, $\Phi : (-\infty, 0] \to \Re$ by $\Phi(y) = \log[p + (1-p) e^y$. Equation 7.2 shows that Φ is strictly convex and it is obvious that it is increasing. Let $u_o = F_o \circ G^{-1}$ and let $u = F \circ G^{-1}$ Then $\hat{u} = \Phi \circ \hat{u}_o$. By assumption, $\hat{u}_o$ is convex. Φ is an increasing convex map, and thus $\Phi \circ \hat{u}_o = \hat{u}$ is strictly convex, i.e., F strictly geometrically dominates G. Exactly the same argument, with U replacing u, works to prove the assertion for geometric convexity on average.

Now consider CSSD. Let $\theta(t) = p/u(t)$, then we can write

$$\Pi[u](t) = \frac{p+(1-p)\,U_o(t)}{p+(1-p)\,u_o(t)} = \frac{\theta(t)+(1-p)\,\Pi[u_o](t)}{\theta(t)+(1-p)}.$$

Because, $\Pi\,[u_o] < 1$, the function

$$\theta \hookrightarrow \frac{\theta+(1-p)\,\Pi[u_o](t)}{\theta+(1-p)} \tag{A-29}$$

is strictly increasing in θ. Moreover, if $t \in (0,1)$, $\theta(t) > \theta(1)$. Thus,

$$\Pi[u](t) = \frac{\theta(t)+(1-p)\,\Pi[u_o](t)}{\theta(t)+(1-p)} > \frac{\theta(1)+(1-p)\,\Pi[u_o](t)}{\theta(1)+(1-p)},\ t \in (0,1). \tag{A-30}$$

By the assumption of CSSD dominance of F_o over G, $\Pi[u_o](t) \geq \Pi[u_o](1)$. Thus,

$$\frac{\theta(1)+(1-p)\,\Pi[u_o](t)}{\theta(1)+(1-p)} \geq \frac{\theta(1)+(1-p)\,\Pi[u_o](1)}{\theta(1)+(1-p)} = \Pi\,[u](1). \tag{A-31}$$

Expressions (A-30) and (A-31) imply that, $\Pi[u_o](t) > \Pi[u_o](1), t \in (0,1)$, i.e., F strictly CSSD dominates G. □

Proof of Theorem 6. We first establish (i). First note that if F places positive mass on $\underline{x}$ then because G places no mass on $\underline{x}$, the result is trivially true. So suppose that F places no mass on $\underline{x}$. In this case, $u(0) = 0$.. Note that $\Pi[u](1) = U(1)/u(1) = U(1)$ and that F being strictly stochastically dominated implies that $U(1) > 1/2$ (See the discussion in Section 3.5. Thus CSSD requires that $\Pi[u](t) = U(t)/u(t) > \Pi[u](1) > 1/2$. Thus there exists $c > 0$, such that $U(t)/u(t) \geq 1/2 + c$. Next note $(t\,U(t))' = u(t)$ and thus $u(t) = t\,U'(t) + U(t)$. Hence,

$$U(t) \geq \left(\frac{1}{2}+c\right)\,u(t) = \left(\frac{1}{2}+c\right)\,\Big(t\,U'(t)+U(t)\Big). \tag{A-32}$$

Letting $K = (\frac{1}{2}+c)/(\frac{1}{2}-c)$, algebraic rearrangement shows that there exists $K > 1$ such that

$$K\,U'(t) \leq \frac{U(t)}{t}, \text{for all } t \in (0,1). \tag{A-33}$$

Because U is differentiable on $(0,1)$ and continuous on $[0,1]$ we can apply the mean value theorem, to show that there exists $\eta(t) \in (0,t)$ such that

$$K\,U'(t) \leq \frac{U(t)}{t} = U'(\eta(t)) \leq \sup_{s\in(0,t)} U'(\eta(t)). \tag{A-34}$$

The fact that F is stochastically dominated by G implies (see discussion in Section 3.5) that $U(t)/t \geq 1/2$. If $U(t)/t$ is bounded on $(0,t)$.

$$\limsup_{t\to 0} U'(t) = L \in (0,\infty) \text{ and } \limsup_{t\to 0} K U'(t) = K\,L = \limsup_{t\to 0} U'(t) = L. \tag{A-35}$$

This is not possible because $K > 1$. Thus it must be the case that $\limsup_{t\to 0} U(t)/t = \infty$.

Next note that

$$\frac{U(t)}{t} = \frac{\int_0^t \frac{u(s)}{s} s\, ds}{t^2} = \frac{\int_0^t \frac{u(s)}{s} s\, ds}{2\int_0^t s\, ds} \tag{A-36}$$

$u(s)/s$ is continuous for $s > 0$ and thus can only be unbounded on a neighborhood of 0. If $u(s)/s$ is bounded on a neighborhood of 0 there exists $B < \infty$ such that $u(s)/s < B$ uniformly on this neighborhood which by equation (27) implies that $U(t)/t \leq B/2$. As this is not possible given that $\limsup U(t)/t \to \infty$ we conclude that $\limsup u(t)/t = \infty$.. Substituting $G(x)$ for t then shows that the assertion in part (i) must hold.

Now consider part (ii). First define the function

$$\hat{\beta}(y) = \hat{U}(y) - y, \quad y \leq 0. \tag{A-37}$$

First note that, because U is strictly geometrically convex, $\hat{U}$ is thus strictly convex and hence $\hat{\beta}$ is strictly convex. Because $\hat{\beta}$ is the conjugate transformation of $U(t)/t$. Because geometric dominance on average implies CSSD dominance, and, thus by part (i) $U(t)/t$ is not increasing everywhere. Thus, $U(t)/t$ must be strictly decreasing on some interval. The conjugate transform is continuous, smooth, and order preserving so it must be the case that $\hat{\beta}$ is also strictly decreasing on some interval. But $\hat{\beta}$ is convex and thus, if it is strictly decreasing at any point, say y^o, it is strictly decreasing for all $y < y^o$. Thus for $y < y^o$, β is strictly decreasing.

Using the definition of the conjugate function given in Lemma 2 we can see that

$$\hat{\beta} \circ \log(t) = \hat{U} \circ \log(t) - \log(t). \tag{A-38}$$

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

$$\frac{U(t)}{t} = \exp[\hat{\beta}(\log(t))]. \tag{A-39}$$

Because the log and exp functions are order preserving, and β is strictly decreasing when $\log t < y^o$, equation (A-39) implies that there exists $t^o \in (0,1)$ such that for all $t < t^o$, $U(t)/t$ is strictly decreasing. Finally note that the function defined in part (ii) is simply the composition of $U(t)/t$ with G. Thus we have established assertion (a) of part (ii).

To prove assertion (b) of part (ii), note that monotonicity of $U(t)/t$ combined with the result in the proof of part (i) which showed that $\limsup_{t\to 0} U(t)/t = \infty$, implies that $\lim_{t\to 0} U(t)/t = \infty$. Next note that because $u \geq U$, $u(t)/t \geq U(t)/t$. Hence,$\lim_{t\to 0} U(t)/t = \infty$ implies that $\lim_{t\to 0} u(t)/t = \infty$. Which implies by the same argument as used in part (i) the assertion in the Theorem that $\lim_{x\to \underline{x}} F(x)/G(x) = \infty$.

Now consider part (iii). Analogously to part (ii) define $\hat{v}(t) = \hat{u}(t) - t$, where $\hat{u}$ is the conjugate function to u, (defined in Lemma 2). Note that $u(1) = 1$, which implies that $\hat{u}(0) = 0$.

The fact that F is strictly stochastically dominated implies that $u(t) > t$, $t \in (0,1)$, thus,

$$\hat{v}(y) \geq 0,\, y \leq 0 \text{ and } \hat{v}(0) = 0. \tag{A-40}$$

Because $\hat{v}$ is strictly convex, (A-40) implies that $\hat{v}$ is strictly decreasing. To see this, note that because $\hat{v} \geq 0$ and $\hat{v}(0) = 0$, $\hat{v}$, and $\hat{v}$ is strictly convex and thus monotone on intervals, it is strictly decreasing in neighborhood of 0. Convexity, then implies that it must be strictly decreasing for all $y < 0$. Using equation (A-26), we can write

$$\frac{u(t)}{t} = \exp[\hat{v}(\log(t))],\; t \in (0,1]. \tag{A-41}$$

Because, exp and log are strictly increasing functions and $\hat{v}$ is strictly decreasing, $t \to u(t)/t$ is strictly decreasing. Thus, $t \to u(t)/t$ is strictly decreasing. Because $\hat{v}$ is strictly convex and strictly decreasing it is bounded from below by support line with a negative slope. Thus $\lim_{y \to -\infty} \hat{v}(y) = \infty$. Hence, (A-41) implies that, $\lim_{t \to 0} u(t)/t = \infty$. This implies for the same reasons as given in part (ii) that, $x \hookrightarrow F(x)/G(x)$ is strictly decreasing. This establishes part (a) of (iii). Part (b) of (iii) follows because geometric dominance implies geometric dominance on average by Lemma 4, and by part (ii.b), part (b) holds. □

Proof of Lemma 11. The distribution of $\tilde{X}_c$ is $G(x-c)$, $x \in \Re$. Thus, the transform function, u, is given by

$$u(t) = G(G^{-1}(t) - c).$$

Using the inverse function theorem we see that

$$u'(t) = \frac{G'(G^{-1}(t) - c)}{G'(G^{-1}(t))} = \frac{g(G^{-1}(t) - c)}{g(G^{-1}(t))}.$$

Thus, the function $R = (tu')/u$ defined in Lemma 3 is given by

$$R(t) = \frac{g(G^{-1}(t) - c)}{G(G^{-1}(t) - c)} \frac{t}{g(G^{-1}(t))}.$$

Strict geometric dominance is equivalent to R being strictly increasing. Make the substitution $x = G^{-1}(t)$. This yields

$$R \circ G(x) = \frac{g(x-c)}{G(x-c)} \frac{G(x)}{g(x)} = \frac{r(x-c)}{r(x)}.$$

The ratio $r(x-c)/r(x)$ is (strictly) increasing for every choice of $c > 0$ if and only if

$$\forall c > 0,\, x \hookrightarrow \log \circ r(x-c) - \log \circ r(x) \text{ is (strictly) increasing over } \Re. \tag{A-42}$$

Because $\log \circ r$ is continuous, Condition (A-42) holds if and only if $\log \circ r$ is concave (Kuczma, 2000, Theorems 7.3.3 and 7.3.4). □

Proof of Lemma 12. The distribution of $\tilde{X}_c$ is $G(x/c)$, $x \in \Re^+$. Let $\theta = 1/c < 1$ Thus, the transform function, u, is given by

$$u(t) = G(\theta\, G^{-1}(t)).$$

Using the inverse function theorem, we see that

$$u'(t) = \frac{\theta\, G'(\theta\, G^{-1}(t))}{G'(G^{-1}(t))} = \frac{\theta\, g(\theta\, G^{-1}(t))}{g(G^{-1}(t))}.$$

Thus, the function $R = (tu')/u$ defined in Lemma 3 is given by

$$R(t) = \frac{\theta\, g(\theta\, G^{-1}(t))}{G(\theta\, G^{-1}(t))} \frac{t}{g(G^{-1}(t))}.$$

(Strict) geometric dominance is equivalent to R being (strictly) increasing. Make the substitution $x = G^{-1}(t)$, this yields

$$R \circ G(x) = \frac{\theta g(\theta\, x)}{G(\theta\, x)} \frac{G(x)}{g(x)} = \frac{\theta\, r(\theta\, x)}{r(x)}.$$

The ratio $\theta\, r(\theta\, x)/r(x)$ is (strictly) increasing for every choice of $\theta \in (0,1)$ if and only if

$$\forall \theta \in (0,1), x \hookrightarrow \log \circ r(\theta\, x) - \log \circ r(x) \text{ is (strictly) increasing over } \Re^{+}. \tag{A-43}$$

Next, let $\hat{r}$ represent the conjugate function to r defined in Lemma 2. Let $y = \log(x)$ and let $\beta = -\log(\theta)$ and note that $\log \circ r = \hat{r} \circ \log$. Thus, condition (A-43) is equivalent to the condition that

$$\forall \beta > 0, y \hookrightarrow \hat{r}(y - \beta) - \hat{r}(y) \text{ is (strictly) increasing over } (-\infty, 0). \tag{A-44}$$

By exactly the same argument given in Lemma 11, condition (A-44) is equivalent to $\hat{r}$ being (strictly) concave and $\hat{r}$ being (strictly) concave is equivalent to r being (strictly) geometrically concave. □

Proof of Lemma 13. First consider part (i). Assume that F is an upscaling of G. By the same argument as given in Lemma 12, for the upscaled distribution to be geometrically dominated it would have to be the case that the reverse hazard rate of G was geometrically convex, which, given differentiability, by Lemma 3, is equivalent to $(xr')/r$ being increasing. Note that the quotient rule for differentiation implies that

$$\frac{x\, r'(x)}{r(x)} = x \left(\frac{f'(x)}{f(x)} - r(x) \right). \tag{A-45}$$

If we differentiate the right-hand side of (A-45), the condition for geometric convexity can be expressed as

$$\left(\frac{g'(x)}{g(x)} - r(x) \right) + x \left(\frac{g'(x)}{g(x)} - r(x) \right)' \geq 0. \tag{A-46}$$

If the distribution G has a logconcave density, then $(g'/g)'$ is nonpositive. Thus,

$$\left(\frac{g'(x)}{g(x)}-r(x)\right)+x\left(\frac{g'(x)}{g(x)}-r(x)\right)'\leq\left(\frac{g'(x)}{g(x)}-r(x)\right)-x\,r(x). \tag{A-47}$$

Thus, a necessary condition for geometric convexity is that

$$\left(\frac{g'(x)}{g(x)}-r(x)\right)-x\,r(x)\geq 0. \tag{A-48}$$

Next note that applying the quotient rule for differentiation yields

$$\frac{r'(x)}{r(x)}=\frac{g'(x)}{g(x)}-r(x). \tag{A-49}$$

Substituting equation (A-49) into (A-48) yields the equivalent condition

$$r'(x)(1-x\,r(x))\geq 0.$$

Because g is strictly logconcave, $r'<0$, except perhaps on a set of zero measure. Thus, for almost all $x>0$, it must be the case that $x\,r(x)\geq 1$, i.e., $r(x)\geq 1/x$. However,

$$G(x)=\exp\left(-\int_x^{\infty}r(z)\,dz\right). \tag{A-50}$$

If $r(x)\geq 1/x$, the integral on the right-hand side of (A-50) is unbounded for all x, and thus $G(x)=1$, $x>0$, contradicting G having support $[0,\infty)$.

Now consider part (ii). We construct and example where G is logconcave yet the reverse hazard rate is logconvex. The example is constructed as follows: define

$$r(x)=x^{-2}\exp(-\operatorname{Ei}(-x)),\quad x>0,$$

were Ei represents the Exponential Integral function. Define G using equation (A-50). Note first that

$$r'(x)=-\frac{e^{-x-\operatorname{Ei}(-x)}\,(1+2\,e^{x})}{x^3}<0.$$

and thus G is strictly logconcave. Next note that $(x\,r'(x))/r(x)=-e^{-x}-2$, and thus $(xr'(x))/r(x)$ is strictly increasing. Hence r is geometrically convex, which implies that G geometrically dominates all of its upscalings.

Now consider part (iii). Again, we establish the result by means of an example, let G be defined by

$$G(x)=\exp\left(1-\exp\left(e^{-x}\right)\right),\quad x\in\Re.$$

Direct calculation using the G function defined above shows that r and g'/g are given by

$$r(x)=\exp\left(e^{-x}-x\right),$$
$$\frac{g'(x)}{g(x)}=-1-e^{-x}\left(1-\exp\left(e^{-x}\right)\right).$$

Inspection shows that r is logconvex and that g'/g is strictly decreasing for all $x \in \Re$. Thus, G has a log concave density yet G strictly geometrically dominates all of its upshifts. $\square$

Appendix B 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

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

where $\eta_o = 3/50$.

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

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

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, F, G, and H over the unit interval:

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

These distributions, as well as their associated selection-dominance functions, Π, defined in equation (16), are graphed in Figure B. Panels B and C of Figure B verify that u_1 and u_2 satisfy the CSSD condition given by expression (17). Theorem 2 shows this condition 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 CSSD condition given in expression (17). 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 dominate H by CSSD. Hence, the CSSD relation is not transitive.

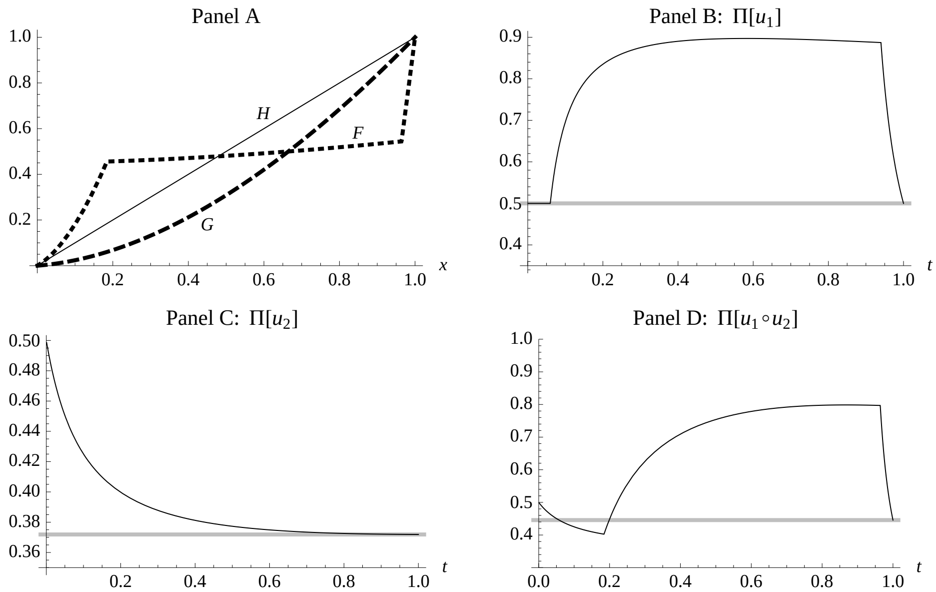


Figure 5: Counterexample to transitivity of selection dominance. Panel A plots the distribution functions, F G and H. Panel B plots the function, $\Pi(u_1)$, (defined in expression (16)) used to test the selection dominance of F over G, Panel C plots the function, $\Pi(u_2)$, used to test the selection 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.

Appendix C First-price auction with minimum bid

Using the inverse bid function derived in Kaplan and Zamir (2012) and the relation between the inverse bid function and the bid distribution developed in the body of this paper, we see that the bid distributions for H and L are given as follows, for $b < m$, which corresponds to case where the bidders valuation is less than the minimum price, bidders bid their value, and their bids are rejected. Thus, $F_i(b) = b/v_i$, $i = H, L$ for $b < m$. Otherwise

$$F_L(b) = \frac{m^2}{v_L\left(b - c\sqrt{b^2 - m^2}\right)}, \quad b \in [m, \bar{b}], \tag{C-1}$$

$$F_H(b) = \frac{m^2}{v_H\left(b - \frac{1}{c}\sqrt{b^2 - m^2}\right)}, \quad b \in [m, \bar{b}], \text{ where} \tag{C-2}$$

$$\bar{b} = \frac{v_L v_H + m^2}{v_H + v_L}, \text{ and} \tag{C-3}$$

$$c = \frac{v_H}{v_L}\sqrt{\frac{v_L^2 - m^2}{v_H^2 - m^2}} < 1. \tag{C-4}$$

Using some algebraic manipulation, it is possible to show that, for $b \in [m, \hat{b}]$.

$$F_H(b) = F_L(b)\,\phi(b), \text{ where} \tag{C-5}$$

$$\phi(b) = \frac{v_L}{v_H}\left(c^2 + \frac{c\left(1-c^2\right)}{c - \sqrt{1-\frac{m^2}{b^2}}}\right). \tag{C-6}$$

Thus, the reverse hazard rate for H, for $b \in [m, \hat{b}]$ can be expressed as follows:

$$r_H(b) = (\log \circ F_H(b))' = (\log \circ (F_L(b)\,\phi(b))' = (\log \circ F_L(b))' + (\log \circ \phi(b))' = r_L(b) + (\log \circ \phi(b))'. \tag{C-7}$$

(C-7) implies that the ratio between the reverse hazard rates is given by

$$\frac{r_H(b)}{r_L(b)} = 1 + \frac{(\log \circ \phi(b))'}{r_L(b)} = 1 + \frac{\left(1-c^2\right)m^2}{\left(b\,c - \sqrt{b^2-m^2}\right)^2}, \quad b \in (m, \bar{b}]. \tag{C-8}$$

By Lemma 7, if we can verify that this ratio is increasing we will establish the geometric dominance of F_H over F_L. Next, note that

$$\frac{d}{db}\left(b\,c - \sqrt{b^2-m^2}\right) = c - \frac{1}{\sqrt{1-m^2/b^2}}. \tag{C-9}$$

The right-hand side of (C-9) is always negative because

$$\frac{1}{\sqrt{1-m^2/b^2}} \geq \frac{1}{\sqrt{1-m^2/\bar{b}^2}} = \frac{v_H\left(m^2+v_H\,v_L\right)}{v_L\left(v_H^2-m^2\right)}\,\frac{1}{c} \geq \frac{1}{c} > 1 > c.$$

Thus, $b \hookrightarrow b\,c - \sqrt{b^2-m^2}$, for $b \in [m, \bar{b}]$, is strictly decreasing, which implies, given (C-8) that r_H/r_L is strictly increasing when $b \in [m, \bar{b}]$. Next note that, at $b = m$,

$$\frac{r_H(m+)}{r_L(m+)} = \frac{1}{c^2} > 1,$$
$$\frac{r_H(m-)}{r_L(m-)} = \frac{1}{c^2} = 1.$$

For $b \in [0, m]$, $r_H/r_L = 1$. Thus, r_H/r_L is increasing and geometric dominance is verified.

Appendix D: Verification of order relations for specific distributions

Supplementary material for

Stochastic orders and the anatomy of competitive selection

Thomas Noe
Balliol College/SBS

1 June, 2015

Preliminary remarks

We will make use of the following theorem frequently and thus, to minimize repetition, we provide it here and, in the sequel, simply refer to it as MLHR.

Theorem 1 (Monotone LĤospital Rule **MLHR** (Proposition 1.1 in Pinelis (2002)))**.** *Let $-\infty \leq a < b \leq \infty$. Let ϕ and ν be real-valued differentiable functions defined on (a,b) and that Suppose that either $\phi(a+) = \nu(a+) = 0$ or $\phi(b-) = \nu(b-) = 0$, and that ν' does not change sign over the interval (a,b), then*

$$x \hookrightarrow \frac{\phi'(x)}{\nu'(x)} \text{ is (strictly) increasing} \Rightarrow x \hookrightarrow \frac{\phi(x)}{\nu(x)} \text{ is (strictly) increasing.}$$

We try to distinguish the specific properties of the given distribution being analyzed from general properties of the distributions and their associated densities and reverse hazard rate. For the short derivations, we do this by labeling the densities, distributions and hazard rates, with a subscript indicating the specific distribution being analyzed. For the longer derivations, e.g., Gompertz and Gamma distributions, this is impractical because it leads to very bulky expressions. Therefore, in sections devoted to these distributions, we simply state that all references to distributions, densities, reverse hazard rates, in the given section refer to the specific densities, distributions, and hazard rates of the distribution being analyzed in that section. Note also that we are dealing with textbook distributions. Only one of the distributions considered, the Laplace distribution, lacks smooth derivatives of all orders. And even the Laplace is smooth at all points in its support save one. Thus, we will frequently use continuous differentiability in the proofs without explicitly invoking continuous differentiability is our arguments.

1 Normal

The Mill's ratio for the Normal distribution is log convex (see, e.g., Sampford (1953)). This implies that the hazard rate, h_{N} of the Normal distribution is log concave. The symmetry of the normal distribution implies that, the hazard rate and reverse hazard rate, r_N, are related by $h_{\mathrm{N}}(-x) = r_{\mathrm{N}}(x)$, $x \in \Re$. Note that $\log \circ h_{\mathrm{N}}(x)$ is concave if and only if $\log \circ h_{\mathrm{N}}(-x) = r_{\mathrm{N}}(x)$ is concave. Thus r_N is logconcave. Geometric dominance follows from Lemma 11.

2 Logistic

Assume without loss of generality that $s = 1$ and $\mu = 0$. Under this assumption, the reverse hazard rate for the logistic is given by

$$r_{\mathrm{Logistic}}(x) = \frac{1}{1+e^{-x}}, x \in \Re.$$

Thus,

$$\log \circ r_{\mathrm{Logistic}}(x) = -\log\left(1+e^{x}\right).$$

Next note that

$$(\log \circ r_{\mathrm{Logistic}}(x))'' = -\frac{e^{x}}{(1+e^{x})^{2}} < 0.$$

Thus r_{Logistic} is logconcave and geometric dominance thus follows from Lemma 11.

3 Laplace

Let G_{Laplace} and g_{Laplace} represent the distribution and density of the standard ($s = 1$, $\mu = 0$) Laplace distribution. First, we show that geometric dominance fails in general and then show that geometric dominance on average holds. Note that the reverse hazard rate for the Laplace, r_{Laplace} is given by $r_{\mathrm{Laplace}} = (2\exp(\max[x,0]-1))^{-1}$. It is easy to verify that, restricted to $x > 0$, r_{Laplace} is log convex and thus by Lemma 11, G is not geometrically dominated by all of its upshifts.

Let F_{Laplace} represent a μ-upshift of G_{Laplace}. We verify geometric convexity on average using the conditions given in Lemma 7 to show that the function

$$\mathscr{R}_{\mathrm{Laplace}}(x) = \frac{\int_{-\infty}^{x} g_{\mathrm{Laplace}}(z-\mu)\, G_{\mathrm{Laplace}}(z)\, dz}{\int_{-\infty}^{x} g_{\mathrm{Laplace}}(z)\, G_{\mathrm{Laplace}}(z-\mu)\, dz}.$$

is increasing. After considerable, tedious, but quite standard, calculus calculations, one can express $\mathscr{R}_{\mathrm{Laplace}}$ in the following somewhat compact form. Let $w = e^{x}$, and let $m = e^{\mu}$, then

$$\mathscr{R}_{\text{Laplace}}(w) = \begin{cases} 1 & \text{if } w \leq 1, \\ T_2(w,m) & \text{if } w \in (1,\mu], \\ T_3(w,m) & \text{if } w > \mu, \end{cases} \tag{D-1}$$

$$T_2(w,m) = \frac{4w-2}{1+2\log(w)} - 1, \tag{D-2}$$

$$T_3(w,m) = \frac{w^2(8m-4-2\text{Log}(m)) - 4m^2 w + m^2}{w^2(4+2\log(m)) - 4mw + m^2}. \tag{D-3}$$

Note that $\mathscr{R}_{\text{Laplace}}$ is continuous and thus, to verify that it is increasing, one needs only verify that it is increasing on each leg of its definition. Since $\mathscr{R}_{\text{Laplace}}$ is constant on the first leg, the assertion is verified for this leg. Now consider the second leg where $w \in (1,\mu]$. Note that

$$\frac{\partial T_2}{\partial w} = \frac{4(2w\log(w) - (w-1))}{w(1+2\text{Log}(w))^2}.$$

$w \hookrightarrow w\log(w)$ is strictly convex for $w \geq 1$. It support line at 1 is $w-1$. Strictly convex functions exceed their support lines. Thus, $\partial T_2/\partial w > 0$ and hence, T_2 is strictly increasing in y.

Now consider the third leg where $y > \mu$. Note that

$$\frac{\partial T_3}{\partial w} = \frac{4mS(w,m)}{\left((2w-m)^2 + 2w^2\text{Log}(m)\right)^2}, \tag{D-4}$$

$$S(w,m) = 2w^2 Q(m) + 2wm(2(m-1) - \log(m)) - (m-1)m^2, \tag{D-5}$$

$$Q(m) = 2((1+m)\log(m) - 2(m-1)). \tag{D-6}$$

First note that differentiation shows that Q is strictly convex for $m > 1$. Next note that $Q'(1) = Q(1) = 0$. This fact combined with the strict convexity of Q implies that $Q(m) > 0$, for all $m > 1$.

Q is positive and$S$ is strictly convex in y. Thus, to show that S is positive on the third leg, we need only show that evaluated at $y = m$, S is nonnegative and increasing.

$$S(y=m,m) = m^2(2m\log(m) - (m-1)), \tag{D-7}$$

$$\frac{\partial S}{\partial w}(y=m,m) = 2m(\log(m) + 2(m\log(m) - (m-1))). \tag{D-8}$$

Because $m > 1$ on the third leg, by the same argument as used in the proof for branch two, $m\log(m) - (m-1) > 0$. Thus, the right-hand sides of (D-7) and (D-8) are both positive and hence, the strict convexity of S in y implies that $S > 0$. Inspection of (D-4) shows that this implies that $\frac{\partial T_3}{\partial w} > 0$, and thus on third leg, $\mathscr{R}_{\text{Laplace}}$ is strictly increasing. Hence, we have established that $\mathscr{R}_{\text{Laplace}}$ is increasing in y. Because $y = e^x$, $\mathscr{R}_{\text{Laplace}}$ is increasing in x. Thus, the translated Laplace distribution always geometrically dominates on average the original distribution but does not dominate it geometrically.

In essence, the kink in the density of the Laplace distribution generates a nonconvexity in the conjugate transform distribution which is always smoothed out sufficiently by the averaging to render the conjugate of the average transform function convex. This averaging effect is illustrated in Figure 1.

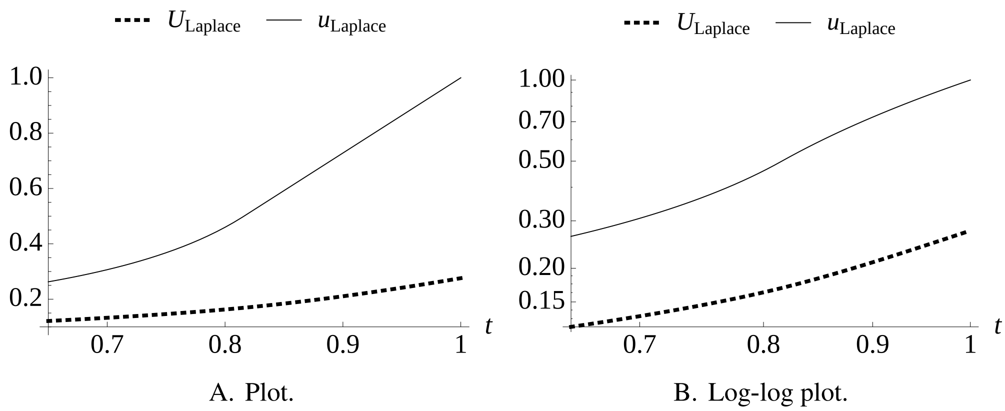


Figure 1: *Convexity and geometric convexity of the transform functions for the Laplace Distribution.* The figure in Panel 1 represent a plot of the transform, u, and average transform, U, functions when the produced by comparing a standard $\mu = 0$, $s = 1$ Laplace distribution to and upscaled Laplace distribution with $\mu = 1$, $s = 1$. The first figure Panel A graphs these functions using standard scaling of the axes. The in the second, Panel B, the axes are log scaled. The domain of the graph, is restricted to $[0.65, 1]$ to make the non-convexity of u when plotted in log-scale more apparent.

The convexity of the transform function, u, indicated by Panel 1A (which implies the concavity of the average transform function, U) indicates by Lemma 7, that the up translated distribution is MLRP dominant. As shown in Section 3.3, geometric dominance requires the transform function to be convex when plotted in a log-scaled graph. Panel B shows that this is not the case in this example. However, averaging smooths out the non-convexity and as Panel B shows, the average transform function, U, is convex in the log scale plot.

4 Gumbel

Selection equivalence is quite easy to show using many of the tests developed in the manuscript. Perhaps the most transparent demonstration follows from noting that if G is standard $s = 1, \mu = 0$ Gumbel, and F is a c-upshift of G then, the transform function, u, is given by $u(t) = t^C$, where $C = e^c$. The function $t \hookrightarrow t^C$ is clearly geometrically linear. Thus, the c-upshift is geometrically equivalent to the original distribution and hence, by Lemma 6, the distributions are selection equivalent.

5 Gamma

In order to make the notation more manageable. Define $\theta = 1/s$. In this section let G represent a standard unit scale Gamma distribution with shape parameter α, i.e.,

$$G(x) = \frac{1}{\Gamma(\alpha)} \int_0^x z^{1-\alpha} e^{-z} dz = \frac{\gamma(\alpha,x)}{\Gamma(\alpha)},$$

where Γ represents the Gamma function and γ represents the lower incomplete Gamma function.

Let F represent an upscaled Gamma distribution with the same shape parameter α but with scale parameter $1/\theta$, $\theta \in (0,1)$. Thus the distribution of F is given by $G(\theta x)$, $x > 0$, and the reverse hazard rates of F and G are given by

$$r_F(x) = \theta\, r(\theta x); \quad r_G(x) = r(x), \quad x > 0.$$

Note that for $x > 0$, the Gamma distribution, its density, and reverse hazard rate are continuously differentiable. We will use this fact without comment frequently in the sequel. Geometric dominance follows from a series of results.

Result 1. The reverse hazard rate, r, for a unit scale Gamma distribution is defined by

$$r(x) = \frac{e^{-x} x^{\alpha-1}}{\gamma(\alpha,x)}, \quad x > 0.$$

The reverse hazard rate for the Gamma, r, has the following basic properties:
(a) r is continuously differentiable on $[0,\infty)$.
(b) r is strictly decreasing.
(c) $\lim_{x\to 0} r(x) = \infty$.
(d) $\lim_{x\to\infty} r(x) = 0$.

Proof. Part (a) follows from the infinite differentiability of the Gamma distribution. Parts (c) and (d) follow from straightforward calculations, and part (b) follows from the log concavity of the Gamma distribution function (Baricz, 2010). □

Result 2. The reverse hazard rate of for the Gamma distribution is geometrically concave on a sufficiently small neighborhood of 0.

Proof. Simple but tedious manipulations show that

$$\frac{x\,r'(x)}{r(x)} = -1 + \left(\frac{e^{-x} x^{\alpha}}{\gamma(\alpha,x)}\right)^2 \left(\frac{\gamma(\alpha,x)}{e^{-x} x^{\alpha}} - \frac{\gamma(\alpha,x)}{e^{-x} x^{\alpha+1}}\right). \tag{D-9}$$

Applying L'Hopital's Rule to the terms of (D-9) shows that

$$\lim_{x\to 0} \frac{e^{-x}x^{\alpha}}{\gamma(\alpha,x)} = \alpha \tag{D-10}$$

$$\lim_{x\to 0} \frac{\gamma(\alpha,x)}{e^{-x}x^{\alpha}} = \frac{1}{\alpha} \tag{D-11}$$

$$\lim_{x\to 0} \frac{\gamma(\alpha,x)}{e^{-x}x^{\alpha+1}} = \frac{1}{1+\alpha}. \tag{D-12}$$

Therefore

$$\lim_{x\to 0} \frac{x\,r'(x)}{r(x)} = -\frac{1}{1+\alpha} < 0.$$

Thus, by Lemma 3, r and satisfies, in some neighborhood of 0, the conditions for geometric concavity □

Result 3. (a) $\lim_{x\to 0} \frac{r_F(x)}{r_G(x)} = \lim_{x\to 0} \frac{\theta\, r(\theta x)}{r(x)} = 1$ and (b) $\lim_{x\to\infty} \frac{r_F(x)}{r_G(x)} = \lim_{x\to\infty} \frac{\theta\, r(\theta x)}{r(x)} = \infty$.

Proof. First consider part (a). An application of L'Hospital rule shows that

$$\lim_{x\to 0} x\, r_F(x) = \lim_{x\to 0} x\, r_G(x) = \alpha > 0.$$

Thus,

$$\lim_{x\to 0} \frac{r_F(x)}{r_G(x)} = \lim_{x\to 0} \frac{x\, r_F(x)}{x\, r_G(x)} = 1.$$

Now consider part (b). As $x \to \infty$, $F(x) \to 1$ and $G(x) \to 1$. Thus, the limiting behavior of the ratio will be determined by the ratio f/g which clearly approaches infinity at $x \to \infty$. □

Define,

$$\rho(x) = \begin{cases} \frac{r_F(x)}{r_G(x)} & \text{if } x > 0 \\ 1 & \text{if } x = 0. \end{cases} \tag{D-13}$$

Result 4. On a sufficiently small neighborhood of 0, $\rho'(x) > 0$.

Proof. Note from the definition of ρ, and the fact that ρ is positive, ρ' will be positive if and only if

$$\frac{r_F'(x)}{r_F(x)} - \frac{r_G'(x)}{r_G(x)} = \frac{\theta\, r'(\theta x)}{r(\theta x)} - \frac{r'(x)}{r(x)} > 0.$$

For $x > 0$, this condition will be satisfied if and only if

$$\frac{x\,\theta\, r'(\theta x)}{r(\theta x)} - \frac{x\,r'(x)}{r(x)} > 0. \tag{D-14}$$

Because, $\theta \in (0,1)$, $\theta x < x$. Thus Result 2 shows that condition (D-14) is satisfied on a sufficiently small neighborhood of 0. □

Note that, Result 3, Result 2, and the continuity and differentiability properties of the Gamma distribution, imply the following result

Result 5. The ρ function defined in equation (D-13) verifies the following conditions:

(a) ρ is continuous on $[0,\infty)$.

(b) ρ is continuously differentiable on $(0,\infty)$.

(c) There exits, $\varepsilon > 0$, such that, for all $x \in (0,\varepsilon)$

  (i) $\rho'(x) > 0$.

  (ii) $\rho(x) > 1$.

(d) $\lim_{x\to\infty} \rho(x) = \infty$.

Define

$$\zeta(x) = \frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}, \quad x > 0. \tag{D-15}$$

Explicit calculation shows that

$$\zeta(x) = (1-\theta) > 0. \tag{D-16}$$

The quotient rule for differentiation shows that ρ verifies the equation

$$\rho'(x) = \rho(x)\left(\zeta(x) - (r_F(x) - r_G(x))\right), \quad x > 0. \tag{D-17}$$

Equations (D-17) and (D-16) thus imply that

$$\rho'(x) = \rho(x)\left((1-\theta) - (r_F(x) - r_G(x))\right), \quad x > 0. \tag{D-18}$$

Result 6. For all $x > 0$, $\rho(x) > 1$.

Proof. Suppose, to obtain a contradiction, that the result is false. Let $x^o = \inf\{x > 0 : \rho(x) \le 1\}$. By the continuity of ρ, $\rho(x^o) = 1$. By the definition of x^o, for all $x \in (0,x^o)$, $\rho(x) > 1$. Thus $\rho(x^0) = \min\{\rho(x) : x \in (0,x^o]\}$. However, because $\rho(x^o) = 1$, $r_F(x) - r_G(x) = 0$ and thus by equation (D-18), in a neighborhood of x^o, $\rho'(x) > 0$. Thus, ρ is increasing to the left of x^o and thus $\rho(x^o)$ cannot be the minimum value of ρ over $(0,x^o]$. □

Result 7. For all $x \ge 0$, $\rho'(x) > 0$.

Proof. Let $B = \{x > 0 : \rho'(x) \le 0\}$. By Results 5.d and 5.c.i, B is bounded and its infimum exceeds 0. By the continuity of ρ', B is open and, if it is not empty, it is thus a countable collection of disjoint open intervals. Select any of these open intervals, say (x_1,x_2), where $0 < x_1 < x_2 < \infty$. By the continuity of ρ', $\rho'(x_1) = \rho'(x_2) = 0$. Because, $\rho' \le 0$ on (x_1,x_2),

$$\rho(x_1) \ge \rho(x_2). \tag{D-19}$$

Because, $\rho'(x_1) = \rho'(x_2) = 0$, equation (D-18), and Result 6 imply that

$$0 < r_F(x_1) - r_G(x_1) = r_F(x_2) - r_G(x_2). \tag{D-20}$$

By Result 1.b, and the fact that $x_1 < x_2$, $r_G(x_1) > r_G(x_2)$. This fact, and equation (D-20) imply that

$$\rho(x_1) - 1 = \frac{r_F(x_1) - r_G(x_1)}{r_G(x_1)} < \frac{r_F(x_2) - r_G(x_2)}{r_G(x_1)} = \rho(x_2), \tag{D-21}$$

contradicting expression (D-19). Thus, B is empty and hence, for all $x > 0$, $\rho' > 0$. Thus, ρ is strictly increasing and hence the condition, given in Lemma 7, the geometric dominance of the upscaled distribution, F, is verified. □

6 Generalized Exponential

The reverse hazard rate for a standard ($s = 1$) Generalized Exponential with shape parameter b is given by

$$r_{\text{GExp}}(x) = b\,\frac{e^{-x}}{1 - e^{-x}}.$$

Applying the differential test for geometric concavity given in Lemma 3, yields

$$R_{\text{GExp}}(x) = \frac{x\, r'_{\text{GExp}}(x)}{r_{\text{GExp}}(x)} = \frac{x\, e^{x}}{1 - e^{x}}.$$

Differentiating R_{GExp} yields

$$R'_{\text{GExp}}(x) = \frac{e^{x}}{(1 - e^{x})^2}\,(1 + x - e^{x}).$$

$1 + x$ is the support line for e^x at $x = 1$. Because e^x is convex, e^x lies above the support line and thus $1 + x - e^x < 0$ for $x \neq 1$ and $1 + x - e^x = 0$ at $x = 1$. Hence, R is strictly decreasing, implying that r_{GExp} is strictly geometrically concave. By Lemma 12 this implies that G is strictly dominated by all of its upscalings.

7 Weibull

Let $a_F > a_G$ and suppose that F is distributed Weibull with size parameter a_F and shape parameter λ and that G is distributed Weibull with size parameter a_G and shape parameter λ. Then a simple direct calculation shows the transform function, u associated with F and G is given by

$$u_{\text{Weibull}}(t) = F \circ G^{-1}(t) = 1 - (1 - t)^{r}, \quad q = \left(\frac{a_G}{a_F}\right)^{\lambda} < 1, \quad t \in [0, 1].$$

Next note that

$$R_{\text{Weibull}} = \frac{t\, u'_{\text{Weibull}}(t)}{u_{\text{Weibull}}(t)} = \frac{q\,(1 - t)^{q-1}\, t}{1 - (1 - t)^{q}}.$$

Both the numerator and denominator of the above expression vanish at $t = 1$ and the derivative of denominator never changes sign. Thus, we can apply MLHR. The ratio of the derivatives of the numerator and denominator in the above expression is given by

$$1+(1-q)\,\frac{t}{1-t}$$

and this expression is clearly strictly increasing in t for $q<1$. Thus, u is strictly geometrically convex, and thus F strictly geometrically dominates G.

8 Pareto Distribution

Let $a_F > a_G > 1$ and suppose that F is distributed with size parameter α_F and shape parameter x_m and that G is distributed Weibull with size parameter α_G and shape parameter x_m. Suppose, without loss of generality that $x_m = 1$. Then a simple direct calculation shows the transform function, u associated with F and G is given by

$$u_{\text{Pareto}}(t) = F\circ G^{-1}(t) = 1-(1-t)^q, \quad q = \frac{\alpha_F\,(\alpha_G-1)}{\alpha_G\,(\alpha_F-1)} < 1, \quad t\in[0,1].$$

This is exactly the same transform function derived for the Weibull case above. Thus, for the same reasons as for the Weibull, u_{Pareto} is geometrically strictly convex that thus F strictly geometrically dominates G.

9 Kumaraswamy

The reverse hazard rate for the Kumaraswamy with parameters α and b is given by

$$\frac{b\,(1-x^{\alpha})^{b-1}}{1-(1-x^{\alpha})^{b}}.$$

We can rewrite this expression as

$$\chi(\alpha\,\log(x)), \quad \chi(y) = \frac{b\,e^{y}\,(1-e^{y})^{b-1}}{1-(1-e^{y})^{b}},\ y<0.$$

Suppose that $\alpha_F > \alpha_G$ the ratio of reverse hazard rates can be expressed as

$$\frac{\chi\,(\alpha_F\,\log(x))}{\chi\,(\alpha_G\,\log(x))}.$$

Because log is a strictly increasing function. This ratio will (strictly) increase if and only if $y\to\chi(\alpha_F\,y)/\chi(\alpha_G\,y)$ is (strictly) increasing over $(-\infty,0)$. By an argument very similar to the argument used to prove Lemma 12, this ratio will be increasing if and only if χ is geometrically convex. Geometric convexity thus depends on the monotonicity properties of

$$R_{\chi}(y) = \frac{y\,\chi'(y)}{\chi(y)}.$$

If R_χ is increasing the stochastically dominant distribution will be geometrically dominant. Similarly, the geometric dominance of the stochastically dominated distribution is equivalent to R_χ being decreasing; stochastic equivalence is equivalent to R_χ being constant. Next, note that

$$R_\chi(y) = \Lambda\left(e^y\right) M(y), \tag{D-22}$$

$$M(y) = \frac{e^y\,(-y)}{1-e^y}, \quad y \le 0, \tag{D-23}$$

$$\Lambda(z) = \frac{(1-z)^b - (1-bz)}{\left(1-(1-z)^b\right) z}, \quad z \in (0,1). \tag{D-24}$$

Inspection and an application of MLHR to the right-hand side of equation (D-23) shows that

$$M \text{ is positive and strictly increasing.} \tag{D-25}$$

The geometric convexity properties of Kumarasway all follow from the following result.

Result 8. If $b > 1$, then Λ is positive and strictly increasing; If $b < 1$, Λ is negative and strictly decreasing; If $b = 1$, Λ is identically 0.

Proof of Result. First note that $1-bz$ is the support line for $(1-z)^b$ at 0. If $b > 1$ then $(1-z)^b)$ is strictly convex in z. Thus, $(1-z)^b > 1-bz$. If $b < 1$ then $(1-z)^b$ is strictly concave in z. Thus, $(1-z)^b < 1-bz$. Since the denominator in the expression defining Λ is clearly positive, these observations imply that if $b > 1$, $\Lambda > 0$ and, if$b < 1$,$\Lambda > 0$. Now consider monotonicity. Evaluated at $z = 0$, both the numerator and denominator of Λ vanish and the derivative of the denominator in the expression defining Λ never changes sign. Thus, we can determine the direction of monotonicity using the MLHR. The ratio of the derivatives of the numerator and denominator in the expression defining Λ is given by

$$\text{DRatioKumar} = b\,(1+\lambda(z))^{-1}, \tag{D-26}$$

$$\lambda(z) = \frac{(1-z)^{b-1}\,z\,(1+b)}{1-(1-z)^{b-1}}. \tag{D-27}$$

Note that

$$\frac{\lambda'(z)}{\lambda(z)} = \frac{1}{z}\,\frac{1}{(1-z)\left(1-(1-z)^{b-1}\right)}\left(z\,(1-b) - \left((1-z)^b - (1-z)\right)\right). \tag{D-28}$$

I claim that if $b \neq 1$, $\lambda'(z)/\lambda(z) < 0$. Note that $z\,(1-b)$ is the support line for $(1-z)^b - (1-z)$ at $z = 0$. If $b > 1$, ($b < 1$) then $(1-z)^b - (1-z)$ is strictly convex (concave). Thus,

$$b > 1 \Rightarrow z\,(1-b) - ((1-z)^b - (1-z)) < 0 \text{ and } b < 1 \Rightarrow z\,(1-b) - ((1-z)^b - (1-z)) > 0. \tag{D-29}$$

Now consider the remaining part of the expression for λ'/λ. Inspection shows that

$$b > 1 \Rightarrow \frac{1}{z\left((1-z)\left(1-(1-z)^{b-1}\right)\right)} > 0 \text{ and } b < 1 \Rightarrow \frac{1}{z\left((1-z)\left(1-(1-z)^{b-1}\right)\right)} < 0. \tag{D-30}$$

Conditions (D-29) and (D-30) verify that $\lambda'/\lambda < 0$. Now consider λ'. $\lambda' = \lambda\,(\lambda'/\lambda)$. When $b > 1$, $\lambda < 0$, thus $\lambda' > 0$. Similarly, if $\beta < 1$, $\lambda' < 0$. Thus, when $b > 1$, λ is λ is strictly decreasing and when $b < 1$, λ is strictly increasing. Inspection of equation (D-26) shows that λ being strictly decreasing (increasing) implies that DRatioKumar is strictly increasing (decreasing). Thus, if $\beta > 1$, DRatioKumar is strictly increasing. Hence, The MLHR implies that the monotonicity of DRatioKumar controls the monotonicity of Λ. □

Given the result, the final demonstration is simple. The definition of R_χ given by equation (D-22), Equation (D-25), and Result 8, imply that

$$b > 1 \Rightarrow R_\chi'(y) > 0 \text{ and } b < 1 \Rightarrow R_\chi'(y) < 0.$$

Thus, for $b > 1$, χ is geometrically convex and for $b < 1$, χ is geometrically concave. This result implies by the argument given at the start of this section, that if $b > 1$ the stochastically dominant up-sized distribution is geometrically dominant and if $b < 1$ the up-sized distribution is geometrically dominated.

10 Lognormal

Suppose, without loss of generality, that the common log variance of the original and upscaled Lognormal equals 1. Let r_{LN} represent the reverse hazard rate of the standard Lognormal distribution and let r_{N} and h_{N} represent the reverse hazard rate and hazard rate respectively of the standard normal distribution. Then, the definitions of the Normal and Lognormal distributions imply that

$$r_{\mathrm{LN}}(x) = \frac{1}{x}\, r_{\mathrm{N}}(\log(x)), \quad x > 0.$$

The symmetry of the Normal distribution implies that

$$r_{\mathrm{N}}(\log(x)) = h_{\mathrm{N}}(-\log(x)), \quad x > 0.$$

Thus,

$$\frac{x\, r_{\mathrm{LN}}'(x)}{r_{\mathrm{LN}}(x)} = -\frac{h_{\mathrm{N}}'(-\log(x))}{h_{\mathrm{N}}(-\log(x))}. \tag{D-31}$$

First note that $x \hookrightarrow -\log(x)$ is strictly decreasing. Next, note that, as shown above in the analysis of the Normal distribution, h_{N} is strictly logconcave and thus $x \hookrightarrow h_{\mathrm{N}}'(x)/h_{\mathrm{N}}(x)$ is also strictly decreasing. Finally, note that $x \hookrightarrow -x$ is strictly decreasing. Thus, the right-hand side of (D-31) is strictly decreasing and, hence by Lemma 12, the upscaled distribution is strictly geometrically dominant.

11 Fréchet

Selection equivalence is quite easy to show using many of the tests developed in the manuscript. Perhaps the most transparent demonstration follows from noting that if G is standard $s=1$ Fréchet with shape parameter α and F is an s-upscaling of G then, the transform function, u, is given by $u(t)=t^S$, where $S=s^\alpha$. The function $t\hookrightarrow t^S$ is clearly geometrically linear. Thus, the s-upscaling is geometrically equivalent to the original distribution and hence, by Lemma 6, the distributions are selection equivalent.

12 Log Logistic

Consider the reverse hazard rate for Log-logistic distribution with shape parameter $\beta>1$ and scale parameter $\alpha=1$. The reverse hazard rate for this distribution is given by

$$r_{\text{LLogistic}(x)}=\frac{\beta}{x+x^{1+\beta}},\quad x>0.$$

Using Lemma 12 to test for geometric concavity we see that

$$\frac{x\,r'_{\text{LLogistic}}(t)}{r_{\text{LLogistic}}(t)}=\left(\frac{1}{1+x^\beta}-1\right)\beta-1. \tag{D-32}$$

This expression is clearly strictly decreasing and thus, $r_{\text{LLogistic}}$ is strictly geometrically concave. Hence, the upscaled distribution always strictly geometrically dominates the original distribution.

13 Gompertz

In this section let G represent the distribution function of a standard unit-scale Gompertz distribution, i.e.,

$$G(x)=1-\exp((1-e^x)\,\eta),\quad x>0.$$

Let r represent the reverse hazard rate of the unit-scaled Gompertz distribution,

$$r(x)=\frac{1}{\eta}\left(e^{-x}\left(\exp\left(\eta\,\left(e^x-1\right)\right)-1\right)\right),\quad x>0.$$

Suppose that F is an upscaling of G, i.e, $F(x)=G(x/s)$, $s>1$. As in the case of the Gamma we define $\theta=1/s$. Expressed in terms of θ, $F=G(\theta x)$, $\theta\in(0,1)$, and the reverse hazard rates of F and G are given by

$$r_F(x) = \frac{f(x)}{F(x)} = \theta\, r(\theta\, x); \quad r_G(x) = \frac{g(x}{G(x)} = r(x), \quad x > 0.$$

13.1 Gompertz distribution when $\eta \geq 1$

When the scale parameter for the Gompertz distribution equals 1, its distribution function is given by

$$G(x) = 1 - \exp\left(\eta\,(1 - e^x)\right), \quad x \in (0, \infty).$$

The reverse hazard rate for this distribution can be written as

$$\begin{aligned} r_{\text{Gompertz}}(x) &= \frac{\eta\,(1-\phi(x))\,(1-\phi(\eta\,\phi(x)))}{\phi(\eta\,\phi(x))}, \\ \phi(x) &= 1 - e^x. \end{aligned}$$

The test for Geometric concavity given in Lemma 12, requires that $x \hookrightarrow xr'(x)/r(x)$ be decreasing. Next note that

$$\begin{aligned} R_{\text{Gompertz}} = \frac{x\,r'_{\text{Gomperz}}(x)}{r_{\text{Gomperz}}} &= -\Lambda(\phi(x))\,M(\phi(x)), \\ \Lambda(y) &= e^{y\,\eta} - (1 + y\,\eta) + \eta, \quad y \leq 0, \\ M(y) &= \frac{\log(1-y)}{1 - e^{y\,\eta}}, \quad y \leq 0. \end{aligned} \tag{D-33}$$

Because ϕ is strictly decreasing in x, equation (D-33) shows that to prove that R_{Gompertz} is strictly decreasing and thus r_{Gomperz} is strictly geometrically concave we need only show that the map $y \hookrightarrow \Lambda(y)\,M(y)$ is strictly decreasing.

First we show that $M$is positive and is also strictly decreasing if $\eta \geq 1$. To see this, first note that because $y < 0$, M is clearly positive. Next note that both the numerator and denominator of the expression defining M vanish as $y \to 0$ and the derivative of denominator never changes signs, thus, by the MLHR, a sufficient condition for M to be (strictly) decreasing is that

$$y \hookrightarrow \frac{\partial_y \log(1-y)}{\partial_y\,(1 - e^{y\,\eta})} = \frac{e^{-y\,\eta}}{(1-y)\,\eta} \text{ is (strictly) decreasing.} \tag{D-34}$$

Simple differentiation shows that $e^{-y\,\eta}/((1-y)\,\eta)$ is strictly decreasing in y if and only if $1 - \eta\,(1-y) < 0, y < 0$. This condition is satisfied if and only if $\eta \geq 1$. Thus M is positive and strictly decreasing for all $y < 0$ if and only if $\eta \leq 1$.

Now consider Λ. Λ is strictly convex over $y \leq 0$ and at $y = 0$, the upper end point of its domain of definition, $(-\infty, 0]$, Λ and its derivative both vanish. Thus, Λ is strictly decreasing in y and is positive. Thus, the product $\Lambda(y)\,M(y)$ is strictly decreasing when $\eta \geq 1$. Thus, when $\eta \geq 1$, r_{Gompertz} is strictly geometrically concave. This implies that when $\eta \geq 1$, the upscaled distribution is is strictly geometrically dominant.

If on the other hand, $\eta < 1$, then explicit calculations show that

$$\lim_{y \to 0} (M(y)\,L(y))' = \frac{1-\eta}{2} > 0.$$

Thus, on a sufficiently small neighborhood of 0. $y \hookrightarrow M(y)\,L(y)$ is increasing. which implies that inverse hazard rate is not geometrically convex. Thus, when $\eta < 1$, the upscaled distribution is not geometrically dominant. We now turn characterizing selection dominance in this case.

13.2 Gompertz distribution when $\eta < 1$

Result 9. The Gompertz distribution function is log-concave and thus r_F and r_G are strictly decreasing.

Proof. The fact that the distribution function of the Gompertz distribution is strictly logconcave, implies that r is strictly decreasing and thus that r_F and r_G are strictly decreasing. The logconcavity of the Gompertz distribution is well known (Sengupta and Nanda, 1999). □

Result 10. The function ρ defined by

$$\rho(x) = \frac{r_F(x)}{r_G(x)}, \quad x > 0,$$

has the following properties.

(i) $\lim_{x \to 0} \rho(x) = 1$

(ii) $\lim_{x \to \infty} \rho(x) = \infty$

Proof. Explicit calculation and, in the case of (i), the use of L'Hospital's Rule. □

Result 11. The derivative of ρ verifies the equation

$$\rho'(x) = \rho(x)\left(\zeta(x) - (r_F(x) - r_G(x))\right), \quad x > 0, \tag{D-35}$$

$$\text{where } \zeta(x) = \frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}, \quad x > 0. \tag{D-36}$$

Proof. Differentiate using the product and quotient rules. □

Result 12. The function ζ defined in Result 11 has the following properties:

(i) $\zeta(0) = -(1-\theta)(1-\eta) < 0$ and $\lim_{x \to \infty} \zeta(x) = \infty$.

(ii) ζ is strictly increasing.

Proof. Explicit calculation in the Gompertz distribution case show that

$$\zeta(x) = e^x\left(\eta - (e^x)^{-(1-\theta)}\,\eta\,\theta\right) - (1-\theta).$$

The results then follow from calculus. □

Result 13. Let $x^o = \inf\{x > 0 : \rho(x) \geq 1\}$, then

(i) $\rho(x^o) = 1$.

(ii) $\zeta(x^o) > 0$.

Proof. Part (i) follows from the definition of x^o, the continuity of ρ, and Result 10. To prove part (ii), note that by the quotient rule for differentiation:

$$\left(\frac{F}{G}\right)' = \frac{g}{G}\left(\frac{f}{g} - \frac{F}{G}\right), \tag{D-37}$$

$$\left(\frac{f}{g}\right)' = \frac{g}{f}\zeta. \tag{D-38}$$

Let x^* be the unique 0 of ζ. Result 12 and (D-38), f/g is U-shaped, strictly decreasing for $x < x^*$ and strictly increasing for $x > x^*$. Evaluating f/g at 0 and applying L'Hopitals's rule to F/G at 0, shows that $\lim_{x\to 0} F(x)/G(x) = f(0)/g(0) = \theta < 1$. Thus, $F/G < 1$ and strictly decreasing whenever $x < x^*$. Hence, $x^* < x^o$, and thus by Result 12, $\zeta(x^o) > 0$. □

Result 14. Let $x^o = \inf\{x > 0 : \rho(x) \geq 1\}$, then

(i) If $x < x^o$, then $\rho(x) < 1$.

(ii) If $x > x^o$, then $\rho(x) > 1$.

Proof. Parts (i) simply follows from the definition of x^o and the continuity of ρ. The proof of part (ii) is essentially identical to the proof of Result 7. The only difference is that in the Gompertz case, x^o plays the role that 0 played in the analysis of the Gamma distribution. □

Result 15. Let $x^o = \inf\{x > 0 : \rho(x) \geq 1\}$, then if $x \geq x^o$, ρ is strictly increasing.

Proof. The proof of part (ii) is essentially identical to the proof of Result 7. The only difference is that in the Gompertz case, x^o plays the role that 0 played in the analysis of the Gamma distribution. □

Define the distribution function M_x over the positive real line by

$$M_x(z) = \begin{cases} \frac{\int_0^z g(w)F(w)\,dw}{\int_0^x g(w)F(w)\,dw} & \text{if } z \leq x \\ 1 & \text{if } z > x. \end{cases} \tag{D-39}$$

Next, note that we can write $\mathscr{R}$ (defined in equation (A-18)) as

$$\mathscr{R}(x) = \int_0^\infty \rho(z)\,dM_x(z), \tag{D-40}$$

where ρ is defined in Result 10. For $x > x^o$, let C_x represent the distribution of M_x conditioned on $z > x^o$, where x^o is defined in Result 14, i.e.,

$$C_x(z) = \begin{cases} 0 & \text{if } z < x^o \\ \frac{1}{1-M_x(x^o)}\left(M_x(z) - M_x(x^o)\right) & \text{if } x \geq x^o. \end{cases}$$

Let $\gamma(x)$ represent the conditional expectation of ρ under C_x, i.e.,

$$\gamma(x) = \mathbb{E}_x[\rho(\tilde{z})|\tilde{z} > x^o] = \int_{x^o}^{\infty} \rho(z)\, dC_x(z).$$

Results 14 and 15 show that if $z \geq x^o$, $\rho > 1$ and is strictly increasing. Note that the conditional measures, C_x are ordered by strict stochastic dominance, i.e., if $x' < x''$, then $C_{x'}$ is strictly stochastically dominated by $C_{x''}$. Therefore, for $x > x^o$,

$$1 < \gamma(x) \text{ and } \gamma \text{ is strictly increasing.} \tag{D-41}$$

Next note that using equation (D-40), we can write

$$\mathscr{R}(x) = \int_0^{\infty} \rho(z) dM_x(z) = \int_0^{x^o} \rho(z) dM_x(z) + (1 - M_x(x^o))\gamma(x). \tag{D-42}$$

Reverting to the definition of M_x given in equation (D-39) yields

$$\mathscr{R}(x) = \frac{\int_0^{x^o} \rho(z)\, g(z)\, F(z) dz + \gamma(x) \int_{x^o}^{x} g(z)\, F(z) dz}{\int_0^x g(z)\, F(z) dz}.$$

Differentiating the log of this expression yields,

$$(\log \circ \mathscr{R}(x))' = \frac{\gamma'(x) \int_0^x g(z)\, F(z) dz + \gamma(x)\, g(x)\, F(x)}{\int_0^{x^o} \rho(z)\, g(z)\, F(z) dz + \gamma(x) \int_{x^o}^{x} g(z)\, F(z) dz} - \frac{g(x)\, F(x)}{\int_0^{x^o} g(z)\, F(z) dz + \int_{x^o}^{x} g(z)\, F(z) dz}. \tag{D-43}$$

Next, divide the numerator and denominator of the first term on the right-hand side of (D-43) by γ to yield

$$(\log \circ \mathscr{R}(x))' = \frac{(\gamma'(x)/\gamma(x)) \int_0^x g(z)\, F(z) dz + g(x)\, F(x)}{\int_0^{x^o} (\rho(z)/\gamma(x))\, g(z)\, F(z) dz + \int_{x^o}^{x} g(z)\, F(z) dz} - \frac{g(x)\, F(x)}{\int_0^{x^o} g(z)\, F(z) dz + \int_{x^o}^{x} g(z)\, F(z) dz}. \tag{D-44}$$

By expression (D-41), $(\gamma'(x)/\gamma(x) \geq 0$. By Results 14.i, and expression (D-41), $\rho(z)/\gamma(x) < 1$ for $z \leq x^o$. Thus, the numerator of the first term on the right-hand side of equation (D-44) exceeds the numerator of the second term and the denominator of the first term is less than the denominator of the second term. Thus, the right-hand side of equation (D-44) is positive, which implies that for $x > x^o$, $\mathscr{R}'(x) > 0$.

Hence (a) $\mathscr{R}$ is strictly increasing for $x > x^o$. (b) For $x < x^o$, $\rho(x) < 1$, and thus for $x < x^o$, $\mathscr{R} < 1$. Because of the stochastic dominance of the upscaled Gompertz, (c) $\lim_{x \to \infty} \mathscr{R}(x) > 1$. Conditions (a), (b), and (c), imply that the supremum of $\mathscr{R}$ equals its limit as $x \to \infty$. Because for the Gompertz distribution, $\bar{x} = \infty$, Lemma 7 implies that the upscaled distribution CSSD dominates the original distribution.

References

BARICZ, Á. (2010): "Geometrically concave univariate distributions," *Journal of Mathematical Analysis and Applications*, 363(1), 182–196.

PINELIS, I. (2002): "L'Hospital Type rules for monotonicity, with applications," *J. Ineq. Pure & Appl. Math*, 3(1), Art. 5, 5 pp.

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

SENGUPTA, D., AND A. K. NANDA (1999): "Log-concave and concave distributions in reliability," *Naval Research Logistics (NRL)*, 46(4), 419–433.