

# Comparing the chosen: Selection bias when selection is competitive

Thomas Noe  
Balliol College/Saïd Business School  
University of Oxford\*

5 November, 2018

## Abstract

Consider a decision maker who selects between paired random draws from two unconditional distributions, always selecting the larger draw in the pair. When will the resulting selection-conditioned distributions be ordered by first-order stochastic or monotone likelihood-ratio dominance? In various guises, this question arises in many economic contexts—tournaments, contests, auctions, cheap-talk games, announcement returns, qualitative choice models, and treatment effects under self-selection. This paper develops simple, applicable characterizations of the properties of unconditional distributions which result in dominance conditioned on selection and uses these characterizations to analyze a number of economic selection problems.

**Key Words:** selection bias, stochastic dominance, all-pay auctions, tournaments, self-selection

I would like thank Alexei Boulatov and the participants in the Higher School of Economics research symposium for many helpful comments on a very early draft of this paper. In addition, I am indebted to participants in the Nuffield Economic Theory Workshop (Oxford), the Saïd Business School Finance Workshop, the Econometric Society, and AEA meetings. Special thanks are extended to Jeffrey Ely, Dawei Fang, Andy Gardner, Thomas Hellmann, Dunhong Jin, Todd Kaplan, Charles Manski, Debrah Noe, the editor, Emir Kamenica, and anonymous reviewers for the *Journal of Political Economy* for many useful comments. The usual disclaimer applies. Because a significant portion of the analysis in this paper was developed while I was a visiting fellow at the Bank of Finland, I am indebted to the bank for its support of my research.

---

\*Saïd Business School, Park End Street, Oxford OX1 1HP UK, email:thomas.noe@sbs.ox.ac.uk

# 1 Introduction

Consider a college admissions officer tasked with awarding one Apex Scholarship each year to the incoming freshmen most likely to have the highest cumulative GPA upon graduation. The college follows a gender-balanced admissions policy and thus admits an equal number of male and female students each year. University policy also mandates that gender should not factor into the award of the Apex. After many years, the university performs an internal audit of gender-balance in scholarship awards. The audit reveals a “troubling” pattern: (a) the vast majority of Apex Scholars were female, (b) the GPAs of male Apex Scholars were significantly higher.

Does this evidence suggest that the administrator is biased toward female students, only offering the Apex to male students when they obviously outshine all their female rivals? The answer to this question turns out to be not necessarily. Even if the administrator is completely unbiased and can perfectly predict the final GPA of admitted students, we could end up with this troubling pattern because of the effects of competitive selection.

This paper develops the logic underlying these sorts of competitive selection questions and provides tools for resolving them: necessary and sufficient conditions on compared unconditional distributions which guarantee that, under unbiased competitive selection, the selected draws from one distribution will dominate selected draws from another. One of these conditions, *geometric dominance* is simple and easy to check.

We view competitive selection as choosing between prospects. Prospects are random variables. Our baseline analysis, extended in Section 7 to multiple prospects and dependent prospects, focuses on comparisons between two statistically independent prospects,  $\tilde{X}$  and  $\tilde{Y}$ . Competitive selection is defined as making the optimal choice between the two prospects, i.e., selecting  $\tilde{X}$  ( $\tilde{Y}$ ) whenever  $\tilde{X} > \tilde{Y}$  ( $\tilde{Y} > \tilde{X}$ ). Selection results in competitive *selection-conditioned distributions* of  $\tilde{X}$  and  $\tilde{Y}$ , represented by  $[\tilde{X} | \tilde{X} > \tilde{Y}]$  and  $[\tilde{Y} | \tilde{Y} > \tilde{X}]$  respectively.

Our analysis will focus primarily on the relation between unconditional superiority and superiority conditioned on competitive selection.<sup>1</sup> In economics research, by far the most ubiquitous criteria for measuring superiority are first-order stochastic dominance (FSD) and monotone likelihood ratio (MLR) orders.<sup>2</sup> Thus, our formal analysis will measure superiority, i.e., dominance, using these two orders. When one distribution,  $F$ , is larger (smaller) than another,  $G$ , under the FSD (MLR) order, we will state that  $F$  FSD (MLR) dominates (is dominated by)

<sup>1</sup>Since, in this paper, we only analyze competitive selection, except when we explicitly discuss of other forms of selection, we will typically simply use “selection” to denote competitive selection.

<sup>2</sup>Given distribution functions  $F_1$  and  $F_2$ ,  $F_1$  is larger than  $F_2$  under the FSD order if  $F_1(x) \leq F_2(x)$  for all  $x \in (-\infty, \infty)$ .  $F_1$  is larger than  $F_2$  under the MLR order if the ratio between the distributions’ density functions,  $f_1/f_2$  increases over the union of the interiors of the supports of  $F_1$  and  $F_2$  (Chapt. 1: Shaked and Shanthikumar, 2007). Surveys of the economic applications of MLR and FSD orders are provided by Jewitt (1991) and Levy (1992).

$G$ .<sup>3</sup>

Our aim is to identify the conditions under which the transformation of the prospects' distributions effected by selection preserves or reverses stochastic order relations. Dominance is *preserved* if, under a given stochastic order relation (FSD or MLR), the unconditionally dominant prospect's selection-conditioned distribution is also (FSD or MLR) dominant. Dominance is *reversed* if the unconditionally dominant prospect's selection-conditioned distribution is dominated.

Questions related to dominance preservation and reversal arise with surprising frequency in both empirical and theoretical economics research. Sometimes, the prospects represent populations with different distributions of some desired attribute, as in the scholarship example above. Sometimes, the prospects are alternative options for a single decision maker whose utility depends on the option with the highest realized value (e.g., self-selection bias and treatment effects). Sometimes, the prospects represent the performance strategies of ex ante asymmetric agents competing for selection (e.g., the winner's bid or effort level in asymmetric auction, contest, and tournament models).

An example in Section 2 illustrates the fundamental driver of dominance reversal—the dependence of each prospect's selection-conditioned distribution on the characteristics of the compared prospect's distribution. If prospects faced a common selection bar, say  $k$ , the sort of dominance reversals we consider could be ruled out by standard stochastic order conditions. For example,  $\tilde{X}$  MLR dominating  $\tilde{Y}$  implies that  $[\tilde{X} | \tilde{X} > k]$  MLR dominates and (*a fortiori*) FSD dominates  $[\tilde{Y} | \tilde{Y} > k]$ .<sup>4</sup>

However, the selection bar under competitive selection is different for each prospect: the draws from one prospect's distribution, say, for example,  $\tilde{X}$ 's distribution, are admitted into  $\tilde{X}$ 's selection-conditioned sample if and only if  $\tilde{X} > \tilde{Y}$ , i.e.,  $\tilde{Y}$  is the selection bar for  $\tilde{X}$  and vice versa. The dependence of the selection bar for each prospect on the character of the rival prospect's distribution produces an effect, which we term the *admission effect*, that favors the selection-conditioned dominance of unconditionally dominated prospects: the low selection bar provided by the unconditionally dominated prospect admits low draws into the unconditionally dominant prospect's selection-conditioned distribution.<sup>5</sup>

Because of this effect, the analysis of dominance under competitive selection requires developing new tools. The framework for this development is provided by Section 3. In Section 4, we develop two *competitive selection relations*: *average elasticity dominance* and *geometric dominance*. Both of these relations impose conditions on the unconditional distributions of

<sup>3</sup>In addition, if  $F_1$  is the distribution of  $\tilde{X}_1$  and  $F_2$  is the distribution of  $\tilde{X}_2$  and  $F_1$  is larger than  $F_2$  in the FSD (MLR) order we will say that  $\tilde{X}_1$  FSD (MLR) dominates  $\tilde{X}_2$ .

<sup>4</sup>See Theorem 1.C.7 in Shaked and Shanthikumar (2007).

<sup>5</sup>The idea that lower draws from the compared prospect's distribution reduce selection-conditioned quality can be formalized as follows: if  $y''$  and  $y'$ , are two draws from prospect  $\tilde{Y}$ , then  $y'' > y'$  implies that  $[\tilde{X} | \tilde{X} > y'']$ , FSD dominates  $[\tilde{X} | \tilde{X} > y']$ .

the compared prospects which are necessary and sufficient for dominance relations between their selection-conditioned distributions: average elasticity dominance characterizes FSD dominance conditioned on selection and geometric dominance characterizes MLR dominance conditioned on selection.

Average elasticity dominance is not a very tractable relation. However, geometric dominance is both tractable and easy to verify, as it simply requires the ratio between the reversed hazard rates of the two compared distributions,  $f/F$  and  $g/G$ , to be increasing.<sup>6</sup> Geometric dominance is neither necessary nor sufficient for MLR dominance. However, as shown in Section 6.2, most standard parametric distributions (Normal, Lognormal, Exponential, Pareto, etc.) are ordered consistently by MLR and geometric dominance, which implies that, for these distributions, selection preserves MLR dominance.

In Section 5, we use these competitive selection relations to characterize dominance reversal and preservation. In order for selection to reverse FSD dominance, the left-tail ratio between the FSD dominated and dominant distributions, which measures the relative likelihood of a draw less than  $x$ , must be unbounded as  $x$  converges to the distributions' lower bound (perhaps  $-\infty$ ). This result shows that the admission effect must be quite strong for selection to reverse FSD dominance. An asymptotically infinite tail ratio implies that the lowest draws from the prospects' distributions almost surely come from the FSD dominated prospect. In this case, because the lowest draws are almost never selected, they do not affect the selection-conditioned distribution of the FSD dominated prospect, but, by providing a very low bar for selection, significantly lower the quality of the selected draws from the FSD dominant prospect's distribution.

If FSD dominance is preserved by selection, geometric dominance implies MLR dominance. This result is a direct consequence of the admission effect. Because of the admission effect, conditioning on selection can never "strengthen" the dominance of an unconditionally dominant prospect. The MLR order is a stronger order than the FSD order. Consequently, if a prospect is FSD dominant but not MLR dominant, conditioning on selection cannot "make" the prospect's selection-conditioned distribution MLR dominant or, equivalently, the prospect's distribution cannot be geometrically dominant. There is no analogous result for dominance reversal: a FSD dominated but geometrically dominant prospect may or may not be unconditionally MLR dominated.

These results address the question of how unconditional prospect distributions affect dominance relations between selection-conditioned distributions. We complete the baseline analysis

---

<sup>6</sup>In this paper, a function,  $f$ , defined over a subset of the real line,  $D$ , is called *increasing* if, for all  $x, y \in D$ ,  $y > x$  implies that  $f(y) \geq f(x)$ . A function is called *strictly increasing* if  $y > x$  implies that  $f(y) > f(x)$ . This usage is adopted for the sake of consistency with the definitions of "increasing" and "strictly increasing" used by the authors of the mathematical theorems cited in this paper. Other authors apply the term "nondecreasing" to the functions we call "increasing," and "increasing" to the functions we call "strictly increasing."

by considering the converse question: the restrictions that the selection-conditioned prospect distributions impose on unconditional prospect distributions. We show that, using only selection-conditioned information, it is possible to identify dominance preservation and reversal.

The economic relevance of these competitive selection relations is illustrated by applying them to a sample of economic selection problems. The first problem we consider is a hiring competition in which applicants, drawn from groups with differing abilities, compete for employment. Selection is determined by an all-pay auction, effort-bidding mechanism (e.g., Baye, Kovenock, and de Vries, 1996).

In this setting, we show that FSD dominance is reversed by selection. Thus, if employers are unbiased, and thus select purely on the basis of applicant ability, the performance distribution of hired (i.e., selected) applicants from the inferior (i.e., FSD dominated) applicant pool will FSD dominate the performance distribution of hired applicants drawn from the superior applicant pool. At the same time, because, under meritocratic selection, applicants from the inferior pool are less likely to be hired, they will be underrepresented in the firm's workforce. Thus, under the effort-bidding mechanism, members of one group being underrepresented yet, on average, performing better on the job, is consistent with merit-based hiring.

Our next application considers the same hiring competition but changes the selection mechanism to a Lazear and Rosen (1981) tournament. Under the tournament mechanism, FSD dominance is preserved by selection. For this reason, under the tournament mechanism, selected job candidates from the better applicant pool perform better on the job. Thus, underrepresentation of a better performing group is not consistent with merit-based hiring policies.

Like MLR ordering, geometric dominance ordering can also be used to ensure that the predictions of theoretical models conform with intuition. We illustrate this role in the context of Che, Dessein, and Kartik (2013) model of costless communication in which an informed self-interested agent makes project recommendations to a less informed principal. We show that, when projects are ordered by geometric dominance, the bias in agent recommendations identified in Che, Dessein, and Kartik (2013) always takes the form of pandering to the principal's prejudices, i.e., recommending the project the less informed principal believed was best before receiving the agent's advice.

Geometric dominance can also be applied to competitive selection problems when selection is produced by optimizing policy choices of decision makers. We illustrate such applications in Section 6.2 by applying geometric dominance to the extensive body of empirical research that documents market reactions to corporate policy announcements. We identify parametric error laws that permit valid inferences of the unconditional effects of corporate policies from market reactions to the announcement (i.e., selection) of these policies.

To further illustrate the usefulness of the competitive selection relations, we develop four more applications in the supplement. The first application, Supplement C.1, illustrates the

computational simplification produced by using geometric dominance to identify dominance conditioned on selection. The second application, Supplement C.2, illustrates the usefulness of geometric dominance for signing comparative statics by providing a geometric dominance condition which ensures that strong bidders pay more than weak bidders in a Maskin and Riley (2000) asymmetric, private value, first-price auction.

The third application, Supplement C.3, uses geometric dominance to show that the Probit model of binary choice imposes a strong restriction on the relation between utility and choice: a perfect positive association between the propensity to choose an action and the selection-conditioned expected utility of the action. This result complements the Anas and Feng (1988) and de Palma and Kilani (2007) result that, in the Logit framework, expected selection-conditioned utility is invariant to the propensity to select an action.

The fourth application, Supplement C.4, considers the problem of identifying the sign of average treatment effects under self-selection (Manski, 1990; Heckman, 2008). Geometric dominance is used to specify the error laws under which the sign of the average treatment effect can be identified by the sign of the treatment effect in a self-selected sample.

Section 7 considers the robustness of the baseline analysis to non-binary choice menus and dependent prospects. Here we show that geometric dominance is robust to non-binary choice menus: pairwise geometric dominance relations characterize MLR dominance conditioned on optimizing selection.

The effect of dependence on competitive selection dominance is a more challenging question. When prospects are dependent, the competitive selection relations (average elasticity and geometric dominance) between unconditional prospect distributions are not equivalent to dominance (FSD or MLR) conditioned on selection. However, when dependence is produced by a common shock, the competitive selection relations imply selection-conditioned FSD dominance. For non-linear copula dependence, a simple sufficient condition, satisfied by copulas commonly used in applied research, ensures that geometric dominance implies selection-conditioned MLR dominance.

The results in this paper are new in the sense that, to our knowledge, the question of relating unconditional stochastic order to selection-conditioned stochastic order has not been addressed. The closest “relative” to our research project is probably Che, Dessein, and Kartik (2013) discussed above. Our results for Extreme Value Type I distributions are related to the analysis in Anas and Feng (1988) and de Palma and Kilani (2007). More generally, the paper is part of a vast research program devoted to using stochastic orders to sign comparative static relations (Quah and Strulovici, 2009; Athey, 2002; Milgrom, 1981, *inter alia*), and to mathematical research related to geometric convexity, (e.g., Niculescu and Persson, 2004).

## 2 Example: Worse is better, conditioned on being selected

We start with a simple example that abstracts from all the complications involved in our more general analysis. Instead of defining superior and inferior in terms of stochastic orders, in this example, we simply define superior and inferior in terms of expectations. We choose the simplest possible prospect distributions. Despite the stripped-down setting, this example reveals the essence of selection's effect on superiority relations between competing prospects.

Consider the problem of a firm that staffs its management positions with students from two universities: University X and University Y. Graduates of the universities come in two flavors: *viable* graduates and *ringers*. The quality of viable graduates is uniformly distributed between 0 and 1. In contrast, ringers resemble university graduates in outward manner and appearance but have zero quality with probability 1. All graduates of University Y, Y-candidates, are viable candidates. University X is less selective: a fraction  $\rho \in (0, 1)$  of X-candidates are ringers and the remaining fraction,  $1 - \rho$ , are viable. Knowing the quality of an X-candidate provides no information about the quality of a Y-candidate and vice versa.

The firm follows the following employee-selection procedure: when a position opens up, the firm interviews one candidate from each university and hires one of them. The interviewer can perfectly discern candidate quality. After a large number of positions have been filled with this procedure, management notices the following “anomaly:”

- (a) Y-candidates are more likely to be hired than X-candidates;
- (b) X-candidates perform better, on average, on the job.

The anomaly raises the question of whether interviewers are actually making meritocratic hiring decisions, i.e., selecting the highest quality candidates.

Formalizing the question, we see that the quality of X-candidates,  $\tilde{X}$ , and the quality of Y-candidates,  $\tilde{Y}$ , are given as follows:  $\tilde{X}$  and  $\tilde{Y}$  are independent random variables with distributions

$$\tilde{X} \stackrel{\text{dist}}{=} \begin{cases} 0 & \text{with probability } \rho \\ \tilde{U}_X & \text{with probability } 1 - \rho \end{cases}, \quad \tilde{Y} = \tilde{U}_Y,$$

where  $\tilde{U}_X$  and  $\tilde{U}_Y$  are independent random variables uniformly distributed over the interval  $[0, 1]$ .

The anomaly described by (a) and (b) can formally can be expressed as follows:

$$(a) \mathbb{P}[\tilde{Y} \text{ is selected}] > \mathbb{P}[\tilde{X} > \text{ is selected}]; \quad (b) \mathbb{E}[\tilde{X} | \tilde{X} \text{ is selected}] > \mathbb{E}[\tilde{Y} | \tilde{Y} \text{ is selected}]. \quad (1)$$

Perhaps surprisingly, under the assumed distributions of candidate quality, equation (1) *must* hold under meritocratic hiring policies. If hiring is meritocratic, then the candidate with the highest quality is always selected. An X-candidate will thus be selected if and only if  $\tilde{X} > \tilde{Y}$

and a Y-candidate selected if and only if  $\tilde{Y} > \tilde{X}$ .<sup>7</sup> Thus, if hiring is meritocratic, conditions (a) and (b) of equation (1) imply that

$$(a) \mathbb{P}[\tilde{Y} > \tilde{X}] > \mathbb{P}[\tilde{X} > \tilde{Y}]; \quad (b) \mathbb{E}[\tilde{X} | \tilde{X} > \tilde{Y}] > \mathbb{E}[\tilde{Y} | \tilde{Y} > \tilde{X}]. \quad (2)$$

Condition (a) of equation (2) holds: Y-candidates are always selected when their X-rival is a ringer and have the same chance of being selected when their X-rival is viable. Thus, overall, they have a higher probability of being selected. It is also easy to see that the expected quality of X-candidates is lower than Y-candidates. In fact, the distribution of X-candidates is FSD dominated by the distribution of Y-candidates.

Now consider condition (b) of equation (2). If a Y-candidate is paired with a ringer, the Y-candidate will always have higher quality than the zero-quality ringer. Consequently, the Y-candidate will always be selected. This implies that the expected quality of Y candidates paired with ringers equals their unconditional expected quality,  $1/2$ . If a Y-candidate is paired with a viable X-candidate, then, in order to be selected, the Y-candidate's quality must top the X-candidate's. Thus, when the X-candidate is viable, because the quality of viable candidates is distributed Uniform[0, 1], the expected quality of a selected Y-candidate equals

$$\mathbb{E}[\tilde{U}_Y | \tilde{U}_Y > \tilde{U}_X] = \mathbb{E}[\tilde{U}_Y | \tilde{U}_Y = \max[\tilde{U}_X, \tilde{U}_Y]] = \mathbb{E}[\max[\tilde{U}_X, \tilde{U}_Y]] = \frac{2}{3},$$

where the second-to-last equality results because  $\tilde{U}_X$  and  $\tilde{U}_Y$  are i.i.d., and the last inequality from the fact that  $\tilde{U}_X$  and  $\tilde{U}_Y$  are Uniform[0, 1] i.i.d. random variables.<sup>8</sup>

Because all Y-candidates are viable, and thus have higher quality than ringers, selected X-candidates are always viable. Thus, in order to be selected, an X-candidate must be viable and top a viable Y candidate. Therefore, for the same reasons as noted above for Y-candidates, the average quality of selected X-candidates is  $2/3$ . Using these observations we compute, in Table 1, the selection-conditioned expected quality of the graduates of the two universities.

As Table 1 shows, for all  $\rho \in (0, 1)$ , condition (b) of equation (2) is also satisfied. Consequently, X-candidates are clearly lower quality on average than Y-candidates; yet, conditioned on being selected, they are, on average, higher quality.<sup>9</sup> An alternative interpretation of this result is that, if a risk-neutral Bayesian CEO learned that the interview resulted in hiring a graduate of the less selective university, X, the CEO's utility would be higher than if CEO learned that a candidate from the more selective university, Y, had been hired.

In this example, the reversal of superiority results from the difference between competitive and fixed-threshold selection. If candidates were selected based on a fixed performance thresh-

<sup>7</sup>Under the assumed distributions, the probability of two candidates having the same quality is zero. So we ignore the case of tied quality.

<sup>8</sup>Equation 1.7.3 in Reiss (1989).

<sup>9</sup>In fact, the selection-conditioned distribution of X-candidates FSD dominates the selection-conditioned distribution of Y-candidates. It is possible to demonstrate this by direct computation. However, as this result follows directly from the subsequent analysis, Proposition 2, we will not provide a demonstration here.



		Y-candidate			X-Candidate		
		X-candidate is		sum	X-candidate is		sum
		ringer	viable		ringer	viable	
(1)	Probability	$\rho$	$1 - \rho$	--	$\rho$	$1 - \rho$	--
(2)	Exp. quality of selected	$\frac{1}{2}$	$\frac{2}{3}$	--	--	$\frac{2}{3}$	--
(3)	Prob[Selection]	1	$\frac{1}{2}$	--	0	$\frac{1}{2}$	--
(4)	$(1) \times (2) \times (3)$	$\rho \frac{1}{2}$	$(1 - \rho) \frac{1}{3}$	$(2 + \rho) \frac{1}{6}$ $= \mathbb{E}[\tilde{Y} I_{\tilde{Y} > \tilde{X}}]$	0	$(1 - \rho) \frac{1}{3}$	$(1 - \rho) \frac{1}{3}$ $= \mathbb{E}[\tilde{X} I_{\tilde{X} > \tilde{Y}}]$
(5)	$(1) \times (3)$	$\rho$	$(1 - \rho) \frac{1}{2}$	$\rho + (1 + \rho) \frac{1}{2}$ $= \mathbb{P}[\tilde{Y} > \tilde{X}]$	0	$(1 - \rho) \frac{1}{2}$	$(1 - \rho) \frac{1}{2}$ $= \mathbb{P}[\tilde{X} > \tilde{Y}]$
				$\mathbb{E}[\tilde{Y}   \tilde{Y} > \tilde{X}] = \frac{2 + \rho}{3 + 3\rho}$			
					$\mathbb{E}[\tilde{X}   \tilde{X} > \tilde{Y}] = \frac{2}{3}$		

Table 1: *Expected candidate quality conditioned on selection.* In the table,  $\mathbb{E}$  represents expectation,  $\mathbb{P}$  represents probability, and  $I_{\tilde{Y} > \tilde{X}}$  ( $I_{\tilde{X} > \tilde{Y}}$ ) is the indicator function for the event that  $\tilde{Y} > \tilde{X}$  ( $\tilde{X} > \tilde{Y}$ ).  $\mathbb{E}[\tilde{X} | \tilde{X} > \tilde{Y}] = \mathbb{E}[\tilde{X} I_{\tilde{X} > \tilde{Y}}] / \mathbb{P}[\tilde{X} > \tilde{Y}]$  and  $\mathbb{E}[\tilde{Y} | \tilde{Y} > \tilde{X}] = \mathbb{E}[\tilde{Y} I_{\tilde{Y} > \tilde{X}}] / \mathbb{P}[\tilde{Y} > \tilde{X}]$ .

old, the expected quality of a selected candidate from one university would not depend on the quality of their rivals from the other university. The proportion of X-candidates who are ringers would not effect the selection-conditioned quality of either X or Y-candidates as ringers would always be weeded out by selection and weeded out candidates would make no contribution to selection-conditioned expectations.

Under competitive selection, in contrast, losers in competition for selection set the “admission bar” for winners. The lower the bar, the lower the average quality of selected winners. Thus, under competitive selection, in contrast to fixed-threshold selection, very low draws from the inferior prospect’s distribution, ringers in this example, lower the superior prospect’s selection-conditioned quality. At the same time, because the very low draws are never selected, they do not lower the inferior prospect’s selection-conditioned quality.

The admission effect favors selection-conditioned superiority of the, on average, inferior X-candidates. In order to overcome the admission effect, the best Y candidates must top the best X candidates with sufficient probability. However, in the example, this is not possible because, conditioned on quality exceeding 0, the candidate quality distributions are identical.

This example suggests that if the worst draws from an inferior prospect’s distribution are much worse than the worst draws from the superior prospect’s distribution, but, as the focus of comparisons moves from comparisons between the worst draws to comparisons between the best, the disadvantage of the inferior prospect rapidly falls, the inferior prospect can be superior conditioned on selection. In fact, as our later results confirm, this very rough, imprecise intuition captures the essence of dominance reversal for quite general prospect distributions and for more sophisticated notions of inferior and superior.

### 3 The problem

So far, we have simply provided an example where the unconditionally inferior prospect's distribution is superior conditioned on selection. However, the compared prospects were quite stylized and the notion of superiority, expectational dominance, was quite crude. To what extent can the insights from this stylized example be extended to provide a general framework for analyzing the effects of selection on superiority?

In our formalization, prospects are random variables, which, in the baseline binary-comparison case, are represented by  $\tilde{X}$  and  $\tilde{Y}$ . In the baseline analysis, we assume that prospects are independent. Thus, we can identify each prospect with its distribution function. We represent the distribution function of  $\tilde{X}$  with  $F$  and the distribution function of  $\tilde{Y}$  with  $G$ , and write  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$ . In Sections 7.1 and 7.2, we will extend the analysis to consider non-binary comparisons and dependence.

We define “superiority” and “inferiority” in terms of first-order stochastic (FSD) and monotone likelihood ratio (MLR) dominance. By “conditioned on selection” we mean conditioning on a prospect's realization exceeding the realization of the compared prospect. Conditioning on selection transforms both prospects' distributions into *selection-conditioned distributions*.

$$(F, G) \xrightarrow{\text{selection}} (F^{\text{sc}}, G^{\text{sc}}),$$

where  $F^{\text{sc}}$  and  $G^{\text{sc}}$  represent the selection-conditioned distribution of  $\tilde{X}$  and  $\tilde{Y}$  respectively, i.e.,

$$F^{\text{sc}}(t) = \mathbb{P}[\tilde{X} \leq t | \tilde{X} > \tilde{Y}] \quad \text{and} \quad G^{\text{sc}}(t) = \mathbb{P}[\tilde{Y} \leq t | \tilde{Y} > \tilde{X}].$$

To simplify the discussion, we will refer to dominance (either FSD or MLR) relations between the selection-conditioned prospects as *competitive selection dominance*. Thus, if the selection-conditioned distribution of  $\tilde{X}$ ,  $F^{\text{sc}}$ , FSD dominates the competitive selection-conditioned distribution of  $\tilde{Y}$ ,  $G^{\text{sc}}$ , we will say that  $\tilde{X}$  *competitive FSD dominates*  $\tilde{Y}$ . Similarly, if  $F^{\text{sc}}$  MLR dominates  $G^{\text{sc}}$ , we will say that  $\tilde{X}$  *competitive MLR dominates*  $\tilde{Y}$ . References to dominance relations in a stochastic order (FSD or MLR) without qualifications related to competitiveness always refer to unconditional dominance.

If  $\tilde{X}$  FSD (MLR) dominates  $\tilde{Y}$  and  $\tilde{X}$  competitive FSD (MLR) dominates  $\tilde{Y}$ , we will say that selection *preserves* FSD (MLR) dominance. If  $\tilde{X}$  FSD (MLR) dominates  $\tilde{Y}$  but  $\tilde{Y}$  competitive FSD (MLR) dominates  $\tilde{X}$ , we will say that selection *reverses* FSD (MLR) dominance. Our plan of attack is to first identify competitive selection relations and then use these relations to characterize dominance preservation and reversal. A *competitive selection relation*,  $R^{\text{cs}}$ , is a relation between the unconditional prospect distributions,  $F$  and  $G$ , which characterizes competitive FSD or MLR dominance. Thus, if  $\succeq$  represents a stochastic order relation, either FSD or MLR, a competitive selection relation is a relation,  $R^{\text{cs}}$  with the property that

$$F^{\text{sc}} \succeq G^{\text{sc}} \iff F R^{\text{cs}} G. \quad (3)$$

When  $F R^{\text{cs}} G$  we will say that  $F$  dominates  $G$  under the given competitive selection relation. When  $\succeq$  represents FSD dominance, we call the competitive selection relation we identify *average elasticity dominance*. When  $\succeq$  represents MLR dominance, we call the competitive selection relation we identify *geometric dominance*. Statements about competitive FSD and MLR dominance are equivalent to statements about dominance under the competitive selection relations (average elasticity dominance and geometric dominance). For example, the assertion that  $\tilde{X}$  FSD dominates  $\tilde{Y}$ , but  $\tilde{Y}$  competitive FSD dominates  $\tilde{X}$  is equivalent to the assertion that  $\tilde{X}$  FSD dominates  $\tilde{Y}$  but  $G$  average elasticity dominates  $F$ .<sup>10</sup>

As well as outlining the logical structure of our problem, we also need to impose some mathematical structure. We aim to impose restrictions on the marginal distributions,  $F$  and  $G$  sufficient to ensure tractability, and, at the same time, avoid excluding economic applications. To this end, we assume that the supports of  $F$  and  $G$  are, possibly unbounded, intervals and that  $G$  is continuous. The continuity of  $G$  rules out the problem of ties, i.e., ensures that  $\mathbb{P}[\tilde{X} = \tilde{Y}] = 0$ . The assumption of interval supports is a simplifying assumption which, given the assumption of continuity, does not significantly restrict the range of application but does greatly simplify the presentation of the results. All derivations supporting our results are deferred to the Supplement to this paper.

## 4 Competitive selection dominance

### 4.1 Average elasticity dominance and competitive FSD dominance

We define competitive FSD dominance by applying the standard definition of FSD dominance to the selection-conditioned distributions of  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$ .

**Definition 1** (Competitive FSD dominance). Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are an admissible pair of prospects.  $\tilde{X}$  competitive FSD dominates  $\tilde{Y}$  if

$$F^{\text{sc}}(x) \leq G^{\text{sc}}(x), \quad x \in (-\infty, \infty). \quad (4)$$

When the inequality in (4) is strictly satisfied for all  $x$  in the interior of the union of the supports of the compared distributions we will say that  $\tilde{X}$  strictly competitive FSD dominates  $\tilde{Y}$ .<sup>11</sup>

<sup>10</sup>Recall that the baseline analysis assumes that the prospects are independent. Absent independence, this equivalence, in general, fails. In Section 7.2, we consider the effects of relaxing the independence assumption.

<sup>11</sup>Admittedly this is a slightly unorthodox use of the term “strict” with respect to first-order stochastic dominance. Strict FSD dominance is typically defined as the FSD condition, expression (4), being satisfied always with strict inequality holding at some point. However, when we develop our most important ordering relation, geometric dominance, it will be defined based on a convexity relation. The conventional definition of strict convexity requires that, on the interior of a function’s support, the convexity inequality is always strictly satisfied, not that the convexity inequality is weakly satisfied at all points and strictly at one. Rather than introduce yet more new terminology, e.g., “strong first-order stochastic dominance,” we choose to define strict first-order stochastic dominance analogously with strict convexity.

Because ties occur with probability 0, the selection-conditioned distributions of the two prospects can be defined as follows:

$$\begin{aligned} F^{\text{sc}}(x) &= \mathbb{P}[\tilde{X} \leq x | \tilde{X} > \tilde{Y}] = \mathbb{P}[\tilde{X} \leq x | \tilde{X} \geq \tilde{Y}], \\ G^{\text{sc}}(x) &= \mathbb{P}[\tilde{Y} \leq x | \tilde{Y} > \tilde{X}] = \mathbb{P}[\tilde{Y} \leq x | \tilde{Y} \geq \tilde{X}]. \end{aligned}$$

Computing these conditional distributions yields

$$\begin{aligned} F^{\text{sc}}(x) &= \frac{\int_{-\infty}^x \mathbb{P}[z \geq \tilde{Y}] dF(z)}{\int_{-\infty}^{\infty} \mathbb{P}[z \geq \tilde{Y}] dF(z)} = \frac{\int_{-\infty}^x G(z) dF(z)}{\int_{-\infty}^{\infty} G(z) dF(z)} = \frac{\int_{-\infty}^x G(z) dF(z)}{\mathbb{P}[\tilde{X} > \tilde{Y}]}, \\ G^{\text{sc}}(y) &= \frac{\int_{-\infty}^y \mathbb{P}[z \geq \tilde{X}] dG(z)}{\int_{-\infty}^{\infty} \mathbb{P}[z \geq \tilde{X}] dG(z)} = \frac{\int_{-\infty}^y F(z) dG(z)}{\int_{-\infty}^{\infty} F(z) dG(z)} = \frac{\int_{-\infty}^y F(z) dG(z)}{\mathbb{P}[\tilde{Y} > \tilde{X}]}.^{12} \end{aligned} \quad (5)$$

*Remark 1.* Note that, if the maximum of the support of  $G$  is less than the maximum of the support of  $F$ , for  $x$  between the two maxima,  $F^{\text{sc}}(x) < 1$  and  $G^{\text{sc}}(x) = 1$ . Thus,  $G$  cannot competitive FSD dominate  $F$ . Note also that if the minimum of the support of  $F$  is strictly less than the minimum of the support of  $G$ , then draws from  $F$  between these minima are never selected. Thus, the selection-conditioned distributions of  $F$  and  $G$  are the same as they would have been had the probability mass between the minima of the two distributions been moved to the lower endpoint of the support of the  $F$  distribution.

Thus, assuming a common support for the two distributions entails little loss of generality and, henceforth, we restrict attention to distributions satisfying this condition:

**Definition 2** (Admissible distributions). Distribution functions  $F$  and  $G$  form an *admissible pair of distributions* if

- (i)  $F$  and  $G$  have common support  $[\underline{x}, \bar{x}]$ ,  $-\infty \leq \underline{x} < \bar{x} \leq \infty$ .
- (ii)  $G$  is continuous and  $F$  is continuous except perhaps at  $\underline{x}$ .

If  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  and  $F$  and  $G$  are an admissible pair of distributions, then  $\tilde{X}$  and  $\tilde{Y}$  are an *admissible pair of prospects*.

Our aim is to develop a competitive selection relation defined in terms of the prospects' unconditional distributions that characterizes competitive FSD dominance, a relation between the prospects' selection-conditioned distributions. The problem with using (5) to develop this relation is that it contains “too much” information about the compared distributions. First-order stochastic dominance between the selection-conditioned distributions depends only on the relative performance of the two compared distributions. The selection-conditioned distribution functions impound properties of the compared distributions, such as length of their common support, which is not relevant to competitive selection dominance. To abstract from these extraneous properties we develop the competitive selection relations in terms of a function,  $u$ , which we call the *quantile transform function*, or simply the *transform function*, that relates the

<sup>12</sup>A detailed derivation of these expressions is provided in Supplement A.2.

quantiles of the compared distributions.

The quantile transform function,  $u$ , is defined as follows: because  $F$  and  $G$  are an admissible pair,  $G$  is continuous and strictly increasing over its support. Thus,  $G$  has a well-defined, continuous and strictly increasing inverse function. We represent this function with  $G^{-1}$ . The quantile transform function is defined by  $u(t) = F(G^{-1}(t))$ ,  $t \in (0, 1)$ . Note that  $u(t)$  represents the probability that  $\tilde{X}$  is less than the  $t$ -th quantile of the distribution of  $\tilde{Y}$ . Also note that  $F = u \circ G$  (here, “ $\circ$ ” represents functional composition) and thus  $u$  transforms  $G$  into  $F$ . Because  $u$  is continuous and strictly increasing on  $(0, 1)$  and  $\lim_{t \uparrow 1} u(t) = 1$  and  $\lim_{t \downarrow 0} u(t) = F(\underline{x})$ , we extend the definition of  $u$  by defining  $u(1) = 1$  and  $u(0) = F(\underline{x})$ . The resulting function,  $u : [0, 1] \rightarrow [0, 1]$ , is strictly increasing and continuous.<sup>13</sup>

*Remark 2.* The transform function,  $u$ , captures both stochastic dominance and probability of selection:  $\tilde{X}$  FSD dominates  $\tilde{Y}$  if and only if, at all quantiles  $t$ , the probability that  $\tilde{X}$  is less than or equal to the  $t$ -th quantile of  $\tilde{Y}$  is less than or equal to  $t$ , i.e., if and only if  $u(t) \leq t$ . Next note that, if a given draw from  $\tilde{Y}$ 's distribution equals  $y$ , the probability that  $\tilde{X}$  is not selected, i.e.,  $\mathbb{P}[\tilde{X} < y]$  equals  $F(y) = u(G(y))$ . Thus, the overall probability that  $\tilde{X}$  is not selected, and therefore  $\tilde{Y}$  is selected, equals  $\mathbb{E}[u(G(\tilde{Y}))]$ . Since  $G$  is continuous,  $G(\tilde{Y})$  is uniformly distributed over the unit interval (Lemma 1.2.4: Reiss, 1989). Hence,

$$\mathbb{P}[\tilde{Y} > \tilde{X}] = \int_0^1 u(s) ds.$$

Based on the intuition developed in the example in Section 2, the rate of change in  $u$  over the lower end of the quantile range should have a significant effect on whether prospects are ordered by competitive FSD dominance. To capture the behavior of the quantile transform function over this range, we define the *average transform function*,  $U$ , as follows:

$$U(t) = \begin{cases} \frac{1}{t} \int_0^t u(s) ds & \text{if } t \in (0, 1], \\ u(0) & \text{if } t = 0. \end{cases} \quad (6)$$

We can easily recover the probability of selection from  $U$  using Remark 2 by noting that

$$U(1) = \int_0^1 u(s) ds = \mathbb{P}[\tilde{Y} > \tilde{X}]. \quad (7)$$

A natural measure of the proportional rate of change of  $U$  is the elasticity of  $U$ , i.e.,

$$R_U(t) = \frac{t U'(t)}{U(t)}. \quad (8)$$

<sup>13</sup>For an example of the application of the quantile transform approach in microeconomics, see MIT open courseware <http://ocw.mit.edu/courses/economics/14-123-microeconomic-theory-iii-spring-2015/lecture-notes-and-slides/MIT14MIT14.123S15.Chapt4.pdf>.

Note that

$$R_U(t) = \frac{(\log(U(t)))'}{\log(t)'} = \frac{U'(t)}{U(t)/t}. \quad (9)$$

The first equality in equation (9) shows that  $R_U$ , in fact, represents an elasticity. The second equality shows that  $R$  is the ratio between the marginal and average change in  $U$ . This ratio can be thought of as a measure of the level of the convexity of  $U$ , and thus  $R_U$  being increasing can be interpreted as  $U$ 's convexity increasing. This interpretation implies that a concave  $U$  can have increasing elasticity if its concavity is decreasing as  $t$  increases, and a convex  $U$  can have decreasing elasticity if its convexity is decreasing as  $t$  increases. We use the elasticity of the average transform function to define average elasticity dominance.

**Definition 3** (Average elasticity dominance). Suppose that  $F$  and  $G$  are an admissible pair of distributions,  $u = F \circ G^{-1}$ , and that  $U$  and  $R_U$  are defined by equations (6) and (8) respectively, then

- (i)  $F$  average elasticity dominates  $G$  if for all  $t \in (0, 1)$ ,  $R_U(t) \leq R_U(1)$ .
- (ii)  $F$  strictly average elasticity dominates  $G$  if for all  $t \in (0, 1)$ ,  $R_U(t) < R_U(1)$ .

$F$  is average elasticity (strictly) dominated by  $G$  if the inequalities in the definition are reversed.

The next theorem shows that average elasticity dominance is a necessary and sufficient condition for competitive FSD dominance.

**Theorem 1.** Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are an admissible pair of prospects then  $\tilde{X}$  (strictly) competitive FSD dominates  $\tilde{Y}$  if and only if  $F$  (strictly) average elasticity dominates  $G$ .

Dominance based on the elasticity of the average transform function maps into the notion of *geometric convexity* which will provide simple sufficient conditions for average elasticity dominance as well as characterizing conditions for competitive MLR dominance. The results from the theory of geometric convexity required to develop these conditions are provided below.

**Definition 4** (Geometric convexity). Let  $I$  and  $J$  be subintervals of  $(0, \infty)$ , then a continuous function  $\phi : I \rightarrow J$  is *geometrically convex* if, for all  $s, t \in I$  and  $\alpha \in (0, 1)$ ,

$$\phi(s^\alpha t^{1-\alpha}) \leq \phi(s)^\alpha \phi(t)^{1-\alpha}. \quad (10)$$

If the weak inequalities are replaced with strict inequalities, we will say that  $\phi$  is *strictly geometrically convex*. Geometric concavity is defined analogously by reversing the inequality in (10).

The following lemma provides the required characterizations of geometric convexity.

**Lemma 1** (Lemma 2.1.1, Niculescu and Persson (2004)). Let  $I$  and  $J$  be subintervals of  $(0, \infty)$  and let  $\phi : I \rightarrow J$  be a continuous function. Then

---

<sup>14</sup>Because  $U$  is the integral of a bounded continuous function, it is differentiable. Thus, the elasticity is well defined.

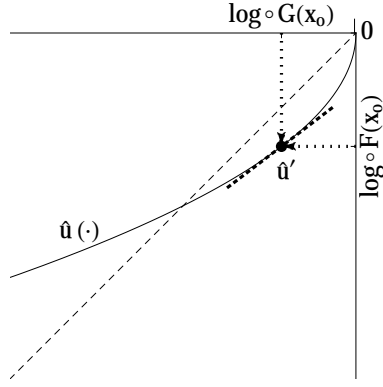


Figure 1: *Geometric convexity of the transform function,  $u$ .* The curve represents the set of points  $(\log \circ G(x), \log \circ F(x))$  produced by varying  $x$  over  $(\underline{x}, \bar{x})$ . This curve traces the conjugate transform function,  $\hat{u}$ . The convexity of  $\hat{u}$  is equivalent to the geometric convexity of  $u$ .

(i) *the following statements are equivalent.*

(a)  $\phi$  is geometrically convex.

(b)  $\log(\phi(t))$  is a convex function of  $\log(t)$ , i.e., there exists a convex function  $\hat{\phi}$ , called the conjugate function, such that  $\hat{\phi}(\log(t)) = \log \circ \phi(t)$ .

(ii) *If  $\phi$  is differentiable, then  $\phi$  is (strictly) geometrically convex if and only if the elasticity function  $R_\phi$  defined by*

$$R_\phi(t) = \frac{t \phi'(t)}{\phi(t)}$$

*is (strictly) increasing.*

*These statements remain valid if we replace “convex” with “concave,” and “increasing” with “decreasing”.*

**Remark 3.** Condition (i.b) of the lemma implies that the quantile transform function is geometrically convex if it is convex when plotted on a log-log scaled grid. In fact, the *conjugate transform function* of  $u$ ,  $\hat{u}$ , defined using Lemma 1.i.b, is simply the graph of a parametric curve, where  $x$  is the parameter and the abscissa and ordinate in the graph are the logarithms of the distribution functions,  $G$  and  $F$  respectively.<sup>15</sup> This parametric log–log representation is depicted in Figure 1.

In addition, Lemma 1.ii shows that geometrically convex functions are characterized by increasing elasticity. Consequently, the next proposition shows that increasing elasticity of the average transform function or the transform function are sufficient conditions for average elasticity dominance.

**Proposition 1.** *Suppose that  $F$  and  $G$  are an admissible pair of distributions then*

$$u \text{ is (strictly) geometrically convex} \Rightarrow U \text{ is (strictly) geometrically convex} \Rightarrow \\ F \text{ (strictly) average elasticity dominates } G.$$

*These statements remain valid if we replace “convex” with “concave” and “dominant” with*

<sup>15</sup>See Cargo (1965) for the derivation of this representation.

“dominated.”

Verifying the elasticity condition for geometric convexity is typically much easier than computing selection-conditioned prospect distributions or verifying the average elasticity condition in Definition 3. Thus, Proposition 1 greatly simplifies verifying average elasticity dominance and thus verifying competitive FSD dominance. We illustrate this simplification in Supplement C.1.<sup>16</sup>

## 4.2 Geometric dominance and competitive MLR dominance

In this section, we develop geometric dominance, the competitive selection relation characterizing MLR ordering relations between selection-conditioned distributions. MLR is a stronger stochastic order than stochastic dominance (Theorem 1.C.1: Shaked and Shanthikumar, 2007) and MLR ordering is typically defined in terms of probability density functions. Thus, we turn our attention to distributions with well-defined and tractable densities.

**Definition 5** (Regular and regularly related distributions). Distributions are *regular* if they are absolutely continuous with respect to the Lebesgue measure with continuous positive densities over the interior of their interval support. Two distributions,  $F$  and  $G$ , are *regularly related* if they are an admissible pair of distributions and both distributions are regular.

Regularity implies that  $F(\underline{x}) = G(\underline{x}) = 0$  and that  $u$  is differentiable over  $(0, 1)$  and thus Lemma 1.ii can be used to determine whether  $u$  is geometrically convex.

In like fashion to the definition of competitive FSD dominance, we define competitive MLR dominance simply by applying the standard definition of MLR dominance to the selection-conditioned distributions of  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  given by  $F^{\text{sc}}(x)$  and  $G^{\text{sc}}(x)$  respectively:

**Definition 6** (Competitive MLR dominance). Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are an regularly related pair of prospects, then  $\tilde{X}$  (strictly) competitive MLR dominates  $\tilde{Y}$  if

$$x \mapsto \frac{f^{\text{sc}}(x)}{g^{\text{sc}}(x)} \text{ is (strictly) increasing, } x \in (\underline{x}, \bar{x}).$$

Next note that, under the regularity assumption, equation (5) shows that the conditional densities of  $F$  and  $G$  are given as follows:

$$f^{\text{sc}}(x) = \frac{f(x)G(x)}{\int_{\underline{x}}^{\bar{x}} f(z)G(z)dz}, \quad g^{\text{sc}}(x) = \frac{g(x)F(x)}{\int_{\underline{x}}^{\bar{x}} g(z)F(z)dz}. \quad (11)$$

<sup>16</sup>As well as being a sufficient condition for average elasticity dominance, as the next section shows, the geometric convexity of  $u$  defines the competitive selection relation characterizing competitive MLR dominance. Supplement Lemma A-3 shows that geometric convexity of  $U$  defines the competitive selection relation characterizing competitive reversed hazard rate order dominance (Chapt. 1.B.6: Shaked and Shanthikumar, 2007). Given the paper’s focus on MLR and stochastic dominance, we will not develop this characterization further. For an application of reversed hazard rate ordering in economics, see Kirkegaard (2012).



Let  $r_F$  ( $r_G$ ) represent the *reversed hazard rate* of  $F$  ( $G$ ), i.e.,

$$r_F(x) = \frac{f(x)}{F(x)}, \quad r_G(x) = \frac{g(x)}{G(x)}.$$

Because the denominators in equation (11) are constants, it is easy to see that  $F$  (strictly) competitive MLR dominates  $G$  if and only if

$$x \mapsto \frac{r_F(x)}{r_G(x)} \text{ is (strictly) increasing.}$$

How is this characterization of competitive MLR dominance related to our earlier characterization of competitive FSD dominance in terms of the transform function,  $u$ ? It is reasonable to expect that the relation is quite tight because MLR dominance between (unconditional) distributions is characterized by the convexity of quantile transform function,  $u$ .

**Lemma 2** (Theorem 2.3 (Chan, Proschan, and Sethuraman, 1990)). *Suppose that  $F$  and  $G$  are regularly related and  $u = F \circ G^{-1}$ . Then  $u$  is (strictly) convex if and only if  $F$  (strictly) MLR dominates  $G$ .*

Motivated by this result, we define our second competitive selection relation: *geometric dominance*.

**Definition 7** (Geometric Dominance). Suppose that  $F$  and  $G$  are an admissible pair of distributions and  $u = F \circ G^{-1}$ .

- (i)  $F$  geometrically dominates  $G$  if  $u$  is geometrically convex.
- (ii)  $F$  strictly geometrically dominates  $G$  if  $u$  is strictly geometrically convex.

$F$  is (strictly) geometrically dominated by  $G$  if “convex” is replaced by “concave” in the definitions given above.

The next result characterizes competitive MLR dominance in terms that are quite analogous to Lemma 2: just as MLR dominance is equivalent to the convexity of  $u$ , competitive MLR dominance is equivalent to the geometric convexity of  $u$ .

**Theorem 2.** *Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$ ,  $F$  and  $G$  are regularly related, and  $u = F \circ G^{-1}$ . Let  $r_F = f/F$  and  $r_G = g/G$  represent the reversed hazard rates of  $F$  and  $G$  respectively. Then the following conditions are equivalent:*

- i.  $\tilde{X}$  (strictly) competitive MLR dominates  $\tilde{Y}$ .
- ii.  $x \mapsto r_F(x)/r_G(x)$  is (strictly) increasing.
- iii.  $F$  (strictly) geometrically dominates  $G$ .

An immediate and interesting consequence of Theorem 2 is that, in quantile space, the conditions for MLR dominance and competitive MLR dominance differ only by a log transformation of the coordinate system. In the space of realizations, the conditions for dominance are also quite analogous: MLR dominance requires the ratio  $f/g = F'/G'$  to be increasing; competitive MLR dominance requires the ratio  $(\log \circ F)' / (\log \circ G)' = r_F/r_G$  to be increasing.

## 5 Conditional vs. unconditional dominance

Having developed the competitive selection relations, we now apply them to identifying the relationship between conditional and unconditional dominance. What restrictions do dominance reversal and preservation impose on the unconditional distributions of the compared prospects? Is it possible to identify dominance reversal and preservation using only selection-conditioned information?

### 5.1 Dominance reversal

First consider FSD dominance reversal, i.e., when one prospect, say  $\tilde{Y}$ , FSD dominates the compared prospect,  $\tilde{X}$ , but  $\tilde{X}$  competitive FSD dominates  $\tilde{Y}$ . In view of Theorem 1, this is equivalent to  $\tilde{Y}$  being FSD dominant and  $\tilde{X}$  being average elasticity dominant.

We start with a simple result, that provides some insight into dominance reversal. It shows that a point mass at the lower endpoint of the support of the compared distributions militates very strongly for reversal: a mixture between a continuous distribution and a point mass at the lower endpoint always dominates the continuous distribution under the stronger competitive selection relation, geometric dominance, and no distribution with a point mass at its lower endpoint can be dominated by a continuous distribution even under the weaker competitive selection relation, average elasticity dominance.<sup>17</sup>

**Proposition 2.** *Let  $F_o$  and  $G$  be an admissible pair of distribution functions such that  $F_o(\underline{x}) = 0$ . Let  $\delta_{\underline{x}}$  represent the distribution function for a point mass at  $\underline{x}$ . Let  $F = p\delta_{\underline{x}} + (1 - p)F_o$ ,  $p \in (0, 1)$ . Then, (i)  $F$  is never average elasticity dominated by  $G$ . (ii)  $F$  strictly geometrically (and a fortiori average elasticity) dominates  $F_o$ .*

#### 5.1.1 Application: Effort-bidding contests

Proposition 2 can be used to characterize the effects of selection in asymmetric, all-pay auction, effort-bidding contests. Consider a firm that receives job applications from two groups of applicants, a strong group  $S$  and a weak group  $W$ . The cost of effort for applicants from group  $i$ ,  $e_i \geq 0$ ,  $i = S, W$ , is given by  $e_i/a_i$ ,  $a_i > 0$ . Thus,  $a$  measures ability. We assume that applicants of type  $S$  are more able, i.e.,  $a_S > a_W$ . The value to applicants of being hired is normalized to 1.

At date 0, applicants make their effort choice. Employee selection occurs at date 1. Each job is filled by head-to-head competition between pairs of applicants from the two groups. The winner of the competition is the applicant with the highest quality, unless both applicants' qual-

<sup>17</sup>Recall that admissibility requires  $G$  to be continuous; thus, in Proposition 2, the hypothesis that  $G$  is continuous is entailed by the admissibility hypothesis.

ity equals 0, in which case, neither applicant is hired.<sup>18</sup> Between date 1 and 2, hired applicants work for the firm and their overall performance is recorded at date 2.

We model the contest mechanism for employee selection in Baye, Kovenock, and de Vries (1996) effort-bidding all-pay auction framework. We make the standard assumption in this literature that output equals effort (e.g., Moldovanu and Sela, 2001). Thus, both the quality of applicants and their subsequent performance, if hired, equals their date 0 effort. Applicants independently select effort distributions,  $F_S, F_W$ , over non-negative effort levels. Because quality equals effort, the applicant with the highest realized effort is hired.

Thus, in the Baye, Kovenock, and de Vries (1996) framework, hiring is determined by an all-pay auction in which applicants bid effort. Using the ability/effort relation described above, the payoffs to applicants of each type,  $v_i$ , are given by

$$\begin{aligned} v_S(e_S | F_W) &= F_W(e_S) - \frac{e_S}{a_S}, \\ v_W(e_W | F_S) &= F_S(e_W) - \frac{e_W}{a_W}. \end{aligned}$$

An equilibrium is a pair of effort distributions satisfying the condition that each effort level in the support of the effort distributions is a best reply to the effort distribution selected by applicants from the other group. The following distribution pair  $(F_S^*, F_W^*)$  is an equilibrium:

$$\begin{aligned} F_S^*(e) &= e/a_W, \\ F_W^*(e) &= p + (1-p)F_S^*(e), \text{ where } p = (a_S - a_W)/a_S, \end{aligned} \quad e \in [0, a_W].^{19}$$

To see that  $(F_S^*, F_W^*)$  is an equilibrium, note that

$$\begin{aligned} v_S(e_S | F_W^*) &= \frac{a_S - a_W}{a_S} - \frac{a_W}{a_S} \max \left[ 0, \frac{e_S}{a_W} - 1 \right], \\ v_W(e_W | F_S^*) &= -\max \left[ \frac{e_W}{a_W} - 1, 0 \right]. \end{aligned}$$

Thus, over the common support of  $F_i$ ,  $i = S, W$ , which equals  $[0, a_W]$ , the payoffs to both  $S$  and  $W$  applicants are constant and their payoffs are strictly decreasing off the support (i.e., when  $e > a_W$ ), verifying the candidate equilibrium.

In this equilibrium, a strong applicant,  $S$ , plays a Uniform distribution over the interval  $[0, a_W]$ ; a weak applicant,  $W$ , randomizes between an effort level of 0, and the same Uniform $[0, a_W]$  distribution as the strong applicant. Because quality equals effort, the quality of strong applicants FSD dominates the quality of weak applicants.

Now consider the performance of hired, i.e., selected, applicants. A strong applicant is

<sup>18</sup>Because in Baye, Kovenock, and de Vries (1996) type all-pay auction models, equilibrium “bid” (effort in our context) distributions never produce a positive probability of tied bids, we ignore tie bids.

<sup>19</sup>In this expression and subsequent expressions, we will only explicitly represent distribution functions over their support intervals. Distribution functions are understood to equal 0 for  $x$  below their support intervals and 1 for  $x$  above their support intervals.

hired if and only if her quality exceeds the quality of her weak competitor, i.e.,  $\tilde{e}_S > \tilde{e}_W$  and vice versa. Because performance equals effort, the performance of hired strong applicants will be distributed  $[\tilde{e}_S | \tilde{e}_S > \tilde{e}_W]$  and the performance of hired weak applicants will be distributed  $[\tilde{e}_W | \tilde{e}_W > \tilde{e}_S]$ . Because the performance distribution of a weak applicant is a mixture between a point mass at 0 and the performance distribution of a strong applicant, Proposition 2.ii implies that the performance of weak applicants strictly geometrically dominates the performance of strong applicants. Thus Proposition 1 and Theorem 1 imply that the selection-conditioned performance of weak applicants strictly FSD dominates the selection-conditioned performance of strong applicants.

Note also that, because strong applicants are hired whenever weak applicants choose zero effort, which occurs with probability  $p$ , and weak and strong applicants are equally likely to be hired when weak applicants play the Uniform $[0, a_W]$  distribution, the probability that a weak applicant will be hired equals  $\frac{1}{2}(1 - p)$  and the probability a strong applicant will be hired equals  $p + \frac{1}{2}(1 - p)$ . This implies that, under meritocratic hiring policies, applicants from the strong group are more likely to be hired yet perform worse, in the sense of FSD dominance, when hired than applicants from the weak group. Thus, in the effort-bidding framework, underrepresentation of workers from a group whose members on average perform better if hired is not evidence that the firm is biased against the underrepresented group.

### 5.1.2 Necessary conditions for dominance reversal

Given Proposition 2, the analysis in Section 2, and the application above, it might appear that a point mass at the lower endpoint of the support of one of the prospects' distributions is required for dominance reversal. Actually, this not the case. However, as the next result shows, "something like" a point mass is indeed the *only* way to induce reversal: a point mass per se is not required but, for reversal to occur, the ratio between the left tails of the two compared distributions must approach infinity.

**Theorem 3.** *Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are an admissible pair of prospects and let  $u \circ G = F$ . Suppose that  $G$  strictly FSD dominates  $F$ . Then,*

- (i) *if  $F$  average elasticity dominates  $G$ ,  $\limsup_{x \rightarrow \underline{x}} \frac{F(x)}{G(x)} = \infty$ .*
- (ii) *If in addition,  $F$  strictly geometrically dominates  $G$ , then (a) for all  $x \in (\underline{x}, \bar{x})$ , the function  $x \mapsto \frac{F(x)}{G(x)}$  is strictly decreasing and (b)  $\lim_{x \rightarrow \underline{x}} \frac{F(x)}{G(x)} = \infty$ .*

The logic for dominance reversal is fairly straightforward: the geometrically dominant distribution is strictly FSD dominated, i.e.,  $u(t) > t$ ,  $t \in (0, 1)$ . At the same time,  $u(1) = 1$ . These two constraints restrict the rate of increase in the quantile transform function,  $u'$ . Thus, in order to exhibit increasing elasticity,  $R_u = u'/(u/t)$  must start from a very low base, which requires that  $u(t)/t \rightarrow \infty$  as  $t \rightarrow 0$ . The tail ratio,  $F/G$ , equals  $(u \circ G)/G$  and thus must be asymptotically unbounded.

When the distributions are regularly related and  $u$  is smooth over the interval  $(0, 1)$ , this rather casual intuition can be buttressed by the following argument which goes a long way toward establishing the theorem: The existence of a finite limit of the ratio  $u(t)/t$  as  $t \rightarrow 0$  entails a contradiction. If  $\lim_{t \rightarrow 0} u(t)/t < \infty$  existed,  $u'(0) < \infty$  would exist. Because the geometrically dominant distribution is strictly FSD dominated,  $u(t) > t$  over  $(0, 1)$ , thus, clearly  $u'(0) > 0$ . Given that  $u'(0) > 0$ , L'Hôpital's rule could be applied to show that  $\lim_{t \rightarrow 0} R_u(t) = 1$ . Strict geometric convexity implies that  $R_u$  is strictly increasing. Thus,  $R_u(t) > 1$ ,  $t \in (0, 1)$ . The geometrically dominated distribution being strictly FSD dominated implies that  $u(t)/t > 1$ . This fact, and the definition of  $R_u$ , would then imply that  $u'(t) > 1$  for all  $t \in (0, 1)$ . However,  $u(0) = 0$  and  $u(1) = 1$ , which, by the mean value theorem, is not consistent with  $u'(t) > 1$  for all  $t \in (0, 1)$ . This contradiction shows that  $\lim_{t \rightarrow 0} u(t)/t < \infty$  does not exist. Thus, because  $F/G = u \circ G/G$ , a finite limit of the tail-ratio cannot exist as  $x \rightarrow \underline{x}$ . The formal proof in Supplement A.4 establishes the precise behavior of the tail ratio,  $F/G$ , asserted in the theorem.

### 5.1.3 Example: Dominance reversal

Thus far, we have only exhibited dominance reversal when the FSD dominated but competitive FSD dominant distribution featured a point mass at the lower bound of its support,  $\underline{x}$ . Are there simple examples of dominance reversal when the compared distributions are continuous? As will be shown in Section 6.2, when the compared distributions are standard textbook continuous distributions, dominance reversal is quite unusual. So, we will present an example that features rather non-standard but absolutely continuous distribution functions.

Our starting point is the function  $u : [0, 1] \rightarrow [0, 1]$  defined by

$$u(t) = \frac{\alpha}{\alpha - \log(t)}, \quad \alpha > 0.$$

And  $u(0) = \lim_{t \rightarrow 0} u(t) = 0$ . Next note that

$$\log(u(t)) = \hat{u}(\log t), \quad t \in (0, 1], \quad \text{where } \hat{u}(y) = -\log(1 - y/\alpha), \quad y < 0, \quad (12)$$

and that  $\hat{u}$  is strictly convex. Thus,  $\log \circ u$  is a strictly convex function of  $\log(t)$  and, hence, by Lemma 1,  $u$  is strictly geometrically convex. It is also easy to verify the following additional properties of  $u$ : (i)  $u$  is concave if and only if  $\alpha \geq 2$ . (ii)  $u(t) > t$  for sufficiently small  $t$ . (iii)  $u(t) > t$  for all  $t \in (0, 1)$  if and only if  $\alpha \geq 1$ .<sup>20</sup>

Suppose that  $G$  is any absolutely continuous distribution function and  $F = u \circ G$ , then Lemma 2, Theorem 2, and Remark 2 imply the following relationships between  $F$  and  $G$ : (a) For all  $\alpha > 0$ ,  $F$  strictly geometrically dominates  $G$ . (b)  $G$  strictly FSD dominates  $F$  if and only if  $\alpha \geq 1$ . (c)  $G$  strictly MLR dominates  $F$  if and only if  $\alpha \geq 2$ . In order to generate a graphical illustration of dominance reversal, we need to pick a distribution for  $G$ . We will as-

<sup>20</sup>Verification of these assertions provided in Supplement A.4

sume that  $G$  is supported by the unit interval and its distribution function over the unit interval is given by  $x^\alpha$ ,  $\alpha > 0$ . Because  $F = u \circ G$ , this implies that

$$F(x) = \frac{1}{1 - \log x}, \quad G(x) = x^\alpha, \quad x \in (0, 1], \alpha > 0, \quad (13)$$

$F(0) = G(0) = 0$ . These distributions, and the behavior of the tail-ratio  $F/G$ , are illustrated in Figure 2.

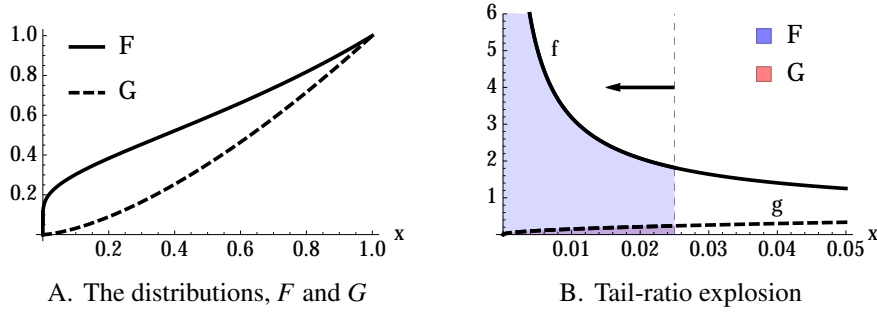


Figure 2: *Dominance reversal*. In the figures, the distributions are as specified in equation (13) with  $\alpha = 3/2$ . In Panel B, the shaded areas beneath the densities  $f$  and  $g$ , represent the distribution functions of  $F$  and  $G$  evaluated at  $x = 0.025$ .

As the dashed line in Panel B is shifted to the left, the ratio between the area beneath the densities of  $F$  and  $G$ ,  $F/G$ , increases without bound.

## 5.2 Dominance preservation

Now consider the case of stochastic dominance preservation, i.e., the FSD dominant prospect is also competitive FSD dominant. Our basic result is that, when dominance is preserved and the prospects are regularly related, if the FSD dominant prospect's distribution is geometrically dominant, it must also be MLR dominant. This implies (Theorem 2) that if a prospect is strictly FSD dominant and strictly competitive MLR dominant, it must also be strictly MLR dominant.

**Theorem 4.** Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are an admissible pair of prospects and let  $u \circ G = F$ . Suppose that  $\tilde{X}$  strictly FSD dominates  $\tilde{Y}$ . If  $F$  also strictly geometrically dominates  $G$ , then (i) the transform function,  $u$ , is strictly convex, and (ii) if  $F$  and  $G$  are regularly related,  $\tilde{X}$  strictly MLR dominates  $\tilde{Y}$ .

The intuition underlying this result is that the admission effect, discussed in Section 2, always favors the competitive selection dominance of the FSD dominated prospect. Thus, if a prospect is FSD dominant and “ends up,” after conditioning for selection, MLR dominant, i.e., is competitive MLR dominant (which, by Theorem 2 is equivalent to its distribution being geometrically dominant), it must “start off” (before conditioning on selection) MLR dominant. Note, as Example 5.1.3 shows, there is no analogous result for dominance reversal. When

selection reverses dominance, the FSD dominated prospect may or may not be competitive MLR dominant.

A simple heuristic proof of Theorem 4 is possible under the following restrictive assumptions: (a) the quantile transform function is twice continuously differentiable over  $[0, 1]$ , and (b)  $u'(0) > 0$ . These assumptions are not innocuous and, in fact, are never satisfied in the case of dominance reversal considered earlier. However, they permit a simple and more intuitive derivation of Theorem 4, which is proven under less restrictive assumptions in Supplement A.4. Consider the elasticity function  $R_u$  defined in Lemma 1.ii. Its derivative, which exists by (a), can be expressed as

$$R'_u(t) = \frac{1}{u(t)} \left( R_u(t) \left( \frac{u(t)}{t} - u'(t) \right) + t u''(t) \right), \quad t \in (0, 1). \quad (14)$$

Applying L'Hôpital's rule, justified by (a) and (b), shows that  $\lim_{t \rightarrow 0} R_u(t) = 1$ . Because strict geometric convexity implies that  $R_u$  is strictly increasing,  $R_u(t) > 1$ ,  $t \in (0, 1)$ , which, by the definition of  $R_u$ , implies that  $u(t)/t - u'(t) < 0$ . Strict geometric convexity requires  $R'_u > 0$ . Inspection of equation (14) shows that  $R'_u > 0$  only if  $u'' > 0$ , i.e.,  $u$  is strictly convex. Thus, by Lemma 2, the FSD dominant distribution is MLR dominant.

### 5.2.1 Application: Hiring tournaments

Consider the economic framework developed for the effort-bidding model of hiring in Section 5.1.1. Maintain all of the assumptions of that model except for those relating to the contest's selection mechanism. Instead of assuming that effort determines applicant quality, assume that effort determines mean applicant quality, and that actual applicant quality depends on mean effort and a random shock as in a Lazear and Rosen (1981) style tournament model. As in the effort-bidding model, the performance of hired workers equals their quality. Specifically, assume that the quality of applicants from group  $i$ ,  $\tilde{Q}_i$ ,  $i = S, W$ , is given by

$$\tilde{Q}_i = e_i \tilde{Z}_i, \quad \tilde{Z}_i \stackrel{d}{\sim} \text{Exponential}(\beta), \quad i = S, W, \quad (15)$$

where  $\tilde{Z}_S$  and  $\tilde{Z}_W$  are independent random variables. Independence is a standard assumption in this literature. The multiplicative error structure is a departure from the literature which typically assumes an additive random shock generated by a distribution which is symmetric, absolutely continuous, centered at 0, and supported by  $(-\infty, \infty)$ . We depart from the literature because the typical assumption leads, as Lazear and Rosen note, to the non-existence of pure strategy equilibria for some parameterizations of the model.

It is sufficient to consider the best responses of group  $S$  applicants and then exploit symmetry to produce the best responses of group  $W$  applicants. So, consider the problem of an  $S$  applicant, given that  $W$  applicants choose effort level  $e_W$ , the probability that the applicant is hired equals the probability that  $\tilde{Q}_S > \tilde{Q}_W$ . Using equation (15) we see that the probability of

hiring is given by  $\mathbb{P}[\tilde{Z}_W/\tilde{Z}_S < e_S/e_W]$ . Next note that  $\tilde{Z}_W/\tilde{Z}_S \stackrel{d}{\sim} x/(1+x)$ ,  $x \geq 0$ . Thus, using the normalization of the value of employment to 1, the assumed effort costs, and the discussion above, we see that the utilities of an  $S$  applicant,  $v_S$ , and  $W$  applicant,  $v_W$ , are given by

$$v_S(e_S, e_W) = \begin{cases} \frac{e_S}{e_S+e_W} - \frac{e_S}{a_S} & \text{if } e_S + e_W > 0, \\ 0 & \text{if } e_S + e_W = 0 \end{cases} \quad v_W(e_S, e_W) = \begin{cases} \frac{e_W}{e_S+e_W} - \frac{e_W}{a_W} & \text{if } e_S + e_W > 0, \\ 0 & \text{if } e_S + e_W = 0. \end{cases}$$

No positive level of effort is a best reply to the other applicant choosing zero effort, and choosing zero effort is also never a best reply to the other applicant choosing zero effort. Thus, in any Nash equilibrium, both applicants choose positive effort levels. Because each applicant's utility is strictly concave in effort, optimal positive effort choices are determined by the first-order conditions:

$$\begin{aligned} \frac{\partial}{\partial e_S} v_S &= \frac{e_W}{(e_S + e_W)^2} - \frac{1}{a_S} = 0, \\ \frac{\partial}{\partial e_S} v_W &= \frac{e_S}{(e_S + e_W)^2} - \frac{1}{a_W} = 0. \end{aligned}$$

It is easy to verify that these equations have a unique solution given by

$$e_i^* = a_i \left( \frac{a_S a_W}{(a_S + a_W)^2} \right), \quad i = S, W. \quad (16)$$

The unconditional distribution of applicant quality for applicants from group  $i$  is Exponential with mean  $a_i^* \beta$ . Because  $a_S > a_W$ , equation (16) reveals that  $e_S^* > e_W^*$ . Thus, the unconditional quality of  $S$  applicants,  $e_S^* \tilde{Z}_S$ , strictly FSD dominates the unconditional quality of  $W$  applicants,  $e_W^* \tilde{Z}_W$ .

If an applicant from group  $i$  is hired, the quality and performance of the applicant will equal  $\tilde{Q}_i = e_i^* \tilde{Z}_i$ . An applicant from group  $i$  is hired if and only if her quality exceeds the quality of the rival from the other group, i.e.,  $e_i^* \tilde{Z}_i > e_j^* \tilde{Z}_j$ ,  $j \neq i$ . Thus, the expected performance of hired applicants from group  $i$  equals  $\mathbb{E}[e_i^* \tilde{Z}_i | e_i^* \tilde{Z}_i > e_j^* \tilde{Z}_j]$ .

In Section 6.2 we will show that, when two Exponential distributions with different means are compared, FSD dominance implies geometric dominance and, *a fortiori* (Proposition 1), average elasticity dominance. Thus, by Theorem 1, the selection-conditioned distribution of  $S$ -applicant's performance,  $[e_S^* \tilde{Z}_S | e_S^* \tilde{Z}_S > e_W^* \tilde{Z}_W]$  strictly FSD dominates the selection-conditioned distribution of  $W$ -applicant's performance,  $[e_W^* \tilde{Z}_W | e_W^* \tilde{Z}_W > e_S^* \tilde{Z}_S]$ . This implies that

$$\mathbb{E}[e_S^* \tilde{Z}_S | e_S^* \tilde{Z}_S > e_W^* \tilde{Z}_W] > \mathbb{E}[e_W^* \tilde{Z}_W | e_W^* \tilde{Z}_W > e_S^* \tilde{Z}_S].$$

For the same reasons as advanced in the effort-bidding model,  $S$  applicants are more likely to be hired than  $W$  applicants. Thus, in contrast to the effort-bidding framework, in the tournament framework, underrepresentation of workers from a group whose members on average perform better if hired is not consistent with merit-based hiring policies.



### 5.3 Competitive selection equivalence

Thus far, we have considered dominance preservation and dominance reversal. Of course, it is possible that the prospects are not ordered by either competitive selection relation even though they are ordered by FSD or MLR dominance. However, there is another interesting possibility: the prospect's selection-conditioned distributions are identical. This possibility can result because, unlike the FSD and MLR orders, the competitive selection relations are not anti-symmetric, i.e., there exist admissible distributions  $F$  and  $G$ ,  $F \neq G$ , such that

EQUIV :  $F$  average elasticity (geometrically) dominates  $G$

and  $G$  average elasticity (geometrically) dominates  $F$ .

When EQUIV holds, we will say that  $F$  and  $G$  are average elasticity (geometrically) equivalent. The next result shows that if the compared distributions are equivalent under one of the competitive selection relations, they are equivalent under both. In which case, they are related by a power transform, and their selection-conditioned distributions are identical.<sup>21</sup>

**Proposition 3.** *For an admissible pair of distributions,  $F$  and  $G$ , (i) the following conditions are equivalent: (a)  $F$  and  $G$  are average elasticity equivalent. (b)  $F$  and  $G$  are geometrically equivalent. (c)  $F(x) = G(x)^p$  for some  $p > 0$ . (ii) When one of these equivalent conditions is satisfied, the selection-conditioned distributions of  $F$  and  $G$ ,  $F^{sc}$  and  $G^{sc}$ , are both equal to  $G(x)^{p+1}$  for the  $p$  defined in (c).*

Because the equivalence classes of the competitive selection relations are the same, we will say that prospects which are equivalent under either competitive selection relation are *competitive selection equivalent*. When compared prospects are competitive selection equivalent, selection neutralizes unconditional dominance. The FSD or MLR dominant prospect is more likely to be selected, but its selection-conditioned distribution is identical to the selection-conditioned distribution of the dominated prospect.

An obvious case of selection equivalence is provided by two power law distributions,  $F(x) = x^{p_1}$  and  $G(x) = x^{p_2}$ ,  $x \in [0, 1]$ ,  $p_1, p_2 > 0$ . As we discuss in Section 6.2, there are less obvious cases, e.g., two Fréchet distributions with the same shape but different scale parameters or two Gumbel (i.e., Generalized Extreme Value Type-I) distributions with the same scale but different location parameters.

<sup>21</sup>Regarding the other conditions for partial ordering: average elasticity dominance is a reflexive but non-transitive relation. Geometric dominance is a reflexive, transitive relation, and thus defines a preorder but not partial order. Since these other conditions for partial ordering play no role in our subsequent development, we defer their discussion and analysis to Supplement A.4.

## 5.4 Identifying dominance reversal and preservation

Consider two prospects  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$ . Suppose that  $F$  strictly geometrically dominates  $G$ . Theorems 4 and 3 then imply that (a) if  $\tilde{X}$  is FSD dominant, i.e., dominance is preserved, it is MLR dominant. (b) If  $\tilde{X}$  is FSD dominated, i.e., dominance is reversed, the lower tail ratio,  $F/G$  is unbounded. Hence,  $F$  is either “much better” or “much worse” than  $G$ . It is possible, using only selection-conditioned information, to distinguish between these two alternatives? This question is important because the implications of being much better versus much worse are very different. The question is relevant because, in many cases, only selection-conditioned information is available. Below, we answer this question affirmatively.

Selection produces selection-conditioned distributions of the two prospects and the probabilities that the prospects are selected, i.e., their *selection propensities*. We represent selection propensity with  $p$ , i.e.,  $p_X = \mathbb{P}[\tilde{X} > \tilde{Y}]$  and  $p_Y = \mathbb{P}[\tilde{Y} > \tilde{X}]$ . Define the *propensity-weighted likelihood ratio*,  $\mathcal{L}$ , as follows:

$$\mathcal{L}(x) = \frac{p_X f^{\text{sc}}(x)}{p_Y g^{\text{sc}}(x)}, \quad x \in (\underline{x}, \bar{x}).$$

From equation (11), we see that the selection-conditioned densities for the two prospects are given by  $f^{\text{sc}}(x) = (f(x)G(x))/p_X$  and  $g^{\text{sc}}(x) = (g(x)F(x))/p_Y$ . Thus, the propensity-weighted likelihood ratio,  $\mathcal{L}$ , equals the ratio between the reversed hazard rates of the compared distributions,  $r_F/r_G$ . This implies that the propensity-weighted likelihood ratio identifies the conditions for geometric dominance (Theorem 2).

Dominance preservation and reversal can also be identified. The demonstration of this assertion relies on a technical result.

**Lemma 3.** *If  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$ ,  $F$  and  $G$  are regularly related and  $F$  strictly geometrically dominates  $G$ , then,*

- (i) *If  $\tilde{X}$  strictly FSD dominates  $\tilde{Y}$ , then  $\mathcal{L}(x) = \frac{r_F(x)}{r_G(x)} > 1$  for  $x \in (\underline{x}, \bar{x})$ .*
- (ii) *If  $\tilde{X}$  is strictly FSD dominated by  $\tilde{Y}$ , then  $\mathcal{L}(x) = \frac{r_F(x)}{r_G(x)} < 1$  for  $x \in (\underline{x}, \bar{x})$ .*

Because the propensity-weighted likelihood function equals the ratio between the reversed hazard rates, the restrictions on the magnitude of this ratio derived in Lemma 3 identify dominance preservation and reversal. Dominance is preserved if and only if the propensity-weighted likelihood ratio,  $\mathcal{L}$ , is strictly increasing and  $\inf\{\mathcal{L}(x) : x \in (\underline{x}, \bar{x})\} \geq 1$ . Dominance is reversed if and only if the propensity-weighted likelihood ratio is strictly increasing and  $\sup\{\mathcal{L}(x) : x \in (\underline{x}, \bar{x})\} \leq 1$ . The propensity-weighted likelihood ratio also appears to be estimable: estimating propensity is straightforward. There is a well developed literature on non-parametric estimation of likelihood functions (e.g., Kitamura, Tripathi, and Ahn, 2004) and on testing for functional monotonicity (e.g., Ghosal, Sen, and Van Der Vaart, 2000).

## 5.5 Origins of geometric dominance

Our primary aim is to uncover the relation between selection-conditioned and unconditional stochastic orderings of prospects. However, in some applications, such as the application developed below, understanding the restrictions that the competitive selection relations impose on unconditional distributions generally, in the absence of unconditional stochastic order relations, is useful. Average elasticity dominance is defined in terms of averages of the quantile transform function and thus cannot produce pointwise restrictions on the behavior of distribution functions. In contrast, geometric dominance imposes very strong and simple restrictions.

**Theorem 5.** *Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are an admissible pair of prospects and let  $u = F \circ G^{-1}$ . Suppose that  $F$  strictly geometrically dominates  $G$ .*

- (i) *If, on some neighborhood of  $\underline{x}$ ,  $F(x) < G(x)$ , then  $\tilde{X}$  strictly FSD dominates  $\tilde{Y}$ .*
- (ii) *If, on some neighborhood of  $\underline{x}$ ,  $F(x) > G(x)$ , then either*
  - (a)  *$\tilde{Y}$  strictly FSD dominates  $\tilde{X}$ , or*
  - (b)  *$\tilde{X}$  is riskier than  $\tilde{Y}$  in the sense that  $F$  crosses  $G$  once from above.*

*Sketch of proof.* The formal proof of this result is presented in Supplement A.4. The intuition behind the proof of this result nicely illustrates how the “geometry” of geometric dominance restricts the behavior of distribution functions. At quantiles where the two distribution functions,  $F$  and  $G$ , cross, i.e., points where  $F(x) = G(x)$ , the conjugate transform function, defined in Lemma 1 and Remark 3, meets the diagonal line,  $\hat{u}(y) = y$  depicted in Figure 1. Points below the diagonal map into values of the distribution functions at which  $F(x) < G(x)$  and points above the diagonal map into points at which  $F(x) > G(x)$ .

Strict geometric dominance implies strict convexity of the conjugate transform function. Strict convexity imposes strong restrictions on how often the conjugate transform function can cross the diagonal. If the conjugate transform function starts below the diagonal, as it will if  $F$  is strictly FSD dominant, it can only meet the diagonal once. However, the conjugate transform function must meet the diagonal at the upper endpoint of its domain,  $y = 0$ , which corresponds to the point where both distribution functions equal 1. Thus, the conjugate transform function cannot meet the diagonal at any other point and hence must lie below the diagonal for all  $y < 0$ . This case corresponds to part (i).

If the conjugate transform function starts out above the diagonal, as it will if  $F$  is strictly FSD dominated, strict convexity implies that it can cross the diagonal at most twice. Again, because one of these crossings must occur at the upper endpoint, the conjugate transform crosses the diagonal at most once from above before reaching the endpoint. One crossing from above before reaching the endpoint corresponds to part (ii.b) and no crossing before reaching the endpoint corresponds to part (ii.a).  $\square$

In essence, Theorem 5 divides geometric dominance relations into three possible config-

urations: In the first case, developed in part (i) of the Theorem, the geometrically dominant distribution is FSD dominant. In the second, developed in part (ii.a), the geometrically dominant distribution is FSD dominated. In the third case, developed in part (ii.b), the geometrically dominant distribution is in some sense more risky as the single crossing of the distribution functions is from above. These observations imply that if a geometrically dominant distribution has a lower mean than the distribution it dominates, the geometrically dominant distribution must be second-order stochastically dominated, i.e., riskier in the Rothschild and Stiglitz (1970) sense.

### 5.5.1 Application: Pandering in cheap-talk games

Che, Dessein, and Kartik (2013) develop a cheap-talk model in which an agent makes a recommendation to a principal about a project. There is no conflict of interest between the principal and the agent except that the agent would like the principal to select one of the projects rather than an outside option, e.g., rejecting all of the projects. For each potential project, the principal has a prior probability distribution over the utility engendered by adopting the project,  $b$ . The agent knows  $b$  and recommends one of the projects to the principal. Che, Dessein, and Kartik (2013) show that, when there are two potential projects, 1 and 2, that the agent can recommend, the agent will always (relative to truthful revelation), be biased toward project 1 if project 1 is *conditionally better looking*, i.e. if

$$\mathbb{E}[b_1 | b_1 > b_2] > \mathbb{E}[b_2 | b_2 > b_1].$$

This bias is interpreted as pandering. Che, Dessein, and Kartik consider the question of the direction of pandering, i.e., what sort of prior distributions will lead to one project being conditionally better looking than another? They state that the answer to this question is subtle.

Given a mild generalization of their framework, the stochastic order conditions developed in this paper can go a long way to answering this question. Interpret  $b$  not as the utility produced by the project but rather its monetary payoff, and assume that the principal is an expected utility maximizer endowed with expected utility function,  $v$ . Under this interpretation, the value of the project to a decision maker equals  $v(b)$ ,  $v$  strictly increasing. If, for any strictly increasing  $v$ ,

$$\mathbb{E}[v(b_1) | b_1 > b_2] > \mathbb{E}[v(b_2) | b_2 > b_1], \quad (17)$$

we will say that project 1 is *uniformly conditionally better looking* than project 2. Condition (17) is equivalent to project 1 competitive FSD dominating project 1. If the projects' distributions are ordered by geometric dominance which, as we show in Section 6.2, is a pervasive feature of textbook distributions, then Theorem 5 shows there are only three possible configurations in which project 1 is uniformly conditionally better looking than project 2. (a) Project 1

FSD dominates project 2, in which case, by Theorem 4, project 1 also MLR dominates project 2. (b) Project 2 FSD dominates project 1, in which case, by Theorem 3, the lower tail of project 1's payoff distribution explodes relative to project 2. (c) Project 1 is riskier, in the sense of one crossing from above of the compared distribution functions, than project 2.

Thus, if we assume that project 1 and 2 are ordered by FSD dominance and that the tail-ratio is bounded, then only alternative (a) remains. The condition that the projects are FSD ordered simply means that the principal has a strong *ex ante* preference for one of the projects. Bounded tail ratios appear to be a reasonable restriction on prior beliefs in cases where the principal does not believe that the probability of disastrous outcome is dramatically larger for one of the projects. Under these assumptions, geometric dominance ordering implies that the uniformly conditionally better looking project is always the project that the principal prefers *ex ante*. Thus, in this case, “pandering” as defined in Che, Dessein, and Kartik (2013) corresponds to our intuitive notion of pandering—recommending actions consistent with the principal's prior beliefs rather than the agent's superior information.

## 6 Applying the competitive selection relations

In this section, we develop a few of the basic properties of the competitive selection relations that are useful for characterizing competitive MLR and FSD dominance within specific families of parametric distributions.

*Remark 4.* Our first observation is fairly obvious: The competitive selection relations are invariant to common strictly increasing continuous transformations of the compared prospects. This is easy to see: suppose  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  is an admissible pair of prospects. Applying a strictly increasing continuous function,  $\phi$  to both prospects, would result in the admissible pair  $\phi(\tilde{X}) \stackrel{d}{\sim} F \circ \phi^{-1}$  and  $\phi(\tilde{Y}) \stackrel{d}{\sim} G \circ \phi^{-1}$ . The associated quantile transform function would equal  $F \circ \phi^{-1}(G \circ \phi^{-1})^{-1} = F \circ G^{-1} = u$ , the quantile transform function relating  $F$  and  $G$ . Since the competitive selection relations depend only on  $u$ , they would not be affected by the transformation. An immediate consequence of this observation is that the competitive selection orders cannot be characterized by restrictions on the tail properties, moments, or supports of compared distributions as such properties are not invariant to all continuous strictly increasing transformations of the compared prospects.

### 6.1 Upscalings and upshifts

Next, we develop simple criteria for geometric dominance within location/scale families of distributions: Suppose that  $G$  is a regular distribution function with support  $(-\infty, \infty)$ . 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, in which case,  $\tilde{X} \stackrel{d}{\sim} G$  implies that  $\tilde{X} + c \stackrel{d}{\sim} F$ . Similarly, suppose that  $G$  is regular distribution function with support  $[0, \infty)$ . 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, in which case,  $\tilde{X} \stackrel{d}{\sim} G$  implies that  $s\tilde{X} \stackrel{d}{\sim} F$ .<sup>22</sup>

It is clear that all upshifts (upscalings) of  $G$  strictly FSD dominate  $G$ . However, in general, upscalings and upshifts of distributions need not MLR or geometrically dominate the original distribution. The necessary and sufficient conditions for upshifts and upscalings to MLR dominate the original distribution are well known and provided below.

**Lemma 4** (Theorems 1.C.66 and Theorem 1.C.29 (Shaked and Shanthikumar, 2007)). *Let  $G$  be a regular distribution,*

- (i) *If the support of  $G$  is  $(-\infty, \infty)$ , all  $c$ -upshifts of  $G$  (strictly) MLR dominate  $G$  if and only if the density of  $G$ ,  $g$ , is (strictly) logconcave.*
- (ii) *If the support of  $G$  is  $[0, \infty)$ , all  $s$ -upscalings of  $G$  (strictly) MLR dominate  $G$  if and only if the density of  $G$ ,  $g$ , is (strictly) geometrically concave.*

The next result shows that, for geometric dominance, The necessary and sufficient conditions for upshifts and upscalings to geometrically dominate the original distribution are completely analogous to the MLR conditions: the same restrictions are imposed, but on the reversed hazard rate rather than the density of  $G$ .

**Lemma 5.** *Let  $G$  be a regular distribution,*

- (i) *If the support of  $G$  is  $(-\infty, \infty)$ , all  $c$ -upshifts of  $G$  (strictly) geometrically dominate  $G$  if and only if the reversed hazard rate function of  $G$ ,  $r = g/G$ , is (strictly) logconcave.*
- (ii) *If the support of  $G$  is  $[0, \infty)$ , all  $s$ -upscalings of  $G$  (strictly) geometrically dominate  $G$  if and only if the reversed hazard rate function of  $G$ ,  $r = g/G$ , is (strictly) geometrically concave.*

Because upscalings and upshifts always result in the strict FSD dominance of the upshifted (upscaled) distribution over the original distribution, the geometric dominance of the upshifted (upscaled) distribution implies its MLR dominance (Theorem 4). Consequently, the conditions imposed Lemma 5 for geometric dominance are strictly stronger than the conditions imposed for MLR dominance.<sup>23</sup>

<sup>22</sup>At this point, a reasonable question the reader might ask is how we can characterize competitive selection dominance with criteria based only the properties of an upscaled or upshifted distribution. Remark 4 shows that dominance under the competitive selection relations is invariant to a strictly increasing continuous transformations of the compared prospects but, obviously, distribution functions are not invariant to such transformations. Such transformations will, indeed, preserve competitive selection dominance but, after the transformation, the compared distributions will no longer, in general, be related by an upshift or an upscaling. In order to retain the upshift/upscale relation between the prospects, the common transformation must be increasing and linear (for upscaling) or a translation (for upshifts). All characterizations provided below are invariant to these sorts of transformations.

<sup>23</sup>The fact that the conditions for geometric dominance in Lemma 5 are *strictly* stronger than the conditions for MLR dominance in Lemma 4 follows from examples B.2 and B.3 in Supplement B, which show that  $G$  can verify

<i>A. Ordered consistently by MLR and competitive MLR dominance</i>	
Normal, Lognormal, Logistic, Log-logistic, Gamma, Pareto, Exponential, Generalized Exponential, Weibull, Rayleigh, Chi-squared, Kumaraswamy ( $b > 1$ ), Gompertz ( $\eta \geq 1$ )	
<i>B. Ordered by MLR dominance but competitive selection equivalent</i>	
Gumbel, Fréchet, Kumaraswamy ( $b = 1$ )	
<i>C. Ordered by MLR dominance but reverse ordered by competitive MLR dominance</i>	
Kumaraswamy ( $b < 1$ )	
<i>D. Ordered consistently by MLR dominance and competitive FSD dominance</i>	
Laplace	
<i>E. Not ordered by MLR</i>	
Gompertz ( $\eta < 1$ )	

Table 2: Dominance relations for parametric distributions.

## 6.2 Distributions

In Table 2 we present the relations between MLR dominance and competitive MLR and FSD dominance for prospects ordered by MLR dominance whose distributions belong to the same parametric family, e.g., Normally distributed prospects with the same variance. We report the strongest dominance relation satisfied. Detailed descriptions of the parameterizations of the distributions are provided by Table D.1 in Supplement D.

The most striking observation about Table 2 is that most of the distribution families in the table are ordered consistently by the MLR and competitive MLR dominance. For these distributions, the bias introduced by competitive selection is not sufficient to reverse MLR dominance.

Most of the exceptional cases are quite easy to rationalize. When  $\eta < 1$ , Gompertz distributions are not even MLR ordered, so the question of consistency is moot. Competitive selection equivalence for Gumbel, Fréchet, and power law distributions (i.e., Kumaraswamy distribution when  $b = 1$ ) was discussed in Section 5.3. In the case of the Laplace distributions, the MLR dominant distribution, although not competitive MLR dominant, is competitive FSD dominant. So, selection does not completely reverse dominance. The only case where selection reverses MLR (and FSD) dominance is the Kumaraswamy distribution when the Kumaraswamy's shape parameter,  $b$ , is less than 1. The rationale behind reversal in this case is the same as the rationale provided in Section 5.1.3.

---

the hypotheses of Lemma 4 but not the hypotheses of Lemma 5.

### 6.3 Application: The implications of stock-price announcement effects

Numerous studies have attempted to test hypotheses about the value effects of corporate policy choices using stock-price announcement returns. Most of these studies condition market responses on information publicly available at the time of announcement, e.g., in the case of acquisitions, the size of the firm (Moeller, Schlingemann, and Stulz, 2004) or the composition of the firm's board (Byrd and Hickman, 1992).

In this section, we consider the problem of inferring the average effects of corporate choices, conditioned on public information, from the selection-conditioned valuation effects of the choices. Consider the choice between two policy alternatives, or prospects,  $\tilde{X}_1$  and  $\tilde{X}_2$ , with the following stochastic structure:

$$\begin{aligned}\tilde{X}_1 &= \tilde{W}_1 + \tilde{\varepsilon}_1, \\ \tilde{X}_2 &= \tilde{W}_2 + \tilde{\varepsilon}_2,\end{aligned}$$

where  $\tilde{W} = (\tilde{W}_1, \tilde{W}_2)$  represents public information about the competing choices and  $\tilde{\varepsilon} = (\tilde{\varepsilon}_1, \tilde{\varepsilon}_2)$  represents idiosyncratic private information that cannot be verified or easily communicated. An optimizing firm, after observing  $\tilde{X}_1$  and  $\tilde{X}_2$ , announces its choice between the two prospects. At the time of announcement  $\tilde{W} = (\tilde{W}_1, \tilde{W}_2)$  is public information. Based on  $\tilde{W}$  and the firm's choice, market participants will revise their valuation of the firm. The market's announcement-conditioned valuations, or *announcement valuations*, when the firm chooses  $\tilde{X}_1$ ,  $\mathcal{A}_1$ , or  $\tilde{X}_2$ ,  $\mathcal{A}_2$ , will be given by

$$\begin{aligned}\mathcal{A}_1(\tilde{W}_1, \tilde{W}_2) &= \mathbb{E}[v(\tilde{W}_1 + \varepsilon_1) | \tilde{W}_1 + \varepsilon_1 > \tilde{W}_2 + \varepsilon_2, \tilde{W}_1, \tilde{W}_2], \\ \mathcal{A}_2(\tilde{W}_1, \tilde{W}_2) &= \mathbb{E}[v(\tilde{W}_2 + \varepsilon_2) | \tilde{W}_2 + \varepsilon_2 > \tilde{W}_1 + \varepsilon_1, \tilde{W}_1, \tilde{W}_2].\end{aligned}$$

Under the assumption of risk-neutral valuation and rational expectations, higher (lower) announcement valuations imply positive (negative) announcement returns.

We assume that  $\tilde{W} = (\tilde{W}_1, \tilde{W}_2)$  has a joint density given by  $h$  but do not assume that  $\tilde{W}_1$  and  $\tilde{W}_2$  are independent of each other. We do assume that  $\tilde{W}$  is independent of  $\tilde{\varepsilon}$  and that  $\tilde{\varepsilon}_1$  and  $\tilde{\varepsilon}_2$  are zero mean i.i.d. random variables supported by the real line.

Consider a hypothesis which predicts that, conditioned on observable characteristics,  $\tilde{X}_1$  is stochastically better than  $\tilde{X}_2$ , i.e.  $\tilde{W}_1$  dominates  $\tilde{W}_2$  under some stochastic order at least as strong as the FSD order, then ignoring the problem of selection and assuming the hypothesis is correct, the predicted announcement value under choice  $\tilde{X}_1$  ( $\tilde{X}_2$ ) would equal  $\mathbb{E}[v(\tilde{X}_1)|W]$  ( $\mathbb{E}[v(\tilde{X}_2)|W]$ ). Because  $\tilde{\varepsilon}$  is independent of  $\tilde{W}$ , announcement values would equal

$$\mathcal{A}_1^{\text{ns}}(\tilde{W}_1) = \int_{-\infty}^{\infty} v(\tilde{W}_1 + \varepsilon_1) f(\varepsilon_1) d\varepsilon_1 \quad \mathcal{A}_2^{\text{ns}}(\tilde{W}_2) = \int_{-\infty}^{\infty} v(\tilde{W}_2 + \varepsilon_2) f(\varepsilon_2) d\varepsilon_2.$$

respectively. Because FSD dominance is preserved when the compared distributions are shocked by an identically distributed independent noise term (Theorem 1.A.3: Shaked and Shanthikumar, 2007),  $\tilde{W}_1 + \varepsilon_1$  would FSD dominate  $\tilde{W}_2 + \varepsilon_2$ . And thus  $\mathbb{E}[\mathcal{A}_1^{\text{ns}}(\tilde{W}_1)] > \mathbb{E}[\mathcal{A}_2^{\text{ns}}(\tilde{W}_2)]$ , i.e.,



ignoring selection, the hypothesis predicts that expected announcement value conditioned on choosing  $\tilde{X}_1$  will be higher than the expected announcement value conditioned on choosing  $\tilde{X}_2$ . Thus, the hypothesis that  $\tilde{W}_1$  dominates  $\tilde{W}_2$  would be falsified by non-positive announcement returns. Hence, ignoring selection, announcement returns seem to be an ideal way to test hypotheses about the average effects of policy choices on firm value.

If we account for selection, can such hypotheses about average unconditional effects ever be verified or falsified by announcement returns? The following result provides a fairly complete answer to this question.

**Proposition 4.**

- (i) *If*
  - (a) *the reversed hazard rate of the distribution of the error terms  $\tilde{\epsilon}_1$  and  $\tilde{\epsilon}_2$  is strictly logconcave, and*
  - (b)  *$h$ , the joint density of public information,  $\tilde{W}$ , satisfies  $x > y \Rightarrow h(x, y) \geq h(y, x)$  and  $h(x, y) \neq h(y, x)$  on a set of positive Lebesgue measure,**then the expected announcement valuation conditioned on the selection of  $\tilde{X}_1$  is greater than the expected announcement valuation conditioned on the selection of  $\tilde{X}_2$ , i.e.,*

$$\mathbb{E}[\mathcal{A}_1(\tilde{W}_1, \tilde{W}_2)] > \mathbb{E}[\mathcal{A}_2(\tilde{W}_1, \tilde{W}_2)].$$
- (ii) *If the reversed hazard rate of the distribution of the error terms is log-linear, then the announcement valuations,  $\mathcal{A}_1$  and  $\mathcal{A}_2$ , are identically distributed and thus  $\mathbb{E}[\mathcal{A}_1(\tilde{W}_1, \tilde{W}_2)] = \mathbb{E}[\mathcal{A}_2(\tilde{W}_1, \tilde{W}_2)]$ .*

Hypothesis (a) is simply the condition in Lemma 5.i for upshifts to engender strict geometric dominance. Condition (b) is a stochastic ordering condition for dependent random vectors analogous to the MLR order (Shanthikumar and Yao, 1991). In fact, when  $\tilde{W}_1$  and  $\tilde{W}_2$  are independent, condition (b) reduces to MLR ordering.

Thus, part (i) of the proposition shows that when one prospect dominates another *and* the error law satisfies condition (a) (e.g., the error law has a Normal or Logistic distribution), the hypothesis that one choice is stochastically dominant unconditionally (in the sense defined by condition (b) of the proposition) predicts positive announcement effects for that choice when announcement returns are conditioned on selection and observable information. However, part (ii) shows that this prediction is sensitive to the choice of error law. Under error laws with log-linear reversed hazard rates, e.g., the Gumbel error law, selection renders announcement returns uninformative.

## 7 Robustness

This paper has focused on binary choices between independent prospects. In this section, we analyze the robustness of these results to non-binary prospect choice and dependence. In the

interest of brevity, we assume that all distributions considered are regularly related.

## 7.1 Multiple prospects and non-optimizing selection

In life, and in many economic problems, agents frequently have more than two choices. Thus, a natural question is the extent to which our results can be extended to non-binary selection competitions between prospects. Below we show that when selection is optimizing selection, i.e., the selection of the best prospect, pairwise geometric dominance characterizes competitive MLR dominance even when more than two prospects are candidates for selection.<sup>24</sup>

This result is fairly easy to develop. Consider  $n \geq 3$  prospects  $\{\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n\}$  with corresponding distribution functions  $\{F_1, F_2, \dots, F_n\}$ . Let  $\tilde{X}^{-(i)}$  represent the vector of  $n - 1$  prospects obtained by excluding prospect  $i$ , and let  $\tilde{X}^{-(i,j)}$  be the vector of  $n - 2$  prospects obtained by excluding both prospects  $i$  and  $j$ . In this setting, the selection-conditioned distributions of any two prospects,  $i$  and  $j$ , are given by

$$[\tilde{X}_i \leq x | \tilde{X}_i > \max(\tilde{X}^{-(i)})], \quad [\tilde{X}_j \leq x | \tilde{X}_j > \max(\tilde{X}^{-(j)})]$$

respectively. Note that, for any choice of  $i$  and  $j$ ,

$$\begin{aligned} \mathbb{P}[x > \max(\tilde{X}^{-(i)})] &= \mathbb{P}[x > \max(\tilde{X}^{-(i,j)})] \mathbb{P}[x > \tilde{X}_j] = H_{\max}^{-(i,j)}(x) F_j(x), \text{ where,} \\ H_{\max}^{-(i,j)}(x) &= \mathbb{P}[x > \max(\tilde{X}^{-(i,j)})] = \prod_{k \neq i,j} F_k(x). \end{aligned} \quad (18)$$

Thus, the selection-conditioned distribution of prospect  $i$  can be obtained from equations (5) and (11) by replacing  $G$  in that formula with  $\mathbb{P}[x > \max(\tilde{X}^{-(i)})]$  and the selection-conditioned distribution of prospect  $j$  can be obtained from equations (5) and (11) by replacing  $F$  in that formula with  $\mathbb{P}[x > \max(\tilde{X}^{-(j)})]$ . Performing these substitutions and using expression (18) yields

$$\begin{aligned} F_i^{\text{sc}}(x) &= \frac{\int_{-\infty}^x F_j(z) H_{\max}^{-(i,j)}(z) f_i(z) dz}{\int_{-\infty}^{\infty} F_j(z) H_{\max}^{-(i,j)}(z) f_i(z) dz}, \quad F_j^{\text{sc}}(x) = \frac{\int_{-\infty}^x F_i(z) H_{\max}^{-(i,j)}(z) f_j(z) dz}{\int_{-\infty}^{\infty} F_i(z) H_{\max}^{-(i,j)}(z) f_j(z) dz}, \\ f_i^{\text{sc}}(x) &= \frac{f_i(x) F_j(x) H_{\max}^{-(i,j)}(x)}{\int_{-\infty}^{\infty} F_j(z) H_{\max}^{-(i,j)}(z) f_i(z) dz}, \quad f_j^{\text{sc}}(x) = \frac{f_j(x) F_i(x) H_{\max}^{-(i,j)}(x)}{\int_{-\infty}^{\infty} F_i(z) H_{\max}^{-(i,j)}(z) f_j(z) dz}. \end{aligned}$$

The constants in the denominators of the density expression will not affect the monotonicity of the density ratios. Thus, we see that the selection-conditioned distribution  $\tilde{X}_i$  will MLR dominate the selection-conditioned distribution of  $\tilde{X}_j$  if and only if

$$x \mapsto \frac{f_i^{\text{sc}}(x)}{f_j^{\text{sc}}(x)} = \frac{f_i(x) F_j(x) H_{\max}^{-(i,j)}(x)}{f_j(x) F_i(x) H_{\max}^{-(i,j)}(x)} = \frac{f_i(x) F_j(x)}{f_j(x) F_i(x)} = \frac{r_i(x)}{r_j(x)} \text{ is increasing.} \quad (19)$$

In equation (19), the distributions of the prospects not included in the comparison cancel out,

<sup>24</sup>Not surprisingly, given that, as pointed out in footnote 21, average elasticity dominance is not even a transitive relation, pairwise average elasticity dominance relations do not characterize competitive FSD dominance relations between multiple prospects.

ensuring that the pairwise geometric dominance condition in Theorem 2 remains necessary and sufficient for competitive MLR dominance.

When selection is not optimizing, i.e., is based on ranks less than the highest rank, the selection-conditioning event is different from the highest-rank conditioning event in our analysis. Thus, unsurprisingly, the average elasticity and geometric dominance conditions, which characterize selection-conditioned distributions under optimizing selection, do not characterize the effect of conditioning on lower ranks. In principle, analogous but different ordering conditions could be developed for selection conditioned on attaining ranks other than the highest rank.

However, in general, the order statistics which determine the selection event will be polynomial functions of the distributions of the compared prospects, and thus, the neat cancellation used in equation (19) to extend geometric dominance to non-binary prospect choices will not be valid. Consequently, in general, the analogous conditions will depend not only on the rank defining selection, but also on the number of compared prospects, and pairwise comparisons will not suffice to determine competitive selection dominance.

The exception is *minimal-rank selection*, i.e., selection conditioned on attaining the lowest rank. Minimal-rank selection does generate a different but very tractable pairwise order relation which is invariant to the number of prospects considered and characterizes MLR dominance conditioned the compared prospects attaining the minimal rank. To see this, note that, under minimal-rank selection, a draw from the  $i$ th prospect's distribution will be selected if and only if  $\{x < \min(\tilde{X}^{-i})\}$ . The probability of this event is just the product of the complementarity distribution functions for all prospects excluding  $i$ . Thus, in this case, the order statistics which determine the selection event will be monomial functions of the distributions of the compared prospects. Consequently, a factor cancellation similar to equation (19) will yield a simple necessary and sufficient condition for MLR dominance conditioned on minimal-rank selection.

## 7.2 Relaxing the independence assumption

Thus far, we have assumed that prospects are independent. In some cases, assuming independence is not problematic. When compared prospects are generated by a mixed strategy equilibrium (as in Section 5.1.1) or are functions of independent random variables (as in Supplement C.2) they are, by definition, independent. Independence between the unobservable error terms is also assumed by the standard Logit and Probit probabilistic choice models (Chapt. 15: Wooldridge, 2002) (analyzed in Supplement C.3). In other settings, such as the tournament and cheap-talk models (Sections 5.2.1 and 5.5.1), randomness is produced by exogenous random shocks affecting the outcomes of endogenous choices. To our knowledge, all such models assume that these random shocks are independent.

However, in some contexts, such as announcement effects analyzed in Sections 6.3 and treatment effects analyzed in Supplement C.4, questions of dependence are more salient. How will dependence affect our analysis? Theorems 1 and 2 show that, when prospects are independent, the competitive selection relations characterize competitive selection dominance. Because the competitive selection relations depend only on the marginal distributions of the compared prospects, for arbitrary dependence structures, these competitive selection relations cannot characterize competitive FSD or MLR dominance.

Thus, our robustness analysis will consider whether the competitive selection relations are sufficient conditions for competitive selection dominance under some commonly employed dependence structures. If under a given dependence structure, average elasticity (geometric dominance) implies competitive FSD (MLR) dominance, we will say that average elasticity (geometric) dominance is *robust to dependence*.

The simplest commonly employed dependence structure results when both prospects are affected by a common random shock, i.e., that the stochastic structure of the prospects  $\tilde{X}$  and  $\tilde{Y}$  is given by

$$\tilde{X} = \tilde{x} + \tilde{z}, \quad \tilde{Y} = \tilde{y} + \tilde{z}, \quad (20)$$

where  $\tilde{x}$ ,  $\tilde{y}$ , and  $\tilde{z}$  are jointly independent. In economics, common shock models are pervasive, e.g., common-value auction and rational expectations models, models of security market micro-structure. By increasing the relative size of the common component,  $\tilde{z}$ , it is possible to generate an arbitrarily large degree of dependence between the prospects. Average elasticity dominance is always robust to this sort of dependence. The argument for robustness is simple. Remark 4 observed that average elasticity dominance is preserved by common, strictly increasing, continuous transformations of the compared prospects. Thus, for all  $z \in (-\infty, \infty)$ ,  $\tilde{x}$ 's distribution average elasticity dominating  $\tilde{y}$ 's distribution implies that  $\tilde{x} + z$ 's distribution average elasticity dominates  $\tilde{y} + z$ 's distribution. This implies (Theorem 1) that the selection-conditioned distribution of  $\tilde{x} + z$ ,  $F_z^{\text{sc}}$ , FSD dominates the selection-conditioned distribution of  $\tilde{y} + z$ ,  $G_z^{\text{sc}}$ . The selection-conditioned distributions of  $\tilde{X}$ , and  $\tilde{Y}$ , are given by

$$F^{\text{sc}}(x) = \int F_z^{\text{sc}}(x) dJ(z), \quad G^{\text{sc}}(x) = \int G_z^{\text{sc}}(x) dJ(z),$$

respectively, where  $J$  is the distribution function of  $\tilde{z}$ . FSD dominance is preserved by mixtures (Theorem 1.A.3: Shaked and Shanthikumar, 2007). Thus,  $F^{\text{sc}}$  FSD dominates  $G^{\text{sc}}$ . Consequently, if  $\tilde{x}$ 's distribution average elasticity dominance  $\tilde{y}$ 's distribution,  $\tilde{X}$  competitive FSD dominates  $\tilde{Y}$ .

The result developed for additive shocks can be extended to multiplicative common shocks and, in fact, to any shock whose functional relationship to the compared prospects is increasing, continuous, and sufficiently regular to permit the calculation of conditional probabilities. Note that an analogous result for geometric dominance cannot be verified because MLR dominance

is not generally preserved by mixtures.

One interesting implication of this result is that, regardless of their degree of dependence, when the dependent prospects are Bivariate Normal random variables with the same variance, average elasticity dominance is always robust to dependence, i.e., FSD dominance implies competitive FSD dominance. This follows because (a) unconditional marginal distributions of the Bivariate Normal are (univariate) Normal and the upshift of a Normal distribution geometrically (and thus *a fortiori* average elasticity) dominates the original distribution (Table 2) and (b) a defining feature of Bivariate Normal distribution is that, for any degree of imperfect correlation, there exist random variables satisfying (20) whose joint distribution is Bivariate Normal.

Outside the common shock setting, the effects of dependence are much more subtle.<sup>25</sup> However, insight into the robustness of the competitive selection orders can be obtained from first examining selection conditioning under perfect dependence, i.e., assuming the compared prospects are perfectly positively dependent (co-monotone dependence) or perfectly negatively dependent (counter-monotone dependence). In these cases, both prospects are functions of the same underlying random variable. Perhaps surprisingly, perfect negative dependence produces the most enlightening results. Results that hearken back to the example in Section 2.

The driver of dominance reversal in the example in Section 2 was the admission effect: unselected zero draws from the stochastically dominated prospect's distribution permitted low draws from the stochastically dominant prospect's distribution to be incorporated into the dominant prospect's selection-conditioned distribution, lowering its stochastic quality. Under perfect negative dependence, the admission effect is turned off: low draws from the stochastically dominated prospect's distribution are matched with high draws from the stochastically dominant prospect's distribution. Turning off the admission effect turns off dominance reversal. Thus, an MLR dominant prospect is always competitive FSD dominant.

This result is quite intuitive: Suppose  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are the compared prospects and that these prospects are perfectly negatively dependent. Because  $G$  is strictly increasing and  $1 - F$  is strictly decreasing over the distributions' support, there exists a unique value of  $x$ ,  $x^*$  such that

$$1 - F(x^*) = G(x^*), \quad x^* \in (\underline{x}, \bar{x}). \quad (21)$$

Under perfect negative dependence, draws less than  $x^*$  are never selected. If a draw from one prospect's distributions is less than  $x^*$ , the draw from the compared prospect's distribution exceeds  $x^*$ . Thus, low draws from the prospects' distributions never affect their selection-conditioned distributions. Dominance hinges simply on which prospect's high draws are better.

<sup>25</sup>To simplify the analysis of these more complex settings, we will assume the probability of ties is zero, i.e.,  $\mathbb{P}[\tilde{Y} = \tilde{X}] = 0$ . Except in the case of perfect dependence, the satisfaction of this condition is assured by the assumption of regularity.

Because the MLR dominant distribution is uniformly “better” than the MLR dominated distribution, its selection-conditioned distribution is always FSD dominant.

**Lemma 6.** *Suppose that  $\tilde{X} \stackrel{d}{\sim} F$  and  $\tilde{Y} \stackrel{d}{\sim} G$  are perfectly negatively dependent (counter-monotone) random variables. Then*

(i)  *$\tilde{X}$  competitive FSD dominates  $\tilde{Y}$  if and only if*

$$\frac{F(x) - F(x^*)}{G(x) - G(x^*)} \leq \frac{G(x^*)}{F(x^*)}, \quad x \in (x^*, \bar{x}),$$

*where  $x^*$  is defined by equation (21).*

(ii)  *$F$  MLR dominating  $G$  is a sufficient condition for (i).*

Lemma 6 suggests that, when the FSD dominated distribution is average elasticity dominant, and thus, under independence, selection reverses dominance, imposing a high but imperfect degree of negative dependence between the prospects might annul the competitive FSD dominance of the FSD dominated prospect. This intuition is confirmed in Supplement B.4 where we provide an example.

The analysis of perfect positive dependence is somewhat less enlightening. If two compared prospects are perfectly positively dependent, then both can be expressed as increasing functions of the same random variable. Thus, if one prospect strictly FSD dominates the other, the FSD dominant prospect is selected with probability 1. Hence, the selection-conditioned distribution of the FSD dominated prospect is undefined and the selection-conditioned distribution of the FSD dominant prospect is its unconditional distribution.<sup>26</sup>

Although perfect positive dependence provides little direct insight into the effects of positive dependence on selection, the basic fixed vs. competitive selection intuition, which has proved a good guide to the effects of selection thus far, does suggest that the effect of positive dependence on competitive selection dominance should depend on the quantiles of the compared distributions at which the dependence is most pronounced. Positive dependence at low quantiles amplifies the admission effect—low draws from the FSD dominated prospect’s distribution, which are unlikely to be selected, are matched with low draws from the FSD dominant prospect’s distribution, admitting these low draws into the FSD dominant prospect’s selection-conditioned distribution. This effect favors dominance reversal. In contrast, positive dependence at high quantiles implies that high draws from the FSD dominant prospect’s distribution, which are likely to be selected, are matched with high draws from the FSD dominated prospect’s distribution, and thus are likely to exclude these high draws from the FSD dominated prospect’s selection-conditioned distribution. This effect favors dominance preservation.

Verifying this intuition requires formally considering imperfect dependence. We will focus

<sup>26</sup>When conditional beliefs are undefined, it is sometimes useful to define conditional beliefs as the limit of beliefs generated by a sequence of approximating distributions (e.g., Kreps and Wilson, 1982). However, this approach is not very fruitful here: arbitrarily small distributional perturbations of perfectly positively dependent prospects can make either prospect MLR dominant. See Supplement A.6 for a detailed discussion.

our analysis on positive dependence because its effects appear to be more difficult to parse and, judging from the applied literature's focus, of much greater practical relevance. We also assume that the marginal distributions of the prospects are strictly ordered by both geometric dominance and MLR dominance. Hence, under independence, the prospects are MLR and competitive MLR ordered (Theorem 2).

We assume, without loss of generality, that  $F$ , the marginal distribution of  $\tilde{X}$ , strictly geometrically dominates  $G$ , the marginal distribution of  $\tilde{Y}$ . We consider the robustness of geometric dominance, i.e., whether geometric dominance of  $F$  implies that  $\tilde{X}$  is competitive MLR dominant.

We impose dependence using the most commonly used approach to imposing non-linear dependence on marginal distributions: the copula approach. Under copula dependence, the joint distribution of the two prospects,  $H$ , is given by

$$H(x, y) = C(F(x), G(y)),$$

where  $C$  is a copula function and  $F$  and  $G$  represent the marginal distributions of the prospects. A copula function  $C : [0, 1]^2 \rightarrow [0, 1]$  is a function that satisfies the following conditions:

For all  $u, v \in [0, 1]$ ,  $C(0, v) = C(u, 0) = 0$ ,  $C(u, 1) = C(1, v) = 1$ ,

$$0 \leq u_1 \leq u_2 \leq 1 \text{ and } 0 \leq v_1 \leq v_2 \leq 1, C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.^{27}$$

In our case, where marginal distributions are continuous, any joint distribution can be represented in the copula framework (Sklar, 1959). However, we will restrict the range of dependence relations we consider somewhat by restricting our analysis to copulas that are continuously twice differentiable over  $(0, 1)^2$  and thus have densities.

As shown in Supplement A.6, the selection-conditioned densities of  $\tilde{X}$ ,  $f^{\text{sc}}$ , and  $\tilde{Y}$ ,  $g^{\text{sc}}$ , are given by

$$f^{\text{sc}}(x) = \frac{f(x) D_1 C(F(x), G(x))}{\mathbb{P}[\tilde{X} > \tilde{Y}]}, \quad g^{\text{sc}}(x) = \frac{g(x) D_2 C(F(x), G(x))}{\mathbb{P}[\tilde{Y} > \tilde{X}]}, \quad x \in (\underline{x}, \bar{x}),$$

where  $D_i$  represents the derivative of the copula with respect to its  $i$ -th argument. Thus, the likelihood ratio between the selection-conditioned distributions under copula dependence,  $f^{\text{sc}}/g^{\text{sc}}$ , is proportional to  $\mathcal{R}_\theta$ , defined below:

$$\mathcal{R}(x) = \frac{r_F(x)}{r_G(x)} \Psi(F(x), G(x)), \quad \text{where } \Psi(u, v) = \frac{u D_1 C(u, v)}{v D_2 C(u, v)}. \quad (22)$$

In equation (22),  $r_F/r_G$  is the ratio of reversed hazard rates of the marginal distributions  $F$  and  $G$  and thus,  $F$ 's geometric dominance over  $G$  implies, by Theorem 2, that  $r_F/r_G$  is strictly increasing. If  $x \mapsto \Psi(F(x), G(x))$  is also increasing then,  $\mathcal{R}$  will be strictly increasing and

<sup>27</sup>Unfortunately, the standard notation for the first and second arguments of the copula function,  $u$  and  $v$ , clashes with our use of  $u$  and  $v$  to represent the quantile transform function and generic payoff functions respectively. The notational collision should not produce any confusion because, in this section, we do not use the quantile transform function or represent any payoff functions. In other sections of the paper, we do not use copula functions.

geometric dominance will be robust to dependence. Hence, the following result is immediate.

**Lemma 7.** *Let  $\tilde{X}$  and  $\tilde{Y}$  be two prospects with marginal distributions  $F$  and  $G$  respectively, such that  $F$  strictly geometrically dominates  $G$ . Suppose the joint distribution of  $\tilde{X}$  and  $\tilde{Y}$  is given by  $H(x, y) = C(F(x), G(y))$ . Then, if*

$$x \mapsto \Psi(F(x), G(x)) \text{ is increasing,} \quad (23)$$

*$\tilde{X}$  competitive MLR dominates  $\tilde{Y}$ , i.e., geometric dominance is robust to dependence.*

An example of a copula satisfying equation (23) is provided by the Gumbel copula, defined by

$$C_{\alpha}^{\text{Gum.}}(u, v) = \exp \left[ -((-\log(u))^{\alpha} + (-\log(v))^{\alpha})^{1/\alpha} \right], \quad \alpha \geq 1.$$

As pointed out by Beare (2010), the Gumbel copula is quite widely used in applied economics and finance research. When  $\alpha = 1$ ,  $\tilde{X}$  and  $\tilde{Y}$  are independently distributed. As  $\alpha \rightarrow \infty$ , dependence approaches perfect positive dependence.

For the Gumbel copula  $\Psi_{\alpha}^{\text{Gum.}}$  is increasing in  $x$  when  $F$  strictly geometrically dominates  $G$ .<sup>28</sup> Thus, under the Gumbel copula, geometric dominance is robust at all levels of imperfect positive dependence.

This result should not be too surprising given the intuition developed in the discussion of positive dependence. This discussion suggested that robustness hinges on the relative effect of dependence on the upper and lower tails of the compared distributions. The Gumbel copula exhibits no asymptotic lower-tail dependence and positive asymptotic upper-tail dependence.

These observations suggest that positive dependence is most likely to produce fragility when the copula's tail dependencies are the reverse of the Gumbel's. One copula which can have this property is the Ali-Mikhail-Haq (AMH) copula (Ali, Mikhail, and Haq, 1978) defined as follows:

$$C_{\theta}^{\text{AMH}}(u, v; \theta) = \frac{uv}{1 - (1 - u)(1 - v)\theta}, \quad \theta \in [-1, 1]. \quad (24)$$

The AMH copula, like the Gumbel copula, is widely used in applied dependence modeling. Like the Gumbel copula, but unlike most copulas used in applied research, the AMH produces tractable closed-form representations of the selection-conditioned densities of the compared prospects. When  $\theta = 1$  (and only for  $\theta = 1$ ) the AMH copula exhibits positive asymptotic lower-tail dependence and no asymptotic upper-tail dependence, and, in this case,  $\Psi_1^{\text{AMH}}(F(x), G(x)) = G(x)/F(x)$ .

If  $F$  is strictly MLR dominant, i.e., dominance is preserved under independence, the MLR condition implies that  $x \mapsto G(x)/F(x)$  is decreasing. Thus,  $\tilde{X}$  might or might not competitive MLR dominate  $\tilde{Y}$ , depending on the rate of increase of  $r_F/r_G$  versus the rate of decrease of  $G/F$ . For some distribution families, e.g., the Exponential, Pareto, Weibull, and Rayleigh,

<sup>28</sup>See Supplement Lemma A-7 in Supplement A.6.



geometric dominance is robust to positive asymptotic lower-tail dependence under the AMH copula when  $\theta = 1$ .<sup>29</sup> For other families, it is not.

In contrast, if  $F$ , is strictly MLR dominated, i.e., dominance is reversed under independence,  $G/F$  is increasing (see Theorem 3). Thus, Lemma 7 shows that  $\tilde{X}$  competitive MLR dominates  $\tilde{Y}$ . Hence, dominance reversal is always robust to positive asymptotic lower-tail dependence under the AMH copula when  $\theta = 1$ .

To provide a bit more insight into the fragility of geometric dominance to positive dependence when dominance is preserved under independence, we choose marginal distributions under which the AMH copula when  $\theta = 1$  breaks competitive MLR dominance—two Normal Distributions with the same variance, and plot  $\mathcal{R}$ , which is proportional to the selection-conditioned likelihood ratio, under the AMH and under the standard Gaussian copula, varying the dependence parameter,  $\theta$ .<sup>30</sup>

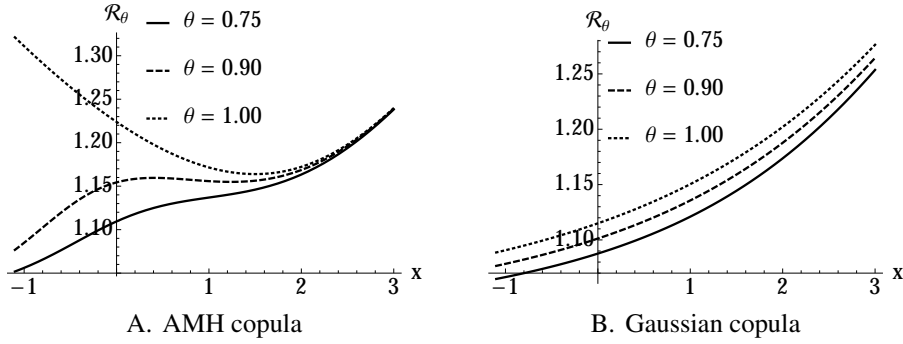


Figure 3: *The effect of positive dependence on competitive MLR dominance under the Gaussian and AMH copulas.* The marginal distribution of  $\tilde{X}$  is Normal[2, 1] and the marginal distribution of  $\tilde{Y}$  is Normal[1.9, 1]. Competitive MLR dominance is equivalent to  $\mathcal{R}_\theta$  being increasing.  $\theta$  is the dependence parameter associated with the AMH copula.

In Figure 3, we see that, under the AMH copula, for very high values of  $\theta$ ,  $\mathcal{R}_\theta$  is not increasing, and thus geometric dominance is not robust. In contrast, under the Gaussian copula, i.e., when the prospects' joint distribution is Bivariate Normal, geometric dominance is completely robust to positive dependence.

In summary, we cannot provide a complete characterization of the effects of all forms of dependence on the robustness of geometric dominance. However, these results suggest that, typically, when dominance is preserved under independence, geometric dominance will imply competitive MLR dominance when dependence is either negative or not disproportionately concentrated in the lowest quantiles of the compared distributions. When dominance is reversed

<sup>29</sup>See Supplement Lemma A-8 in Supplement A.6.

<sup>30</sup>In the figure, we match the correlation coefficients defining the Gaussian copulas with the three  $\theta$  parameters defining the AMH copulas as follows: Kumar (Theorem 1: 2010) is used to determine the Kendall's Tau,  $\tau$ , associated each of the three  $\theta$  parameters; then the correlation coefficients are fixed to match each  $\tau$  using Table 1 in Edwardes (1993).

under independence, geometric dominance will imply competitive MLR dominance when dependence is positive.

## 8 Conclusion

This paper developed competitive selection relations between unconditional prospect distributions, relations that identify dominance relationships conditioned on selection. One of these relations, geometric dominance, is easy to verify and, in many ways, resembles monotone likelihood ratio dominance. These relations were then applied to determine the conditions under which the bias introduced by competitive selection is so large that it reverses dominance. The analysis showed that the necessary conditions for such reversals are quite strong and that, for most parametric textbook distributions, selection bias does not reverse dominance. The utility of this analysis was illustrated by applying it to a sample of economic selection problems. Because selection questions frequently arise in economics research, and many of the relations and characterizations developed in this paper are quite tractable, the scope for their application is broad and by no means exhausted by this sample.

## References

- Ali, Mir M, NN Mikhail, and M Safiul Haq. 1978. "A class of bivariate distributions including the bivariate logistic." *Journal of Multivariate Analysis* 8 (3):405–412.
- Anas, Alex and Cheng Min Feng. 1988. "Invariance of expected utilities in logit models." *Economics Letters* 27 (1):41–45.
- Athey, Susan. 2002. "Monotone comparative statics under uncertainty." *Quarterly Journal of Economics* 117 (1):187–223.
- Baye, Michael R., Dan Kovenock, and Casper G. de Vries. 1996. "The all-pay auction with complete information." *Economic Theory* 8 (2):291–305.
- Beare, Brendan K. 2010. "Copulas and temporal dependence." *Econometrica* 78 (1):395–410.
- Byrd, John W and Kent A Hickman. 1992. "Do outside directors monitor managers?: Evidence from tender offer bids." *Journal of Financial Economics* 32 (2):195–221.
- Cargo, GT. 1965. "Comparable means and generalized convexity." *Journal of Mathematical Analysis and Applications* 12 (3):387–392.

- Chan, Wai, Frank Proschan, and Jayaram Sethuraman. 1990. "Convex-ordering among functions, with applications to reliability and mathematical statistics." *Lecture Notes-Monograph Series* 16:121–134.
- Che, Yeon-Koo, Wouter Dessein, and Navin Kartik. 2013. "Pandering to persuade." *American Economic Review* 103 (1):47–79.
- de Palma, André and Karim Kilani. 2007. "Invariance of conditional maximum utility." *Journal of Economic Theory* 132 (1):137–146.
- Edwardes, Michael. 1993. "Kendall's  $\tau$  is equal to the correlation coefficient for the BVE distribution." *Statistics & Probability Letters* 17 (5):415–419.
- Ghosal, Subhashis, Arusharka Sen, and Aad W Van Der Vaart. 2000. "Testing monotonicity of regression." *Annals of Statistics* 28 (4):1054–1082.
- Heckman, James J. 2008. "Selection bias and self-selection." In *The New Palgrave Dictionary of Economics*, edited by Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.
- Jewitt, Ian. 1991. "Applications of likelihood ratio orderings in economics." *Institute of Mathematical Statistics: Lecture Notes-Monograph Series* 19:174–189.
- Kirkegaard, René. 2012. "A mechanism design approach to ranking asymmetric auctions." *Econometrica* 80 (5):2349–2364.
- Kitamura, Yuichi, Gautam Tripathi, and Hyungtaik Ahn. 2004. "Empirical likelihood-based inference in conditional moment restriction models." *Econometrica* 72 (6):1667–1714.
- Kreps, David M. and Robert Wilson. 1982. "Sequential equilibria." *Econometrica* 50 (4):863–894.
- Kumar, Pranesh. 2010. "Probability distributions and estimation of Ali-Mikhail-Haq copula." *Applied Mathematical Sciences* 4 (14):657–666.
- Lazear, Edward P. and Sherwin Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy* 89 (5):841–864.
- Levy, Haim. 1992. "Stochastic dominance and expected utility: Survey and analysis." *Management Science* 38 (4):555–593.
- Manski, Charles F. 1990. "Nonparametric bounds on treatment effects." *American Economic Review* 80 (2):319–323.

- Maskin, Eric and John Riley. 2000. "Asymmetric auctions." *Review of Economic Studies* 67 (3):413–438.
- Milgrom, Paul R. 1981. "Good news and bad news: Representation theorems and applications." *Bell Journal of Economics* 12 (2):380–391.
- Moeller, Sara B, Frederik P Schlingemann, and Rene M Stulz. 2004. "Firm size and the gains from acquisitions." *Journal of Financial Economics* 73 (2):201–228.
- Moldovanu, Benny and Aner Sela. 2001. "The optimal allocation of prizes in contests." *American Economic Review* 91 (3):542–558.
- Niculescu, Constantin P. and Lars-Erik Persson. 2004. *Convex Functions and Their Applications: A Contemporary Approach*. Springer-Verlag.
- Quah, John K-H and Bruno Strulovici. 2009. "Comparative statics, informativeness, and the interval dominance order." *Econometrica* 77 (6):1949–1992.
- Reiss, R.-D. 1989. *Approximate Distributions of Order Statistics*. Springer-Verlag.
- Rothschild, Michael and Joseph E Stiglitz. 1970. "Increasing risk: I. A definition." *Journal of Economic Theory* 2 (3):225–243.
- Shaked, Moshe and J. George Shanthikumar. 2007. *Stochastic Orders*. Springer-Verlag.
- Shanthikumar, J George and David D Yao. 1991. "Bivariate characterization of some stochastic order relations." *Advances in Applied Probability* 23 (3):642–659.
- Sklar, M. 1959. "Fonctions de repartition an dimensions et leurs marges." *Publ. Inst. Statist. Univ. Paris* 8:229–231.
- Wooldridge, Jeffery M. 2002. *Econometric Analysis of Cross Section and Panel Data*, chap. 15. MIT Press.