Log-Supermodular Contests*

René Kirkegaard†

November 2024

Abstract

A large and unifying family of noisy or non-deterministic contests is proposed. The defining characteristic is that the marginal return to effort exhibits a log-supermodularity property. The model nests both the usual rank-order tournament and the microfoundations for the Tullock contest. With homogeneous technologies, strategic incentives and comparative statics are qualitative similar across the entire family. Robust comparative statics include collective discouragement and disparity effects. The effects of precommitment are also robust. Moreover, the model provides a framework for studying the role of heterogeneous technologies. Sufficient conditions are provided under which the comparative statics and their policy implications are preserved.

JEL Classification Numbers: C72, D72
Keywords: Contests, discouragement effect, disparity effect, log-supermodular contests, stochastic performance, tournaments.

*I thank SSHRC (grant 435-2022-0360) for funding this research.
†Department of Economics and Finance, University of Guelph. Email: rkirkega@uoguelph.ca.

1 Introduction

The literature on contests and tournaments is vast, well-established, and still rapidly growing. However, it remains dominated by a small set of tractable models. In the case of static contests with "noisy" outcomes, the workhorse models are the Tullock contest and the rank-order tournament, due originally to Tullock (1980) and Lazear and Rosen (1981), respectively. The outcome of a contest is noisy whenever the identity of the winner is not a deterministic function of the action or effort profile.

The Tullock contest and the rank-order tournament exemplify but do not exhaust all the ways in which noise may impact a contest. This observation immediately invites some overdue questions: Are there general and robust properties or characteristics of noisy contests, and to what extent are the workhorse models representative? This paper aims to answer these foundational questions in contest theory. The main message is that a number of comparative statics, and their policy implications, are remarkably robust along one dimension but more fragile along another, more subtle, dimension. Roughly speaking, the faultline is whether agents are subjected to noise in a homogeneous or heterogeneous manner. In the former case, the exact nature of the noise is unimportant as long as it falls within a very permissible family.

The workhorse models have particularly stark properties in contests with two agents. Dixit (1987) shows that the favorite – i.e. the agent most likely to win the contest – views actions as strategic complements whereas the underdog views actions as strategic substitutes. Such properties enable a deep understanding of the economic incentives in contests and games more generally, as explored in e.g. Bulow, Geanakoplos, and Klemperer (1985). Hence, it becomes possible to predict, and to intuitively understand, the consequences of changes in the contest environment. Dixit (1987) uses these properties to study the effects of precommitment in sequential contests. Whether the resulting insights are robust depends on whether or not the strategic incentives are qualitatively similar in a broader family of contests.

This paper considers contests in which noise takes the form of stochastic performance. That is, the agent's action combines with random factors to produce a non-degenerate distribution over her performance. The winner is the agent with the best realized performance, but this need not be the agent with the highest effort.

Lazear and Rosen's tournament is a special case in which the agent's action shifts the location of a symmetric distribution function. Similarly, Tullock's contest has been microfounded as a contest with stochastic performance. In the microfoundation proposed by Fullerton and McAfee (1999), higher effort makes the distribution more convex in a particular way. In Hirschleifer and Riley's (1992) microfoundation, the agent manipulates the scale parameter of an exponential distribution.

To fix ideas, let $F^i(x_i|a_i)$ denote the cumulative distribution function of agent i's performance, x_i, when her action is a_i. The distribution $F^i(x_i|a_i)$ describes the agent's "performance technology" and it is the most important primitive of the model. Unlike existing models, this paper imposes no functional forms on $F^i(x_i|a_i)$ but instead identifies basic properties that unify and generalize the workhorse models.

The central property is log-supermodularity. It is imposed on the model in two ways. First, the monotone likelihood-ratio property (MLRP) is assumed to hold, but this is equivalent to assuming that the density is log-supermodular in (x_i, a_i). Second, and more importantly, the distribution is assumed to have the no-upward-crossing (NUC) property introduced by Chade and Swinkels (2020) for problems with a single agent. To interpret NUC, recall that the probability that agent i's performance exceeds some threshold, x_i, is $1 - F^i(x_i|a_i)$. The latter increases at a rate of $-F^i_a(x_i|a_i)$ in the agent's action, and NUC is now precisely that $-F^i_a(x_i|a_i)$ is log-supermodular. Chade and Swinkels (2020) note that "NUC is satisfied for every distribution that we are aware of that is commonly used in economic applications."

This paper is the first to use log-supermodularity to systematically study contests. Contests in which performance technologies satisfy MLRP and NUC are henceforth referred to as *log-supermodular contests*, or LSM contests for short. The contests in Fullerton and McAfee (1999) and in Hirschleifer and Riley (1992) are LSM contests. So are rank-order tournaments under the common assumption that densities are log-concave. Thus, the family of contests under consideration nests the standard models.

The analysis starts with contests in which agents have access to fully homogeneous performance technologies. The easiest interpretation is that $F^i(x_i|a_i)$ is identity-independent. In the literature mentioned above, agents independently convexify identical distributions, manipulate the scale of independent exponential distributions, or shift the location of identical distributions. Hence, any agent can perfectly replicate the performance distribution of any other agent by simply taking the same action. However, agents may differ in their effectiveness at, or cost of, manipulating the performance technology. They may also attach different values to winning the contest.

The first, but central, insight concerns the strategic incentives in LSM contests with homogeneous technologies and two contestants. As in Dixit's (1987) contests, for any action profile, the favorite (underdog) views actions as strategic complements (substitutes). The next step is to understand equilibrium properties and comparative statics. These are in turn determined by the reaction functions. An iconic image in contest theory is the hump-shaped reaction function that arise in Tullock contests. The hump-shape is a direct consequence of how the favorite and underdog view actions as strategic complements and substitutes, respectively. Thus, the reaction functions have the same shape in all LSM contests with fully homogeneous technologies. Comparative statics with respect to changes in such characteristics as valuations and cost functions are therefore qualitatively similar in all these contests. Dixit's (1987) conclusions regarding precommitment are similarly robust.

Three comparative statics are worth highlighting. Their resulting policy implications are detailed later. The first comparative static is an *individual discouragement effect*. Any individual agent works hardest if the equilibrium is such that she wins with probability $\frac{1}{2}$, as is the case when she faces a symmetric opponent. On the other hand, if she is matched with an asymmetric opponent then her incentives are lessened and she is discouraged from working as hard, although the reason is different depending on whether the agent is the favorite or the underdog.

The second and third comparative statics relate to the overall effects of an increase in asymmetry across agents. The *collective discouragement effect* reveals that both agents work less hard in equilibrium if the underdog becomes weaker, or values winning less. In comparison, the *collective disparity effect* says that if the favorite becomes stronger, then equilibrium actions move even further apart – the disparity widens – and the favorite becomes even more likely to win. The two effects capture different adverse consequences of a more uneven playing field.

However, the assumption that technologies are fully homogeneous is strong. Since any agent can perfectly copy the distribution of her competitor, there is no room for personal idiosyncrasies. In a contest between product developers, one may be more effective at thinking outside the box, meaning that her performance is likely to have higher variance. If salespeople in a promotion contests are assigned to different regions, then small differences in income distributions or tastes across regions carry over to the agents' performance distributions. Importantly, heterogeneous technologies are more subtle than heterogenous cost functions, since the former maps actions into different functions (distributions) rather than simply into different real numbers (costs). Consequently, the problem is more nuanced with heterogeneous technologies.

The most important consequence of heterogeneous technologies is that the favorite may not view actions as strategic complements. However, sufficient conditions are identified under which this is the case in equilibrium. Roughly speaking, the sufficient conditions hold when one agent has a performance technology that is both more productive and more sensitive to effort than the other. In such "ordered LSM contests", the first agent is enticed to work harder and ends up being the favorite not only due to her higher effort but also because she is more productive to start with. Versions of the collective discouragement and disparity effects hold in ordered LSM contests, but not more generally. Moreover, even in ordered LSM contests, the individual discouragement effect changes qualitatively, which as discussed later has important implications for biased contests. These findings are significant because they highlight that the received wisdom may be sensitive to heterogeneity in performance technologies.

Noisy contests with more agents are less well-studied. Still, LSM contests with many agents have several properties in common. These results and other extensions are detailed in the discussion section. It is shown that the standard results and intuition are turned on their heads in the rare contests that are log-submodular.

Policy implication: Discouragement effects of various kinds take a central place in the literature. An early example is the celebrated "exclusion principle" discovered by Baye, Kovenock, and de Vries (1993) in deterministic contests, or all-pay auctions. Here, the presence of a very strong agent can be so discouraging to everyone else that excluding the former may lead to higher aggregate effort. Kirkegaard (2024) shows that it may instead be optimal to exclude a weaker agent in a specific contest model with noise. In comparison, the current paper investigates a large class of noisy contests and shows that certain discouragement effects are common to many of them.[1]

Discouragement effects are integral to the debate on the merits of policies meant to level the playing field. In their survey, Chowdhury, Esteve-González, and Mukherjee (2023) ask for more work to "help scholars form conclusions about the generalizability of the different policy outcomes observed" in existing models. The current paper represents a step towards answering this call for generality.

Imagine that there is some regulator with the ability to influence the stakes of

[1] Fang, Noe, and Strack (2020) study different manifestations of discouragement effects in all-pay auctions with symmetric contestants when e.g. prizes become more unequal or when more agents enter. In the present paper, discouragement effects refer to properties of asymmetric contests.

the contest. If his objective is to achieve more even outcomes, then the collective disparity effect supports a policy of disincentivizing the favorite, e.g. by making it less desirable for her to win or by making the consolation prize more desirable to her. On the other hand, the collective discouragement effect implies that improving the underdog's incentives, e.g. by promising her an additional bonus if she wins, causes both agents to work harder, which is beneficial for any regulator whose payoff is increasing in the performances or actions of the two agents.

The model also helps understand and extend key results in the literature on biases in Tullock contests, culminating in Fu and Wu (2020). They show that if the regulator is able to bias the impact of the two agents' actions, then it is typically optimal to do so in a way that leads to a completely level playing field. The current paper argues that this conclusion is ultimately due to the individual discouragement effect. Thus, the result may generalize to other LSM contests with homogeneous technologies, although this depends in part on the degree of flexibility the regulator has. Conversely, with heterogeneous technologies, a level playing field is generally not optimal.

Although the collective disparity and discouragement effects are robust, they predict the direction but not the magnitude of equilibrium responses. It is possible that aggregate outcomes can be sensitive to the performance technologies. For instance, aggregate effort can increase or decrease when valuations become more asymmetric. Thus, there is no robust aggregate discouragement effect within LSM contests.[2]

Related literature: The literature on generalized contests is scant. Focusing on two-agent contests, Malueg and Yates (2005) examine the role of the contest success function (CSF) in comparative statics. However, their CSF is a black box that maps the action profile into winning probabilities and no microfoundations are offered. Malueg and Yates (2005) mainly concentrate on (potentially asymmetric) CSFs that are homogenous of degree zero in actions. For such CSFs, the reaction functions are hump-shaped. The comparative statics are again determined by who views actions as strategic substitutes or complements in equilibrium, but there are no insights into how this relate to who is the underdog or favorite in equilibrium. In the current paper, sufficient conditions on the primitives are provided for contests with heterogeneous technologies under which it can be concluded that the favorite views actions as complements in equilibrium. Thus, one advantage of the current approach

[2] See e.g. the examples in Kirkegaard (2024), whose model is also a special case of LSM contests. Drugov and Ryvkin (2022) thoroughly discuss the literature on aggregate discouragement effects.

is that there are clearly defined primitives – namely the performance technologies – that give structure to the CSF and which lend themselves to economic interpretation.[3]

Using the "state-space approach", Bastani, Giebe, and Gürtler (2022) consider tournaments in which agents have identical production functions but identity-dependent noise or state distributions. These are contests with heterogeneous technologies. They assume that agents have the same valuations and cost functions. Their focus is on ranking equilibrium actions and they do not consider the shape of the reaction functions or the comparative statics of changes to individual valuations or cost functions.

In two-agent contests, Bastani, Giebe, and Gürtler (2022) find that the two agents have the same equilibrium action.[4,5] The current paper shows that this surprising result is due to the precise way in which they capture technological heterogeneity. There are two important features of their model. First, as in Lazear and Rosen (1981), they assume that noise is one-dimensional. However, it is easy to imagine that there are multiple sources of noise in the production process. It turns out that Bastani, Giebe, and Gürtler's (2022) result does not hold with multivariate noise. The current paper makes use of the "Mirrlees approach", which takes the performance distribution $F^i(x_i|a_i)$ as the starting point. As explained in Section 5, this approach can easily encompass what corresponds to multiple sources of noise in the state-space approach.

Second, Bastani, Giebe, and Gürtler (2022) assume that agents have the same production function but, depending on the interpretation, different distributions of noise, luck, or skill. It is equally possible that agents have access to different resources, opportunities, machinery, or technology more broadly. With such identity-dependent production functions, Bastani, Giebe, and Gürtler's (2022) result does not hold.

Finally, considering biased contests, Kirkegaard (2023a) argues that it can be unclear how to directly bias actions if what is observed is performance. Thus, he derives general design principles when the designer can bias how each agent's performance is scored. A substantial improvement over Fu and Wu (2020) can be achieved in this manner. However, the optimal bias rarely leads to a completely level playing field.

[3] Relatedly, Ewerhart and Serena (2024) ask what can be said about the structure of the CSF in rank-order tournaments with two agents when the only assumptions are that noise is additive and i.i.d. In comparison, it turns out that the CSF is log-supermodular in all LSM contests.

[4] Bastani, Giebe, and Gürtler (2022) also examine how the equilibrium action changes when the noise distributions change. Such comparative statics are omitted from the current paper, in large part because it is not generally speaking the case in LSM contests that agents take the same action.

[5] Giebe and Gürtler (2024) allow for more agents. Some of their results are generalized in Section 6. However, Giebe and Gürtler (2024) consider more varied and general prize structures.

2 Model and preliminaries

2.1 Contests with stochastic performance

A total of $n \geq 2$ risk neutral agents compete in a winner-take-all contest. Agent i's value of winning the contest is $v_i > 0$, $i = 1, 2, ..., n$. She exerts effort $a_i \geq 0$ in pursuit of winning. The cost of action a_i is described by a twice continuously differentiable cost function $c_i(a_i)$, with $c_i'(a_i) > 0$ for all $a_i > 0$. The cost function is unbounded above. These characteristics are common knowledge.

Let $P^i(\mathbf{a})$ denote the probability that agent i wins the contest if the action profile is $\mathbf{a} = (a_1, a_2, ..., a_n)$.[6] Thus, agent i's expected utility from action profile $\mathbf{a}$ is

$$u^i(\mathbf{a}) = v_i P^i(\mathbf{a}) - c_i(a_i). \tag{1}$$

Let $\overline{a}_i$ denote the action for which $v_i - c_i(\overline{a}_i) = -c_i(0)$. Following Siegel (2009), $\overline{a}_i$ is referred to as the agent's reach. Any action above $\overline{a}_i$ is strictly dominated by the action $a_i = 0$. Hence, attention is at times restricted to the set of actions $[0, \overline{a}_i]$.

The function $P^i(\mathbf{a})$ is known as the contest success function (CSF). Rather than treating the CSF as a black box, the current paper unpacks it by assuming that performance is stochastic and that the agent with the highest realized performance is the winner. In this story, agent i's action matters because it shapes the distribution from which her performance is drawn. In particular, agent i's stochastic performance, X_i, is described by some distribution function, $F^i(x_i|a_i)$, which is parameterized by her action, a_i. The distribution function can be thought of as the agent's performance technology. Given $\mathbf{a}$, performance is statistically independent across agents.

If $a_i > 0$ then F^i is atomless and has full support on $[\underline{x}, \overline{x}]$, with $\overline{x} > \underline{x}$. Let $f^i(x_i|a_i)$ denote the density function, assumed to be strictly positive on $(\underline{x}, \overline{x})$. In order to invoke some results in Chade and Swinkels (2020), assume that F^i and f^i are C^2. Assume that $F^i_a < 0$ for all $x \in (\underline{x}, \overline{x})$. Thus, the performance distribution improves in the sense of first order stochastic dominance when the agent works harder. The support $[\underline{x}, \overline{x}]$ is the same for all $a_i > 0$ and all agents $i = 1, 2, ..., n$.

If $a_i = 0$ then F^i either has the above properties, or it is degenerate at $\underline{x}$. In the

[6] Subscripts on variables and univariate functions indicate the identity of the agent. For multivariate functions, the identity of the agent is captured by a superscript, such that subscripts can be used to indicate partial derivatives.

latter case, the agent's performance is the worst possible with probability one. Hence, she has no chance of outperforming an agent whose distribution is not degenerate. If $\mathbf{a} = \mathbf{0}$ and all F^i are degenerate, then all agents deliver the same performance, $\underline{x}$. In this case, let $P^i(\mathbf{0}) = \frac{1}{n}$. In all other cases, the best performance is greater than $\underline{x}$ with probability one. Thus, when F^i is not degenerate, agent i wins with probability

$$P^i(\mathbf{a}) = \int \prod\nolimits_{m \neq i} F^m(x|a_m) f^i(x|a_i) dx. \tag{2}$$

If F^i is degenerate when $a_i = 0$, it is still assumed that the distribution is continuous in a_i, i.e. that $\lim_{a_i \to 0} F(x_i|a_i) = 1$ for all $x \in [\underline{x}, \overline{x}]$. The implication is that $P^i(\mathbf{a})$ is continuous in a_i whenever at least one opponent has a non-degenerate distribution.

The next examples demonstrate that existing contest and tournament models can be viewed as special cases of contests with stochastic performance.

Example 1 (Rank-order tournaments and probit CSFs): Lazear and Rosen (1981) consider a rank-order tournament with additive noise. Agent i's performance is $x_i = a_i + \varepsilon_i$, where $\varepsilon_i \in \mathbb{R}$ is the realization of a random variable that is unknown to the agent when she decides her action. If ε_i has distribution $Q_i(\varepsilon_i)$, then $F^i(x_i|a_i) = Q_i(x_i - a_i)$ captures the performance technology. If $n = 2$, agent i wins if $a_i - a_j > \varepsilon_j - \varepsilon_i$ and it follows that $P^i(\mathbf{a})$ is a function of $a_i - a_j$. Hence, rank-order tournaments with additive noise provide a microfoundation for the so-called probit CSF.

Lazear and Rosen (1981) follow the state-space approach, in which ε_i is the "state" and the agent's performance is described by a state-contingent production function. This approach requires the specification of both a production function, $x_i = \pi^i(\varepsilon_i, a_i)$, and a distribution over states, $Q_i(\varepsilon_i)$. Lazear and Rosen (1981) assume that $\pi^i(\varepsilon_i, a_i) = a_i + \varepsilon_i$ and that $Q_i(\varepsilon_i)$ is identity-independent. In the current paper, the "Mirrlees approach" is used instead, meaning that $F^i(x_i|a_i)$ is taken to be the primitive, as in the next example. ▲

Example 2 (best-shot contests and logit CSFs): Tullock's (1980) CSF is oftentimes thought of as a black box, but it has also been microfounded as a contest with stochastic performance. A slight variation of the microfoundation in Fullerton and McAfee (1999) is presented here.

Agent i's performance is independently drawn from the distribution function $F^i(x_i|a_i) = H(x_i)^{p_i(a_i)}$, where $H(x_i)$ is some atomless and twice differentiable distribution function with full support and where $p_i(a_i) \geq 0$ is a strictly increasing and twice differentiable function. The function $p_i(a_i)$ is known as the impact function. It is as if agent i's action buys her $p_i(a_i)$ independent draws from the distribution $H(x_i)$, and her performance equals the best of those draws. Since H is identity-independent, any draw by any agent has an equal chance of being the best draw. Thus, agent i wins with probability

$$P^i(\mathbf{a}) = \frac{p_i(a_i)}{\sum_{j=1}^n p_j(a_j)}, \tag{3}$$

whenever $\sum_{j=1}^n p_j(a_j) > 0$. This can also be verified directly from (2). In the special case where $\sum_{j=1}^n p_j(a_j) = 0$, it must hold that $a_j = 0$ and that $p_j(0) = 0$ for all j. However, note that when $p_j(0) = 0$, the distribution of agent j's performance is degenerate. Thus, when $\sum_{j=1}^n p_j(a_j) = 0$, all agents tie for the same performance, $\underline{x}$.

Following Dixit (1987), the CSF in (3) is referred to as the logit CSF. Fullerton and McAfee (1999) assume that $p_i(a_i) = a_i$, but this is clearly not important for the argument that led to (3). However, when $p_i(a_i) = a_i$, (3) reduces to the lottery CSF. If $p_i(a_i) = a_i^r$ for some $r > 0$, then (3) is known as the Tullock CSF. ▲

The primitives of contests with stochastic performance are $\{v_i, c_i(\cdot), F^i(\cdot|\cdot)\}_{i=1}^n$. It is common in contest theory to allow agents to have identity-dependent valuations and cost functions, but rare to allow identity-dependent technologies. However, the proposed model provides a framework for studying heterogeneous technologies in detail. A more precise taxonomy is needed.

Definition 1 (Homogeneous and heterogeneous technologies) *Agents have homogeneous technologies if F^i can be written as $F^i(x_i|a_i) = G(x_i|p_i(a_i))$ for all i. Technologies are fully homogeneous if the range of $p_i(a_i)$ is the same for all i. Agents have heterogeneous technologies in all other cases.*

Technologies are fully homogeneous if and only if any agent can perfectly reproduce the performance distribution of any other agent, regardless of the latter's action. In particular, agent i can mimic $F^j(x_j|a_j)$ simply by choosing an action a_i for which $p_i(a_i) = p_j(a_j)$. However, the cost of said performance distribution may be different for agent i and agent j. Hence, this situation is identical to one in which the p_i functions are normalized to be the same for all agents, but the c_i functions are not. In this case, technologies are identity-independent.

For homogeneous technologies, $F_a^i(x_i|a_i) = G_p(x_i|p_i(a_i))p_i'(a_i)$. To conform to previous assumptions, assume that $G_p(x_i|p_i(a_i)) < 0$ for $x \in (\underline{x}, \overline{x})$ and that $p_i'(a_i) > 0$.

Thus, $p_i(a_i)$ can be interpreted as an impact function, with higher impacts translating into stronger distributions of performance. Let $g(x_i|p_i(a_i))$ denote the density.

Technologies are fully homogeneous in Example 2 if e.g. $p_i(a_i) = a_i^{r_i}$ with potentially identity-dependent $r_i > 0$, but not fully homogeneous if $p_i(a_i) = i + a_i^r$. In the latter case, agent 2 cannot reproduce the poor performance technology of agent 1 if a_1 is small enough. This is similar to a situation in which the action set is different for different agents. Note that all homogeneous technologies can be ranked in a first-order stochastic dominance sense, based on how large $p_i(a_i)$ is.

In comparison, technologies are heterogeneous if $H(x_i)$ is made identity-dependent in Example 2. Bastani, Giebe, and Gürtler (2022) and Giebe and Gürtler (2024) consider rank-order tournaments with noise that have identity-dependent distributions. These are also heterogeneous technologies. Heterogeneous technologies can not necessarily be ranked across agents in a first-order stochastic dominance sense.

2.2 Log-supermodular contests

The distribution functions are important because they shape the CSF, $P^i(\mathbf{a})$. A main contribution of the paper is to identify general conditions on the performance technologies that in turn give structure to the CSF. Key to these developments is the notion of log-supermodularity. Recall that a non-negative and multivariate function $t(z_1, z_2, ..., z_m)$ is log-supermodular if, for all $\mathbf{z}'$ and all $\mathbf{z}''$, it holds that

$$t(\mathbf{z}' \vee \mathbf{z}'')t(\mathbf{z}' \wedge \mathbf{z}'') \geq t(\mathbf{z}')t(\mathbf{z}''), \tag{4}$$

where $\mathbf{z}' \vee \mathbf{z}''$ (the "join") is the component-wise maximum and $\mathbf{z}' \wedge \mathbf{z}''$ (the "meet") is the component-wise minimum. Thus, log-supermodularity is a complementarity condition. If t is strictly positive and differentiable, then log-supermodularity obtains if and only if

$$\frac{\partial^2 \ln t(z_1, z_2, ..., z_m)}{\partial z_i \partial z_j} \geq 0$$

for all i and j, with $i \neq j$. In the following, t is said to be strictly log-supermodular if the above inequality is strict whenever t is strictly positive. Athey (2002) uses log-supermodularity for comparative statics in decision problems and some games.

To start on familiar ground, the commonly invoked monotone likelihood ratio property (MLRP) is precisely that the density $f^i(x_i|a_i)$ is log-supermodular in (x_i, a_i).

A strict version of MLRP will sometimes be required.

Definition 2 (MLRP) *The distribution F^i has the (strict) monotone likelihood ratio property (MLRP) if f^i is (strictly) log-supermodular on $[\underline{x}, \overline{x}] \times (0, \infty)$.*

The MLRP implies that $F_a^i \leq 0$ but, as mentioned earlier, the stronger assumption that $F_a^i < 0$ for all $x \in (\underline{x}, \overline{x})$ is made. The next step adds further structure to F_a^i. In particular, a regularity property from Chade and Swinkels (2020) is imposed. They also assume that $F_a^i < 0$ for all $x \in (\underline{x}, \overline{x})$. Their no-upward-crossing (NUC) condition then says that for any $\gamma \in \mathbb{R}$, $F_{aa}^i(\cdot|a_i) - \gamma F_a(\cdot|a_i)$ never crosses zero from below on $(\underline{x}, \overline{x})$. However, they prove that this is equivalent to assuming that $-F_a^i$ is log-supermodular. A strict version is often required for this paper.

Definition 3 (NUC) *The distribution F^i satisfies NUC if $-F_a^i$ is log-supermodular on $[\underline{x}, \overline{x}] \times (0, \infty)$. It satisfies strict NUC if $-F_a^i$ is strictly log-supermodular on $(\underline{x}, \overline{x}) \times (0, \infty)$.*

The NUC condition has a natural economic interpretation, although it takes a few steps to develop the intuition. Fix some performance threshold level, x_i, and note that the probability that the agent successfully exceeds the threshold is $1 - F^i(x_i|a_i)$, given her action is a_i. Thus, a marginal increase in a_i increases the probability of success by $-F_a^i$. Hence, NUC disciplines how the marginal return to extra effort depends on the threshold and the starting level of effort. In the following, an explanation that complements the one in Chade and Swinkels (2020) is provided.

First, consider two actions a_H and a_L and two interior thresholds x_H and x_L, with $a_H > a_L$ and $x_H > x_L$. Then, it follows from (4) that

$$\frac{-F_a^i(x_H|a_H)}{-F_a^i(x_H|a_L)} \geq \frac{-F_a^i(x_L|a_H)}{-F_a^i(x_L|a_L)}.$$

That is, the relative increase in the success probability from extra effort is more sensitive to the starting effort level the higher the threshold is. If the threshold is high, then it is hard to succeed and a small increase in effort is unlikely to make much difference if effort is low to begin with, but it may make a difference if effort is high. Thus, the left hand side is large. On the other hand, if the threshold is low then it is easy to succeed and a small increase in effort is of little consequence if effort is high to start with, but it may be significant if effort was low originally. Hence, the right hand side is small.

Some of the following results rely only on MLRP, others only on NUC, but many rely on both. Since both properties are log-supermodularity properties, contests that satisfy both MLRP and NUC will be referred to as log-supermodular (LSM) contests.

Definition 4 (LSM Contests) *If F^i satisfies (strict) MLRP and (strict) NUC for all i, then the contest is said to be a (strict) log-supermodular (LSM) contest.*

Chade and Swinkels (2020) observe that MLRP and NUC are identical requirements in the location families of distributions where $F^i(x_i|a_i) = Q_i(x_i - a_i)$. In this case, $f^i(x_i|a_i) = -F^i_a(x_i|a_i) = q_i(x_i - a_i)$, where q_i is the density. Hence, the two properties hold if and only if the density is log-concave. It follows that rank-order tournaments with additive noise are LSM contests as long as densities are log-concave. The best-shot contest in Example 2 is also a LSM contest. Similarly, Hirschleifer and Riley (1992) use an exponential distribution to microfound the lottery contests with two agents. This distribution also satisfies both MLRP and NUC. In this sense, LSM contests nest and unify existing models.

As proven in e.g. Athey (2002), the MLRP implies that F^i is log-supermodular. Hence, f^i, F^i, and $-F^i_a$ are all log-supermodular in LSM contests. Some of the paper's main insights rely on the structure that these properties lend to the CSF. The first proposition summarizes some of these properties. The main take-away is that the log-supermodularity properties of the primitives lead the CSF to have certain log-supermodular properties as well. In other words, CSFs with log-supermodular properties are microfounded. The underlying reason is that the CSF and its partial derivatives are obtained by integrating products of functions like f^i, F^i, and $-F^i_a$ across agents, but such operations preserve log-supermodularity. The derivative of $P^i(\mathbf{a})$ with respect to a_j is denoted $P^i_j(\mathbf{a})$.

Proposition 1 *The CSF in any LSM contest has the following properties:*

1. *$P^i(\mathbf{a})$ is log-supermodular in $\mathbf{a}$ on $(0, \infty)^n$.*
2. *For $j \neq i$, $-P^i_j(\mathbf{a})$ is log-supermodular in $\mathbf{a}$ on $(0, \infty)^n$.*
3. *For $n = 2$, $P^i_i(\mathbf{a})$ is log-supermodular in $\mathbf{a}$ on $(0, \infty)^2$.*

3 Incentives in small LSM contests

This section considers LSM contests with two agents. The first objective is to understand the strategic incentives. The second, related, aim is to describe the shape of the reaction functions, which in turn opens the door for comparative statics later on.

3.1 Strategic substitutes and complements

Two important definitions help frame the analysis and discussion.

Definition 5 (Complements/substitutes) *Fix any interior action profile. With two agents, agent i views actions as strategic complements if $P^i_{12}(\mathbf{a}) > 0$ and strategic substitutes if $P^i_{12}(\mathbf{a}) < 0$.*

Definition 6 (Favorite/underdog) *Fix any action profile. With two agents, agent i is the favorite if $P^i(\mathbf{a}) > \frac{1}{2}$ and the underdog if $P^i(\mathbf{a}) < \frac{1}{2}$.*

These definitions are local, i.e. specific to the action profile. To understand why agent i views actions as strategic complements if $P^i_{12}(\mathbf{a}) > 0$, note that her return to extra effort is proportional to $P^i_i(\mathbf{a})$. If this is increasing in agent j's action – or $P^i_{12}(\mathbf{a}) > 0$ – then her incentives to invest extra effort is higher the higher a_j is. For homogeneous technologies, agent i is the favorite if and only if $p_i(a_i) > p_j(a_j)$.

Since $P^1(\mathbf{a}) + P^2(\mathbf{a}) = 1$ for all $\mathbf{a}$, the cross-partial derivatives must cancel out, or $P^1_{12}(\mathbf{a}) + P^2_{12}(\mathbf{a}) = 0$. Hence, if one agent views actions as strategic complements then the other views actions as strategic substitutes, and vice versa.

Dixit (1987) reasons that: "little can be said about [the cross-partial derivative] in general. Therefore, I will consider special functional forms [the logit and probit CSFs]." In these settings, Dixit (1987) establishes that the sign of $P^i_{12}(\mathbf{a})$ depends only on whether the agent is the favorite or the underdog. Contrary to Dixit (1987), the current paper solves the problem without special functional forms. In fact, with homogeneous technologies, a generalization follows immediately from strict NUC.

Assume that technologies are identity-independent, or $F^i(x_i|a_i) = F(x_i|a_i)$, $i = 1, 2$. Then,

$$P^i(\mathbf{a}) = \int F(x|a_j) f(x|a_i) dx,$$

where the integrand $F(x|a_j)f(x|a_i)$ is the probability that agent i wins contingent on performance x, weighted by the likelihood that her performance is indeed x. With this in mind, note that strict NUC is equivalent to

$$\frac{\partial}{\partial a}\frac{f_a(x|a)}{F_a(x|a)} > 0 \text{ for all } x \in (\underline{x}, \overline{x}),$$

meaning that

$$F_a(x|a_j)f_a(x|a_i) > F_a(x|a_i)f_a(x|a_j) \tag{5}$$

if $a_i > a_j$. In other words, for any performance level x, there is "greater complementarity" between a_i and a_j in $F(x|a_j)f(x|a_i)$ than in $F(x|a_i)f(x|a_j)$ whenever $a_i > a_j$. Since this holds for all performance levels, there must be greater complementarity between a_i and a_j in $P^i(\mathbf{a})$ than in $P^j(\mathbf{a})$, or $P^i_{12}(\mathbf{a}) > P^j_{12}(\mathbf{a})$. It follows that $P^i_{12}(\mathbf{a}) > 0 > P^j_{12}(\mathbf{a})$. The argument easily extends to all homogeneous technologies.

Theorem 1 *Consider any strict LSM contest with homogeneous technologies and two agents. Fix any interior action profile with $p_1(a_1) \neq p_2(a_2)$. Then, the favorite views actions as strategic complements and the underdog views actions as strategic substitutes.*

The MLRP is not required for Theorem 1. However, it is shown later on that the conclusion in Theorem 1 does not generally hold for heterogeneous technologies. Thus, what is important for Dixit's (1987) results is that technologies are homogeneous.

3.2 The iconic hump-shaped reaction function

By continuity, agent i has at least one best response to any $a_j > 0$ and she may or may not have a best response to $a_j = 0$. In the following, attention is narrowed to LSM contests in which best responses are unique. Let $b_i(a_j) = \arg\max_{a_i} u^i(\mathbf{a})$ denote the reaction function, defined on $(0, \infty)$ or $[0, \infty)$, depending on whether $F^1(x|0)$ and $F^2(x|0)$ are both degenerate or not.

Assumption A1 (Best responses) Whenever a best response exists, it is unique and $u^i_i(b_i(a_j), a_j) = 0$ implies that $u^i_{ii}(b_i(a_j), a_j) < 0$.[7]

A standard justification for unique best responses is that $c_i(a_i)$ is "sufficiently convex". In the spirit of Rogerson's (1985) and Jewitt's (1988) approach to the

[7] Strict local concavity, or $u^i_{ii}(b_i(a_j), a_j) < 0$, simplifies some of the proofs. Note that if $b_i(a_j) = 0$ then $u^i_i(b_i(a_j), a_j) < 0$ is possible, in which case A1 does not impose local concavity.

moral hazard problem, restrictions can alternatively be imposed on the performance technologies to ensure that $P^i(\mathbf{a})$ is concave in a_i, which in turn implies that $u^i(\mathbf{a})$ is concave in a_i as long as $c_i(a_i)$ is weakly convex. Rogerson's (1985) condition is satisfied in the best-shot contest described in Example 2 when $p_i(a_i)$ is concave.

Concavity of $u^i(\mathbf{a})$ is stronger than required. Building on Chade and Swinkels (2020), the supplementary material shows that the structure of LSM contests makes it easier to verify quasiconcavity. In particular, it is possible to more precisely quantify "how convex" the cost function must be in order to ensure quasiconcavity. For instance, if F^j is concave in x, then agent i's problem is quasiconcave as long as her cost function is "more convex" in a_i than her expected performance, $\mathbb{E}[X_i|a_i]$.

The problem is uninteresting if $b_i(a_j) = 0$ always. Thus, assume that the contest is non-trivial in the sense that agents finds it optimal to actively participate in the contest in response to at least some rival actions.

Assumption A2 (Non-trivial contests) For any $i = 1, 2$, there is some $a_j > 0$ for which $b_i(a_j) > 0$.

Let $a_j^{\text{sup}} = \sup\{a_j | b_i(a_j) > 0\}$, such that $b_i(a_j) = 0$ if $a_j > a_j^{\text{sup}}$. Hence, it is the properties of the reaction function on $(0, a_j^{\text{sup}})$ that is of interest. Note that since $u^i(\mathbf{a})$ is continuous in a_j for all $a_j > 0$, $b_i(a_j)$ is continuous in a_j when $a_j > 0$.

Given these assumptions, the aim of this subsection is to investigate whether the iconic hump-shaped reaction functions known from Tullock and rank-order tournaments arise in LSM contests. For fully homogeneous technologies, the question is answered in the affirmative. Outside of fully homogeneous technologies, reaction functions are at least single-peaked. That is, they are either hump-shaped or monotonic.

Definition 7 *The reaction function $b_i(a_j)$ is (strictly) hump-shaped on an interval if it is first (strictly) increasing and then (strictly) decreasing on the interval. It is (strictly) single-peaked if it is either (strictly) hump-shaped or (strictly) monotonic.*

3.2.1 Homogeneous technologies

Theorem 1 implies that agent i's reaction function is increasing when she is the favorite and decreasing when she is the underdog. Other things equal, agent i is the favorite when her opponent is not working very hard and the underdog when her opponent is working hard. It is therefore intuitive that agent i's reaction function is first increasing and then decreasing in a_j. This intuition is broadly speaking correct, but there are some subtle differences between contests with homogeneous and fully homogeneous technologies, respectively.

Geometrically, Theorem 1 splits the action space into two regions, neatly divided by the increasing locus of points along which $P^i(\mathbf{a}) = \frac{1}{2}$ and $P^i_{12}(\mathbf{a}) = 0$. In one region, agent i is the favorite and her reaction function is locally increasing. In the other region, she is the underdog and her reaction function is locally decreasing. Thus, Theorem 1 essentially produces a "slope field" for the reaction function. In the case of fully homogeneous technologies, the $P^i(\mathbf{a}) = \frac{1}{2}$ locus starts from the origin and diverges to infinity. One implication is that the reaction function enters both regions.

Proposition 2 *Consider any strict LSM contest with fully homogeneous technologies and two agents. Assume that A1 and A2 hold. Then,* $b_i(a_j) > 0$ *on* $(0, a_j^{\sup})$ *and it is **strictly hump-shaped** in* a_j *on this interval. At any point on the increasing (decreasing) part, agent* i *is the favorite (underdog).*

When technologies are homogeneous but not fully homogeneous, the $P^i(\mathbf{a}) = \frac{1}{2}$ locus need not start at the origin or diverge to infinity. The first implication is that it is possible that the best response is zero for a non-degenerate interval before becoming strictly positive. The second implication is that the reaction function may be monotonic. In this case, the agent is either the favorite or the underdog everywhere along her reaction function.[8]

Proposition 3 *Consider any strict LSM contest with homogeneous technologies and two agents. Assume that A1 and A2 hold. Then,* $b_i(a_j) > 0$ *on an interval, and it is **strictly single-peaked** in* a_j *on the interior of this interval. At any point where the reaction function is increasing (decreasing), agent* i *is the favorite (underdog).*

EXAMPLE 2 (CONTINUED): Assume that $p_1(a_1) = 2 - e^{-a_1}$ and that $p_2(a_2) = a_2$ in the best-shot model. Assume that $c_i(a_i) = a_i$, $i = 1, 2$. The impact functions are weakly concave and the costs are weakly convex. Hence, given (3), agent i's payoff is concave in her own action. The defining characteristic of this example is that $p_1(a_1)$ has limited range, $[1, 2)$, which is a proper subset of the range of $p_2(a_2)$. In other words, agent 1 cannot access as many performance technologies as agent 2.

[8] Imagine as a special case that the range of the impact functions do not overlap, or $p_1(0) > \lim_{a_2 \to \infty} p_2(a_2)$. Then, regardless of the action profile, agent 1 is always the favorite. Her reaction function is globally increasing, whereas agent 2's reaction function is globally decreasing.

Agent 2's reaction function is

$$b_2(a_1) = \max\{0, \sqrt{v_2\,(2 - e^{-a_1})} - \left(2 - e^{-a_1}\right)\},$$

the properties of which depends on v_2. For $v_2 > 8$, $b_2(a_1)$ is strictly positive always, and globally increasing. For $v_2 \in (4, 8)$, $b_2(a_1)$ is hump-shaped and always strictly positive. For $v_2 \in (2, 4)$, it is decreasing and always strictly positive. For $v_2 \in (1, 2)$, it is decreasing and becomes zero for large enough a_1. For $v_2 \in (0, 1]$, the best response is always zero. Figure 1(a) depicts $b_2(a_1)$ for $v_2 \in \{3, 6, 9\}$ in (a_1, a_2) space.

Agent 1's reaction function is

$$b_1(a_2) = \max\{0, -\ln(a_2 + \frac{1}{2}a_2 v_1 + 2 - \frac{1}{2}\sqrt{a_2 v_1\,(4a_2 + a_2 v_1 + 8)})\}.$$

Regardless of v_1, the best response is zero if a_2 is small. Indeed, if $v_1 \leq 4$, then $b_1(a_2) = 0$ for all a_2. For higher v_1, the reaction function is first increasing and then decreasing. Figure 1(b) depicts $b_1(a_2)$ in (a_2, a_1) space when $v_1 = 6$. ■

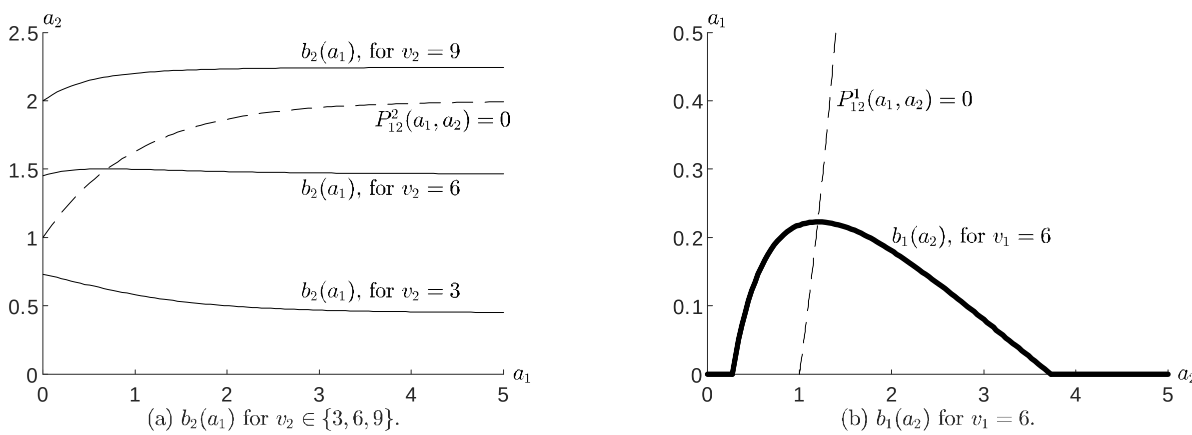

Figure 1: Single-peaked reaction functions.

3.2.2 Heterogeneous technologies

Theorem 1 and the implications in Propositions 2 and 3 rely on strict NUC. The additional restriction to homogeneous technologies gives enough structure to the problem, meaning that the MLRP is in fact not required. In the following, heterogeneous technologies are permitted, at the cost of utilizing the MLRP instead.

The last two properties in Proposition 1 turn out to imply that if $P^i_{12}(\mathbf{a}) < 0$ then it remains the case that $P^i_{12}(\mathbf{a}) < 0$ if a_i decreases and/or a_j increases. Thus, the action space is again split into two regions, but the curve that does so need not be strictly increasing in the interior. Consequently, the reaction function must still be single-peaked, but not necessarily strictly single-peaked.

Proposition 4 *Consider any LSM contest with two agents and assume that A1 and A2 hold. Then, $b_i(a_j)$ is single-peaked on $(0, \infty)$.*

The general property in Proposition 4 does not rule out that the reaction function is flat on an interval where the best response is positive. However, such a situation can be ruled out in strict LSM contests. The idea is to use the insight in Chade and Swinkels (2020) that NUC can be helpful in establishing strict quasiconcavity. In particular, it can be shown that $-P^i_j(\mathbf{a})$ is strictly quasiconcave in a_i in any strict LSM contest. Indeed, $P^i_{i12}(\mathbf{a}) > 0$ whenever $P^i_{12}(\mathbf{a}) = 0$. Thus, for any a_j, $P^i_{12}(\mathbf{a}) = 0$ for at most one a_i. A counterpart to Theorem 1 is now clear.

Theorem 2 *Consider any strict LSM contest with two agents. If $P^i_{12}(a'_1, a'_2) = 0$ for some interior (a'_1, a'_2), then $P^i_{12}(a_1, a_2) > 0$ ($P^i_{12}(a_1, a_2) < 0$) for all interior (a_1, a_2) with $a_i \geq a'_i$ and $a_j \leq a'_j$ ($a_i \leq a'_i$ and $a_j \geq a'_j$) and at least one strict inequality.*

Theorem 2 implies that the action space once again consists of two regions, divided by a strictly increasing (when interior) locus of points along which $P^i_{12} = 0$. However, the dividing line need not coincide with the line where $P^i = \frac{1}{2}$. In other words, with heterogenous technologies it is no longer necessarily the case that it is the favorite (underdog) that views actions as strategic complements (substitutes). Nevertheless, the general shape of the reaction functions are unchanged.

Proposition 5 *Consider any strict LSM contest with two agents and assume that A1 and A2 hold. Then, $b_i(a_j) > 0$ on an interval, and it is* ***strictly single-peaked*** *in a_j on the interior of this interval.*

3.3 Equilibrium

In any game, the properties of the reaction functions help determine both the number of equilibria and the comparative statics. Thus, attention now turns to equilibria of strict LSM contests. The next sections are devoted to comparative statics.

A pure-strategy equilibrium is *interior* if both agents take strictly positive actions in equilibrium. Given the shape of the reaction functions, it is straightforward to show that there is at most one interior pure-strategy equilibrium when A1 is satisfied. Clearly, if A2 is violated then the contest has no interior equilibrium. Similarly, if valuations across agents are too asymmetric, an interior equilibrium need not exist. The agent with the stronger valuation may work so hard that she completely deters the other agent. To illustrate, Figure 2 revisits the set-up in Figure 1. It describes how the valuations determine the nature of the equilibrium. Finally, if there is no interior equilibrium, then it is possible that no pure-strategy equilibrium exists at all, due to the discontinuity that arises in cases where $F^1(x|0)$ and $F^2(x|0)$ are degenerate.

Theorem 3 (Existence and uniqueness) *Consider any strict LSM contest with two agents in which A1 holds. Then, there is at most one interior pure-strategy Nash equilibrium. If $F^1(x|0)$ and $F^2(x|0)$ are non-degenerate, then a unique pure-strategy equilibrium exists for any (v_1, v_2).*

Consider the case in which $F^1(x|0)$ and $F^2(x|0)$ are degenerate. Perturb the game by defining some $\varepsilon > 0$ and restrict the action set to $[\varepsilon, \infty)$ rather than $[0, \infty)$. It follows from the proposition that as long as $\varepsilon > 0$, the perturbed game has a pure-strategy Nash equilibrium. Similar observations can be found in the literature on logit CSFs. See Ewerhart (2014) and the references therein.

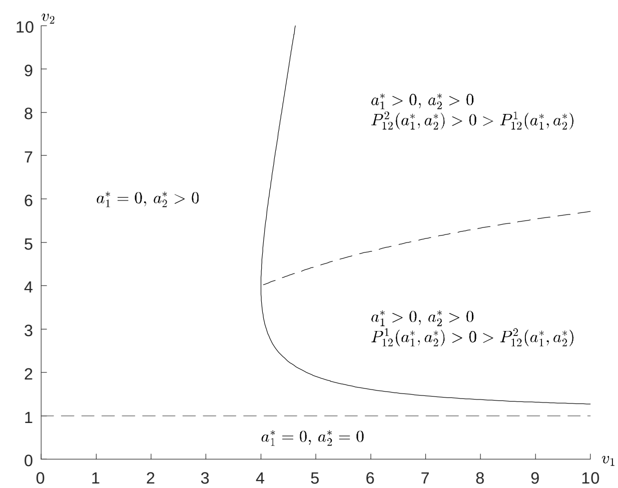

Figure 2: Equilibrium properties as a function of (v_1, v_2).

4 Comparative statics: Homogeneous technologies

This section examines the comparative statics of strict LSM contests with homogeneous technologies. Technologies are held fixed throughout, and focus is entirely on the consequences of changes to the agents' other characteristics, such as their valuations.[9] This approach makes it possible to determine whether "classical" comparative statics are robust within the class of LSM contests with homogeneous technologies.

For expositional simplicity, the comparative statics take as the starting point a contest in which an interior pure-strategy Nash equilibrium exists, which implies that A2 is satisfied. Moreover, A1 is assumed to be satisfied in the original contest.

4.1 Discouragement and disparity effects

As noted by Drugov and Ryvkin (2022), it is often argued that asymmetry is "detrimental for incentives in contests". To evaluate this claim from the point of view of an individual agent, consider the following question: Does an agent exert lower effort when matched with an opponent who is different from herself?

To answer the question, assume for now that technologies are fully homogeneous. Hence, the reaction functions are strictly hump-shaped and, importantly, reach their peak where $P^i(\mathbf{a}) = \frac{1}{2}$. Thus, the very highest possible equilibrium action for agent i is if the contest is such that agent i wins with probability $\frac{1}{2}$, i.e. if the outcome of the contest is even. This occurs when she is matched with an agent who has the exact same characteristics as herself, or $v_1 = v_2$, $p_1(\cdot) = p_2(\cdot)$, and $c_1(\cdot) = c_2(\cdot)$. All else equal, if the two agents have different valuations, say, then the reaction functions are different, and the two will not intersect at the peak of agent i's reaction function.

Corollary 1 (Individual discouragement effect) *Consider a strict LSM contest with two agents and fully homogeneous technologies. Assume that A1 and A2 hold when agents are symmetric. Then, an interior and symmetric pure-strategy equilibrium exists in the symmetric contest. Moreover, asymmetries discourage effort: When a pure-strategy Nash Equilibrium exists, agent i's equilibrium action is* ***strictly*** *lower when agent j's characteristics change in such a way that $P^i(\mathbf{a}) \neq \frac{1}{2}$ in equilibrium.*

[9] A rescaling of the cost function by a constant factor is isomorphic to a change in the agent's valuation. More general changes in the cost function can be dealt with as well, but at greater expositional cost. The consequences of changing the impact function is similar to the effects of changing the cost function, since what matters to the agent is the cost of attaining a given impact.

To illustrate, if $v_j > v_i = v$ then agent i is competing against someone who is "stronger". The competitor works hard, and agent i views actions as strategic substitutes. Hence, she is discouraged. On the other hand, if $v_j < v_i$, then agent i is competing against a "weaker" opponent. This opponent works less hard and agent i views actions as complements. Hence, agent i is not encouraged to work very hard.

Corollary 1 does not extend beyond fully homogeneous technologies. To illustrate, recall that reaction functions can be monotonic in general LSM contests. When that occurs, the agent's equilibrium action is monotonic in the other agent's valuation. Nevertheless, it is argued in Section 6 that the individual discouragement effect is responsible for some key results in the literature on optimal biases in logit contests. Section 6 contains a discussion of the robustness of these insights.

Drugov and Ryvkin's (2022) main focus is on the aggregate effect of asymmetry. They begin by elucidating what they term the "standard story." The story is that the weaker agent works less hard, which in turn means that the stronger agent has little incentive to work hard either. Their main objective is to clarify that the story is true only in some contest models. However, the standard story has merit in any LSM contests with homogeneous technologies in the sense that if the underdog becomes even weaker, then she lowers her effort as a direct result of which the favorite – who views actions as strategic complements – responds by working less hard as well.

Corollary 2 (Collective discouragement effect) *Consider a strict LSM contest with two agents and homogeneous technologies. Assume that A1 holds and that there is an interior pure-strategy Nash Equilibrium in which agent i is the underdog, or* $p_i(a_i) < p_j(a_j)$ *in equilibrium. Then, a decrease in* v_i *causes both* a_i *and* a_j *to decrease in equilibrium as long as a pure-strategy Nash Equilibrium exists.*

The asymmetry between players also increases if the favorite becomes even stronger. In this case, the favorite works even harder and in doing so she discourages her opponent (who views actions as strategic substitutes). Hence, the disparity between impacts widens. As a result, the disparity in winning probabilities widens as well, meaning that the stronger agent wins even more often than before.

Corollary 3 (Collective disparity effect) *Consider a strict LSM contest with two agents and homogeneous technologies. Assume that A1 holds and that there is an interior pure-strategy Nash Equilibrium in which agent i is the favorite, or* $p_i(a_i) >$

$p_j(a_j)$ in equilibrium. Then, an increase in v_i causes a_i to increase and a_j to decrease in equilibrium as long as a pure-strategy Nash Equilibrium exists. The disparity in winning probabilities increases and agent i wins even more often than before.

In summary, Corollaries 2–3 describe different adverse effects of increased asymmetry. If the underdog becomes weaker, then both agents work less hard. If the favorite becomes stronger, then the outcome of the contest becomes even more uneven. These insights are robust across all LSM contests with homogeneous technologies. In contrast, and as discussed in the introduction, there is no robust result on whether total effort increases or decreases when agents become more asymmetric.[10]

4.2 Precommitment in sequential contests

Dixit (1987) is interested in sequential contests in which the leader commits to an action that is observed by the follower before the latter makes her move. Dixit's aim is to understand whether the opportunity for such precommitment gives the leader an incentive to increase or decrease her action. He assumes that the CSF is of the probit or the logit form, while allowing for identity-dependent impact functions.

Dixit articulates the intuition as follows: since "each player's effort level harms the other [it is] strategically desirable for [the leader] to precommit his effort level in such as way as to induce a lower effort from [the follower] in response." How to induce a lower effort by the follower depends on whether she considers actions to be strategic substitutes or complements. By Theorem 1, starting from the equilibrium action profile in the simultaneous-move contests, the favorite (underdog) views actions as strategic complements (substitutes) whenever technologies are homogeneous. Thus, Dixit's (1987) main result can be generalized, as stated in the next Corollary.

Corollary 4 *Consider a strict LSM contest with two agents and homogeneous technologies. Assume that A1 holds and that there is an interior pure-strategy Nash Equilibrium, (a_1^*, a_2^*), in the simultaneous contest. Then, in a sequential contest, if the leader was the favorite (underdog) in the simultaneous contest, she has an incentive*

[10] However, total effort is not always the most relevant metric in contests with stochastic performance. For instance, expected performance is typically non-linear in effort, meaning that the distribution of effort across agents matters if the regulator cares about total expected performance. Likewise, the policy maker may be interested in the best performance, but the distribution of this statistic again depends on the whole action profile and not just total effort.

to increase (decrease) her action in the sequential contest, or

$$\frac{\partial u^i(a_i, b_j(a_i))}{\partial a_i}_{|a_i=a_i^*} \gtreqless 0 \Longleftrightarrow P^i_{12}(a_1^*, a_2^*) \gtreqless \frac{1}{2}.$$

Corollary 4 is a local result. It is hard to formally rule out that a very large departure from a_i^* in the opposite direction could be better. Dixit (1987) himself is quite careful to focus on actions in a "neighborhood of the Nash equilibrium".[11]

5 Ordered LSM contests

Corollaries 2–4 do not necessarily hold in contests with heterogeneous technologies. The reason is that the link between who views actions as complements and who is the favorite may break down. The purpose of this section is to examine a subset of heterogeneous technologies in which the link can be reestablished.

Definition 8 (Ordered LSM contests) *Any strict LSM contest with two agents is an ordered LSM contest if*

(T1) $\frac{f^1(x|a_1)}{F^1(x|a_1)} \geq \frac{f^2(x|a_2)}{F^2(x|a_2)}$ *for all* $x \in (\underline{x}, \overline{x})$ *when* $a_1 = a_2 > 0$*,*

(T2) $\frac{f^1_a(x|a_1)}{F^1_a(x|a_1)} \geq \frac{f^2_a(x|a_2)}{F^2_a(x|a_2)}$ *for all* $x \in (\underline{x}, \overline{x})$ *when* $a_1 = a_2 > 0$*,*

(T3) $\frac{-F^1_a(x|a_1)}{f^1(x|a_1)} \geq \frac{-F^2_a(x|a_2)}{f^2(x|a_2)}$ *for all* $x \in (\underline{x}, \overline{x})$ *when* $a_1 = a_2 > 0$*, and*

(C1) $v_1 \geq v_2$ *and the two agents have the same cost function,* $c(\cdot)$*.*

Properties (T1)–(T3) relate to the agents' technologies and property (C1) to their other characteristics. To illustrate and motivate, consider the case of homogeneous technologies with identity-dependent impact functions. Assume that agent 1 is stronger than agent 2 in the sense that $p_1(a) \geq p_2(a)$. Then, MLRP implies (T1) and NUC implies (T2). Property (T3) is more model-specific and generally requires $p_1(a)$ to be sufficiently steeper than $p_2(a)$. For rank-order tournaments with production function $\pi^i(\varepsilon_i, a_i) = p_i(a_i) + \varepsilon_i$ and i.i.d. noise with log-concave density, (T3) holds if $p_1'(a) \geq p_2'(a)$. In homogeneous best-shot contests, $\frac{p_1'(a)}{p_1(a)} \geq \frac{p_2'(a)}{p_2(a)}$ is required.

[11] There are other technical complications. Baye and Shin (1999) note that if agents are completely symmetric, or $P^i(\mathbf{a}^*) = \frac{1}{2}$, then (consistent with Corollary 4) there is no first-order effect of a change in the leader's action, but this does not rule out a second-order effect. Baik and Shogren (1992) endogenize the order of moves.

Property (T1) says that agent 1 has a more productive performance technology than agent 2 in the sense of reverse hazard rate dominance. Thus, agent 1 is the favorite if $a_1 > a_2$. Property (T2) is related to NUC. By combining (T2) and strict NUC,

$$\frac{f_a^1(x|a_1)}{F_a^1(x|a_1)} > \frac{f_a^1(x|a_2)}{F_a^1(x|a_2)} \geq \frac{f_a^2(x|a_2)}{F_a^2(x|a_2)} \tag{6}$$

when $a_1 > a_2$. Using the same argument as in the proof of Theorem 1 then implies that agent 1 views actions as strategic complements if $a_1 > a_2$.

The last two properties are used to rank incentives. Since $P^1(\mathbf{a}) = 1 - P^2(\mathbf{a})$,

$$u_1^1(\mathbf{a}) = v_1 \int \left(-F_a^1(x|a_1)\right) f^2(x|a_2)dx - c(a_1).$$

Hence, (T3) and (C1) imply that $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ if $a_1 = a_2$. At this point, more explicit structure is imposed on the agent's problem. In particular, assume in addition to A1 that $u^i(\mathbf{a})$ is quasiconcave in a_i. Then, $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ at $a_1 = a_2$ implies that if agent 2's best response to some a' is a', then agent 1's best response to a' is no smaller than a'. This property is sufficient to conclude that $a_1 \geq a_2$ in equilibrium.

Proposition 6 *Consider any ordered LSM contest in which A1 holds and $u^i(\mathbf{a})$ is quasiconcave in a_i, $i = 1, 2$. Assume that there is an interior equilibrium. Then, $a_1 \geq a_2$ in equilibrium, with $a_1 > a_2$ if the inequality in (T3) and/or (C1) is strict. Whenever $a_1 > a_2$ in equilibrium, agent 1 is the favorite and views actions as complements.*

Note that agent 1 is weakly stronger than agent 2 along two dimensions in ordered LSM contests. She has both the more productive technology by (T1) and the higher valuation by (C1). The significance of Proposition 6 is that it re-connects the identity of the favorite (underdog) with the identity of the agent who considers actions to be complements (substitutes). Hence, ordered LSM contests have many of the same key features as LSM contests with homogeneous technologies.

Corollary 5 *Consider any ordered LSM contest in which A1 holds and $u^i(\mathbf{a})$ is quasiconcave in a_i, $i = 1, 2$. Assume that the inequality in (T3) and/or (C1) is strict and that there is an interior equilibrium. Then, agent 1 (2) is the favorite (underdog) in equilibrium, and the conclusions of Corollaries 2–4 hold.*

Three examples of ordered LSM contests follow. The examples are selected to show how ordered LSM contests relate to, and extend, the standard models.

Example 1 (continued): Consider a rank-order tournament with identity-independent production function, $\pi(\varepsilon_i, a_i)$, which is assumed to be strictly increasing in both arguments. Hence, given her action a_i, agent i's production is below x_i if ε_i is below some threshold, $T(x, a_i)$. As in Bastani, Giebe, and Gürtler (2022), let the distribution of ε_i be identity-dependent, and described by $Q_i(\varepsilon_i)$. Then, $F^i(x|a_i) = Q_i(T(x, a_i))$.[12] Assume that Q_i and T is such that the contest is a strict LSM contest.

Bastani, Giebe, and Gürtler (2022) assume that agents have the same valuations and cost functions. They conclude that there is a symmetric equilibrium even though performance technologies are heterogeneous. To understand why, note that

$$\frac{F_a^i(x|a_i)}{f^i(x|a_i)} = \frac{T_a(x, a_i)}{T_x(x, a_i)}$$

is identity-independent. Hence, (T3) trivially holds.

Assume now also that Q_1 dominates Q_2 in terms of the likelihood-ratio, which is exactly identical to (T2) in the current example, and stronger than (T1). In this case, Bastani, Giebe, and Gürtler's (2022) equilibrium characterization can be further extended. First, assuming quasiconcavity as in Proposition 6, equilibrium is unique. Second, if $v_1 > v_2$ then $a_1 > a_2$ in equilibrium and Corollary 5 applies. If Q_1 strictly dominates Q_2 in terms of the likelihood-ratio, then agent 1 is the favorite and views actions as complements even if $v_1 = v_2$, and $a_1 = a_2$ in equilibrium. ▲

Example 2 (continued): Consider a version of the best-shot model in which

$$F^i(x|a_i) = H_i(x)^{a_i}, \quad x \in [\underline{x}, \overline{x}],$$

where $H_i(x)$ is now allowed to be an identity-dependent distribution. Note that the support is the same for both agents, however. This setting describes a strict LSM contest. Generally speaking, the CSF no longer reduces to a logit CSF. An exception

[12] An isomorphic contest arises if the function $T(x, a_i)$ is itself a distribution function. For instance, assume that $F^i(x|a_i) = T(x, a_i)^{k_i}$, $k_i \geq 1$. In this setting, agent i has an exogenous number, k_i, of team members. Each member independently delivers some performance, the distribution of which is $T(x, a_i)$ and which depends on the agent's effort at motivating the team, a_i. However, the agent is judged based only on the best performance within her team, the distribution of which is exactly $T(x, a_i)^{k_i}$.

is when $H_i(x) = H(x)^{\alpha_i}$, where $\alpha_i > 0$ is an identity-dependent parameter, but in this case agents have fully homogeneous technologies.

The first three properties of ordered LSM contests are satisfied under two further assumptions. First, assume that H_1 dominates H_2 in terms of the reverse hazard rate, or equivalently that $\frac{H_1(x)}{H_2(x)}$ is increasing. This directly implies (T1), and (T2) can be seen to follow. The interpretation is that any given draw that agent 1 makes is more likely to be of high quality than any given draw that agent 2 makes. The second assumption is that $\frac{\ln H_1(x)}{\ln H_2(x)}$ is increasing as well, which implies (T3). ▲

The final example provides a different perspective on the somewhat surprising result in Bastani, Giebe, and Gürtler (2022, henceforth BGG) that heterogeneous noise distributions do not lead to heterogeneous incentives when the production functions are homogeneous. The example establishes that BGG's result is due to the fact that they assume that noise is univariate. The result does not hold more generally, i.e. when more sources of noise is permitted.

Example 3 (rank-order tournaments with multivariate noise): Assume that agent i's production function can be written as $\pi^i(\varepsilon_i, \theta_i, a_i) = \theta_i a_i + \varepsilon_i$, $i = 1, 2$. If one of θ_i or ε_i is deterministic or degenerate and the other is stochastic, then the model reduces to the one in BGG. In the remainder, both θ_i and ε_i are assumed to be stochastic. In other words, the agent is subjected to multivariate noise. For concreteness, assume that θ_i follows an exponential distribution with rate parameter $r_i > 0$ and mean $\frac{1}{r_i}$. Hence, $\theta_i a_i$ is exponentially distributed with rate parameter $\frac{r_i}{a_i}$ and mean $\frac{a_i}{r_i}$. Assume that ε_i follows an exponential distribution with rate parameter $\lambda_i > 0$, and mean $\frac{1}{\lambda_i}$. Note that the agent's expected performance is lower the higher r_i and λ_i are.

Finally, assume that θ_i and ε_i are independent. The agent's performance is the convolution of two exponential distributions, which gives rise to the hypoexponential distribution

$$F^i(x_i|a_i) = 1 - \frac{\lambda_i e^{-\frac{r_i x_i}{a_i}} - \frac{r_i}{a_i} e^{-\lambda_i x_i}}{\lambda_i - \frac{r_i}{a_i}}$$

when the rate parameters are distinct, or $\lambda_i \neq \frac{r_i}{a_i}$. When the rate parameters coincide, the distribution is an Erlang distribution, which can be obtained from the hypoexponential distribution via an application of L'Hôpital's rule.

It is straightforward to verify that the contest is a strict LSM contest. Similarly, it can be verified that T1 and T2 hold if agent 1 has a more productive technology than agent 2 in the sense that $r_1 \leq r_2$ and $\lambda_1 \leq \lambda_2$. The more interesting condition is T3, which as explained above holds in a trivial sense in BGG's setting with univariate noise. However, with multivariate noise, the inequality in T3 is *strict* if $r_1 < r_2$ and $\lambda_1 = \lambda_2$. Thus, the contest is an ordered LSM contest if C1 holds and $r_1 \leq r_2$ and $\lambda_1 = \lambda_2$. On the other hand, when $r_1 = r_2$ and $\lambda_1 < \lambda_2$, the inequality in T3 is *reversed*, in which case it is possible that the agent with the less productive technology (agent 2) has stronger incentives (depending on her valuation). In either case, it is to be expected that the two agents take different actions even if their valuations are identical, unlike in BGG.

To understand the intuition, recall that a higher r_i implies that θ_i has lower mean, which in turn means that an increase in a_i does less to change the agent's production. Hence, incentives are weaker and the agent is betting on a high realization of ε_i compensating for her low effort. On the other hand, a higher λ_i implies that ε_i has lower mean. Hence, the agent has to bet on $a_i\theta_i$ for a good performance, and incentives are therefore larger. This intuition is not valid in BGG, because there is no interaction or substitutability in that setting between different sources of noise. ▲

Example 3 illustrates the parsimonious character of the Mirrlees approach. The agent ultimately cares about the distribution of her performance, and how her action changes it. These characteristics are directly captured in the distribution function that the Mirrlees approach relies on. Whether there is one or more source of noise in the background is not important per se. Indeed, the best-shot model can also be seen as the result of a production function with multivariate noise. Imagine that there is an infinite sequence of i.i.d. random variables in the background. The action a_i is the number of these that agent i gets to inspect. She picks the best of the realizations available to her, and submits this as her entry into the contest. In this sense, the best-shot model with identity-dependent H_i described earlier can already be seen as an illustration that BGG does not hold with multivariate noise.[13]

Example 3 also cautions that not all LSM contests are ordered LSM contests. Hence, it should be expected that the received wisdom in Corollaries 2–4 may fail

[13] The best-shot model can also be forced into a univariate state-space representation. Consider $\pi^i(\varepsilon_i, a_i) = \varepsilon_i^{1/a_i}$ and assume that ε_i is uniformly distributed on $[0,1]$. Then, $F^i(x_i|a_i) = x_i^{a_i}$, $x_i \in [0,1]$. The production function is not concave in a_i, yet $F^i(x_i|a_i)$ satisfies Rogerson's (1985) regularity condition that leads expected utility to be concave in a_i when $c_i(a_i)$ is weakly convex.

in contest with heterogeneous technologies more generally. Section 6.2 discusses a tractable model that can be used to illustrate such failures.

6 Discussion and extensions

The first part of this section discusses some of the assumptions, beginning with Assumption A1. It it also explained and demonstrated why (strict) NUC is crucial to the results when the support is the same for all agents. The supplementary material establishes that Theorem 1 holds when the support shifts with the agent's action. In fact, NUC can be relaxed in some such cases. The supplementary material also shows that some existing CSFs cannot be microfounded as LSM contests. This is the case for the "relative difference CSF" with two or more agents in Beviá and Corchón (2015) and the "difference-form CSFs" in Cubel and Sanchez-Pages (2016) with at least three agents. The structure of these CSFs is not consistent with what is implied by the LSM model.

The second part of this section examines other extensions. Properties of larger LSM contests are identified. A discussion of biased contests concludes the section. In both cases, details are in the supplementary material.

6.1 Quasiconcavity and unique best responses

Assumption A1 is stronger than required in practice. It implies that the properties of the reaction function, $b_i(a_j)$, can be characterized for all $a_j \in (0, \infty)$. However, recall that agent j's equilibrium action can never exceed $\overline{a}_j$. Thus, for equilibrium existence and uniqueness, as well as for most of the comparative statics, it is enough that a unique best response exists to any $a_j \in (0, \overline{a}_j]$. A sufficient condition is that $u^i(\mathbf{a})$ is strictly quasiconcave in a_i for any $a_j \in (0, \overline{a}_j]$. Now, as alluded to earlier, LSM contests have special properties that make the latter condition easier to check.

Recall that $P^i_i(\mathbf{a})$ is log-supermodular, or that $\frac{P^i_{ii}(\mathbf{a})}{P^i_i(\mathbf{a})}$ is increasing in a_j. The implication is that $P^i(\mathbf{a})$ in a sense becomes more convex in a_i when a_j increases, thus suggesting that the agent's utility maximization problem is less likely to be well-behaved the larger a_j is. Hence, it is sufficient to check that $u^i(\mathbf{a})$ is strictly quasiconcave in a_i when $a_j = \overline{a}_j$. This is the case if $P^i(a_i, \overline{a}_j)$ is "more concave" than $c_i(a_i)$ is convex at any stationary point. A sufficient condition is provided in the supplementary material, along with extensions to several agents.

6.2 The importance of (strict) NUC

Chade and Swinkels (2020) remark that "NUC is satisfied for every distribution that we are aware of that is commonly used in economic applications." Nevertheless, they provide counterexamples, including

$$F^i(x_i|a_i) = -\frac{1}{3}a_i^2x_i^3 + \frac{1}{2}a_i^2x_i^2 - \frac{1}{6}a_i^2x_i + 2a_ix_i^2 - 2a_ix_i + x_i,\ \ x_i \in [0,1],$$

defined on $a_i \in [0, 0.48]$. Here,

$$-F^i_{a_i}(x_i|a_i) = \frac{1}{3}x_i\left(1 - x_i\right)\left(a_i - 2a_ix_i + 6\right)$$

which is strictly *log-submodular* in (x_i, a_i). Thus, in a two-agent contests in which both agents have this kind of distribution function, the exact opposite conclusions of those in Theorem 1 obtain. That is, the favorite (underdog) views actions as strategic substitutes (complements).[14] As a result, the reaction functions are either monotonic or u-shaped, rather than hump-shaped. For instance, when the cost function is $c_i(a_i) = a_i^2$, the reaction function is

$$b_i(a_j) = \min\left\{0.48, \frac{30v_i + a_j^2v_i}{2a_jv_i + 180}\right\}.$$

Depending on v_i, $b_i(a_j)$ is either monotonic or first decreasing and then increasing.

Kirkegaard (2024) analyzes the mixture model of contests with homogeneous technologies. This is a contest in which distribution functions take the form

$$F^i(x_i|a_i) = p_i(a_i)H(x_i) + (1 - p_i(a_i))G(x_i),$$

where H and G are distribution functions, and where H first-order stochastically dominates G in the strict sense that $H(x_i) < G(x_i)$ on the interior. The impact function $p_i(a_i) \in [0,1]$ is strictly increasing, with $p_i'(a_i) > 0$. Then, $\ln(-F^i_a(x_i|a_i)) = \ln p_i(a_i) + \ln(G(x_i) - H(x_i))$ is additively separable. In other words, the mixture

[14] More directly, agent 1's CSF is $P^1(a_1, a_2) = -\frac{1}{90}a_1^2a_2 + \frac{1}{90}a_1a_2^2 + \frac{1}{3}a_1 - \frac{1}{3}a_2 + \frac{1}{2}$. Hence, $P^1_{12}(a_1, a_2) = \frac{a_2 - a_1}{45}$, which is negative when agent 1 is the favorite, or $a_1 > a_2$.

model is *log-modular* and is the boundary case in which NUC is satisfied in the weakest possible sense.

Retracing the steps that led to Theorem 1 proves that $P^i_{12}(\mathbf{a}) = 0$ when $n = 2$. Thus, agent i's best response is independent of a_j and the reaction function is horizontal. That is, each agent has a dominant action. See Kräkel (2010) and Gürtler and Kräkel (2012) for applications of the mixture model with two agents or teams.

Kirkegaard's (2024) focus is on contests with three or more agents, in which case he proves that actions are strategic substitutes in a global sense. The model has so much structure that it is easy to incorporate multivariate incomplete information about valuations, impact functions, cost functions, and action sets. Thus, he is able to study the comparative statics of changes in the multivariate dependence structure (and thus correlation) between such characteristics as valuations and budget constraints.

Kirkegaard (2023b) considers a two-agent mixture model with heterogeneous technologies, i.e. where H and G are identity-dependent. Then, reaction functions are monotonic. Among other things, he considers contests that violate the first three properties of ordered LSM contest and where Corollaries 2–4 need not hold. Thus, the mixture model illustrates that the standard results do not always carry over to environments with heterogeneous technologies.

6.3 Larger LSM contests

The paper concentrates on LSM contests with two agent, in large part because the predictions and implications of the model are so clear-cut in this case. However, contests with more agents also have systematic properties, some of which are outlined in the following. The results are formalized in the supplementary material.

As a point of reference, Dixit (1987) notes that is it is not necessarily the case that any agent takes actions to be (pairwise) strategic complements when there are many agents. With this in mind, consider any pair of agents in a LSM contest with homogeneous technologies. Within the pair, the agent who is weakly less likely to win can be shown to view the actions of the two agents as strategic substitutes. An implication is that in a symmetric pure-strategy equilibrium of a completely symmetric contest, agents views actions as strategic substitutes in equilibrium. Thus, if utility is concave in the agent's own action in such a contest, then there is at most one symmetric pure-strategy equilibrium, and the equilibrium action is increasing in the common valuation.

It is impossible in general to determine how the agent who is more likely to win views actions. However, utilizing Proposition 1, it can be shown that any agent is more likely to consider the actions of the pair to be strategic substitutes the harder their common rivals work. This holds even if technologies are heterogeneous.

Giebe and Gürtler (2024) examine a rank-order tournament with identity-independent production functions but heterogeneous noise distributions. They rank the equilibrium actions of any two agents, say agents 1 and 2, if (T1) and (C1) from Section 5 hold (they assume that $v_1 = v_2$, but their proof carries through to $v_1 \geq v_2$). Recall that (T3) is trivially satisfied in their setting. Their proof that $a_1 \geq a_2$ in any interior equilibrium relies, in part, on the fact that $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ whenever $a_1 = a_2$.

To generalize, consider any strict LSM contest in which the relationship between agents 1 and 2 satisfies (T1), (T3), and (C1). Then, regardless of the total number of agents, it remains the case that $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ if $a_1 = a_2$. However, since less can be said about the structure of the reaction functions, the proof of Proposition 6 does not directly extend. Nevertheless, it turns out that Giebe and Gürtler's (2024) method of proof also applies to the more general setting. However, as alluded to, this proof relies on more than just $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ whenever $a_1 = a_2$. It is also required that the cost function is "sufficiently convex", or in other words that $\min_{a \in [0, \overline{a}_i]} c''(a_i)$ is sufficiently large. However, with this assumption, the result in Giebe and Gürtler (2024) generalizes.[15] Note that if (T2) is added, then the weaker of the two (agent 2) views the pair's actions as substitutes, even though nothing can be said in general about the stronger agent (agent 1).

6.4 Biased LSM contests

The rules of the game have been held fixed. Fu and Wu (2020), and the references therein, instead consider logit contests in which a regulator can somehow manipulate the impact of an agent's action. For a host of natural objective functions, their results imply that it is optimal in two-agent contests to ensure the bias is such that a completely level playing field is created.[16] Their biases amount to affine transformations of the impact functions.

[15] Giebe and Gürtler (2024) also rank equilibrium actions with different prize structures.

[16] This is just one implication. Fu and Wu's (2020) main contribution is to develop a method to handle the case with several agents.

The result should be seen in light of the individual discouragement effect in Corollary 1. After all, any agent works hardest if the equilibrium is such that she wins with probability $\frac{1}{2}$ in equilibrium. The supplementary material formalizes this intuition and extends it to LSM contests more generally. It is shown that the affine transformations that Fu and Wu (2020) use work only because the logit CSF is homogeneous of degree zero in impacts. More flexible transformations are required in general.

Moreover, the individual discouragement effect changes when technologies are heterogeneous. The peak of the reaction function is still where $P^i_{12}(\mathbf{a}) = 0$, but this is not necessarily where $P^i(\mathbf{a}) = \frac{1}{2}$. Thus, optimal biases generally lead to an uneven playing field.

Kirkegaard (2023a) advocates for another approach to the optimal design of contests with stochastic performance. He argues that since performances, but not necessarily actions, are observable, biases are more likely to be applied to the former. The two approaches are not generally equivalent. Kirkegaard (2023a) provides an example in which a large improvement over Fu and Wu (2020) is obtained by biasing how performance is scored and evaluated across agents. Generally, this approach also leads to an uneven playing field. Thus, this section has presented two reasons, based on general contests with stochastic performance, to be cautious of the claim that biases should be used to level the playing field.

7 Conclusion

Currently, the literature on contests – and the ensuing economic insights – rely on rather specific contest models. There is no general or unifying model. However, in a recent survey Fu and Wu (2019) conclude by noting that "theoretical predictions can be sensitive to modelling nuances. It is important to examine more thoroughly the robustness of previous results and the logic for their robustness/fragility."

The current paper presents a unifying model of noisy or non-deterministic contests. It is assumed that noise is due to stochastic performance. Without imposing specific functional forms, a large family of such contests is examined. The resulting log-supermodular (LSM) contests are defined only by weak regularity conditions on the performance technologies. The results for all the contests within this family share the same "logic for their robustness/fragility". However, in this respect it turns out to be important whether technologies are homogeneous or heterogeneous.

The strategic incentives and the comparative statics are robust across all LSM contests with homogeneous technologies. This conclusion is important because it supports the vast literature on noisy contests that relies on two of the workhorse models, both of which are nested within the current framework. Most significantly, reaction functions are single-peaked and incentives are determined by whether the agent is the favorite or the underdog. It is straightforward to identify the mechanisms responsible for these unified results, as they are in fact all rather easily traced back to the weak log-supermodularity properties that the performance technologies are assumed to have.

The model also provides a foundation for challenging the standard assumption that technologies are homogeneous across agents. With heterogeneous technologies, reaction functions are still single-peaked. However, incentives are no longer solely determined by whether the agent is the favorite or the underdog. Nevertheless, a set of "ordered contests" with heterogenous technologies is identified in which the usual comparative statics are recovered.

The paper mostly focuses on contests with two agents. A natural next step is to extend the comparative statics to contests with several agents. Likewise, the technologies were held fixed in the comparative statics. It is left for future research to explore the comparative statics of changing technologies in general LSM contests.

References

Athey, S. (2002): "Monotone Comparative Statics under Uncertainty," *The Quarterly Journal of Economics*, 117: 187–223.

Baye, M.R., Kovenock, D. and C.G. de Vries (1993): "Rigging the Lobbying Process: An Application of the All-Pay Auction," *The American Economic Review*, 83: 289–294.

Beviá, C. an L.C. Corchón (2015): "Relative difference contest success function," *Theory and Decision*, 78: 377–398.

Bulow, J., Geanakoplos, J. and P. Klemperer (1985): "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy*, 93: 488–511.

Baye, M.R. and O. Shin (1999): "Strategic Behavior in Contests: Comment," *The American Economic Review*, 89: 691–693.

Baik, K.H. and J.F. Shogren (1992): "Strategic Behavior in Contests: Comment," *The American Economic Review*, 82: 359–362.

Bastani, S., Giebe, T. and O. Gürtler (2022): "Simple equilibria in general contests," *Games and Economic Behavior*, 134: 264–280.

Chade, H. and J.M. Swinkels (2020): "The no-upward-crossing condition, comparative statics, and the moral-hazard problem," *Theoretical Economics*, 15: 445–476.

Chowdhury, S.M, Esteve-González, P. and A Mukherjee (2023): "Heterogeneity, leveling the playing field, and affirmative action in contests." *Southern Economic Journal*, 89: 924–974.

Cubel, M. and S. Sanchez-Pages (2016): "An axiomatization of difference-form contest success functions, " *Journal of Economic Behavior & Organization*, 131: 92–105.

Dixit, A. (1987): "Strategic Behavior in Contests," *The American Economic Review*, 77: 891–898.

Drugov, M. and D. Ryvkin (2020): "Tournament rewards and heavy tails," *Journal of Economic Theory*, 190, article 105116.

Drugov, M. and D. Ryvkin (2022): "Hunting for the discouragement effect in contests," *Review of Economic Design*, https://doi.org/10.1007/s10058-022-00308-4

Drugov, M., Ryvkin, D., and J. Zhang (2024): "Tournaments with a Standard," mimeo, March 2024.

Ewerhart, C. (2014): "Unique equilibrium in rent-seeking contests with a continuum of types," *Economics Letters*, 125: 115–118.

Ewerhart, C. and M. Serena (2024): "On the (Im-)Possibility of Representing Probability Distributions as a Difference of I.I.D. Noise Terms," *Mathematics of Operations Research*, forthcoming, https://doi.org/10.1287/moor.2023.0081.

Fang, D., Noe, T., and P. Strack (2020): "Turning Up the Heat: The Discouraging Effect of Competition in Contests," *Journal of Political Economy*, 128: 1940–1975.

Fu, Q. and Z. Wu (2019): Contests: Theory and Topics. *Oxford Research Encyclopedia of Economics and Finance.*

Fu, Q. and Z. Wu (2020): "On the optimal design of biased contests," *Theoretical Economics*, 15: 1435–1470.

Fullerton, R.L. and R.P. McAfee (1999): "Auctioning Entry into Tournaments," *Journal of Political Economy*, 107 (3): 573–605.

Giebe, T. and O. Gürtler (2024): "Player strength and effort in contests," mimeo, July 2024.

Gürtler, O. and M. Kräkel (2012): "Dismissal Tournaments," *Journal of Institutional and Theoretical Economics*, 168: 547–562.

Hirschleifer, J. and J. G. Riley, 1992, "*The Analytics of Uncertainty and Information*," Cambridge University Press.

Jewitt, I. (1988): "Justifying the first order approach to principal-agent problems," *Econometrica*, 56: 1177–1190.

Kirkegaard, R. (2023a): "Contest Design with Stochastic Performance," *American Economic Journal: Microeconomics*, 15: 201–238.

Kirkegaard, R. (2023b): "On Technological Heterogeneity in Contests," mimeo.

Kirkegaard, R. (2024): "Multivariate Incomplete Information in the Mixture Model of Contests," *The Rand Journal of Economics*, forthcoming.

Kräkel, M. (2010): "Shutdown Contests in Multi-Plant Firms and Governmental Intervention," IZA Discussion Paper No. 4852.

Lazear, E. P. and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89 (5): 841–864.

Malueg, D.A. and Yates, A.J. (2005): "Equilibria and comparative statics in two-player contests," *European Journal of Political Economy*, 21: 738– 752.

Rogerson, W.P. (1985): "The First-Order Approach to Principal-Agent Problems," *Econometrica*, 53 (6): 1357–1367.

Siegel, R. (2009): "All-pay contests," *Econometrica*, 77: 71–92.

Tullock, G. (1980) "*Efficient Rent Seeking*," in Toward a Theory of the Rent Seeking Society, edited by Buchanan, J.M, R.D Tollison, and G. Tulluck, Texas A&M University Press.

Appendix A: Proofs

Proof of Proposition 1. As in Athey (2002), two important properties are used repeatedly. First, the product of log-supermodular functions is itself log-supermodular. Second, log-supermodularity is preserved under integration. In particular, let $\mathbf{z} = (x, \mathbf{a})$ and think of x as a variable and $\mathbf{a}$ as parameters. Then, if $t(\mathbf{z})$ or $t(x, \mathbf{a})$ is log-supermodular in $(x, \mathbf{a})$, the function $T(\mathbf{a}) = \int t(x, \mathbf{a})dx$ is log-supermodular in $\mathbf{a}$.

In LSM contests, when $\mathbf{a} \in (0, \infty)^n$ the integrand in (2) is log-supermodular in $(x, \mathbf{a})$ since it is the product of log-supermodular functions. Then, since log-supermodularity is preserved under integration, $P^i(\mathbf{a})$ is log-supermodular in $\mathbf{a}$ on $(0, \infty)^n$. Note that this conclusion relies only on the MLRP.

Next, note that $P^i_j(\mathbf{a}) < 0$ for $j \neq i$ when $\mathbf{a} \in (0, \infty)^n$. That is, agent i is less likely to win when agent j works harder. For $j \neq i$, consider then

$$-P^i_j(\mathbf{a}) = \int (-F^j_a(x|a_j)) \prod\nolimits_{m \neq i,j} F^m(x|a_m) f^i(x|a_i)dx.$$

By the MLRP and NUC, the integrand is again log-supermodular in $(x, \mathbf{a})$. Hence, $-P^i_j(\mathbf{a})$ is log-supermodular in $\mathbf{a}$ on $(0, \infty)^n$.

Finally, consider the derivative with respect to a_i. Since winning probabilities sum to one, or $P^i(\mathbf{a}) = 1 - \sum_{j \neq i} P^j(\mathbf{a})$, it follows that $P^i_i(\mathbf{a}) = \sum_{j \neq i}(-P^j_i(\mathbf{a}))$ is the sum of log-supermodular functions. However, the sum of log-supermodular functions need not be log-supermodular. Hence, $P^i_i(\mathbf{a})$ is not guaranteed to be log-supermodular when $n \geq 3$. On the other hand, $P^i_i(\mathbf{a})$ must be log-supermodular when $n = 2$ since in this case $P^i_i(\mathbf{a}) = -P^i_j(\mathbf{a})$. ■

Proof of Theorem 1. Note that

$$P^i_{12}(\mathbf{a}) = \int F^j_a(x|a_j) f^i_a(x|a_i)dx$$

where $j \neq i$ is agent i's competitor. For expositional clarity, consider first the case where technologies are identity-dependent, or $F^i(x|a) = F^j(x|a) = F(x|a)$. Then,

$$P^i_{12}(\mathbf{a}) = \int F_a(x|a_j) f_a(x|a_i)dx.$$

Assume strict NUC is satisfied, i.e. that $-F_a$ is strictly log-supermodular. Then, for any interior action,

$$\frac{\partial}{\partial a}\frac{f_a(x|a)}{F_a(x|a)} > 0 \text{ for all } x \in (\underline{x}, \overline{x}).$$

Thus, for any interior action profile for which agent i is the favorite, or $a_i > a_i$, it must hold that

$$\frac{f_a(x|a_i)}{F_a(x|a_i)} > \frac{f_a(x|a_j)}{F_a(x|a_j)} \text{ for all } x \in (\underline{x}, \overline{x}), \tag{7}$$

or

$$F_a(x|a_j)f_a(x|a_i) > F_a(x|a_i)f_a(x|a_j) \text{ for all } x \in (\underline{x}, \overline{x}).$$

Hence, $P^i_{12}(\mathbf{a}) > P^j_{12}(\mathbf{a})$, which in turn implies that $P^i_{12}(\mathbf{a}) > 0$ and $P^j_{12}(\mathbf{a}) < 0$. Thus, in *any* strict LSM contest with identity-independent technologies, the agent who is the favorite (underdog) views actions as strategic complements (substitutes). The argument holds for all strict LSM contests with homogeneous technologies because, as explained previously, different impact functions amount to different cost functions, and the cost functions are irrelevant to the argument.

More formally, for homogeneous contests, or $F^i(x_i|a_i) = G(x_i|p_i(a_i))$, $i = 1, 2$, it holds that

$$P^i_{12}(\mathbf{a}) = p_i'(a_i)p_j'(a_j)\int G_p(x|p_j(a_j))g_p(x|p_i(a_i))dx.$$

Strict NUC means that $-F^i_a(x_i|a_i) = -G_p(x_i|p_i(a_i))p_i'(a_i)$ is strictly log-supermodular, which is equivalent to strict log-supermodularity of $-G_p(x_i|p_i(a_i))$. Hence, for any interior action,

$$\frac{\partial}{\partial p_i}\frac{g_p(x|p_i(a_i))}{G_p(x|p_i(a_i))} > 0 \text{ for all } x \in (\underline{x}, \overline{x}).$$

Agent i is the favorite if $p_i(a_i) > p_i(a_i)$, in which case

$$\frac{g_p(x|p_i(a_i))}{G_p(x|p_i(a_i))} > \frac{g_p(x|p_j(a_j))}{G_p(x|p_j(a_j))} \text{ for all } x \in (\underline{x}, \overline{x}),$$

and the rest of the proof is the same as before. ■

Proof of Proposition 2. Let $\tau_i(a_j)$ denote the a_i value for which $P^i(\mathbf{a}) = \frac{1}{2}$, as a function of a_j. In other words, $\tau_i(a_j) = p_i^{-1}(p_j(a_j))$ describes the curve along which the impacts are the same and along which the agents win with equal probability. Since p_i and p_j are strictly increasing functions with the same range, $\tau_i(a_j)$ is strictly increasing and satisfies $\tau_i(0) = 0$ and $\tau_i(a_j) \to \infty$ as $a_j \to \infty$.

If the best response to a_j is interior, it must satisfy the first order condition that

$$v_i P_i^i(b_i(a_j), a_j) - c_i'(b_i(a_j)) = 0.$$

By Assumption A1, it then holds that the second-order condition holds in a strict sense, or

$$u_{ii}^i(b_i(a_j), a_j) = v_i P_{ii}^i(b_i(a_j), a_j) - c_i''(b_i(a_j)) < 0.$$

Consider now a marginal change in a_j. The first-order condition implies that

$$b_i'(a_j) = \frac{v_i P_{12}^i(b_i(a_j), a_j)}{-u_{ii}^i(b_i(a_j), a_j)}$$

has the same sign as $P_{12}^i(b_i(a_j), a_j)$. Thus, if $P_{12}^i > 0$ – or, by Theorem 1, $b_i(a_j) > \tau_i(a_j)$ – then $b_i'(a_j) > 0$. Similarly, if $P_{12}^i < 0$, or $b_i(a_j) > \tau_i(a_j)$, then $b_i'(a_j) < 0$. In words, if $b_i(a_j)$ is above (below) $\tau_i(a_j)$ for some $a_j > 0$, then the reaction function must be locally increasing (decreasing) because agent i considers actions to be strategic complements (substitutes), by Theorem 1. Finally, if $P_{12}^i = 0$, or $b_i(a_j) = \tau_i(a_j)$, then $b_i'(a_j) = 0$. Thus, there can be at most one interior point at which $b_i(a_j)$ intersects the strictly increasing $\tau_i(a_j)$ function.

At the same time, there must be some interior a_j value for which $b_i(a_j) = \tau_i(a_j)$. By contradiction, assume there is no such value, or in other words that $b_i(a_j)$ is always either strictly above or strictly below $\tau_i(a_j)$. Now, assume that $b_i(a_j') > \tau_i(a_j') > 0$ for some a_j' and consider $a_j > a_j'$. If $b_i(a_j) > \tau_i(a_j)$ for all $a_j > a_j'$, then $b_i(a_j) \to \infty$ as $a_j \to \infty$ because $\tau_i(a_j) \to \infty$ as $a_j \to \infty$. However, this contradicts the fact that the best response is bounded above by $\overline{a}_i < \infty$. Thus, $b_i(a_j)$ must intersect the $\tau_i(a_j)$ curve to the right of a_j'. This argument also implies that there must be some a_j value for which $b_i(a_j) < \tau_i(a_j)$. However, $b_i(a_j)$ cannot be always below $\tau_i(a_j)$ either. If $0 < b_i(a_j'') < \tau_i(a_j'')$ for some a_j'', then $b_i(a_j)$ increases as a_j is reduced. Hence, $b_i(a_j)$ must intersect the $\tau_i(a_j)$ curve to the left of a_j''. This intersection is guaranteed because $\tau_i(0) = 0$ and $\tau_i(a_j)$ is an increasing function. In conclusion, there exists precisely one interior a_j value, $\widehat{a}_j$ for which $b_i(\widehat{a}_j) = \tau_i(\widehat{a}_j)$.

Finally, $b_i(a_j) > \tau_i(a_j)$ for all $a_j < \widehat{a}_j$ implies that $b_i(a_j) > 0$ for all $a_j \in (0, \widehat{a}_j)$. Along with the hump-shape of the reaction function, this implies that if $b_i(a_j') > 0$ then $b_i(a_j) > 0$ for all $a_j \in (0, a_j')$. In other words, $b_i(a_j) > 0$ on an interval of a_j values. The last part of the proposition follows from Theorem 1 and the fact that

$b_i(a_j) > \tau_i(a_j)$ on the increasing part and $b_i(a_j) < \tau_i(a_j)$ on the decreasing part. ■

Proof of Proposition 3. When technologies are homogeneous but not fully homogeneous, $\tau_i(a_j)$ need not be defined for all a_j and, if it is, its range need not be $[0, \infty)$. Nevertheless, the first steps of the proof of Proposition 2 still apply. Thus, if $P^i_{12} > 0$ then $b'_i(a_j) > 0$ and if $P^i_{12} < 0$ then $b'_i(a_j) < 0$. Likewise, there is at most one a_j value for which $b_i(a_j) = \tau_i(a_j)$.

If there is a point at which $b_i(a_j) = \tau_i(a_j)$, then the reaction function is strictly single-peaked whenever the best-response is strictly positive, by the same argument as in Proposition 2. If there is no such point, then it is because either $P^i_{12} > 0$ or $P^i_{12} < 0$ everywhere along the reaction function, but in this case the reaction function is strictly monotonic whenever the best response is strictly positive.

If the reaction function is decreasing, then it is still the case that $b_i(a_j) > 0$ on $(0, a_j^{\text{sup}})$. To see this, $P^i_{12} < 0$ along the reaction function if the reaction function is decreasing. Hence, $b'_i(a_j) < 0$. Then, $b_i(a_j)$ increases as a_j decreases, which means that it must remain strictly positive.

If the reaction function is increasing, then $P^i_{12} > 0$ along the reaction function. Since $b_i(a_j)$ decreases as a_j decreases, it may therefore become zero for small values of a_j. This is possible when $P^i_{12} > 0$ or $P^i > \frac{1}{2}$ near the origin (see the example following the proposition). Once it has become zero in this manner, it cannot become strictly positive for smaller but strictly positive values of a_j. The reason is that $P^i_{12} > 0$ would hold at such a point, implying that b_i is increasing, but this contradiction the property that b_i is zero for larger a_j.

If the reaction function is non-monotonic, then it is increasing on an interval, in which case the same argument as in the previous paragraph implies that $b_i(a_j)$ may be zero for small a_j values.

The last part of the proposition follows the same logic as in Proposition 2. ■

Proof of Proposition 4. For two-agent LSM contests, Proposition 1 reveals that $P^i_i(\mathbf{a})$ is log-supermodular in $\mathbf{a}$ on $(0, \infty)^2$. Hence, $\frac{P^i_{12}(\mathbf{a})}{P^i_i(\mathbf{a})}$ is weakly increasing in a_i. The implication is that if $P^i_{12}(\mathbf{a}) < 0$ then it remains strictly negative as a_i decreases, holding a_j fixed. Similarly, since $-P^i_j(\mathbf{a})$ is log-supermodular, $\frac{P^i_{12}(\mathbf{a})}{P^i_j(\mathbf{a})}$ is weakly increasing in a_j. As $P^i_j(\mathbf{a}) < 0$, it follows that if $P^i_{12}(\mathbf{a}) < 0$ then it remains strictly negative as a_j increases, holding a_i fixed. In sum, if $P^i_{12}(\mathbf{a}) < 0$ then this is also the case as a_i decreases or a_j increases. Thus, if $b'_i(a_j) < 0$ for some a_j, then not only is $P^i_{12}(\mathbf{a}) < 0$, it is also the case that $b_i(a_j)$ decreases in a_j, thus moving the best response further into the region where $P^i_{12}(\mathbf{a}) < 0$. Hence, once $b_i'(a_j) < 0$, the best response keeps decreasing as a_j increases further. Hence, $b_i(a_j)$ is either monotonic or it is first increasing and then decreasing. ■

The proof of Theorem 2 relies on the following lemma.

Lemma 1 *Consider a distribution function $F^i(x|a_i)$ that satisfies strict NUC and a differentiable function $v(x)$ that is strictly quasiconcave and has at most one interior stationary point. Then,*

$$V(a_i) = \int v(x) f^i(x|a_i)dx$$

is strictly quasiconcave in a_i, with $V''(a_i) < 0$ whenever $V'(a_i) = 0$. As an application, in any strict LSM contest with two agents, $-P^i_j(\mathbf{a})$ is strictly quasiconcave in a_i, with $-P^i_{i12}(\mathbf{a}) < 0$ whenever $-P^i_{12}(\mathbf{a}) = 0$

Proof of Lemma 1. Using integration by parts,

$$\begin{aligned} V'(a_i) &= \int \left(-v'(x) F^i_a(x|a_i)\right) dx \\ &= \int \left(-v'(x) \frac{F^i_a(x|a_i)}{f^i(x|a_i)}\right) f^i(x|a_i)dx \end{aligned}$$

and

$$\begin{aligned} V''(a_i) &= \int \left(-v'(x)\right) F^i_{aa}(x|a_i)dx \\ &= \int \left(-v'(x) \frac{F^i_a(x|a_i)}{f^i(x|a_i)}\right) \frac{F^i_{aa}(x|a_i)}{F^i_a(x|a_i)} f^i(x|a_i)dx. \end{aligned}$$

By strict NUC, the term $\frac{F^i_{aa}(x|a_i)}{F^i_a(x|a_i)}$ is strictly increasing in x. The expectation of the term $-v'(x)\frac{F^i_a(x|a_i)}{f^i(x|a_i)}$ is zero at any stationary point, i.e. whenever $V'(a_i) = 0$. For interior x, the sign of the term is determined by $v'(x)$. Since $v(x)$ is strictly quasiconcave and has at most one interior stationary point, either (i) $v'(x) < 0$ or $v'(x) > 0$ almost always, or (ii) it is first strictly positive and then strictly negative. In case (i), $V'(a_i)$ is either strictly positive or strictly negative, meaning that is is strictly quasiconcave. In case (ii), $v'(x)$ strictly single-crosses zero from above, which in turn means that $-v'(x)\frac{F^i_a(x|a_i)}{f^i(x|a_i)}$ strictly single-crosses zero from above on the interior. As in

Chade and Swinkels (2020, Proposition 4), a version of Beesack's inequality can now be invoked: If $h(x)$ is a function that strictly single-crosses zero from above and $q(x)$ is a strictly increasing function, then the expectation of $h(x)q(x)$ is strictly negative when the expectation of $h(x)$ is zero. This result can be verified using integration by parts. Letting $h(x) = -v'(x)\frac{F_a^i(x|a_i)}{f^i(x|a_i)}$ and $q(x) = \frac{F_{aa}^i(x|a_i)}{F_a^i(x|a_i)}$, the conclusion is that $V''(a_i) < 0$ whenever $V'(a_i) < 0$. Hence, $V(a_i)$ has at most one stationary point, and this stationary point (if it exists) must be a maximum. Hence, $V(a_i)$ is strictly quasiconcave.

For the application mentioned in the lemma, $-P_j^i(\mathbf{a})$ is

$$-P_j^i(\mathbf{a}) = \int(-F_a^j(x|a_j))f^i(x|a_i)dx,$$

and strict MLRP implies that $-F_a^j(x|a_j)$ has the same properties as $v(x)$ in the lemma. To see this, note that

$$\frac{\partial(-F_a^j(x|a_j))}{\partial x} = -\frac{f_a^j(x|a_j)}{f^j(x|a_j)}f^j(x|a_j)$$

is first strictly positive and then strictly negative on $(\underline{x}, \overline{x})$. Hence, $-P_j^i(\mathbf{a})$ is strictly quasiconcave in a_i in any strict LSM contest, with $-P_{i12}^i(\mathbf{a}) < 0$ whenever $-P_{12}^i(\mathbf{a}) = 0$. Thus, for any a_j, $P_{12}^i(\mathbf{a}) = 0$ for at most one a_i. ■

Proof of Theorem 2. Consider some interior (a_1', a_2') for which $P_{12}^i(a_1', a_2') = 0$. Then, $P_{i12}^i(a_1', a_2') > 0$, by the argument following Lemma 1, and it follows that $P_{12}^i(a_1, a_2') > 0$ when $a_i > a_i'$. By a similar argument, $P_{j12}^j(a_1', a_2') > 0$, meaning that $P_{12}^j(a_1', a_2) < 0$ when $a_j < a_j'$ and therefore that $P_{12}^i(a_1', a_2) > 0$ when $a_j < a_j'$. As mentioned already, once $P_{12}^i > 0$ it remains strictly positive as a_i increases further. This proves that $P_{12}^i(a_1, a_2) > 0$ for all interior (a_1, a_2) with $a_i \geq a_i'$ and $a_j \leq a_j'$ and at least one strict inequality. The other part is proven in a similar manner.

To see more directly that the locus of points in question is increasing, let $\kappa_i(a_j)$ denote the a_i value (when it exists) for which $P_{12}^i(\kappa_i(a_j), a_j) = 0$. When $\kappa_i(a_j)$ is interior, implicit differentiation implies that

$$\kappa_i'(a_j) = \frac{-P_{j12}^i(\kappa_i(a_j), a_j)}{P_{i12}^i(\kappa_i(a_j), a_j)} = \frac{P_{j12}^j(\kappa_i(a_j), a_j)}{P_{i12}^i(\kappa_i(a_j), a_j)} > 0.$$

■

Proof of Proposition 5. By the usual arguments, if $P_{12}^i > 0$ then $b_i'(a_j) > 0$ and if $P_{12}^i < 0$ then $b_i'(a_j) < 0$. The latter implies that if the reaction function is in the region where $P_{12}^i < 0$, then it remains in that region. In other words, once $b_i(a_j)$ is below $\kappa_i(a_j)$, defined in the proof of Theorem 2, it remains below $\kappa_i(a_j)$. Thus, $b_i(a_j)$ can cross $\kappa_i(a_j)$ at most once and, if it does, the crossing must be from above. It is not necessary to prove that there is at most one a_j value for which $b_i(a_j) = \kappa_i(a_j)$. For instance, if the two functions "touch" $m > 1$ times (they could also coincide on an interval), then they must be tangent in at least the first $m-1$ of those instances.[17] In those cases, however, $b_i'(a_j) = \kappa_i'(a_j) > 0$, so it remains the case that $b_i(a_j)$ is strictly increasing until it crosses $\kappa_i(a_j)$. The proof that $b_i(a_j) > 0$ on an interval is the same as in Proposition 3. ■

Proof of Theorem 3. **Unique interior equilibrium:** For the first part of the proposition, assume that an interior pure-strategy equilibrium exists, which necessitates that A2 is satisfied. Let (a_1^*, a_2^*) denote the equilibrium action profile. Assume to start that $P_{12}^2(a_1^*, a_2^*) > 0 > P_{12}^1(a_1^*, a_2^*)$. In (a_1, a_2) space, (a_1^*, a_2^*) is thus in the area above the locus of points where $P_{12}^i(a_1, a_2) = 0$ (see Figure 1(a)). In this area $b_2(a_1)$ is strictly increasing and $b_1(a_2)$ strictly decreasing. Hence, the two reaction functions do not intersect at any other point in the interior of this area, nor at any interior point where $P_{12}^i(a_1, a_2) = 0$. Thus, if one of the two reaction functions does not cross the $P_{12}^i(a_1, a_2) = 0$ locus, then (a_1^*, a_2^*) must be the unique interior equilibrium. If both reaction functions cross the $P_{12}^i(a_1, a_2) = 0$ locus, then agent 2's reaction function does so further to the north-east than agent 1's reaction function. At the latter point, $b_1(a_2)$ is maximized, meaning that this is the right-most point of agent 1's reaction function in (a_1, a_2) space. Since agent 2's reaction function intersect the $P_{12}^i(a_1, a_2) = 0$ curve later and becomes decreasing thereafter, the two reaction functions cannot intersect in the region where $P_{12}^2(a_1, a_2) < 0 < P_{12}^1(a_1, a_2)$. Hence, (a_1^*, a_2^*) is the unique interior equilibrium. Similar arguments apply if $P_{12}^2(a_1^*, a_2^*) \leq 0 \leq P_{12}^1(a_1^*, a_2^*)$ to start with.

For the second part, assume that $F^1(x|0)$ and $F^2(x|0)$ are non-degenerate.

[17] In Propositions 2 and 3, $b_i(a_j)$ and $\tau_i(a_j)$ touch only once (which is when they cross each other). The significance of this property is that whenever $b_i(a_j)$ is strictly increasing, $b_i(a_j) > \tau_i(a_j)$ and so agent i is the favorite. However, Proposition 5 with heterogeneous technologies is silent on when agent i is the favorite or underdog.

Existence. Since $F^1(x|0)$ and $F^2(x|0)$ are non-degenerate, agent i has a best response to $a_j = 0$. By A1, there is a unique best response to any $a_j \geq 0$, and the resulting reaction function is continuous.

Consider the rectangle $[0, \overline{a}_1] \times [0, \overline{a}_2]$ in (a_1, a_2) space. Any pure-strategy equilibrium must fall within this rectangle. As mentioned, $b_i(a_j)$ is continuous, and its range on $[0, \overline{a}_j]$ is a closed and convex subset of $[0, \overline{a}_i]$. It follows that the reaction functions intersect at least once on $[0, \overline{a}_1] \times [0, \overline{a}_2]$.

Uniqueness. Given that reaction functions are continuous everywhere, including along the boundaries of $[0, \overline{a}_1] \times [0, \overline{a}_2]$, the proof of the first part of the proposition generalizes to all possible equilibria, both interior and those on the boundary.

However, with the existing literature in mind, it is instructive to discuss the possibility of an equilibrium at $(0, 0)$. Consider a lottery contest, microfounded as a best-shot contests with $p_i(a_i) = a_i$ for $i = 1, 2$. Reaction functions are hump shaped and $b_i(a_j) \to 0$ as $a_j \to 0$. There is a unique (interior) equilibrium, but the two reaction functions also both converge towards the point $(0, 0)$. The latter is not an equilibrium because $F^1(x|0)$ and $F^2(x|0)$ are degenerate and agent i has no best response to $a_j = 0$. Nevertheless, this example may raise the question of whether it is possible to have both an interior equilibrium and an equilibrium at $(0, 0)$ when $F^1(x|0)$ and $F^2(x|0)$ are non-degenerate? The answer is no.

First, if $P^i_{12}(0, 0) \neq 0$, as in Figure 1, then previous arguments show that the two reaction functions cannot intersect at $(0, 0)$ as long as A2 is satisfied. On the other hand, if A2 is violated, then at least one agent's best response is always zero. Then, there must be a unique equilibrium since the other agent has a unique best response to zero effort. Second, if $P^1_{12}(0, 0) = 0 = P^2_{12}(0, 0)$, then reaction reaction functions are strictly hump-shaped on $[0, a^{\sup})$ when A2 is satisfied. Now, if $b_i(0) = 0$ then either $u^i_i(\mathbf{0}) < 0$ or $u^i_i(\mathbf{0}) = 0$. In the first case, $b_i(a_j) = 0$ even if a_j is marginally above 0, but this is not consistent with the strict hump-shape. In the second case, the agent's first-order condition is satisfied at $(0, 0)$ and A1 can be invoked. Since $P^i_{ij}(0, 0) = 0$, the implication is that $b'_i(a_j) = 0$. Hence, $b_i(a_j)$ is below the curve where $P^i_{12}(a_1, a_2) = 0$ as a_j increases further, but this is possible only if $b_i(a_j) = 0$ always, thus violating A2. ■

Proof of Corollary 1. In the symmetric contest, reactions functions are identity-independent and, by Proposition 2, strictly hump-shaped. The peak occurs at the unique a_j value where $b_i(a_j) = \tau_i(a_j)$. Thus, there is a symmetric equilibrium (a^S, a^S)

where $b_i(a^S) = a^S = \tau_i(a^S)$, $i = 1, 2$. By symmetry $P^i(a^S, a^S) = \frac{1}{2}$.

Assume now that agent j's characteristics, i.e. v_j, $p_j(\cdot)$, or $c_j(\cdot)$, change. Changes in v_j and $c_j(\cdot)$ have no impact on agent i's reaction function, which therefore still peaks at a^S. Moreover, the point (a^S, a^S) is the only point on agent i's reaction function where $P^i(\mathbf{a}) = \frac{1}{2}$. Hence, if the equilibrium changes in such a way that $P^i(\mathbf{a}) \neq \frac{1}{2}$, then it must hold that $a_i < a^S$ in the new equilibrium. Regarding changes in $p_j(\cdot)$, note that agent i cares about the impact of agent j's action rather than the action per se. Hence, agent i's reaction is still maximized at the unique a_j where $p_j(a_j) = p_i(a^S)$ and where agent i wins with probability $\frac{1}{2}$. Thus, if $P^i(\mathbf{a}) \neq \frac{1}{2}$ in equilibrium, then $a_i < a^S$. ■

Proof of Corollary 2. When v_i decreases, $b_i(a_j)$ strictly decreases for all $a_j > 0$ where $b_i(a_j) > 0$ originally. To see this, let b_i^* denote the unique best response to a_j in the original contest, and let b_i^{**} denote a (not necessarily unique) best response when agent i's valuation changes from v_i to $v_i' < v_i$. Then,

$$\begin{aligned} v_i P^i(b_i^*, a_j) - c_i(a_i^*) &\geq v_i P^i(b_i^{**}, a_j) - c_i(a_i^{**}) \\ v_i' P^i(b_i^{**}, a_j) - c_i(a_i^{**}) &\geq v_i' P^i(b_i^*, a_j) - c_i(a_i^*), \end{aligned}$$

which implies that

$$v_i \left(P^i(b_i^*, a_j) - P^i(b_i^{**}, a_j) \right) \geq c_i(a_i^*) - c_i(a_i^{**}) \geq v_i' \left(P^i(b_i^*, a_j) - P^i(b_i^{**}, a_j) \right).$$

Since $v_i > v_i'$, $P^i(b_i^*, a_j) \geq P^i(b_i^{**}, a_j)$, or $b_i^* \geq b_i^{**}$. At the same time, $b_i^{**} = b_i^*$ cannot occur since the same action cannot satisfy the first order condition in both the original and new contest. Hence, $b_i^* > b_i^{**}$.

Let (a_1^*, a_2^*) denote the original interior equilibrium action profile. Since agent i views actions as strategic substitutes at (a_1^*, a_2^*), $b_i'(a_j^*) < 0$ and $b_j'(a_i^*) > 0$ in the original contest. In the new contest, $b_j(a_i)$ is unchanged and is increasing in the region where the original equilibrium occurs. Since agent i's best response(s) have decreased, any pure-strategy Nash equilibrium must occur at a point below her old reaction function. Since $b_j(a_i)$ is increasing, the implication is that any pure-strategy equilibrium (when one exists) occurs at point where $a_i < a_i^*$ and $a_j < a_j^*$. ■

Proof of Corollary 3. The first part of the corollary follows from the same type of argument as in the proof of Corollary 2. The second part follows from the fact that the favorite by definition won with probability greater than $\frac{1}{2}$ in the original contest, and wins with even greater probability in the new contest due to the change in equilibrium actions. ■

Proof of Corollary 4. Simple differentiation and $b_j(a_i^*) = a_j^*$ yields

$$\frac{\partial u^i(a_i, b_j(a_i))}{\partial a_i}_{|a_i=a_i^*} = u_i^i(\mathbf{a}^*) + u_j^i(\mathbf{a}^*)b_j'(a_i^*)$$
$$= u_j^i(\mathbf{a}^*)b_j'(a_i^*),$$

where the last inequality follows from the first-order condition. The corollary then follows from $u_j^i(\mathbf{a}^*) < 0$ and Theorem 1. ■

Proof of Proposition 6. Let (a_1^*, a_2^*) denote the equilibrium action profile. If $a_2^* > a_1^*$ then $a_2^* = b_2(a_1^*) > a_1^*$. However, it is not possible for $b_2(a_1) > a_1$ for all a_1, since the best response is bounded above by $\overline{a}_2$. Thus, by continuity, there exists some $a' > a_1^*$ such that $b_2(a') = a'$. It then holds that $b_1(a') \geq a'$, as explained in the main text. If $b_1(a') = a'$ then (a', a') is an equilibrium, and since equilibrium is unique it contradicts that (a_1^*, a_2^*) is an equilibrium. If $b_1(a') > a'$ then agent 1 views actions as complements at $(b_1(a'), a')$, again as explained in the main text. Hence, $b_1(a_2)$ is increasing in a_2 for $a_2 \leq a'$. Since agent 2 views actions as (weak) substitutes at (a', a'), $b_2(a_1)$ is decreasing for $a_1 > a'$. Putting these observations together yields the conclusion that $b_1(a_2)$ and $b_2(a_1)$ intersect at some point where $a_1 \in [a', b_1(a')]$ and $a_2 \in (0, a')$. This intersection is an equilibrium, which contradicts that (a_1^*, a_2^*) is an equilibrium. Hence, there is no interior equilibrium with $a_2^* > a_1^*$.

In conclusion, $a_1^* \geq a_2^*$ in equilibrium. If $a_1^* = a_2^*$ then $u_1^1(\mathbf{a}^*) > u_2^2(\mathbf{a}^*)$ if $v_1 > v_2$, which contradicts that both players are best responding. Hence, $a_1^* > a_2^*$ if $v_1 > v_2$. The last statement follows from the first two properties of ordered LSM contests. ■

Proof of Corollary 5. By Proposition 6, agent 1 (2) views actions as complements (substitutes) in equilibrium. The proofs of Corollaries 2–4 now carry over. ■

Appendix B: Supplementary material

This supplementary material first provides an example that demonstrates sufficient conditions for NUC. Conditions for concave and quasiconcave payoff functions are then described. These conditions lean heavily on the literature on principal-agent models, like Rogerson (1985) and Jewitt (1988), as well as on Chade and Swinkels' (2020) paper. The assumption that the support is the same for all actions is discussed, and relaxed, in the case of homogeneous technologies. Next, properties of larger LSM contests are examined. It is then observed that certain CSFs in the literature are inconsistent with the properties of LSM contests and therefore cannot be microfounded as such. Finally, a discussion of biases in contest is provided, which links the relevant literature to the individual discouragement effect.

B.1 Another example

Chade and Swinkels (2020) offer sufficient conditions for NUC. The next example makes use of one of their sufficient conditions, which has the advantage that the starting point is the density $f^i(x_i|a_i)$ rather than the distribution function $F^i(x_i|a_i)$.

Example 4 (k^{th}-shot contests): Let k be a positive integer. Let $H(x)$ be some atomless distribution with full support and strictly positive and differentiable density $h(x)$. The performance technology is described by a distribution function with density

$$f^i(x|a_i) = B(p_i(a_i))H(x)^{p_i(a_i)-1} \left(1 - H(x)\right)^{k-1} h(x),$$

where $B(\cdot)$ is such that $f^i(x|a_i)$ integrates to one over $[\underline{x}, \overline{x}]$. The impact function $p_i(a_i)$ is non-negative, twice differentiable, and satisfies $p_i'(a_i) > 0$. It is as if agent i makes $p_i(a_i) + k - 1$ draws from $H(x)$ and the agent's performance is taken to be the k^{th} highest of those draws. If $k = 1$, the contest reduces to the best-shot contest in Example 2. If $k = 2$, the agent's performance is determined by the second-best draw out of $p_i(a_i) + 1$ draws. As in best-shot contests, the analogy is not perfect as $p_i(a_i)$ is not restricted to be an integer. The k^{th}-shot contest can e.g. describe a setting in which competing agents are required to develop a product with k critical components. The agent puts together a product consisting of the k best components that she is able to make, out of a total of $p_i(a_i) + k - 1$. However, the quality of the product is determined by the weakest of the selected components.

For any k,

$$\frac{\partial^2 \ln f^i(x|a_i)}{\partial x \partial a_i} = p_i'(a_i)\frac{h(x)}{H(x)}$$

is strictly positive, implying that the distribution has the strict MLRP. Moreover, $p_i'(a_i)\frac{h(x)}{H(x)}$ is trivially log-supermodular. Thus, the least demanding of the three sufficient conditions for NUC in Chade and Swinkels (2020, Proposition 2) applies. Hence, the contest is a LSM contest.

For future reference, when $n = 2$ and $k = 2$ the CSF reduces to

$$P^1(a_1, a_2) = \frac{p_1(a_1)(p_1(a_1)+1)(p_1(a_1)+3p_2(a_2)+2)}{(p_1(a_1)+p_2(a_2))(p_1(a_1)+p_2(a_2)+1)(p_1(a_1)+p_2(a_2)+2)}, \tag{8}$$

using (2). ▲

B.2 Single-valued best responses and the first-order approach

B.2.1 Concave payoff functions

This subsection assumes that $c_i''(a_i) \geq 0$ and asks what properties the performance technologies must have in order for $u^i(\mathbf{a})$ to be strictly concave in a_i, or $u^i_{ii}(\mathbf{a}) < 0$. This exercise is near-identical to how the literature on principal-agent models justify the first-order approach, as pioneered by Rogerson (1985) and Jewitt (1988). In that setting, the performance-contingent reward takes the form of a wage schedule, $w(x)$, whereas in the contest setting it takes the form of a winning probability, $\prod_{j \neq i} F^j(x|a_j)$. The difference in interpretation aside, the same techniques apply.

Rogerson's (1985) result applies directly, since the only structure he requires on $w(x)$ is that it is increasing. Clearly, $\prod_{j \neq i} F^j(x|a_j)$ has that property. Rogerson then shows that $u^i(\mathbf{a})$ is concave in a_i as long as $F^i(x|a_i)$ is convex in a_i. This condition is satisfied in e.g. the best-shot model. A strict version of his result is stated below.

Jewitt's (1988) result requires that $w(x)$ is increasing and concave. In the $n = 2$ case, this holds if $F^j(x|a_j)$ is concave in x, which is satisfied if e.g. $F^j(x|a_j)$ is an exponential distribution, as in Hirschleifer and Riley (1992). In addition, Jewitt (1988) imposes a condition on $F^i(x|a_i)$ that is weaker than Rogerson's convexity condition. It turns out that NUC adds structure to $F^i(x|a_i)$ that makes it easier to check the latter. The exponential distribution $F^i(x|a_i) = 1 - e^{-\frac{x}{p_i(a_i)}}$, $x \in [0, \infty)$, satisfies this condition as long as the mean $p_i(a_i) > 0$ is concave. Let $\mathbf{a}_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$.

Proposition 7 *Consider a strict LSM contest with $n \geq 2$ agents. If $c_i''(a_i) \geq 0$, then agent i's problem is strictly concave in a_i for any $\mathbf{a}_{-i} \neq 0$ if either:*

1. *$F^i(x|a_i)$ is strictly convex in a_i for all a_i and all $x \in (\underline{x}, \overline{x})$, or*

2. *$\prod_{j \neq i} F^j(x|a_j)$ is strictly concave in x and $\mathbb{E}[X_i|a_i]$ is weakly concave in a_i.*

Proof. The condition that $\mathbf{a}_{-i} \neq 0$ means that $\prod_{j \neq i} F^j(x|a_j)$ is not degenerate, but is strictly increasing in x. With this in mind, the first result is a strict version of Rogerson's (1985) result and can be proven by using integration by parts.

Using integration by parts twice, the second result similarly follows from Jewitt (1988) if $\int_{\underline{x}}^{x} F^i(s|a_i)ds$ is strictly convex in a_i for all $x \in (\underline{x}, \overline{x})$. It will now be shown that this property follows from strict NUC and the condition that $\mathbb{E}[X_i|a_i]$ is weakly concave in a_i.

Strict NUC is equivalent to the requirement that $\frac{\partial}{\partial x} \frac{F^i_{aa}}{F^i_a} > 0$ for $x \in (\underline{x}, \overline{x})$. Since

$$\frac{\partial^2}{\partial a_i^2} \int_{\underline{x}}^{x} F^i(s|a_i)ds = \int_{\underline{x}}^{x} F^i_a(s|a) \frac{F^i_{aa}(s|a)}{F^i_a(s|a)} ds,$$

strict NUC and $F^i_a(s|a) < 0$ imply that the integrand is either of constant sign for all s or that it is first positive and then negative. Hence, if $\frac{\partial^2}{\partial a_i^2} \int_{\underline{x}}^{x} F^i(s|a_i)ds \geq 0$ for some $x' \in (\underline{x}, \overline{x}]$, then $\frac{\partial^2}{\partial a_i^2} \int_{\underline{x}}^{x} F^i(s|a_i)ds > 0$ for all $x \in (\underline{x}, x')$. Hence, it is sufficient that $\frac{\partial^2}{\partial a_i^2} \int_{\underline{x}}^{\overline{x}} F^i(s|a_i)ds \geq 0$, but this is the same as requiring that $\mathbb{E}[X_i|a_i]$ is weakly concave in a_i. ■

Example 4 (continued): The second part of the proposition is relevant to Example 4. First, note that it is without loss of generality to let $H(x) = x$, $x \in [0, 1]$. Formally, this can be seen through a change of variable. Intuitively, the reason is that rather than thinking about performance on a scale between $\underline{x}$ and $\overline{x}$, it can be thought of in terms of the corresponding quantiles of $H(x)$. For $n = 2$, it can then be shown that the second part of the proposition is satisfied $p_i(a_i)$ is concave and $p_j(a_j) < 1$. Of course, these are merely sufficient conditions. To illustrate, if $k = 2$ and impact and cost functions are linear, then it can be verified from (8) that $u_i(a_i, a_j)$ is concave in a_i whenever $a_j < 1.8944$. ▲

B.2.2 Quasiconcave payoff functions

Following the logic oulined in Section 6.1, the next result provides a sufficent condition that ensures $u^i(\mathbf{a})$ that is strictly quasiconcave on the relevant range of actions.

Proposition 8 *Consider a LSM contest with two agents. For any $a_j \in (0, \overline{a}_j]$, $u^i(\mathbf{a})$ is strictly quasiconcave in $a_i \in [0, \overline{a}_i]$, with $u^i_{ii}(\mathbf{a}) < 0$ whenever $u^i_i(\mathbf{a}) = 0$, if*

$$\frac{P^i_{ii}(a_i, \overline{a}_j)}{P^i_i(a_i, \overline{a}_j)} - \frac{c''_i(a_i)}{c'_i(a_i)} < 0 \text{ for all } a_i \in (0, \overline{a}_i].$$

Proof. Fix $a_j \in (0, \overline{a}_j]$. If $u^i(a_i, a_j)$ does not have a stationary point on $(0, \overline{a}_i)$, then it is monotonic and thus quasiconcave on $[0, \overline{a}_i]$. Assume now that there is a stationary point, or some $a_i \in (0, \overline{a}_i)$ for which

$$v_i P^i_i(a_i, a_j) = c'_i(a_i).$$

Then, the second derivative of $u^i(a_i, a_j)$ with respect to a_i is proportional to

$$\frac{P^i_{ii}(a_i, a_j)}{P^i_i(a_i, a_j)} - \frac{c''_i(a_i)}{c'_i(a_i)} \leq \frac{P^i_{ii}(a_i, \overline{a}_j)}{P^i_i(a_i, \overline{a}_j)} - \frac{c''_i(a_i)}{c'_i(a_i)} < 0$$

where the first inequality follows from log-supermodularity of $P^i_i(a_i, a_j)$ and the second from the assumption in the proposition. Hence, the second derivative is strictly negative, and it follows that the stationary point is a local maximum. Thus, there is at most one stationary point. In conclusion $u^i(a_i, a_j)$ is either monotonic in a_i or it is first-increasing-then-decreasing on $[0, \overline{a}_i]$. In either case, $u^i(a_i, a_j)$ is quasiconcave in a_i. ■

To explain Proposition 8, consider the following exercise: Hold the technologies and $\overline{a}_i$ fixed. Fix $c_i(0)$ and $c_i(\overline{a}_i)$. Then, the proposition quantifies how much the cost function must be "curved" or "convexified" on the interval $[0, \overline{a}_i]$ in order to achieve quasiconcavity. Finally, note that if the condition in Proposition 8 holds for some (v_1, v_2), then it also holds if the valuations decrease. The reason is that $\overline{a}_1$ and $\overline{a}_2$ then decrease.

The proof of Proposition 8 is self-contained and relies only on log-supermodularity of $P^i_i(\mathbf{a})$. However, the proposition can alternatively be shown to be an implication of a result in Chade and Swinkels (2020). To this end, note that Proposition 8 requires calculating $P^i(a_i, \overline{a}_j)$ and examining its curvature properties. It may be convenient to have a "looser" condition, if such a condition is easier to check. Proposition 5 in Chade and Swinkels (2020) is highly useful in this regard. A version, adapted to the current setting, is provided next.

Proposition 9 (Chade and Swinkels (2020)) *Consider a LSM contest with two agents. Fix $a_j > 0$ and assume that there is a function $q(x)$ that satisfies*

$$(i)\ q'(x) > 0 \ \textit{and}\ \frac{\partial}{\partial x}\frac{f^j(x|a_j)}{q'(x)} \leq 0,\ \textit{and}$$
$$(ii)\ \frac{\mathbb{E}_{aa}[q|a_i]}{\mathbb{E}_a[q|a_i]} - \frac{c_i''(a_i)}{c_i'(a_i)} < 0 \ \textit{for all}\ a_i \in (0, \overline{a}_i].$$

Then $u^i(\mathbf{a})$ is strictly quasiconcave in $a_i \in (0, \overline{a}_i]$.

Proof. This is Proposition 5 in Chade and Swinkels (2020), with Γ degenerate and $v(x) = v_i F^j(x|a_j)$. ■

Proposition 8 can in fact be seen as a corollary of this last proposition. Let $q(x) = F^j(x|\overline{a}_j)$. The MLRP implies that (i) is satisfied for all $a_j \in (0, \overline{a}_j]$. Moreover, since $\mathbb{E}[q|a_i]$ coincides with $P^i(a_i, \overline{a}_j)$, (ii) reduces to the condition in Proposition 8.

Another application arises in contests in which $F^j(x|a_j)$ is concave in x. Letting $q(x) = x$ then implies that (i) is satisfied. In this case $\mathbb{E}[q|a_i] = \mathbb{E}[X_i|a_i]$. Hence, (ii) is satisfied if

$$\frac{\mathbb{E}_{aa}[X_i|a_i]}{\mathbb{E}_a[X_i|a_i]} - \frac{c_i''(a_i)}{c_i'(a_i)} < 0 \text{ for all } a_i \in (0, \overline{a}_i],$$

or in other words if the cost function is more convex than $\mathbb{E}[X_i|a_i]$, which in turn implies that $\mathbb{E}[X_i|a_i] - c_i(a_i)$ is strictly quasiconcave. The same logic is behind Corollary 1 in Chade and Swinkels (2020).

The following example compares Propositions 8 and 9.

Example 4 (continued): Consider the two-shot model in (8). To focus on the role of the cost function, assume that $p_i(a_i) = a_i$, $i = 1, 2$. In this case, it can be verified that

$$\lim_{\overline{a}_2 \to \infty} \frac{P^1_{11}(a_1, \overline{a}_2)}{P^1_1(a_1, \overline{a}_2)} = \frac{2}{2a_1 + 1}.$$

Note that if $c_1(a_1)$ is proportional to $a_1^2 + a_1$, then $\frac{c_1''(a_1)}{c_1'(a_1)} = \frac{2}{2a_1+1}$. Hence, if the cost function is "more convex" than (or a convex transformation of) $a_1^2 + a_1$, then agent

1's problem is quasiconcave by Proposition 8. This holds if, for example, $c_1(a_1) = a_1^{\gamma_1}$ with $\gamma_1 \geq 2$.

For an application of Proposition 9, consider the general k^{th}-shot contest and normalize $H(x) = x$, $x \in [0,1]$. Assume again that $p_i(a_i) = a_i$, $i = 1, 2$. Let $q(x|a_2) = x^{\lceil \overline{a}_2 \rceil}$, where $\lceil \overline{a}_2 \rceil$ is the smallest integer greater than or equal to $\overline{a}_2$. Condition (i) is satisfied for all $a_2 \in (0, \overline{a}_2]$. Note that $\mathbb{E}[q|a_1]$ is the expected value of the $\lceil \overline{a}_2 \rceil^{th}$ moment of the distribution $F^1(x|a_1)$. However, the latter is a beta distribution, for which the $\lceil \overline{a}_2 \rceil^{th}$ moment is known to equal

$$\mathbb{E}[q|a_1] = \prod_{r=0}^{\lceil \overline{a}_2 \rceil - 1} \frac{a_1 + r}{a_1 + k + r}.$$

For instance, assume that $\lceil \overline{a}_2 \rceil = 2$. Then,

$$\frac{\mathbb{E}_{aa}[q|a_1]}{\mathbb{E}_a[q|a_1]} = \frac{-2\left(2a_1^3 + 3a_1^2 k + 3a_1^2 + 3a_1 k + 3a_1 - k^3 + 2k + 1\right)}{\left(a_1^2 + 2a_1 k + a_1 + k^2 + k\right)\left(2a_1 + k + 2a_1 k + 2a_1^2 + 1\right)},$$

which is increasing in k. Hence, condition (ii) is hardest to satisfy for $k \to \infty$. However, the limit is once again $\frac{2}{2a_1+1}$. Thus, as before, if the cost function is a_1^2 or a convex transformation thereof, then agent 1's problem is quasiconcave for all $a_2 \in (0, \overline{a}_2]$. ▲

An extension to Proposition 8 that allows for $n \geq 2$ agents is provided next. First note that in the $n = 2$, case $P^i(\mathbf{a}) = 1 - P^j(\mathbf{a})$ implies that the condition in Proposition 8 is equivalent to

$$\frac{P^j_{ii}(a_i, \overline{a}_j)}{P^j_i(a_i, \overline{a}_j)} - \frac{c''_i(a_i)}{c'_i(a_i)} < 0 \text{ for all } a_i \in (0, \overline{a}_i].$$

If a version of the this inequality holds for all $j \neq i$ in larger contests, then agent i's payoff is quasiconcave in a_i. The point is that it is sufficient to check quasiconcavity only at $a_{-i} = \overline{\mathbf{a}}_{-i} = (\overline{a}_1, ..., \overline{a}_{i-1}, \overline{a}_{i+1}, ..., \overline{a}_n)$.

Proposition 10 *Consider a LSM contest with $n \geq 2$. For any $\mathbf{a}_{-i} \in \times_{j \neq i}(0, \overline{a}_j]$, $u^i(\mathbf{a})$ is strictly quasiconcave in $a_i \in [0, \overline{a}_i]$, with $u^i_{ii}(\mathbf{a}) < 0$ whenever $u^i_i(\mathbf{a}) = 0$, if*

$$\frac{P^j_{ii}(a_i, \overline{\mathbf{a}}_{-i})}{P^j_i(a_i, \overline{\mathbf{a}}_{-i})} - \frac{c''_i(a_i)}{c'_i(a_i)} < 0 \textit{ for all } a_i \in (0, \overline{a}_i] \textit{ and all } j \neq i.$$

Proof. Since $-P_i^j(\mathbf{a})$ is log-supermodular by Proposition 1, $\frac{P_{ii}^j(\mathbf{a})}{P_i^j(\mathbf{a})}$ is increasing in all elements of $\mathbf{a}_{-i}$. Hence,

$$\frac{P_{ii}^j(a_i, \overline{\mathbf{a}}_{-i})}{P_i^j(a_i, \overline{\mathbf{a}}_{-i})} \geq \frac{P_{ii}^j(a_i, \mathbf{a}_{-i})}{P_i^j(a_i, \mathbf{a}_{-i})} \text{ for any } \mathbf{a}_{-i} \in \times_{j\neq i}(0, \overline{a}_j]$$

and the condition in the proposition then implies that for any such $(a_i, \mathbf{a}_{-i})$,

$$\frac{P_{ii}^j(a_i, \mathbf{a}_{-i})}{P_i^j(a_i, \mathbf{a}_{-i})} - \frac{c_i''(a_i)}{c_i'(a_i)} \leq \frac{P_{ii}^j(a_i, \overline{\mathbf{a}}_{-i})}{P_i^j(a_i, \overline{\mathbf{a}}_{-i})} - \frac{c_i''(a_i)}{c_i'(a_i)} < 0.$$

Since

$$u^i(\mathbf{a}) = v_i \left(1 - \sum\nolimits_{j\neq i} P^j(\mathbf{a})\right) - c_i(a_i),$$

it holds that at any point where $u_i^i(\mathbf{a}) = 0$,

$$v_i \sum\nolimits_{j\neq i} \left(-P_i^j(\mathbf{a})\right) = c_i'(a_i)$$

and therefore that

$$\begin{aligned} u_{ii}^i(\mathbf{a}) &= v_i \sum\nolimits_{j\neq i} \left(-P_{ii}^j(\mathbf{a})\right) - \frac{c_i''(a_i)}{c_i'(a_i)} c_i'(a_i) \\ &= v_i \sum\nolimits_{j\neq i} \left(-P_{ii}^j(\mathbf{a})\right) - \frac{c_i''(a_i)}{c_i'(a_i)} v_i \sum\nolimits_{j\neq i} \left(-P_i^j(\mathbf{a})\right) \\ &= v_i \sum\nolimits_{j\neq i} \left(-P_i^j(\mathbf{a})\right) \left(\frac{P_{ii}^j(\mathbf{a})}{P_i^j(\mathbf{a})} - \frac{c_i''(a_i)}{c_i'(a_i)}\right), \end{aligned}$$

which is strictly negative whenever $\mathbf{a}_{-i} \in \times_{j\neq i}(0, \overline{a}_j]$. The proposition follows. ■

The condition in Proposition 10 implies that

$$\frac{P_{ii}^i(a_i, \mathbf{a}_{-i})}{P_i^i(a_i, \mathbf{a}_{-i})} - \frac{c_i''(a_i)}{c_i'(a_i)} < 0 \text{ for all } \mathbf{a} \in \times_{j=1}^n(0, \overline{a}_j], \tag{9}$$

which directly implies quasiconcavity of $u^i(\mathbf{a})$ in a_i. However, note that

$$\frac{P_{ii}^i(a_i, \overline{\mathbf{a}}_{-i})}{P_i^i(a_i, \overline{\mathbf{a}}_{-i})} - \frac{c_i''(a_i)}{c_i'(a_i)} < 0 \text{ for all } a_i \in (0, \overline{a}_i]$$

does not necessarily imply (9). The reason is that $P_i^i(a_i, \mathbf{a}_{-i}) = \sum_{j \neq i}(-P_i^j(a_i, \mathbf{a}_{-i}))$ is not guaranteed to be log-supermodular when $n > 2$, because the sum of log-supermodular functions need not itself be log-supermodular.

B.3 Shifting supports

It has been assumed that the support of the performance is the same for both agents and independent of the action, at least as long as the action is strictly greater than zero. This subsection briefly explores the consequences of relaxing this assumption. Assume that technologies are identity-independent and let the support of agent i's performance be denoted $[l(a_i), u(a_i)]$. Assume that both $l(\cdot)$ and $u(\cdot)$ are differentiable and non-decreasing, or $l'(\cdot), u'(\cdot) \geq 0$. This is consistent with the assumption that $F(x|a_i)$ improves in the sense of first-order stochastic dominance when a_i increases.

Consider an action profile for which $a_i > a_j$. Hence, agent i is the favorite. Assume that $u(a_j) \geq l(a_i)$. Otherwise, agent i wins with probability one, meaning that she could lower her action without lowering her chance of winning. Modifying (2),

$$P^i(\mathbf{a}) = \int_{l(a_i)}^{u(a_j)} F(x|a_j) f(x|a_i) dx + 1 - F(u(a_j)|a_i).$$

Hence, the cross-partial derivative with respect to a_i and a_j is

$$P_{12}^i(\mathbf{a}) = -F_a(l(a_i)|a_j) f(l(a_i)|a_i) l'(a_i) + \int_{l(a_i)}^{u(a_j)} F_a(x|a_j) f_a(x|a_i) dx. \tag{10}$$

The second term is familiar from the case in which the support does not shift. The first term is strictly positive whenever $l'(a_i) > 0$ and $l(a_i) > l(a_j)$, since $F_a < 0$ in that case. This reflects the fact that there is an additional source of complementarity when the support shifts with the action. Starting from $a_j < a_i$, an increase in a_j means that agent j's support eats its way into the support of agent i's performance. Hence, agent i is less likely to win if her performance is the lowest possible, $l(a_i)$. This provides her with an extra incentive to work harder herself, to further separate the supports.

Thus, Theorem 1 extends to the case where the support shifts with the agent's action. In fact, even if $-F_a$ is log-submodular rather than log-supermodular and the second term in (10) is negative, the first term may dominate and ensure that the favorite views actions as strategic complements.

Similarly, the cross-partial derivative of agent j's winning probability,

$$P_{12}^{j}(\mathbf{a}) = F_a(u(a_j)|a_i)f(u(a_j)|a_j)u'(a_j) + \int_{l(a_i)}^{u(a_j)} F_a(x|a_i)f_a(x|a_j)dx, \qquad (11)$$

contains an extra negative term. This is because the underdog's incentive to work hard diminishes when the favorite's support shifts to the right. The underdog then has a smaller chance of influencing the outcome, meaning the return to effort is lower.

Thus, the main message is that the conclusions of Theorem 1 is robust to shifting supports and may hold even without NUC. The following example proves this assertion in a setting where MLRP is violated and $-F_a$ is log-submodular.

Example 5 (Pareto Distributions): Consider a rank-order tournament with additive noise. Assume that the noise term follows a Pareto distribution with scale and shape parameters that are both equal to one. In other words, the distribution of the noise term ε_i is $Q(\varepsilon_i) = 1 - \frac{1}{\varepsilon_i}$ with density $q(\varepsilon_i) = \frac{1}{\varepsilon_i^2}$, $\varepsilon_i \in [1, \infty)$. Agent i's performance is $a_i + \varepsilon_i$, which has distribution $F(x|a_i) = Q(x - a_i)$, $x \in [1 + a_i, \infty)$. This example is inspired by a leading example in Drugov and Ryvkin (2020) and Drugov, Ryvkin, and Zhang (2024). Their focus is on rank-order tournaments with homogenous agents, but they endogenize the prize schedule and a minimum standard, respectively. The Pareto distribution is one of the distributions with a "heavy tail." It has the unusual property that the likelihood-ratio is decreasing, thus violating MLRP, but positive on the entire support. Thus, the second term in (10) is negative, which is consistent with the fact that $-F_a$ is log-submodular. Nevertheless, it can be verified that the first term in (10) dominates, meaning that the favorite does indeed view actions as strategic complements. Intuitively, $F_a(l(a_i)|a_j)$ in the first term is bounded away from zero and $f(l(a_i)|a_i)$ is large, since $f(l(a_i)|a_i) = g(1)$ is the highest value that the density ever takes. Thus, there is a significant incentive to take a higher action in order to shift the support. Alternatively, note that $u(a)$ is constant, meaning that the first term in (11) is zero while the second term is negative. Hence, the underdog views actions as strategic substitutes, which in turn implies that the favorite views actions as complements. ▲

B.4 Larger contests

B.4.1 Basic properties of large LSM contests

With $n \geq 3$ agents,

$$P_{12}^3(\mathbf{a}) = \int F_a^1(x|a_1)F_a^2(x|a_2)\prod\nolimits_{j\in\{4,\dots,n\}} F^j(x|a_j)f^3(x|a_3)dx.$$

Note that $P_{12}^3(\mathbf{a}) > 0$ if all actions are interior. The same property holds for $P_{12}^4(\mathbf{a}), ..., P_{12}^n(\mathbf{a})$. Hence,

$$0 = \sum\nolimits_{i=1}^n P_{12}^i(\mathbf{a}) > P_{12}^1(\mathbf{a}) + P_{12}^2(\mathbf{a}),$$

which implies that at least one of $P_{12}^1(\mathbf{a})$ or $P_{12}^2(\mathbf{a})$ must be strictly negative. That is, within any given pair of agents, at least one views the actions of the two agents as strategic substitutes. It is possible that neither views actions as strategic complements.

Next, note that

$$P_{12}^1(\mathbf{a}) - P_{12}^2(\mathbf{a}) = \int F_a^1(x|a_1)F_a^2(x|a_2)\prod\nolimits_{j\in\{3,\dots,n\}} F^j(x|a_j)\left(\frac{f_a^1(x|a_1)}{F_a^1(x|a_1)} - \frac{f_a^2(x|a_2)}{F_a^2(x|a_2)}\right)dx. \tag{12}$$

Hence, if technologies are identity dependent and satisfy strict NUC, then $P_{12}^1(\mathbf{a}) > P_{12}^2(\mathbf{a})$ if $a_1 > a_2$. Since, $P_{12}^1(\mathbf{a}) + P_{12}^2(\mathbf{a}) < 0$, it follows that $P_{12}^2(\mathbf{a}) < 0$. Thus, within the pair, the agent who is less likely to win must view actions as strategic substitutes. The argument extends to contests with homogeneous technologies, even if they are not fully homogenous.

Recall also that $-P_2^1(\mathbf{a})$ is log-supermodular, by Proposition 1. Hence,

$$\frac{\partial^2 \ln(-P_2^1(\mathbf{a}))}{\partial a_1 \partial a_3} = \frac{\partial}{\partial a_3}\frac{P_{12}^1(\mathbf{a})}{P_2^1(\mathbf{a})} \geq 0.$$

Since $P_2^1(\mathbf{a}) < 0$, the implication is that if $P_{12}^1(\mathbf{a})$ is negative then it remains negative if a_3 increases. In other words, agent 1 is more likely to view a_1 and a_2 as strategic substitutes the harder any of their common rival works. Note that the argument allows for heterogenous technologies.

B.4.2 Ordered LSM contests with $n \geq 3$ agents

As in Section 5, consider the relationship between agents 1 and 2 and their equilibrium actions. The characteristics of the remaining agents $3, ..., n$ are not important. To rank incentives, assume that (T1), (T3), and (C1) are satisfied. Note that (T1) and (T3) together imply that

$$\frac{-F_a^1(x|a_1)}{F^1(x|a_1)} \geq \frac{-F_a^2(x|a_2)}{F^2(x|a_2)} \text{ for all } x \in (\underline{x}, \overline{x}) \text{ when } a_1 = a_2 > 0. \tag{13}$$

Consider now

$$\begin{aligned}
P^1(\mathbf{a}) &= 1 - P^2(\mathbf{a}) - \sum\nolimits_{i=3}^n P^i(\mathbf{a}) \\
&= 1 - \int F^1(x|a_1) f^2(x|a_2) \prod\nolimits_{j \in \{3,...,n\}} F^j(x|a_j) dx \\
&\quad - \sum\nolimits_{i=3}^n \int F^1(x|a_1) F^2(x|a_2) f^i(x|a_i) \prod\nolimits_{j \in \{3,...,n\} \setminus \{i\}} F^j(x|a_j) dx
\end{aligned}$$

and the analogous expansion of $P^2(\mathbf{a})$. Combining the two and letting $\Delta(\mathbf{a}) = P_1^1(\mathbf{a}) - P_2^2(\mathbf{a})$ yields

$$\begin{aligned}
\Delta(\mathbf{a}) &= \int \left[\frac{-F_a^1(x|a_1)}{f^1(x|a_1)} - \frac{-F_a^2(x|a_2)}{f^2(x|a_2)} \right] f^1(x|a_1) f^2(x|a_2) \prod\nolimits_{j \in \{3,...,n\}} F^j(x|a_j) dx \\
&\quad + \sum\nolimits_{i=3}^n \int \left[\frac{-F_a^1(x|a_1)}{F^1(x|a_1)} - \frac{-F_a^2(x|a_2)}{F^2(x|a_2)} \right] F^1(x|a_1) F^2(x|a_2) f^i(x|a_i) \prod\nolimits_{j \in \{3,...,n\} \setminus \{i\}} F^j(x|a_j) dx,
\end{aligned}$$

which is weakly positive by (T3) and the property in (13). In other words, under (T1), (T3), and (C1), $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ at any interior action profile where $a_1 = a_2$. The inequality is strict if one or more of the inequalities in (T1), (T3), and (C1) are strict. Giebe and Gürtler (2024) rely on (T1) in their study of rank-order tournaments with $n \geq 3$ agents. Recall from Section 5 that (T3) is automatic in their setting. Hence, the first term in $\Delta(\mathbf{a})$ is nil, and the sign of $\Delta(\mathbf{a})$ is driven entirely by (T1) though (13).

As mentioned in Section 6, a trick from Giebe and Gürtler (2024) is borrowed to conclude the proof that $a_1 \geq a_2$ in any interior equilibrium. The proof of their Lemma 1 holds for any contest with the property that $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ if $a_1 = a_2$ and that $\min_{a \in [0, \overline{a}_i]} c''(a_i)$ is sufficiently large. To explain, assume by contradiction that there exists an interior equilibrium with $a_1 < a_2$ and write

$$\begin{aligned}
u_1^1(\mathbf{a}) - u_2^2(\mathbf{a}) &= u_1^1(a_1, a_1, a_3, ..., a_n) + \int_{a_1}^{a_2} u_{12}^1(a_1, x, a_3, ..., a_n)dx \\
&\quad - \left(u_2^2(a_1, a_1, a_3, ..., a_n) + \int_{a_1}^{a_2} u_{22}^2(a_1, x, a_3, ..., a_n)dx \right) \\
&= \left(u_1^1(a_1, a_1, a_3, ..., a_n) - u_2^2(a_1, a_1, a_3, ..., a_n) \right) \\
&\quad + \left(\int_{a_1}^{a_2} u_{12}^1(a_1, x, a_3, ..., a_n)dx - \int_{a_1}^{a_2} u_{22}^2(a_1, x, a_3, ..., a_n)dx \right)
\end{aligned}$$

The first part is weakly positive since $u_1^1(\mathbf{a}) \geq u_2^2(\mathbf{a})$ if $a_1 = a_2$. The second part is strictly positive if $\min_{a \in [0, \overline{a}_i]} c''(a_i)$ is sufficiently large. The reason is that u_{12}^1 is independent of $c''(a_i)$ while u_{22}^2 is decreasing in $c''(\cdot)$. In conclusion, $u_1^1(\mathbf{a}) - u_2^2(\mathbf{a}) > 0$, but this violates the requirement that the first-order conditions are satisfied.

Note also that if (T2) holds and $a_1 \geq a_2$, then it follows from (6) and (12) that agent 2 must necessarily view a_1 and a_2 as strategic substitutes.

B.5 LSM contests and CSFs

Proposition 1 signifies that the CSF in any LSM contests has some very specific properties. Moving in the other direction, if the starting point is some postulated CSF, then it can be checked whether the CSF in question satisfies the properties in Proposition 1. If not, then the CSF cannot be microfounded as a LSM contest with constant support. For instance, the CSF in Beviá and Corchón (2015) is not log-supermodular when $\beta(1 + s) < 1$ in their parameterization with two agents. Hence, it cannot be microfounded as a contest with stochastic performance that satisfies the MLRP and the assumption of identity-independent support.

Cubel and Sanchez-Pages (2016) provide an axiomatic justification for difference-form CSFs with several agents. However, in addition to the properties in Proposition 1, it is easily seen that $P_{23}^1(\mathbf{a}) > 0$ in LSM contests with $n \geq 3$ agents. Hence, difference-form CSFs – which have the property that $P_{23}^1(\mathbf{a}) = 0$ – cannot be microfounded as LSM contests when $n \geq 3$.[18]

For completeness, it is shown next that the difference-form CSF can be microfounded as a contest with stochastic performance even when $n \geq 3$, but the setting is

[18] For $n = 2$, the mixture model with homogeneous technologies provides a microfoundation for the difference-form CSF. This explains why there is a dominant action in the former.

somewhat artificial and requires violating MLRP. It it therefore not a LSM contest.

Consider a contest in which distributions have the same support, but where agents' actions effect disjoint subsets of the support. For concreteness, consider

$$F^i(x|a_i) = \begin{cases} \frac{x}{n} & \text{if } x \in [0,n]/[i-1,i] \\ \frac{i-1}{n} + \frac{(x-(i-1))^{1+a_i}}{n} & \text{if } x \in [i-1,i] \end{cases},$$

for $i = 1, 2, ..., n$. Here F^i first-order stochastically dominates the uniform distribution on $[0,n]$ and weakly improves in a first-order stochastic dominance sense when a_i increases. However, the MLRP does not hold (nor is the density continuous in this specific example).

Note that $P^i(\mathbf{a})$ can be written as a sum of terms, where each term depends only on the action of one player, or

$$P^i(\mathbf{a}) = \sum_{k=1}^{n} \int_{k-1}^{k} \prod\nolimits_{j \neq i} F^j(x|a_j) f^i(x|a_i) dx.$$

One term is increasing in a_i and each of the other terms is decreasing in the action of precisely one of the other agents. Thus, this additive CSF is a difference-form CSF.

B.6 Biases in contests

This section explains Fu and Wu's (2020) result on biased contests with two agents and examines whether it is generalizable.

Consider the following general setting. There are two agents with identity-independent technology $F(x|a_i)$. Valuations and cost functions can be identity-dependent. Assume the contest is a strict LSM contest. The regulator can transform, or bias, agent i's impact from a_i to $t_i(a_i)$, where $t_i(\cdot)$ belongs to a set of feasible transformations, T, with the property that $t_i'(a_i) \geq 0$. Note that the transformed contest, with $F^i(x|a_i) = F(x|t_i(a_i))$, remains a strict LSM contest with homogeneous technologies. Moreover, agent i's reaction function depends only on agent j's transformed impact, $t_j(a_j)$.

Assume the regulator has a payoff function that is increasing in actions. Let (t_1^*, t_2^*) and (a_1^*, a_2^*) denote the optimal transformations and the induced equilibrium action profile, respectively. The conjecture is that if T is sufficiently flexible, then the optimal design is such that the playing field is completely level in equilibrium, or

$P^i(\mathbf{a}^*) = \frac{1}{2}$. In other words, equilibrium impacts are the same, or $t_1^*(a_1^*) = t_2^*(a_2^*)$. The intuition stems from the individual discouragement effect: Any agent works hardest if the equilibrium is such that equilibrium impacts are the same, since this is where the reaction functions reach their peaks.

Consider first the following heuristic argument. Assume to start that $P^2(\mathbf{a}^*) < \frac{1}{2}$. Keep $t_1(\cdot)$ fixed, but if possible change t_2 to some $\widehat{t}_2$, where $\widehat{t}_2(a_2^*) = t_2(a_2^*)$ and $\widehat{t}_2'(a_2^*) = t_2'(a_2^*) + \varepsilon$. Given $t_1(\cdot)$ and $\widehat{t}_2'(\cdot)$

$$u_2^2(a_1^*, a_2^*|t_1, \widehat{t}_2) = \int F(x|t_1(a_1^*)) f_a(x|t_2(a_2^*)) \left(t_2'(a_2^*) + \varepsilon\right) dx - c_2'(a_2^*).$$

Hence, agent 2's best response to a_1^* increases when $\varepsilon > 0$. Agent 1's reaction to $\widehat{t}_2(a_2^*)$ and $t_2(a_2^*)$ are by definition the same, and since $P^1(\mathbf{a}^*) > \frac{1}{2}$, her reaction function is strictly increasing locally. Hence, when ε is small, *both* agents work strictly harder in the new equilibrium. In other words, if T is sufficiently flexible, there is an opportunity to "kill two birds with one stone" and entice both agents to work harder, which is beneficial to the regulator.

Fu and Wu (2020) essentially assume that T consists of affine transformations, $t_i(a_i) = \alpha_i a_i + \beta_i$ with $\alpha_i, \beta_i \geq 0$. Using the logic above suggests that $\beta_i > 0$ cannot be optimal (since it would enable a lowering of β_i and increase of α_i), which is indeed what Fu and Wu find. Once $\beta_i = 0$, there is no longer the required flexibility to keep $t_2(a_2^*)$ fixed while changing $t_2'(a_2^*)$. Nevertheless, in their specific contest, the optimal α_i's still ensure $P^i(\mathbf{a}^*) = \frac{1}{2}$. However, as explained next, this turns out to be the case because their CSF is homogeneous of degree zero in actions.

Assume, as in Fu and Wu (2020), that $P^i(\mathbf{a})$ is strictly concave in a_i and $c_i(a_i)$ is convex. Recall that the former is more broadly justified in Section B.2.1. Then, agent i's problem is strictly concave in a_i for all $\alpha_1, \alpha_2 > 0$. Now, normalize $\alpha_1 = 1$, such that a_1 and a_2 are implicit functions of α_2 and derived from the first-order conditions

$$\begin{aligned} v_1 P_1^1(a_1, \alpha_2 a_2) - c_1'(a_1) &= 0 \\ v_2 P_2^2(a_1, \alpha_2 a_2)\alpha_2 - c_2'(a_2) &= 0. \end{aligned}$$

By implicit differentiation,

$$
\begin{aligned}
a_1'(\alpha_2) &= \frac{P_{12}^1\alpha_2 v_1 k - a_2 v_1 P_{12}^1\left[v_2 P_{22}^2\alpha_2^2 - c_2''(a_2)\right]}{\left[v_1 P_{11}^1 - c_1''(a_1)\right]\left[v_2 P_{22}^2\alpha_2^2 - c_2''(a_2)\right] - v_1 v_2\alpha_2^2 P_{12}^1 P_{12}^2}\\
a_2'(\alpha_2) &= \frac{v_1 v_2\alpha_2 a_2 P_{12}^1 P_{12}^2 - \left[v_1 P_{11}^1 - c_1''(a_1)\right]k}{\left[v_1 P_{11}^1 - c_1''(a_1)\right]\left[v_2 P_{22}^2\alpha_2^2 - c_2''(a_2)\right] - v_1 v_2\alpha_2^2 P_{12}^1 P_{12}^2},
\end{aligned}
$$

where

$$k = v_2 P_{22}^2(a_1, \alpha_2 a_2)\alpha_2 a_2(\alpha_2) + v_2 P_2^2(a_1, \alpha_2 a_2).$$

The terms in square brackets are evidently negative. Hence, given that $P_{12}^1 P_{12}^2 \leq 0$, the denominator is positive. Now consider an α_2 for which $P_{12}^i(\mathbf{a}) = 0$ in equilibrium. Then, $a_1'(\alpha_2) = 0$, whereas the sign of $a_2'(\alpha_2)$ is the same as the sign of k. Thus, if $k \neq 0$, a_2 can be increased without any first-order effect on a_1. In this case, as long as the regulator's payoff is increasing in actions, $P_{12}^i(\mathbf{a}) = 0$ cannot be optimal. Remember that if technologies are homogeneous, then $P_{12}^i(\mathbf{a}) = 0$ if and only if $P^i(\mathbf{a}) = \frac{1}{2}$, i.e. if the playing field is perfectly balanced. Thus, in general, a balanced playing field is not optimal when the design instrument is as in Fu and Wu (2020).

However, assume now that $P^i(\mathbf{a})$ is homogeneous of degree zero in actions, as is the case in Fu and Wu (2020). Then, k can be shown to be precisely equal to

$$k = -v_2 P_{12}^2(a_1, \alpha_2 a_2)a_1,$$

which is of course zero when $P_{12}^i(\mathbf{a}) = 0$. Moreover, it holds that

$$
\begin{aligned}
P_{11}^1(a_1, \alpha_2 a_2)a_1 + P_{12}^1(a_1, \alpha_2 a_2)\alpha_2 a_2 &= -P_1^1(a_1, \alpha_2 a_2) \text{ and}\\
P_{22}^2(a_1, \alpha_2 a_2)\alpha_2^2 a_2 + P_{12}^2(a_1, \alpha_2 a_2)\alpha_2 a_1 &= -P_2^2(a_1, \alpha_2 a_2)\alpha_2.
\end{aligned}
$$

Since $P_{12}^2(\mathbf{a}) = -P_{12}^1(\mathbf{a})$, it now holds that

$$
\begin{aligned}
a_1'(\alpha_2) &= P_{12}^1\frac{v_1 v_2 P_2^2\alpha_2 + a_2 v_1 c_2''(a_2)}{\left[v_1 P_{11}^1 - c_1''(a_1)\right]\left[v_2 P_{22}^2\alpha_2^2 - c_2''(a_2)\right] - v_1 v_2\alpha_2^2 P_{12}^1 P_{12}^2}\\
a_2'(\alpha_2) &= P_{12}^1\frac{v_1 v_2 P_1^1 + a_1 v_2 c_1''(a_1)}{\left[v_1 P_{11}^1 - c_1''(a_1)\right]\left[v_2 P_{22}^2\alpha_2^2 - c_2''(a_2)\right] - v_1 v_2\alpha_2^2 P_{12}^1 P_{12}^2},
\end{aligned}
$$

meaning that both have the same sign as $P_{12}^1(\mathbf{a})$. Hence, if $P_{12}^1(\mathbf{a}) > 0$ (< 0) then an increase (decrease) in α_2 leads to an increase in both a_1 and a_2, to the regulator's benefit. In other words, the optimal bias is such that $P^i_{12}(\mathbf{a}) = 0$ in equilibrium. Clearly, this conclusion depends on the combination of linear transformations and a CSF that is homogeneous of degree zero.

In the case where the CSF is asymmetric as a consequence of heterogeneous technologies, but still homogeneous of degree zero, it is typically not the case that $P^i_{12}(\mathbf{a}) = 0$ coincides with $P^i(\mathbf{a}) = \frac{1}{2}$. That is, it is not generally the case that the optimal linear bias implements a perfectly balanced playing field.