# Financial Constraints and Multivariate Incomplete Information in the Mixture Model of Contests*

René Kirkegaard†

December 2023

**Abstract**

A general mixture model of contests is introduced. The contest combines multivariate adverse selection with moral hazard in the form of stochastic performance. Incomplete information does not add to the dimensionality of the problem, and actions are unambiguously strategic substitutes when technologies are homogenous. These properties facilitate tractable comparative statics with respect to changes in the multivariate type distribution. Hence, the role of the dependence structure between different characteristics can be explored. To illustrate, if valuations and budgets become more positively dependent, then each type's expenditure decreases but expected performance increases. With asymmetric agents, the exclusion principle can be reversed.

JEL Classification Numbers: C72, D72, D82.

Keywords: Contests, exclusion principle, supermodular order.

*I thank SSHRC for funding this research.

†Department of Economics and Finance, University of Guelph. Email: rkirkega@uoguelph.ca.

# 1 Introduction

Contests are strategically complicated environments. Realistically, they often feature two distinct sources of uncertainty. First, agents face incomplete information about the characteristics of rivals. These characteristics can be multivariate and correlated. For instance, rivals' valuations, productivity, and financial resources may be unknown, and it is not unreasonable to think that a rival with more resources may also be more likely to have a higher valuation or willingness to pay. Second, there is outcome uncertainty in the sense that performance is stochastic and only imperfectly related to the agent's effort or expenditure. Incentives are shaped by both considerations. It is obvious that characteristics and beliefs impact incentives, but it is also true that the nature of the randomness involved in the agent's performance has a role to play in determining his return to effort and thereby his incentives.

Unfortunately, a completely general analysis of contests is still not within reach. Hence, the contest literature relies to a large extent on a limit set of tractable "workhorse models," chief among them the Tullock contest, the rank-order tournament, and the all-pay auction.[1] These models differ in how actions determine or influence who wins the contests – in the parlance of contest theory, the contest success functions (CSFs) are different in each model. This paper studies a complementary model of contests, termed the mixture model, in which the CSF takes yet another form.

The chief advantage of the mixture model is that unlike existing models it easily accommodates a combination of multivariate incomplete information and stochastic performance.[2] This makes it possible to explore the role of the statistical dependence between different characteristics that was alluded to in the first paragraph. The analysis of the relationship between valuations and budgets and between valuations and starting advantages are particularly clean. Moreover, the mixture model challenges the robustness of some of the comparative statics in the existing literature, thus demonstrating that the workhorse models do not give a full picture of the kind of comparative statics that may arise more generally in contests. This is important because comparative statics inform design choices and policy recommendations.

[1] See Konrad (2009) and Vojnović (2016) for book-length surveys of contest theory. For other recent surveys, see Corchón and Serena (2018) and Fu and Wu (2019).

[2] The Tullock contest (Tullock 1975, 1980) can be microfounded as a contest with stochastic performance. The rank-order tournament also features stochastic performance. However, it is hard to handle even univariate incomplete information in those models. The all-pay auction is deterministic.

The mixture model is as follows. Agents have independent private information about valuations, impact functions, cost functions, and action sets. The impact function influences how productive an agent's action is. To be specific, the distribution of performance is a mixture distribution with two mixture components and endogenous weights that are determined by the agent's impact. Thus, the agent's costly action amounts to purchasing a compound lottery. The higher his impact is, the more likely it is that his performance is drawn from a good rather than a bad component or distribution. Indeed, the agent's expected performance is proportional to his impact.

The mixture distribution may arise naturally in settings where the agent is hiring a subcontractor, manager, or research team, or alternatively procuring equipment, any of which aids, or is responsible for, the agent's output. Imagine that the subcontractor can be good or bad. A good (bad) subcontractor draws performance from the good (bad) mixture component. Then, the agent's action represents his search effort and his impact the probability that he correctly identifies a good subcontractor.

Private information about types implies that the agent's impact function and equilibrium action are not known to his competitors. Hence, the competitors are uncertain about the weights that are placed on the two mixture components. When they attempt to assess the distribution of the agent's performance, the uncertainty over types gives rise to a lottery over the compound lottery chosen by each type. What matters in this doubly-compounded lottery is simply the expected cumulative weight that is placed on the good and bad components, respectively. This implies that the expected value of the impact function is a sufficient measure of the ex ante distribution of the agent's performance – knowing this single number is sufficient to describe everything that is of relevance about the agent to his competitors.

The sufficient uncertainty measure drastically reduces the dimensionality of the problem. Any agent is just best responding to the profile of his competitors' expected impacts. Thus, the optimization problem is no more difficult to handle with incomplete information than with complete information. The lack of such a measure in e.g. the Tullock contest makes it much harder to add private information to that model.

Starting with Section 3, focus is on contests with homogenous technologies or, more precisely, where mixture components are identity-independent. In such contests, actions are always strategic substitutes in the sense that the agent, regardless of his type, lowers his action when a competitor's expected impact increases. This monotonicity property also adds to the relative tractability of the mixture model compared to e.g. the Tullock contest in which reaction functions are hump-shaped.

Section 4 assumes that agents are ex ante symmetric and develops general comparative statics results on how changes to the type distribution influences the unique symmetric equilibrium. Such results are feasible because equilibrium characterization hinges on solving a one-dimensional fixed-point problem that nails down the equilibrium value of the summary uncertainty measure. The fact that actions are strategic substitutes makes it possible in a range of applications to say how the fixed-point problem and its solution is affected by changes in the type distribution.

Different kinds of changes in the type distribution are considered. Roughly speaking, the first assumes that "strong" types become more likely, with the understanding that this can be modeled is several different ways with multivariate types. Nevertheless, the general conclusion is that any given type works less hard but, since stronger types are more likely, the expected equilibrium impact and performance increase.

The second and more novel comparative statics examine the role of the dependence structure between different characteristics. This is an underexplored issue in contests, in no small part simply because it is harder to accommodate multivariate types in existing models. To isolate the role of the dependence structure, the marginal distributions are held fixed while the joint distribution changes. The supermodular order is used to discipline these changes. The correlation between any two characteristics increases when the characteristics become more positively dependent in the sense of the supermodular order. Equivalently, the expected value of any supermodular payoff function – i.e. those for which inputs are complements – increases.

The resulting comparative statics are best explained in the context of applications. In the leading application, the importance of the relationship between private valuations and budgets (or time constraints on the agent's search for a subcontractor) is examined. The supermodular order is ideally suited to this application (Section 5). The reason is that the two characteristics are complements because the agent is willing and able to achieve a high impact only if both his valuation and budget are high at the same time. As this is more likely to occur the more positively dependent the two are, the expected impact increases, other things equal. However, this observation does not entirely settle matters, since other things are not equal when equilibrium changes. Nevertheless, the nature of the fixed-point problem ensures that equilibrium changes in the anticipated direction. Thus, the expected performance increases.

The opposite conclusion obtains if types capture valuations and starting advantages. The reason is that these characteristics are substitutes since high impacts are justified or automatic if the valuation or the starting advantage is high, respectively.

In an application with additive and multiplicative productivity parameters, the two characteristics have complementary effects on actions (Section 6). However, this interaction washes out in the agent's impact. Hence, greater dependence does not change the expected performance, but it does increase the expected action.

In the above applications, the equilibrium impact is influenced in an unambiguous direction by changes in the dependence structure. The conclusion is less clear-cut when types reflect valuations and productivity. Sufficient conditions are identified under which the characteristics are substitutes or complements, respectively.

Lastly, the consequences of a mean-preserving spread of a univariate type distribution in symmetric contests is considered (Section 7). The main application is to uncertainty about valuations. Here, expected performance may increase or decrease depending on the primitives. Sufficient conditions are given for either possibility. It is possible that agents expend more resources, yet perform worse.

These results are also relevant to complete information contests in which agents have different valuations. Increasing the asymmetry is similar to a mean-preserving spread and may, depending on the primitives, encourage competition. This is in contrast to the discouragement effect that is found in much of the existing literature. Similarly, the exclusion principle may be turned on its head. Baye et al (1993) show that excluding a strong competitor in an all-pay auction may increase total effort. In the mixture contest, it may instead be optimal to exclude a weak agent. The intuition is simple. In a two-player mixture contest, equilibrium is in strictly dominant strategies. Hence, a stronger agent works harder than a weaker agent. Thus, starting with a three-player contest, if it is profitable to exclude anyone, it must be the weaker agent in order to retain the stronger agents for the dominant-strategy contest.

Section 8 presents extensions and discusses the empirical falsification of competing contest models, including the mixture model. Section 9 concludes. Proofs are in Appendix A. Appendix B contains supplementary material.

Related Literature: Gürtler and Kräkel (2012) consider a mixture model with two agents and homogenous technologies. Although they do not explicitly phrase it this way, it is clear from their analysis that equilibrium is in dominant strategies. Kräkel (2010) consider a variant with two teams. While there is the usual coordination problem within each team, there is no strategic interaction across the two teams.

Kräkel (2010) and Gürtler and Kräkel (2012) note that the mixture distribution is an application of the spanning condition used in Grossman and Hart's (1983) principal-agent model. Kirkegaard (2017) solves the principal-agent contracting problem when the spanning condition holds but the first-order approach is invalid. In Kirkegaard (2023a), contest design is viewed as a contracting problem with multiple agents. He derives optimal design principles for a general complete information model with stochastic performance. Design takes the form of biasing the allocation rule, and the winner is not necessarily the agent with the best performance. Among the examples is one that satisfies the spanning condition. The current paper focuses on comparative statics in unbiased contests while permitting incomplete information.

There is a small literature on incomplete information in Tullock contests. Malueg and Yates (2004) characterize equilibrium with two agents and binary types. Fey (2008) and Ryvkin (2010) derive equilibrium numerically with a continuum of univariate types and up to four agents. Ewerhart and Quartieri (2020) consider private information in very general Tullock-style contests, but their focus is on equilibrium existence and uniqueness. They provide a thorough literature review. See Hammond and Zheng (2013) for a discussion of private information in rank-order tournaments.

In the mixture model with homogenous technologies, all types agree that actions are strategic substitutes. In other contest models it is usually the case that best-response functions are non-monotonic and that different types disagree about whether actions are strategic substitutes or complements. Hopkins and Kornienko (2007) show that in all-pay auctions, different types react differently to changes in the distribution of valuations. This is also the case in the numerical examples in Ryvkin (2010).

The literature on private valuations and budgets in auctions is also relevant. Che and Gale (1998) compare standard deterministic auctions with correlated valuations and budgets. However, they do not study the effects of changes in the dependence structure. Kotowski and Li (2014) consider the all-pay auction and the war of attrition. They assume that valuations are affiliated across agents but that budgets are independent of valuations. A change in the distribution of budgets may be met with different reactions depending on the agent's valuation. The mixture model assumes independence across agents but allows correlated valuations and budgets in a setting with stochastic performance. The current paper also identifies a stochastic order (the supermodular order) that is particularly well-suited to study how equilibrium is influenced by the dependence between budgets and valuations in the mixture model.

## 2 The mixture model of contests

Consider a contest with $n$ agents and a single prize. Agents are independently and privately informed about their types. Agent $i$'s type is denoted $\boldsymbol{\theta}_i$ and it is allowed to be multi-dimensional. The non-empty type-space is $\Theta_i$. Each agent's type distribution is common knowledge. Complete information arises if $\Theta_i$ is a singleton.

Agent $i$'s type determines his valuation, impact function, cost function, and action set. The action set is a unidimensional, non-empty and compact interval, $A_i(\boldsymbol{\theta}_i)$. Note that the action set may be type-dependent. Different upper bounds may capture different and privately known time or budget constraints. Different lower bounds may describe different "starting advantages," in the language of Siegel (2009). A generic element of $A_i(\boldsymbol{\theta}_i)$ is denoted $a_i$. Actions are costly, with cost function $c_i(a_i|\boldsymbol{\theta}_i)$. The valuation of the prize is $v_i(\boldsymbol{\theta}_i) \geq 0$. Agent $i$ earns net payoff $v_i(\boldsymbol{\theta}_i) - c_i(a_i|\boldsymbol{\theta}_i)$ if he wins and $-c_i(a_i|\boldsymbol{\theta}_i)$ otherwise. There is no option to exit the contest.

Agent $i$'s performance is determined in part by his impact function, $p_i(a_i|\boldsymbol{\theta}_i) \in [0,1]$. For all $a_i \in A_i(\boldsymbol{\theta}_i)$ and all $\boldsymbol{\theta}_i \in \Theta_i$, the cost and impact functions satisfy

$$\begin{aligned}
\frac{\partial c_i(a_i|\boldsymbol{\theta}_i)}{\partial a_i} &> 0,\ \frac{\partial p_i(a_i|\boldsymbol{\theta}_i)}{\partial a_i} \geq 0, \text{ and} \\
\frac{\partial^2 c_i(a_i|\boldsymbol{\theta}_i)}{\partial a_i^2} &\geq 0,\ \frac{\partial^2 p_i(a_i|\boldsymbol{\theta}_i)}{\partial a_i^2} \leq 0, \text{ with at least one strict inequality.}
\end{aligned}$$

Some redundancy is built into this model formulation as both $p_i$ and $c_i$ can be nonlinear in the action. Depending on the application, it may be meaningful to normalize one of the two to be linear. For instance, if the aim is to study budget constraints in contests where actions are monetary expenditures, then it is sensible to assume that $c_i$ is linear. However, the redundancy also facilitates comparison to different strands of the literature. In Tullock contests, it is often assumed that the cost function is linear, while in rank-order contests it is often assumed that the impact is linear.

The contest is won by the agent with the best performance. Agent $i$'s performance, $X_i$, is a unidimensional random variable. In particular, if agent $i$ has type $\boldsymbol{\theta}_i$ and takes action $a_i$, then his performance follows the mixture distribution

$$F_i(x_i|a_i, \boldsymbol{\theta}_i) = p_i(a_i|\boldsymbol{\theta}_i) H_i(x_i) + (1 - p_i(a_i|\boldsymbol{\theta}_i))\, G_i(x_i), \tag{1}$$

where the publicly known mixture components $H_i$ and $G_i$ are continuous and atomless distribution functions with densities $h_i$ and $g_i$, respectively. It is crucial that $H_i$ and $G_i$ are independent of $\boldsymbol{\theta}_i$ and $a_i$. The mixture components are also independent of one another. Moreover, $H_i$ and $G_i$ have the same support, and this is the same for all agents, $i = 1, ..., n$. Thus, the performance of all agents have the same support regardless of actions. Consequently, there is always a chance of winning the contest but winning is never guaranteed. The probability of ties is zero since distributions are atomless. When needed, the support is denoted $[\underline{x}, \overline{x}]$. It can be bounded or unbounded. Finally, $H_i$ first order stochastically dominates $G_i$ in the strict sense that $H_i < G_i$ on $(\underline{x}, \overline{x})$. Thus, $H_i$ is a more productive technology than $G_i$. Let $\mu_{H_i}$ and $\mu_{G_i}$ denote the expected values of $H_i$ and $G_i$ respectively, with $\mu_{H_i} > \mu_{G_i}$.

The agent's action essentially buys him a compound lottery. His performance is drawn from the good and bad component with probability $p_i(a_i|\boldsymbol{\theta}_i)$ and $1 - p_i(a_i|\boldsymbol{\theta}_i)$, respectively. Recall that action sets, valuations, impact functions, and cost functions can all be private information at the same time. Applications are in Sections 5–7.

The mixture distribution is separable in $x_i$ and $p_i$, and only the latter depends on $a_i$ and $\boldsymbol{\theta}_i$. This is why it is important that the mixture components are publicly known, and it has two indispensable consequences. The first relates to the structure of an agent's beliefs about the performance of other agents. The second concerns the structure of the agent's maximization problem, given his beliefs, which in turn gives structure to the best response functions.

## 2.1 A summary uncertainty measure

Agent $i$ faces uncertainty when thinking about the distribution of agent $j$'s performance, $j \neq i$. Agents $j$'s type is unknown, and he could be using a mixed strategy (a complication that arises in Section 8.1). A mixed strategy, $\sigma_j$, is a mapping from types into a probability distribution over $A_j(\boldsymbol{\theta}_j)$. The two sources of uncertainty produce a joint distribution over types and actions. This can be used to compute the ex ante "expected impact" of agent $j$'s action, which is just the expected value of $p_j(a_j|\boldsymbol{\theta}_j)$. Let this be denoted $\overline{p}_j(\sigma_j)$. Then, the ex ante distribution of agent $j$'s performance is

$$F_j^{\sigma_j}(x_j) = \overline{p}_j(\sigma_j)H_j(x_j) + \left(1 - \overline{p}_j(\sigma_j)\right) G_j(x_j).$$

Everything that is relevant to agent $i$ about agent $j$'s strategy is captured by the single number $\overline{p}_j(\sigma_j)$, which is effectively a summary and sufficient uncertainty measure.

## 2.2 Best responses

It is not relevant to agent $i$ what $\sigma_j$ is, beyond the $\overline{p}_j$ value that it implies. Now, fix some profile $\overline{\mathbf{p}}_{-i} = \left(\overline{p}_1, ..., \overline{p}_{i-1}, \overline{p}_{i+1}, ...\overline{p}_n\right)$. Given type $\boldsymbol{\theta}_i$ and action $a_i$, agent $i$'s interim winning probability is

$$q_i\left(a_i, \overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i\right) = \int \left(\prod_{j \neq i} \left(\overline{p}_j H_j(x) + \left(1 - \overline{p}_j\right) G_j(x)\right)\right) \left(p_i(a_i|\boldsymbol{\theta}_i) h_i(x) + \left(1 - p_i(a_i|\boldsymbol{\theta}_i)\right) g_i(x)\right) dx.$$

This can be decomposed into

$$q_i\left(a_i, \overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i\right) = t_i\left(\overline{\mathbf{p}}_{-i}\right) + p_i(a_i|\boldsymbol{\theta}_i) k_i\left(\overline{\mathbf{p}}_{-i}\right), \tag{2}$$

where

$$\begin{aligned} t_i\left(\overline{\mathbf{p}}_{-i}\right) &= \int \left(\prod_{j \neq i} \left(\overline{p}_j H_j(x) + \left(1 - \overline{p}_j\right) G_j(x)\right)\right) \times g_i(x) dx \\ k_i\left(\overline{\mathbf{p}}_{-i}\right) &= \int \left(\prod_{j \neq i} \left(\overline{p}_j H_j(x) + \left(1 - \overline{p}_j\right) G_j(x)\right)\right) \left(h_i(x) - g_i(x)\right) dx. \end{aligned}$$

Given $\overline{\mathbf{p}}_{-i}$, $t_i\left(\overline{\mathbf{p}}_{-i}\right)$ measures the base probability that agent $i$ wins and $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ the return to effort. The latter is the difference between the probability that agent $i$ wins with a draw from $H_i$ and $G_i$, respectively. First order stochastic dominance means that the winning probability is higher under $H_i$ than $G_i$. Thus, $k_i\left(\overline{\mathbf{p}}_{-i}\right) > 0$. Likewise, $k_i(\overline{\mathbf{p}}_{-i}) < 1$ because it is the difference between two probabilities.

In the language of contest theory, (2) is an interim contest success function (CSF). It takes the strategy profile of agent $i$'s competitors along with his own action (and type) and outputs a winning probability. An ex ante CSF is obtained by taking the expectation of (2) with respect to agent $i$'s strategy. This yields the ex ante probability that agent $i$ wins the contest as a function of the profile $\overline{\mathbf{p}} = (\overline{p}_1, \overline{p}_2, ..., \overline{p}_n)$,

$$\overline{q}_i(\overline{\mathbf{p}}) = t_i\left(\overline{\mathbf{p}}_{-i}\right) + \overline{p}_i k_i\left(\overline{\mathbf{p}}_{-i}\right).$$

Hence, the mixture model yields separable CSFs.

Given $\overline{\mathbf{p}}_{-i}$ and type $\boldsymbol{\theta}_i$, agent $i$'s objective is to maximize

$$u_i\left(a_i, \overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i\right) = v_i(\boldsymbol{\theta}_i)\left(t_i\left(\overline{\mathbf{p}}_{-i}\right) + p_i(a_i|\boldsymbol{\theta}_i) k_i\left(\overline{\mathbf{p}}_{-i}\right)\right) - c_i(a_i|\boldsymbol{\theta}_i). \tag{3}$$

Since $t_i\left(\overline{\mathbf{p}}_{-i}\right)$ is independent of agent $i$'s action, the problem is equivalent to maximizing

$$U_i(a_i, \overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i) = v_i(\boldsymbol{\theta}_i)p_i(a_i|\boldsymbol{\theta}_i)k_i(\overline{\mathbf{p}}_{-i}) - c_i(a_i|\boldsymbol{\theta}_i). \tag{4}$$

Recall that $k_i\left(\overline{\mathbf{p}}_{-i}\right) > 0$. Thus, if $v_i(\boldsymbol{\theta}_i) = 0$ then the unique solution is to take the lowest possible action. If $v_i(\boldsymbol{\theta}_i) > 0$, then the problem is strictly concave and the solution is again unique. Let $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ denote type $\boldsymbol{\theta}_i$'s unique best response to $\overline{\mathbf{p}}_{-i}$.

Note that $k_i(\overline{\mathbf{p}}_{-i})$ describes all that is relevant about the aggregate uncertainty faced by agent $i$. The higher $k_i(\overline{\mathbf{p}}_{-i})$ is, the higher is the best response. Thus, how $k_i(\overline{\mathbf{p}}_{-i})$ depends on $\overline{\mathbf{p}}_{-i}$ determines the nature of the strategic interaction. If $\frac{\partial}{\partial \overline{p}_j} k_i\left(\overline{\mathbf{p}}_{-i}\right) < 0$, then the incentive to increase $p_i$ is lower the higher $\overline{p}_j$ is, in which case there is a sense in which agent $i$ considers actions to be strategic substitutes, $j \neq i$. More precisely, an increase in the action by a set of agent $j$'s types of positive mass increases $\overline{p}_j$, which in turn makes agent $i$ decrease his action. Conversely, agent $i$ considers actions to be strategic complements if $\frac{\partial}{\partial \overline{p}_j} k_i\left(\overline{\mathbf{p}}_{-i}\right) > 0$. The sign of $\frac{\partial}{\partial \overline{p}_j} k_i\left(\overline{\mathbf{p}}_{-i}\right)$ depends only on the properties of the mixture components. Finally, since $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ is linear in $\overline{p}_j$, the sign of $\frac{\partial}{\partial \overline{p}_j} k_i\left(\overline{\mathbf{p}}_{-i}\right)$ does not depend on $\overline{p}_j$. Thus, agent $i$'s "reaction function," $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$, is *globally* increasing or decreasing in $\overline{p}_j$, and it moves in the *same* direction for all $\boldsymbol{\theta}_i \in \Theta_i$.

## 2.3 Equilibrium

Let $\overline{p}_i^{\min} = \mathbb{E}_{\boldsymbol{\theta}_i}\left[\min_{a_i \in A_i(\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)\right]$ and $\overline{p}_i^{\max} = \mathbb{E}_{\boldsymbol{\theta}_i}\left[\max_{a_i \in A_i(\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)\right]$ denote the smallest and largest possible values of $\overline{p}_i$. Given $\overline{\mathbf{p}}_{-i}$ and type-dependent best responses, agent $i$'s ex ante expected impact is just $\mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right] \in \left[\overline{p}_i^{\min}, \overline{p}_i^{\max}\right]$. In equilibrium, agents are mutually best responding. Hence, in equilibrium, the profile $\overline{\mathbf{p}} = (\overline{p}_1, \overline{p}_2, ..., \overline{p}_n) \in \times_{i=1}^n \left[\overline{p}_i^{\min}, \overline{p}_i^{\max}\right]$ must satisfy

$$\overline{p}_i = \mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right] \text{ for all } i = 1, ..., n. \tag{5}$$

Characterizing equilibrium amounts to solving this fixed-point problem. Once the fixed-point $\overline{\mathbf{p}}$ has been found, the equilibrium strategy is given by $s_i(\boldsymbol{\theta}_i) = B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$.

**Proposition 1** *A solution $\overline{\mathbf{p}} = (\overline{p}_1, \overline{p}_2, ..., \overline{p}_n) \in \times_{i=1}^n \left[\overline{p}_i^{\min}, \overline{p}_i^{\max}\right]$ to (5) exists. For any such solution, there exists an associated pure-strategy Bayesian Nash Equilibrium of the contest game, in which strategies are given by $s_i(\boldsymbol{\theta}_i) = B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$, $i = 1, ..., n$.*

## 2.4 Interpretation and the organizer's objective function

The equilibrium fixed-point contains important information. First, $\overline{\mathbf{p}}$ is sufficient to describe the ex ante joint distribution of the performance profile, which is

$$J(x_1, x_2, ..., x_n|\overline{\mathbf{p}}) = \prod_{i=1}^{n} \left(\overline{p}_i H_i(x_i) + (1-\overline{p}_i) G_i(x_i)\right).$$

Hence, if the contest organizer has a Bernoulli utility function that depends only on the performance profile, then $\overline{\mathbf{p}}$ is enough to calculate her expected utility. In fact, expected payoff is increasing in each $\overline{p}_i$ as long as the Bernoulli utility function is increasing; see Section 4.1 for a formal argument. To illustrate, agent $i$'s expected performance,

$$\mathbb{E}[X_i|\overline{p}_i] = \overline{p}_i \mu_{H_i} + (1-\overline{p}_i)\mu_{G_i},$$

is proportional to $\overline{p}_i$. Thus, if the organizer is risk neutral and benefits from the agents' total performance or output – her Bernoulli utility is $\sum_{i=1}^{n} x_i$ – then her expected utility is proportional to a weighted sum of the components of $\overline{\mathbf{p}}$. If the mixture components are identity-independent as in Sections 3–7, then total expected output is proportional to the "total impact," $\sum_{i=1}^{n} \overline{p}_i$.

Finally, $\overline{\mathbf{p}}$ determines agent $i$'s ex ante winning probability, $\overline{q}_i(\overline{\mathbf{p}})$. Hence, $\overline{\mathbf{p}}$ reveals who is favorite to win and how unevenly distributed winning probabilities are. Agent $i$ of a given type is worse off when $\overline{\mathbf{p}}_{-i}$ is higher, because this leads the distribution of the highest rival performance to improve in the sense of first-order stochastic dominance.

# 3 Contests with homogeneous technologies

The majority of the paper assumes that technologies are homogeneous across agents, or $H_i = H$ and $G_i = G$ for all $i = 1, .., n$. Hence, if two agents take actions that lead to the same impact, then they have the same distribution of performance.[3]

This section establishes a robust property of the nature of the strategic considerations in contests with homogeneous technologies. The property is robust in the sense that it is invariant not only to the details of $H$ and $G$, but also to agents' valuations,

[3] In contrast, it is generally impossible for one agent to replicate the distribution of performance of another agent when technologies are not homogenous. The companion paper, Kirkegaard (2023b), considers such contests with two agents. The resulting model challenges some common perceptions about the strategic considerations of the favorite and the underdog, respectively.

impact functions, cost functions, action sets, and type distributions.

It was observed after (4) that whether agent $i$ considers the actions of agents $i$ and $j$ to be strategic substitutes or complements is determined by how $k_i(\overline{\mathbf{p}}_{-i})$ depends on $\overline{p}_j$. Thus, the sign of $\frac{\partial}{\partial \overline{p}_j} k_i(\overline{\mathbf{p}}_{-i})$ is of interest. However, note that $k_i(\overline{\mathbf{p}}_{-i}) = \frac{\partial}{\partial \overline{p}_i} \overline{q}_i(\overline{\mathbf{p}})$. Therefore, $\frac{\partial}{\partial \overline{p}_j} k_i(\overline{\mathbf{p}}_{-i}) = \frac{\partial}{\partial \overline{p}_j \partial \overline{p}_i} \overline{q}_i(\overline{\mathbf{p}})$. In other words, the question is whether the ex ante winning probability $\overline{q}_i(\overline{\mathbf{p}})$ is submodular or supermodular in $\overline{p}_i$ and $\overline{p}_j$. This question is easy to answer. Consider agents 1 and 2. Since ex ante winning probabilities sum to one for all $\overline{\mathbf{p}}$, it holds that $\frac{\partial^2}{\partial \overline{p}_1 \partial \overline{p}_2} \sum \overline{q}_i(\overline{\mathbf{p}}) = 0$. Moreover, since technologies are homogenous, it is easy to see that $\frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_1(\overline{\mathbf{p}}) = \frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_2(\overline{\mathbf{p}})$. Combining these two facts implies that if agents 1 and 2 are the only participants, or $n = 2$, then $\frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_1(\overline{\mathbf{p}}) = \frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_2(\overline{\mathbf{p}}) = 0$. Thus, $k_1$ and $k_2$ are constants and there is no strategic interaction whatsoever. Hence, equilibrium is in strictly dominant strategies.

Assume now that $n \geq 3$. Consider agents 1 and 2, along with their common rival, agent 3. Agent 3 obviously wins less often in expectation when $\overline{p}_1$ increases, or $\frac{\partial}{\partial \overline{p}_1} \overline{q}_3(\overline{\mathbf{p}}) < 0$, but it is also easily verified that the change is less pronounced the larger $\overline{p}_2$ is, or $\frac{\partial^2}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_3(\overline{\mathbf{p}}) > 0$. The reason is that agent 3 is already unlikely to win if $\overline{p}_2$ is large, so a marginal increase in $\overline{p}_1$ has less of an effect on agent 3. This argument applies to all of the rivals that agents 1 and 2 have in common. Hence

$$\begin{aligned} 0 &= \frac{\partial^2}{\partial \overline{p}_1 \partial \overline{p}_2} \sum \overline{q}_i(\overline{\mathbf{p}}) \\ &> \frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_1(\overline{\mathbf{p}}) + \frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_2(\overline{\mathbf{p}}) \end{aligned}$$

Since $\frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_1(\overline{\mathbf{p}}) = \frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_2(\overline{\mathbf{p}})$ when technologies are homogeneous, it follows that actions must be strategic substitutes, or $\frac{\partial}{\partial \overline{p}_1 \partial \overline{p}_2} \overline{q}_1(\overline{\mathbf{p}}) = \frac{\partial}{\partial \overline{p}_2} k_1(\overline{\mathbf{p}}_{-1}) < 0$ when $n \geq 3$. More precisely, if $\overline{p}_j$ increases, then agent $i$'s best response is lower type-for-type.[4]

**Proposition 2** *There is no strategic interaction in any mixture contest with homogenous technologies and $n = 2$ agents, and equilibrium is therefore in strictly dominant strategies. Actions are strategic substitutes in any mixture contest with homogeneous technologies and $n \geq 3$ agents.*

Appendix B provides a more careful description of the properties of $k_i\left(\overline{\mathbf{p}}_{-i}\right)$. Note that $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ is a polynomial in the $n-1$ variables in $\overline{\mathbf{p}}_{-i}$. By expanding $k_i\left(\overline{\mathbf{p}}_{-i}\right)$,

[4] Heterogeneous technologies are more complicated because whether actions are strategic substitutes or complements for a given pair of agents may depend on the actions of the other agents.

it is shown in Appendix B that the coefficient to the term $\prod_{j \neq i} \overline{p}_j$ is zero when technologies are homogeneous. Hence, $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ is only a polynomial of degree $n-2$. Thus, as already discussed, it is constant when $n = 2$. When $n = 3$, $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ depends only on the aggregate expected impact of the competitors, or $\sum_{j \neq i} \overline{p}_j$.[5]

# 4 Symmetric contests and comparative statics

Consider now the case in which agents are completely symmetric ex ante. Thus, the type spaces, type distributions, actions sets, valuations, cost functions, and impact functions are symmetric across agents.

**Proposition 3** *Consider a mixture contest with $n \geq 2$ symmetric agents. Then, there exists a unique symmetric equilibrium.*

For completeness, Appendix B contains an example in which there are asymmetric equilibria as well. In that example, actions are at the corners of the action set. The contest organizer may prefer the asymmetric equilibria.

This section and the next two examine the comparative statics of the unique symmetric equilibrium with respect to changes in a multivariate type distribution. The main focus is on the expected equilibrium impact, $\overline{p}^*$. By (5), this satisfies

$$\overline{p}^* = \mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right]. \tag{6}$$

Thus, comparative statics are determined by how the right hand side is impacted by changes in the environment. The central argument is simple. First, $p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ depends only on $\boldsymbol{\theta}_i$ and $\overline{p}^*$. Then, imagine that the type distribution changes in such a way that the right hand side of (6) increases for all $\overline{p}^*$. Since actions are strategic substitutes, a decrease in $\overline{p}^*$ further increases the right hand side, but the left hand side decreases at the same time. This contradiction implies that $\overline{p}^*$ must increase instead. Note that strategic substitutability plays an important role in this argument.

[5]It is worth pointing out a special feature of contests with homogeneous technologies, $n = 3$ agents, and complete information. Such contests are linear aggregative games with aggregator $\sum_{j=1}^{3} \overline{p}_j$. If agent $i$ knows $\overline{p}_i$ and $\sum_{j=1}^{3} \overline{p}_j$, then $\sum_{j \neq i} \overline{p}_j$ can be calculated, which in turn means that $k_i\left(\overline{p}_{-i}\right)$ can be inferred. Thus, everything that is payoff-relevant to agent $i$ is captured in $\overline{p}_i$ and $\sum_{j=1}^{3} \overline{p}_j$. See Jensen (2018) for a survey on aggregative games. Acemoglu and Jensen (2013) discuss comparative statics in aggregative games with strategic substitutes.

Another way to see this is to observe that strategic substitutability implies that the right hand side of (6) is monotonically decreasing in $\overline{p}^*$ (which is also what proves Proposition 3). It is also worth reiterating the implication that the equilibrium action of any given type moves in the opposite direction to the equilibrium expected impact and performance.

Only contests in which types can be ordered in a natural and easily interpretable manner are considered. To this end, assume that types can be ordered in such a way that $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly increasing in $\boldsymbol{\theta}_i$, regardless of $\overline{\mathbf{p}}_{-i}$. In other words, higher types have best responses that lead to higher impacts and therefore a stronger distribution of performance. Examples and applications are in Sections 5 and 6.

To keep notation at a minimum, assume that the type-space is a hyper-rectangle (when continuous) or a grid (when discrete) that is held fixed when the distribution of types changes. If types are $d$-dimensional, write $\boldsymbol{\theta}_i = (\theta_i^1, \theta_i^2, ..., \theta_i^d)$. Let $F_\theta^j(\theta_i^j)$ denote the marginal distribution of $\theta_i^j$ and let $F_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ denote the joint distribution, i.e. the probability that the type is component-wise smaller than $\boldsymbol{\theta}_i$. Let $\overline{F}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ denote the survival function, which measures the probability that the type is component-wise larger than $\boldsymbol{\theta}_i$. It is sometimes convenient to think of the joint distribution as being a combination of the marginal distributions and a dependence structure. By Sklar's lemma, there exists a function, known as the copula, such that $F_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) = C(F_\theta^1(\theta_i^1), F_\theta^2(\theta_i^2), ..., F_\theta^d(\theta_i^d))$. Thus, the copula captures the dependence structure. See Joe (1997), Müller and Stoyan (2002), and Shaked and Shanthikumar (2007) for more on copulas and multivariate stochastic orders.

## 4.1 Stochastic improvements and stronger competition

To begin, consider a "stochastic improvement" of the type distribution.[6] In the familiar univariate case ($d = 1$) it is well known that there are three equivalent ways to define first-order stochastic dominance. Specifically, distribution $G_{\boldsymbol{\theta}}$ dominates distribution $F_{\boldsymbol{\theta}}$ in terms of first-order stochastic dominance if $(i)$ $G_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) \leq F_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ for all $\boldsymbol{\theta}_i$, $(ii)$ $\overline{G}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) \geq \overline{F}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ for all $\boldsymbol{\theta}_i$, or $(iii)$ any expected-utility maximizer with a non-decreasing payoff function weakly prefers $G_{\boldsymbol{\theta}}$ to $F_{\boldsymbol{\theta}}$.

The equivalence between $(i)$ and $(ii)$ is due to the fact that $\{\boldsymbol{\theta}_i' \in \Theta_i | \boldsymbol{\theta}_i' > \boldsymbol{\theta}_i\}$ is the complement to $\{\boldsymbol{\theta}_i' \in \Theta_i | \boldsymbol{\theta}_i' \leq \boldsymbol{\theta}_i\}$ when $\boldsymbol{\theta}_i$ is one-dimensional. However, this is no

[6] The discussion in the next three paragraphs follows Müller and Stoyan (2002).

longer true in higher dimensions. Then, the first set describes an "upper orthant" ($\boldsymbol{\theta}'_i$ is large along all dimensions) and the second set a "lower orthant" ($\boldsymbol{\theta}'_i$ is small along all dimensions), but neither accounts for the possibility that $\boldsymbol{\theta}'_i$ is large along some dimensions and small along others. Thus, it is not the case that $F_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) = 1 - \overline{F}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ unless $d = 1$. In other words, $(i)$ and $(ii)$ are distinct and can each be used to pursue a different multivariate extension of first-order stochastic dominance. These stochastic orders are know as the lower orthant order and the upper orthant order, respectively. When they apply, $G_{\boldsymbol{\theta}}$ is preferred to $F_{\boldsymbol{\theta}}$ for a *subset* of non-decreasing payoff functions. For instance, when $(i)$ applies, $G_{\boldsymbol{\theta}}$ is preferred to $F_{\boldsymbol{\theta}}$ for all payoff functions that are indicator functions of the form $-1_{\{\boldsymbol{\theta}'_i \in \Theta_i | \boldsymbol{\theta}'_i \leq \boldsymbol{\theta}_i\}}$ (punishing every realization in a lower orthant), and when $(ii)$ applies for all indicator functions of the form $1_{\{\boldsymbol{\theta}'_i \in \Theta_i | \boldsymbol{\theta}'_i \geq \boldsymbol{\theta}_i\}}$ (rewarding every realization in an upper orthant).

A third extension to multivariate distributions is based on $(iii)$. Distribution $G_{\boldsymbol{\theta}}$ dominates distribution $F_{\boldsymbol{\theta}}$ in the "usual stochastic order" if the expectation of *any* payoff function that is non-decreasing in $\boldsymbol{\theta}_i$ is higher under $G_{\boldsymbol{\theta}}$ than $F_{\boldsymbol{\theta}}$. Thus, the usual stochastic order is stronger than the orthant orders. It can be shown that $G_{\boldsymbol{\theta}}$ dominates distribution $F_{\boldsymbol{\theta}}$ in the usual stochastic order if and only if it is the case that for any increasing set, $G_{\boldsymbol{\theta}}$ yields a weakly higher probability of a realization somewhere in that set compared to $F_{\boldsymbol{\theta}}$. A set $S$ is increasing if $\mathbf{x} \in S$ implies $\mathbf{x}' \in S$ for all $\mathbf{x}'$ that are componentwise greater than $\mathbf{x}$. Intuitively, a change from $F_{\boldsymbol{\theta}}$ to $G_{\boldsymbol{\theta}}$ means that any set of high realizations become more likely, which in turn increases the expectation of any function that is non-decreasing.

Note that if $G_{\boldsymbol{\theta}}$ dominates $F_{\boldsymbol{\theta}}$ in the usual stochastic order, then all the $d$ marginal distributions of $G_{\boldsymbol{\theta}}$ first-order stochastically dominates the corresponding marginal distributions of $F_{\boldsymbol{\theta}}$. The converse is not true in general but Scarsini (1988) showed that if $G_{\boldsymbol{\theta}}$ and $F_{\boldsymbol{\theta}}$ have the same copula, then $G_{\boldsymbol{\theta}}$ dominates $F_{\boldsymbol{\theta}}$ in the usual stochastic order if and only if the marginal distributions under $G_{\boldsymbol{\theta}}$ first-order stochastically dominates their counterparts under $F_{\boldsymbol{\theta}}$.

Since types are ordered such that $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly increasing in $\boldsymbol{\theta}_i$, an improvement of the type distribution in the usual stochastic order leads the expected impact to increase, other things equal. By the argument following (6), the equilibrium value of $\overline{p}^*$ therefore increases. Since actions are strategic substitutes, all types lower their action in response. This balances out in equilibrium because although each type works less hard, higher types with higher impacts are now more likely.

**Proposition 4** *If the common type distribution improves in the sense of the usual stochastic order, then each type's equilibrium action is weakly lower, yet the expected equilibrium impact is weakly higher because higher types are more likely.*

Proposition 4 implies that the ex ante equilibrium distribution of individual performance improves in the sense of first-order stochastic dominance. Since the performance of agents are independently distributed, it then follows from Scarsini's (1988) result that the joint distribution of the performance of all agents improves in the sense of the usual stochastic order. Thus, as long as the contest organizers' Bernoulli utility function is increasing in the performance profile, she is made better off.

Note that the type ordering does not imply that $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is increasing in $\boldsymbol{\theta}_i$. In part for this reason, it cannot be ruled out that the expected action decreases even though the expected equilibrium impact increases. The application in Section 6 demonstrates this possibility. To reiterate, the mixture model is better geared to study aggregate performance than aggregate effort in generality.

In applications, more detailed information about the properties of $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ may be known. Those properties may potentially be used to relax the concept of stochastic dominance in Proposition 4. The upcoming applications focus on bivariate contests, in which case the orthant orders are particularly tractable and relevant. Indeed, in the bivariate case, it can be shown that $G_{\boldsymbol{\theta}}$ dominates $F_{\boldsymbol{\theta}}$ in the lower orthant order ($G_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) \leq F_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ for all $\boldsymbol{\theta}_i$) if and only if $G_{\boldsymbol{\theta}}$ is weakly preferred to $F_{\boldsymbol{\theta}}$ for any weakly increasing and submodular payoff function. Similarly, $G_{\boldsymbol{\theta}}$ dominates $F_{\boldsymbol{\theta}}$ in the upper orthant order ($\overline{G}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) \geq \overline{F}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ for all $\boldsymbol{\theta}_i$) if and only if $G_{\boldsymbol{\theta}}$ is weakly preferred to $F_{\boldsymbol{\theta}}$ for any weakly increasing and supermodular payoff function. Supermodular and submodular functions are discussed in detail in the next subsection. The joint distribution in Table 1(b) dominates the one in Table 1(a) in both the upper and lower orthant order, but not in the usual stochastic order even though each marginal distribution has improved in the sense of first order stochastic dominance.[7]

**Proposition 5** *Assume that types are bivariate and that the common type distribution becomes stronger in the upper (lower) orthant order. If $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is supermodular (submodular) in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$, then each type's equilibrium action is weakly lower but the expected equilibrium impact is weakly higher.*

[7] The increasing set $S = \{(M, H), (H, M), (H, H)\}$ is less likely in (b) than in (a). Hence, the expected value of the weakly increasing payoff function that gives payoff 1 from an event in $S$ and zero otherwise is lower in (b) than in (a). This function is neither supermodular or submodular.

| | L | M | H |
|---|---|---|---|
| $H$ | 0.10 | 0.11 | 0.12 |
| $M$ | 0.11 | 0.12 | 0.11 |
| $L$ | 0.12 | 0.11 | 0.10 |

(a) A joint distribution

| | L | M | H |
|---|---|---|---|
| $H$ | 0.12 | 0.05 | 0.22 |
| $M$ | 0.13 | 0.16 | 0.05 |
| $L$ | 0.02 | 0.13 | 0.12 |

(b) "Improved" orthants

| | L | M | H |
|---|---|---|---|
| $H$ | 0.10 | 0.10 | 0.13 |
| $M$ | 0.09 | 0.15 | 0.10 |
| $L$ | 0.14 | 0.09 | 0.10 |

(c) Supermodular order

**Table 1:** Bivariate distributions with low (L), medium (M), and high (H) outcomes.

## 4.2 Dependence orders and supermodularity

The previous subsection considered the consequences of increased competition stemming from certain sets of high types becoming more likely. This subsection turns to the interaction and dependence between different dimensions of the type space. To isolate the role of the dependence structure, the marginal distributions are held fixed.

Motivated by the upcoming applications, it is particularly important to understand contests in which $p_i(B_i(\overline{p}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is either supermodular or submodular. Thus, it is worthwhile to pursue a stochastic order known as the supermodular order. By definition, $G_{\boldsymbol{\theta}}$ dominates or is greater than $F_{\boldsymbol{\theta}}$ in the supermodular order if and only if the expected value of any supermodular function is weakly higher under $G_{\boldsymbol{\theta}}$ than $F_{\boldsymbol{\theta}}$. Note that the function is not required to be monotonic. It is an implication that $G_{\boldsymbol{\theta}}$ and $F_{\boldsymbol{\theta}}$ share the same marginal distributions. Hence, it is only their copula or dependence structure that differ.

To help visualize how $G_{\boldsymbol{\theta}}$ and $F_{\boldsymbol{\theta}}$ compare under the supermodular order, consider the bivariate case, or $d = 2$. Then, a function $u(x, y)$ is supermodular if

$$u(x', y') + u(x, y) \geq u(x', y) + u(x, y')$$

whenever $x' > x$ and $y' > y$. Thus, the function is on average higher if both $x$ and $y$ are small or large at the same time compared to when one is large and the other is small. Now, for the case of discrete supports, Epstein and Tanny (1980) consider an aptly named "correlation increasing transformation" that removes an equal amount of mass from $(x', y)$ and $(x, y')$ and moves it to $(x', y')$ and $(x, y)$ instead. This transformation quite clearly increases the expected value of $u$, meaning that the newly created distribution dominates the original distribution in the supermodular order. Indeed, the transformation does not change the marginal distributions of $x$ and

$y$, but it does increase their correlation. The distribution in Table 1(c) is obtained through two correlation increasing transformations of the one in Table 1(a).[8] Meyer and Strulovici (2015) elegantly extend the characterization of the supermodular order to $d \geq 2$ dimensions through a sequence of similar two-dimensional transformations.

Continuing with the bivariate case, allow the supports to be continuous. In the spirit of correlation increasing transformations, it can be shown that $G_{\boldsymbol{\theta}}$ is greater than $F_{\boldsymbol{\theta}}$ in the supermodular order if and only if $G_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) \geq F_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ and $\overline{G}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i) \geq \overline{F}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_i)$ for all $\boldsymbol{\theta}_i$. In words, $G_{\boldsymbol{\theta}}$ assigns more mass to all the lower and upper orthants than $F_{\boldsymbol{\theta}}$ does. For example, bivariate normal distributions with fixed means and variances become greater in the supermodular order if the covariance increases.

Interestingly, the supermodular order is the *unique*, appropriately defined, dependence order in the bivariate case.[9] That is, it is the only stochastic order that satisfies a set of axioms that are reasonable for any notion of stochastic dependence.

A counterpart to the results in the previous subsection is now immediate.

**Proposition 6** *Assume that the common type distribution becomes greater in the supermodular order. If $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is supermodular (submodular) in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$, then each type's equilibrium action is weakly lower (higher) but the expected equilibrium impact is weakly higher (lower).*

Supermodularity of $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ in $\boldsymbol{\theta}_i$ does not generally imply that $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is supermodular in $\boldsymbol{\theta}_i$, or vice versa. However, in most of the following applications, the two functions are supermodular or submodular at the same time.

# 5 Valuations and action sets

This section studies symmetric contests with private information about valuations and either budget constraints or starting advantages. The orthant orders and the supermodular order is particularly well-suited to study such contests.

[8] Probability 0.01 is added to $(M, M)$ and $(H, H)$ and taken away from $(M, H)$ and $(H, M)$. At the same time, probability 0.02 is added to $(L, L)$ and $(M, M)$ and removed from $(L, M)$ and $(M, L)$. In general, however, the events to which mass is added do not have to be on the diagonal. Note that the marginal distributions are the same in Table 1(a) and 1(c).

[9] See the discussion leading up to Theorem 3.8.3 in Müller and Stoyan (2002). There are other dependence orders, such as the concordance order, but they coincide when $d = 2$. Meyer and Strulovici (2012) discuss applications of positive dependence in economics and explore the links between different notions of greater dependence when $d > 2$.

## 5.1 The bivariate valuation/budget contest

Start by assuming that types are bivariate and that $v_i(\boldsymbol{\theta}_i) = \theta_i^1$ and $A_i(\boldsymbol{\theta}_i) = [0, \theta_i^2]$, with $\theta_i^1, \theta_i^2 \geq 0$. Here, $\theta_i^2$ represents a privately known budget constraint or some other resource constraint like a time constraint. Let $\overline{\theta}_i^2$ denote the highest possible budget. Assume that the impact and cost functions are type-independent, and write these as $p(a_i)$ and $c(a_i)$, respectively. Any such environment is from now on referred to as a *bivariate valuation/budget contest*.

The budget constraint may prevent the agent from taking the action that he would ideally want. Thus, best responses take the form

$$B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i) = \min\left\{\arg\max_{a_i \in [0, \overline{\theta}_i^2]} \theta_i^1 p(a_i) k_i\left(\overline{\mathbf{p}}_{-i}\right) - c(a_i), \theta_i^2\right\},$$

and the resulting impacts are

$$p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i) = \min\left\{p\left(\arg\max_{a_i \in [0, \overline{\theta}_i^2]} \theta_i^1 p(a_i) k_i\left(\overline{\mathbf{p}}_{-i}\right) - c(a_i)\right), p(\theta_i^2)\right\}. \quad (7)$$

The two arguments in the min function are increasing in one dimension of $\boldsymbol{\theta}_i$ and independent of the other. Thus, since the min function is supermodular, $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ and $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ are not only weakly increasing in $\boldsymbol{\theta}_i$ but also supermodular in $\boldsymbol{\theta}_i$. Hence, Propositions 4–6 can be applied.

To understand why $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ and $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ are supermodular, note that there is little good in having a high valuation and a low budget, or vice versa. In either case, one of the two dimensions drags down the action and impact. In other words, the agent needs to have both a high budget *and* a high valuation in order to be able and willing to achieve a high impact. Thus, what matters are the upper orthants. Recall that realizations in the upper orthants become more likely if the distribution improves in the upper orthant order or becomes greater in the supermodular order.

**Corollary 1** *Consider any bivariate valuation/budget contest. If the type distribution either improves in the upper orthant order or becomes greater in the supermodular order, then each type's equilibrium action is weakly lower and the budget constraint binds for fewer types. However, the expected equilibrium impact is weakly higher.*

Note that there is no change in the expected valuation and the expected budget if the type distribution becomes greater in the supermodular order (as in Proposition 6).

The reason is again that the marginal distributions are unaffected. This is another way of demonstrating that the comparative statics are driven by the dependence structure. If the distribution improves in the upper orthant order (Proposition 5), then there is an added effect from the fact that marginal distributions improve in a first-order stochastic sense, which increases the expected valuation and the expected budget and spurs a further increase in the expected equilibrium impact.

The upper orthant order and the supermodular order can potentially be used to study budget constraints even beyond the mixture model of contests. Consider a second-price auction with bivariate types that reflect valuations ($\theta_i^1$) and budgets ($\theta_i^2$). In the dominant strategy equilibrium, the bidding strategy is $\min\{\theta_i^1, \theta_i^2\}$ independently of the distribution. When the type distribution improves in the upper orthant order or becomes greater in the supermodular order, the probability of high bids therefore increases. Consequently, the distribution of bids improves in the sense of first-order stochastic dominance and the expected revenue increases.

In the second-price auction, beliefs do not impact the equilibrium strategy and the comparative statics are therefore particularly straightforward. In the mixture contest, the simple form of (6) similarly leads to easy comparative statics.

## 5.2 The bivariate valuation/advantage contest

Assume next that $v_i(\boldsymbol{\theta}_i) = \theta_i^1$ and $A_i(\boldsymbol{\theta}_i) = [\theta_i^2, \overline{a}]$, with $\theta_i^1, \theta_i^2 \geq 0$ and $\overline{a} > \max \theta_i^2$. Here, $\theta_i^2$ captures a privately known starting advantage. As before, the impact function $p(a_i)$ and cost function $c(a_i)$ are type-independent. Such environments are termed *bivariate valuation/advantage contests*. Comparing the action sets in bivariate valuation/budget contests with those in bivariate valuation/advantage contests, it is clear that these kinds of contests are in some sense the opposite of each other.

The agent's problem is

$$\max_{a_i \in A_i(\theta_i)} \theta_i^1 p(a_i) k_i(\overline{\mathbf{p}}_{-i}) - \left(c(a_i) - c(\theta_i^2)\right).$$

Note that the agent incurs a cost only if his effort is above his starting advantage. The starting advantage puts a lower bound on the agent's action. Hence,

$$\begin{aligned} B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i) &= \max\left\{\arg\max\nolimits_{a_i \in [0,\overline{a}]} \theta_i^1 p(a_i) k_i\left(\overline{\mathbf{p}}_{-i}\right) - c(a_i), \theta_i^2\right\} \\ p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i) &= \max\left\{p\left(\arg\max\nolimits_{a_i \in [0,\overline{a}]} \theta_i^1 p(a_i) k_i\left(\overline{\mathbf{p}}_{-i}\right) - c(a_i)\right), p\left(\theta_i^2\right)\right\} \quad (8) \end{aligned}$$

both of which are increasing in $\boldsymbol{\theta}_i$. Since the max function is submodular, $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ and $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ are submodular in $\boldsymbol{\theta}_i$. Proposition 4–6 can be applied.

To understand why valuations and starting advantages are substitutes, note that a high action can be justified if the valuation is high and that it is automatic if the starting advantage happens to be large. Thus, it is enough that one of the two characteristics is large. Conversely, it is only when both characteristics are small that actions are small, which suggests that mass in the lower orthants dampens competition.

**Corollary 2** *Consider any bivariate valuation/advantage contest. If the type distribution improves in the lower orthant order then each type's equilibrium action is weakly lower, meaning that more types are dissuaded from exerting effort above their starting advantage. However, the equilibrium impact is weakly higher. The equilibrium effects are the opposite if the type distribution becomes greater in the supermodular order.*

# 6 Valuations and impact functions

This section shifts the focus onto the impact function. The first application gives clean comparative statics. The predictions of the second application depends more critically on the primitives. These applications also illustrate the importance of how types are ordered: Actions may be decreasing along one dimension of the type space, but impacts never are.

## 6.1 The bivariate productivity contest

Consider a symmetric bivariate contest where types capture additive and multiplicative productivity types, respectively. Thus, $p_i(a_i|\boldsymbol{\theta}_i)$ can with some abuse of notation be written as $p\left(\theta_i^1 + a_i\theta_i^2\right)$. Valuations, $v$, are common knowledge and the cost of effort $a_i$ is $a_i$. Such contests are termed *bivariate productivity contests.*

Assume that best responses are always interior. Agent $i$'s first-order condition is

$$vp'\left(\theta_i^1 + a_i\theta_i^2\right)\theta_i^2 k_i(\overline{\mathbf{p}}_{-i}) = 1.$$

Note that the optimal value of $\theta_i^1 + a_i\theta_i^2$ is independent of $\theta_i^1$. An increase in the additive type $\theta_i^1$ is a bonus to the agent which allows him to proportionally downgrade his action while maintaining the same marginal product of $a_i$. Thus, $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$

depends only on, and is increasing in, $\theta_i^2$. It follows that the expected equilibrium impact is determined entirely by the marginal distribution of $\theta_i^2$. The dependence structure is irrelevant in this regard.

On the other hand, $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is supermodular in $\boldsymbol{\theta}_i$. As mentioned, an increase in $\theta_i^1$ leads the agent to downgrade his action. However, the incentive to downgrade is smaller the higher $\theta_i^2$ is. The reason it that the impact is more sensitive to such a downgrade the higher $\theta_i^2$ is. This explains why the two characteristics are complements in the best response function. Consequently, the expected equilibrium action is sensitive to the dependence structure. The next corollary summarizes.

**Corollary 3** *Consider a bivariate productivity contest in which best responses are always interior. If the type distribution becomes greater in the supermodular order, then the expected equilibrium impact does not change. However, the expected equilibrium action and the expected cost of effort increases.*

## 6.2 The bivariate valuation/impact contest

Consider a *bivariate valuation/impact contest* in which valuations and impact functions are private information. Types are bivariate, with valuations $v_i(\boldsymbol{\theta}_i) = \theta_i^1 \geq 0$ and impact functions $p_i(a_i|\boldsymbol{\theta}_i)$ that do not depend on $\theta_i^1$ and can be written $p(a_i|\theta_i^2)$. Assume that $p(a_i|\theta_i^2)$ is strictly increasing in each argument, or $p_a, p_{\theta_i^2} > 0$ where the subscripts on $p$ now refer to derivatives. The impact function is strictly concave in $a$ and the cost function is linear in $a$ and independent of $\boldsymbol{\theta}_i$.

While $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly increasing in $\theta_i^1$, stronger assumptions are required to ensure monotonicity in $\theta_i^2$. As $\theta_i^2$ increases, the action can be lowered to keep the impact $p(a_i|\theta_i^2)$ constant. However, incentives are determined by $p_a(a_i|\theta_i^2)$. Holding $p(a_i|\theta_i^2)$ constant, does $p_a(a_i|\theta_i^2)$ increase in $\theta_i^2$? If so, an incentive is created to increase $a_i$ and thereby the impact $p(a_i|\theta_i^2)$. This occurs if $p_{a\theta_i^2} \geq 0$ since the increase in $\theta_i^2$ increases $p_a$, which in turns spurs an increase in $a_i$. Of course, this is a standard monotone comparative statics result in the vein of Milgrom and Shannon (1994). The increase in $\theta_i^2$ and $a_i$ then pulls in the same direction, to increase $p(a_i|\theta_i^2)$.

However, when $p_{a\theta_i^2} < 0$ the action decreases in $\theta_i^2$, but as long as it decreases slowly enough compared to $\theta_i^2$, the impact nevertheless increases. Thus, it can be shown that

$$\frac{\partial}{\partial a}\left(\frac{p_{\theta_i^2}}{p_a}\right) \geq 0 \tag{9}$$

is sufficient for $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ to be weakly increasing in $\theta_i^2$ (a formal proof is in the upcoming Corollary 5). This condition therefore implies that the type-space is ordered in the desired manner, and Proposition 4 applies to all such environments.

Taking a step back, as noted above the best response is decreasing in $\theta_i^2$ when actions and types are substitutes in the impact function, or $p_{a\theta_i^2} < 0$. Consequently, it is possible that the equilibrium expected performance and the equilibrium expected action move in opposite directions when the distribution of types improves in the usual stochastic order. For completeness, the next result describes a case in which this occurs. Appendix B contains a fully solved example.

**Corollary 4** *Consider a bivariate valuation/impact contest that satisfies (9) and $p_{a\theta_i^2} < 0$. Assume that $\theta_i^1$ and $\theta_i^2$ are independently distributed. Holding fixed the marginal distribution of $\theta_i^1$, the expected equilibrium impact increases but the expected equilibrium action decreases if the marginal distribution of $\theta_i^2$ improves in the sense of first order stochastic dominance.*

Another way of stating the result is that the expected equilibrium performance may increase while the expected equilibrium costs to agents decrease.[10]

Turning next to the importance of the dependence structure, assume that $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is always interior. Then, the agent's first-order condition implies that

$$\frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial \theta_i^2} = -\frac{p_{a\theta_i^2}}{p_{aa}},$$

which as mentioned can be positive or negative. Since $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is increasing in $\theta_i^1$, it follows that whether $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is supermodular or submodular depends on whether $-\frac{p_{a\theta_i^2}}{p_{aa}}$ increases or decreases in $a_i$. Finally, it is not hard to show that $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ inherits the modularity properties of $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ in this setting.

**Corollary 5** *The impact $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly increasing in $\boldsymbol{\theta}_i$ in any bivariate valuation/impact contest that satisfies (9). Moreover, if $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is always interior then $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ and $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ are submodular (supermodular) in $\boldsymbol{\theta}_i$ if*

$$\frac{\partial}{\partial a}\left(-\frac{p_{a\theta_i^2}}{p_{aa}}\right) < (>)0. \tag{10}$$

[10] If the impact function is normalized to be linear and the cost function is $c(a|\theta_i^2)$ then the best response is weakly increasing in $\theta_i^2$ if and only if marginal costs are weakly decreasing in type, or $c_{a\theta_i^2} \leq 0$. However, the costs of the best response are decreasing in $\theta_i^2$ if $c_{\theta_i^2} - c_a c_{a\theta_i^2}/c_{aa} < 0$.

The corollary describes conditions under which Propositions 5 and 6 apply. It is instructive to consider a special case. Imagine that $\theta_i^2$ is a multiplicative productivity type in the sense that $p(a_i|\theta_i^2)$ takes the form $z(a_i\theta_i^2)$, with $a_i, \theta_i^2 \geq 0$ and $z'(x) > 0 > z''(x)$. Any such function satisfies (9) and a tight characterization of (10) is possible. The derivative in (10) is proportional to $z'(x)z'''(x) - 2\left(z''(x)\right)^2$. Hence, $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is submodular in $\boldsymbol{\theta}_i$ as long as $z'(x)$ is not "too convex" or, stated differently, when $z'''(x)$ is not too large. With this observation in mind, consider the transformation $-\left(z'(x)\right)^{-1}$, which is a concave transformation of $z'(x)$ and is related to a notion of "$\rho$-concavity" that is explored in more depth in the next section. As long as the transformed function is strictly concave in $x$, $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is submodular. This is a weaker condition than log-concavity of $z'(x)$ and is satisfied in the fully solved example in Appendix B. By Proposition 6, the expected impact then decreases if the joint distribution becomes greater in the supermodular order.

A delineating case arises if $z(x) = \ln(\alpha + \beta x)$, with parameter restrictions that imply $\alpha, \beta > 0$ and $\alpha + \beta x \in [1, e]$. Even though $z'(x)$ is convex, $-\left(z'(x)\right)^{-1}$ is linear. Indeed, $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i) = \theta_i^1 k_i(\overline{\mathbf{p}}_{-i}) - \alpha\left(\beta\theta_i^2\right)^{-1}$ and $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i) = \ln\theta_i^1 + \ln\theta_i^2 + \ln\beta k_i(\overline{\mathbf{p}}_{-i})$ are additively separable in $\theta_i^1$ and $\theta_i^2$. Hence, the dependence structure between $\theta_i^1$ and $\theta_i^2$ is irrelevant for the expected performance and the expected action.

It is equally possible to construct examples in which $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is supermodular. This is the case if $z(x) = \frac{1}{16}\left(18x - 3x^2 + x^3\right)$, $x \in [0,1]$. There are admittedly simpler examples, such as when $z(x) = x^\alpha$, $x \in [0,1]$ and $\alpha \in (0,1)$. However, this example hardly captures multivariate incomplete information in an interesting way. The reason is that $z(x)$ is a homogeneous function, and the agent's problem is then to maximize $\theta_i^1\left(\theta_i^2\right)^\alpha z(a_i)k_i(\overline{\mathbf{p}}_{-i}) - a_i$. This problem depends on the single number $\theta_i^1\left(\theta_i^2\right)^\alpha$, which can be thought of as an adjusted valuation. Note that $\theta_i^1\left(\theta_i^2\right)^\alpha$ is supermodular, so when the joint distribution of $\boldsymbol{\theta}_i$ becomes greater in the supermodular order, the expected value of the adjusted valuation increases and this alone helps explain why the equilibrium impact increases.

# 7 Valuations and (a)symmetric mixture contests

Consider a symmetric contest with univariate types (written $\theta_i$ rather than $\boldsymbol{\theta}_i$) and imagine that the "amount" of uncertainty increases in the sense that the type distribution undergoes a mean-preserving spread. If $p_i(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)|\theta_i)$ is concave (convex) in $\theta_i$, then the mean-preserving spread causes the expected impact to decrease (increase), other things equal. This effect pushes down (up) the expected equilibrium impact, which then causes actions to move in the opposite direction.

**Proposition 7** *Consider a symmetric contest with univariate types. Assume that the common type distribution undergoes a mean-preserving spread. If $p_i(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)|\theta_i)$ is concave (convex) in $\theta_i$ for every $\overline{\mathbf{p}}_{-i}$ then each type's equilibrium action is weakly higher (lower), yet the expected equilibrium impact is weakly lower (higher).*

To illustrate, assume that budgets are private information but that everything else is known. Then, $p_i(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)|\theta_i)$ in (7) is concave in the budget. Consequently, a mean-preserving spread of the budget causes each type's equilibrium action to weakly increase, meaning that the budget constraint binds for more types. However, the expected impact is weakly lower.

The remainder of the section zeroes in on the role of valuations in symmetric and asymmetric contests.

## 7.1 Symmetric contests with private valuations

Assume that agents are symmetric and that types are univariate, with private information only about the valuation, and $v_i(\theta_i) = \theta_i$. Action sets, impact functions, and cost functions are type-independent. Assume that the impact function is thrice differentiable and write it as $p(a_i)$. Assume for simplicity that the cost function is $c(a_i) = a_i$. Such contests are termed *private values contests*.

Although a mean-preserving spread does not change the expected valuation, Proposition 7 implies that the expected performance may increase or decrease, depending on the curvature of $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$. Since $p(\cdot)$ is a concave transformation, $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$ is concave in $\theta_i$ if $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ is concave in $\theta_i$. On the other hand, it is also possible that $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ is so convex that $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$ is convex in $\theta_i$. Either case can arise, depending on the curvature of the marginal impact function, $p'(a)$.

The relevant measure of concavity is $\rho$-concavity. Recall that the positive function $p'(a)$ is $\rho$-concave if the transformation $T^\rho(a) = \frac{1}{\rho}(p'(a))^\rho$ is concave in $a$, $\rho \neq 0$ (and log-concave if $\rho = 0$). This is satisfied if and only if $p'(a)p'''(a) - (1-\rho)(p''(a))^2 \leq 0$. Note that higher values of $\rho$ are more stringent, with e.g. $\rho = 0$ coinciding with log-concavity and $\rho = 1$ with regular concavity. For the present problem, it is $\rho$-concavity of order $\rho = -1$ and $\rho = -2$ that are important. See Ewerhart (2013) for an introduction to $\rho$-concavity and some applications.

**Lemma 1** *Consider any private values contest in which best responses are always interior. Then, $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ is concave (convex) in $\theta_i$ if $T^{-1}(a)$ is concave (convex) and $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$ is concave (convex) in $\theta_i$ if $T^{-2}(a)$ is concave (convex). Hence, $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ is convex in $\theta_i$ but $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$ is concave in $\theta_i$ whenever $T^{-1}(a)$ is convex and $T^{-2}(a)$ is concave.*

Thus, Proposition 7 applies if $T^{-2}(a)$ is concave or convex. As demonstrated shortly in an example, either case is possible. Hence, the expected impact may decrease or increase when the type distribution undergoes a mean-preserving spread.

The last part of Lemma 1 identifies conditions under which $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ and $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$ have opposing curvatures. This suggests that the expected impact and the expected action can move in opposite direction when valuations become more spread out. However, there is also a level effect to account for. As it turns out, the level effect sometimes works in the same direction: If $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$ is concave in $\theta_i$, then $\overline{p}^*$ decreases following the mean-preserving spread, which then causes $B_i(\overline{p}^*, ..., \overline{p}^*|\theta_i)$ to increase for all $\theta_i$. Hence, if the conditions in the last part of the lemma are satisfied then $(i)$ convexity of $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ in $\theta_i$ means that the expected action increases for any given $\overline{\mathbf{p}}_{-i}$ and $(ii)$ the equilibrium level of $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ increases for any $\theta_i$ because $\overline{p}^*$ decreases. The two effects reinforce each other.[11] Thus, the expected equilibrium action and the expected equilibrium impact move in opposite directions. In other words, agents expend more resources in expectations, yet are less productive.

**Corollary 6** *Consider any private values contest in which best responses are always interior and where $T^{-1}(a)$ is convex and $T^{-2}(a)$ is concave. Then, the expected equilibrium impact weakly decreases but the expected action weakly increases if the common type distribution undergoes a mean-preserving spread.*

Thus, it should not be taken for granted that total expected performance moves in the same direction as total effort when the competitive environment changes. Of course, this insight is not necessarily unique to mixture contests. In the standard

[11] The proof of Corollary 4 is in the same vein and also relies on two reinforcing effects.

formulation of rank-order tournaments a la Lazear and Rosen (1981), the agent's action, impact, and expected productivity all coincide. However, non-linear cost functions open the door for the possibility that higher total output may or may not come at higher total costs. Relatedly, Fang et al (2020) consider a symmetric, complete information, all-pay auction with convex costs but several prizes. Making the prizes more unequal does not change expected costs, but it does reduce expected total effort.

Example 1: Consider a symmetric private values contest with $n \geq 3$ agents and $\Theta_i \subseteq [0,1]$, $p(a_i) = a_i^r$, $r \in (0,1)$, and action set $A = [0,1]$ for all $\theta_i \in \Theta_i$. Note that $p'(a) = ra^{r-1}$ and $-(p'(a))^{-1}$ are convex functions but that $\frac{-1}{2}(p'(a))^{-2}$ is concave if $r \leq \frac{1}{2}$. It is readily verified that $B_i(\overline{\mathbf{p}}_{-i}|\theta_i) = \left(\theta_i r k_i\left(\overline{\mathbf{p}}_{-i}\right)\right)^{\frac{1}{1-r}}$. Although the best response is convex in $\theta_i$, the resulting impact

$$p_i(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)|\theta_i) = \left(\theta_i r k_i\left(\overline{\mathbf{p}}_{-i}\right)\right)^{\frac{r}{1-r}} \tag{11}$$

is convex in $\theta_i$ when $r > \frac{1}{2}$ but concave when $r < \frac{1}{2}$. In either case, from (11),

$$\overline{p}^* = \left(r k_i\left(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*\right)\right)^{\frac{r}{1-r}} \mathbb{E}_{\theta_i}\left[\theta_i^{\frac{r}{1-r}}\right]. \tag{12}$$

The solution depends on the type distribution only through the last term.

Assume first that $r > \frac{1}{2}$. Then, $\overline{p}^*$ increases if the type distribution undergoes a mean-preserving spread, since the last term in (12) increases. The case in which $r < \frac{1}{2}$ is more interesting. Here, the argument is reversed and $\overline{p}^*$ decreases. By the argument leading to Corollary 6, the expected equilibrium action must increase. In summary, when $r < \frac{1}{2}$, the expected action increases, but the expected impact and the expected performance decreases in response to the mean-preserving spread. ▲

## 7.2 Asymmetric contests and the discouragement effect

This subsection allows agents to be heterogeneous ex ante but leans on the examples and insights in the previous subsection. In particular, the effects of making a common type distribution more spread out in symmetric contests is echoed when agents are made more asymmetric ex ante. To facilitate comparison with the existing literature, assume that information is complete.

The contest literature has long asked whether increased heterogeneity discourages competition. Drugov and Ryvkin (2022) note that the answer depends on how heterogeneity is defined. If the asymmetry takes the form of different valuations then the literature on Tullock contests suggests a discouragement effect. In particular, Konrad (2009) considers a two-player Tullock contest with known valuations $v_1 = v - z$ and $v_2 = v + z$, respectively, where $v > 0$ and $z \in [-v, v]$. The impact function takes the form $a_i^r$. In his setting, total effort decreases in $|z|$, meaning that total effort is *maximized* when agent are symmetric, or $z = 0$.

In the mixture contest, there is no strategic interaction with just two players. Lemma 1 implies that there are impact functions for which the agent's best response is concave in his valuation, and others for which it is convex in his valuation. In the former case, increasing $z$ leads to the same conclusion as in Konrad (2009), i.e. that total effort decreases. However, the opposite occurs when the best response is convex in valuations. Then, total effort is larger when valuations are more spread out. In other words, total effort is *minimized* when the agents are symmetric, or $z = 0$. This occurs with the impact function $a_i^r$, as in Example 1. However, that example also illustrates that increased asymmetry may increase or decrease the total impact, depending on whether $r > \frac{1}{2}$ or $r < \frac{1}{2}$.

Appendix B considers three agents with asymmetric valuations and $r = \frac{1}{2}$. Although strategic interaction complicates the analysis, it is shown that if $(v_1, v_2, v_3) = (v - z, v, v + z)$ then total effort is minimized when agents are symmetric, or $z = 0$.

## 7.3 Asymmetric contests and the exclusion principle

Baye et al (1993) consider a complete information all-pay auction where agents differ only in their valuations. They demonstrate an "exclusion principle" – a contest organizer who is interested in maximizing aggregate effort may want to exclude a number of the "strongest" agents. By excluding strong agents, a more evenly balanced contest among the remaining agents may be created, to the benefit of the organizer. Fang (2002) shows that exclusion can never improve aggregate effort in a Tullock contest in which the impact and cost of an action coincide.

In contrast, in the mixture model with homogeneous technologies it can be optimal to exclude "weaker" agents, i.e. those with low valuations. This is most obvious in a contest with $n = 3$ agents. Excluding any agent means that the remaining two have dominant strategies, and so they work harder the higher their valuations are.

Therefore, if it is optimal to exclude anyone, it must be the weakest agent.

**Proposition 8** *Consider a complete information mixture contest with three agents that differ only in their valuations, with $v_1 \geq v_2 > v_3$. Excluding agent 3 leads to a weakly larger increase (or smaller decrease) in total expected effort and total expected impact than excluding either agent 1 or agent 2.*

Example 2: Consider a complete information contests with three agents with valuations $v_1 = 120$, $v_2 = 9$, and $v_3 = 6$, respectively. In an all-pay auction, total effort is then 4.8375, but this increases to 5 if agent 1 is excluded. Assume now instead that the contest is a mixture contest with impact function $p(a_i) = a_i^r$ with $r = \frac{1}{2}$, and cost function $c(a_i) = a_i$. Hence, the equilibrium analysis in Appendix B applies. Finally, assume that $H(x) = G(x)^\gamma$, where $\gamma = \frac{17}{16}$. The role of $\gamma$ is explored in much more detail in Appendix B. Briefly, the fact that $\gamma$ is close to 1 means that $H$ is "close to" $G$, which in turn means that incentives are weak. It also implies that performance is mostly noise, and that agent 3 has a good chance of winning the contest even with minimal effort. Hence, agent 3's presence discourages agents 1 and 2. In equilibrium, the total impact is 0.92, but this increases to 0.98 when agent 3 is excluded. The total action increases from 0.67 to 0.83.[12] In summary, it is optimal to exclude agent 1 in the all-pay auction but to exclude agent 3 in the mixture contest. ▲

# 8 Extensions and discussion

## 8.1 Non-concave mixture contests

It is possible to extend the equilibrium characterization to contests in which $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is not necessarily single-valued. That is, $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is a correspondence. This may occur if the utility maximization problem is not strictly concave ($p_i(a_i|\boldsymbol{\theta}_i)$ is not concave or $c_i(a_i|\boldsymbol{\theta}_i)$ is not convex) or if the action set is not an interval. Equilibrium may now involve mixed strategies.

The details are in Appendix B, along with an example. The main novelty is that there may exist multiple equilibria with the same expected impact. However, all such equilibria are payoff equivalent.

[12] However, note that excluding an agent at a minimum means foregoing his base level of output stemming from $G$, $\mu_G$. In other words, excluding agent 3 can increase the total impact and total effort but decrease total performance if $\mu_G$ is large enough.

## 8.2 Enlarging the contest

Increasing the number of agents may have surprisingly subtle effects in the mixture model. Other things equal, adding more symmetric competitors lowers the chance of any given agent winning the contest. However, incentives are determined by the agent's ability to change the probability that he wins, rather than by the level of said winning probability itself. In the mixture model, the winning probability may be more sensitive to the agent's effort in a larger contest, thus *heightening* his incentives.

In fact, incentives may be non-monotonic in the number of ex ante symmetric agents. Whether this is the case is determined by the relationship between the mixture components. Starting from a small contest, adding more symmetric agents may intensify competition to such a large extent that it forces both the action and the expected performance of any individual agent to increase. Appendix B demonstrates this possibility in contests in which the expected performance is low to start with.

In contrast, in symmetric complete-information all-pay auctions, the distribution of effort in the mixed strategy equilibrium worsens in the sense of first-order stochastic dominance when more agents participate. In complete-information Tullock contests, equilibrium effort likewise declines under standard assumptions.

However, the comparative statics are richer in rank-order tournaments, and more closely related to the results of the mixture model. Ryvkin and Drugov (2020, Corollary 1) show that individual equilibrium effort may be hump-shaped in the number of agents. They also show that total effort may decrease in the number of agents.

## 8.3 Endogenous entry

The base model studied thus far assumes that participation in the contest is mandatory or alternatively that the set of participants is known before actions are taken.

Consider instead a contest with endogenous and simultaneous entry from among a fixed set of potential participants. Endogenous entry is accommodated by expanding the action set to include an option to "stay out". It is well known that endogenous entry complicates the analysis of contests by adding an extra dimension to agents' decisions. See e.g. Fu et al (2015) for a discussion.

In the mixture model with endogenous entry, the summary uncertainty measure becomes two-dimensional. More precisely, $\overline{p}_j = (\overline{p}_j^{out}, \overline{p}_j^{in})$ where $\overline{p}_j^{out}$ is the probability of exit and $\overline{p}_j^{in}$ is the probability of entry and a performance draw from $H_j$. Agent $j$

enters and draws from $G_j$ with probability $1-\overline{p}_j^{out}-\overline{p}_j^{in}$. Hence, if agent $i$ enters, the ex ante probability that he outperforms agent $j$ is

$$\overline{p}_j^{out}+\overline{p}_j^{in}H_j(x)+\left(1-\overline{p}_j^{out}-\overline{p}_j^{in}\right)G_j(x)$$

when agent $i$'s own performance is $x$. Note that as in the base model, incomplete information does not add to the dimensionality of the problem.

The interim winning probability for agent $i$ upon entry takes the same form as in (2), but with

$$\begin{aligned} t_i\left(\overline{\mathbf{p}}_{-i}\right) &= \int \left(\prod_{j\neq i}\left(\overline{p}_j^{out}+\overline{p}_j^{in}H_j(x)+\left(1-\overline{p}_j^{out}-\overline{p}_j^{in}\right)G_j(x)\right)\right)\times g_i(x)dx \\ k_i\left(\overline{\mathbf{p}}_{-i}\right) &= \int \left(\prod_{j\neq i}\left(\overline{p}_j^{out}+\overline{p}_j^{in}H_j(x)+\left(1-\overline{p}_j^{out}-\overline{p}_j^{in}\right)G_j(x)\right)\right)\left(h_i(x)-g_i(x)\right)dx. \end{aligned}$$

Incentives upon entry are still determined by $k_i\left(\overline{\mathbf{p}}_{-i}\right)$, but note that $t_i\left(\overline{\mathbf{p}}_{-i}\right)$ influences the utility from entering.

With homogenous technologies and two potential entrants, agent $i$'s optimal action upon entry is still independent of $\overline{p}_j^{in}$. Similarly, with more potential entrants it remains the case that productive actions are strategic substitutes in the sense that agent $i$'s best response contingent on entering is decreasing in $\overline{p}_j^{in}$. In either case, agent $i$ is less likely to enter the larger $\overline{p}_j^{in}$ is.

The role of $\overline{p}_j^{out}$ is more nuanced. With two agents, agent $i$'s best response upon entry is decreasing in $\overline{p}_j^{out}$ for the intuitive reason that there is little reason to take costly actions if the sole opponent is likely to stay out. However, as noted above for the case of exogenous participation, incentives are not necessarily monotonic in the number of competitors. For similar reasons, agent $i$'s best response upon entry is not necessarily monotonic in $\overline{p}_j^{out}$ when the number of potential entrants exceeds two.

To illustrate how the analysis is enriched with endogenous entry, Appendix B describes a mixture contest in which the value of the outside option is private information. A solution method as well as an example is provided.

## 8.4 Empirical falsification of competing contest models

With a set of contest models with stochastic performance like the rank-order tournament, the Tullock contest, and the mixture model, it is natural to ask which (if any) model is appropriate. This is likely to depend on the application, and the following discussion therefore centres around the empirical falsification of these various models.

To fix ideas, assume that information is complete and that agents are symmetric. A change in the contest environment occurs, either because the prize increases or marginal costs decrease. Thus, the prediction is that equilibrium actions increase, although the magnitude of the increase cannot be observed. Assume that performance in several contests have been observed both before and after the change. Thus, the outside observed has obtained empirical distributions of performances. Now, different contest models yield different predictions on how these distributions compare.

In the rank-order tournament, noise is additive and a change in the equilibrium action simply shifts the location of the distribution of performance. This testable implication also means that the variance is unchanged. Hirschleifer and Riley (1992) and Jia (2008) provide microfoundations for the Tullock contest. In their settings, noise is multiplicative. Thus, the distribution of performance is scaled up when the equilibrium action and mean performance increases. If the mean performance increases from $\mu$ to $t\mu$, then the variance increases from $\sigma^2$ to $t^2\sigma^2$, $t > 1$.

Fullerton and McAfee (1999) provide another microfoundation for the Tullock contest. Here, the distribution of the performance of an agent with impact $p$ takes the form $F(x|p) = H(x)^p$, where $H$ is some distribution function (unknown to the outside observer). A change in the equilibrium action or impact does not necessarily change the variance of the performance in the same direction as the mean. However, a testable implication is that

$$\frac{\ln F(x|p)}{\ln F(x|p')} = \frac{p}{p'}$$

is independent of the performance.

In the mixture model, the variance of performance does not necessarily move in the same direction as the mean either. Given impact $p$, the distribution of performance is $F(x|p) = pH(x) + (1-p)G(x)$. Imagine now that *three* different contest environments are observed, and denote the respective equilibrium impacts $p$, $p'$, and $p''$ (unknown to the outside observer). In this case,

$$\frac{F(x|p) - F(x|p')}{F(x|p) - F(x|p'')} = \frac{p - p'}{p - p''}$$

is independent of performance.

# 9 Conclusion

This paper introduces a general version of the mixture model of contests. A chief benefit of the model is that any uncertainty about a competitor's action as a consequence of private information is captured by a single uncertainty measure. In this sense, private information does not increase the dimensionality of the problem. Moreover, the structure of the mixture model is such that all types react the same way to changes in the competitive environment, whereas this is not the case in other models. Naturally, this feature also adds to the tractability of the model.

It is challenging to incorporate incomplete information into existing contest models in a tractable manner. Moreover, in part because different types typically react differently to changes in the contest environment, it is hard to reach conclusions about the aggregate effects of such changes. In contrast, the mixture model is ideally suited to shed light on how one particularly important aggregate statistics – namely expected performance – is affected by changes to the contest.

There is a simple trick to determine whether the expected performance in a symmetric mixture contest increases or decreases following a change in the contest environment. In particular, it is sufficient to determine whether the change in the type distribution would cause the expected impact to increase, other things equal. This is a counterfactual question because other things are not equal when equilibrium changes, but it is a question that is usually fairly easy to answer. Indeed, it boils down to understanding how the impact of best responses depends on types.

Likewise, it is possible to predict the direction in which any given type changes his action, but it is rarely possible to determine how expected effort is impacted. Nevertheless, it is demonstrated repeatedly that expected performance and expected effort can move in opposite directions. Depending on the stochastic order involved, this can happen for several reasons, e.g. because the mapping from types into best responses has different modularity properties (Corollary 3), different slopes (Corollary 4), or different curvature (Corollary 6) than the mapping from types into equilibrium impacts. Thus, it should not be taken for granted that higher expected actions or higher expected expenditures necessarily implies higher expected performance.

A main contribution of the mixture model is that it allows for a systematic exploration of the role of the dependence structure between multivariate characteristics. The bivariate case is particularly interesting, because in this case there is a unique stochastic order that captures positive dependence in the "correct" way. Thus, a series of bivariate applications are examined.

Perhaps the most clear-cut application is to contests with private valuations and budgets. Here, the two characteristics are complements in the most archetypical way in the sense that the impact of best responses take a Leontief form. That is, the impact is the minimum of what the budget allows and what an agent with the given valuation would choose if not budget constrained. It is thus intuitive that the more positively dependent the two characteristics are, the higher is the expected performance. This intuition is correct, at least in the mixture model.

Similar intuition explains the conclusions of other applications. The key is to ask whether characteristics are complements or substitutes. In an application with private valuations and starting advantages, the two characteristics are substitutes. On the other hand, in a setting with bivariate productivity characteristics, the two dimensions do not interact in the agent's equilibrium impact. In the former application, greater positive dependence thus decreases expected performance, whereas the dependence structure is irrelevant in the latter.

However, not all applications lead to unambiguous comparative statics. This is demonstrated in an contest with private valuations and impact functions. Here, greater positive dependence can push expected performance in either direction, because the properties of the impact function determines whether the characteristics are complements or substitutes in equilibrium.

Finally, the mixture model offers a way to check the robustness of the common workhorse models. The rank-order tournament and the Tullock contest (or its microfoundations) merely describe different ways to map an agent's action into stochastic performance via a parameterized distribution function. It is difficult a priori to say whether one such mapping is more accurate than another. Absent a completely general analysis of all distribution functions, it is therefore beneficial to have a portfolio of tractable models. A bigger portfolio hopefully contributes to a better understanding of when to expect different comparative statics. Understanding comparative statics are important because they feed into design choices and policy recommendations. To name but one example, it is known that capping the number of agents may intensify competition, but it remains important to know which agents to admit and which ones to exclude, lest the intervention backfires. The mixture model cautions that existing models may identify the wrong agents to exclude when agents are asymmetric.

# References

Acemoglu, D. and M.K. Jensen, 2013, "Aggregate comparative statics," *Games and Economic Behavior*, 81: 27–49.

Baye, M.R., D. Kovenock, and C.G. De Vries, 1993, "Rigging the lobbying process: an application of the all-pay auction," *American Economic Review*, 83: 289–294.

Che, Y.-K. and I. Gale, 1998, "Standard Auctions with Financially Constrained Bidders," *Review of Economic Studies*, 65: 1–21.

Corchón, L.C. and M. Serena, 2018, "*Contest theory*," in Handbook of Game Theory and Industrial Organization, Volume II, edited by Luis C. Corchón and Marco A. Marini, Edward Elgar Publishing.

Drugov, M. and D. Ryvkin, 2022, "Hunting for the discouragement effect in contests," *Review of Economic Design*, forthcoming.

Epstein, L.G. and S.M. Tanny, 1980, "Increasing Generalized Correlation: A Definition and Some Economic Consequences," *The Canadian Journal of Economics*, 13: 16–34.

Ewerhart, C., 2013, "Regular type distributions in mechanism design and $\rho$-concavity," *Economic Theory*, 53: 591–603.

Ewerhart, C. and F. Quartieri, 2020, "Unique equilibrium in contests with incomplete information," *Economic Theory*, 70: 243–271.

Fang, H., 2002, "Lottery versus all-pay auction models of lobbying," *Public Choice*, 112: 351–371.

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

Fey, M., 2008, Rent-seeking contests with incomplete information," *Public Choice*, 135: 225–236

Fu, Q., Q. Jiao, and J. Lu, 2015, "Contests with endogenous entry," *International Journal of Game Theory*, 44: 387-424.

Fu, Q. and Z. Wu, 2019, "*Contests: Theory and Topics*," in Oxford Research Encyclopedia of Economics and Finance, DOI: 10.1093/acrefore/9780190625979.013.440.

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

Grossman, S.J. and O.D. Hart, 1983, "An Analysis of the Principal-Agent Problem," *Econometrica*, 51 (1): 7–45.

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

Hammond, R.G. and X. Zheng, 2013, "Heterogeneity in tournaments with incomplete information: An experimental analysis," *International Journal of Industrial Organization*, 31: 248–260.

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

Holmström, B., 1979, "Moral Hazard and Observability," *The Bell Journal of Economics*, 10: 74–91.

Hopkins, E. and T. Kornienko, 2007, "Cross and Double Cross: Comparative Statics in First Price and All Pay Auctions," *The B.E. Journal of Theoretical Economics.* Vol. 7: Iss. 1 (Topics), Article 19.

Jensen, M.K., 2018, "*Aggregative Games*," in Handbook of Game Theory and Industrial Organization, Volume I, edited by Corchón, L.C. and M.A. Marini, Edward Elgar Publishing.

Jia, H., 2008, "A stochastic derivation of the ratio form of contest success functions," *Public Choice*, 135: 125–130.

Joe, H., 1997, "*Multivariate Models and Dependence Concepts*," Chapman & Hall/CRC.

Kirkegaard, R., 2017, "Moral Hazard and the Spanning Condition without the First-Order Approach", *Games and Economic Behavior*, 102: 373-387.

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

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

Konrad, K.A., 2009, "*Strategy and Dynamics in Contests*," Oxford University Press.

Kotowski, M.H. and F. Li, 2014,"On the continuous equilibria of affiliated-value, all-pay auctions with private budget constraints," *Games and Economic Behavior*, 85: 84-108.

Kräkel, M., 2010, "Shutdown Contests in Multi-Plant Firms and Governmental Intervention," mimeo.

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

Malueg, D.A. and A.J. Yates, 2004, "Rent Seeking with Private Values," *Public Choice*, 119: 161–178.

Meyer, M. and B. Strulovici, 2012, "Increasing interdependence of multivariate distributions," *Journal of Economic Theory*, 147: 1460–1489.

Meyer, M. and B. Strulovici, 2015, "Beyond Correlation: Measuring Interdependence Through Complementarities," mimeo.

Milgrom, P. and C. Shannon, 1994, "Monotone Comparative Statics," *Econometrica*, 62: 157–180.

Müller, A. and D. Stoyan, 2002, "*Comparison Methods for Stochastic Models and Risks*," Wiley.

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

Ryvkin, D., 2010, "Contests with private costs: Beyond two players," *European Journal of Political Economy*, 26: 558–567.

Ryvkin, D. and M. Drugov, 2020, "The shape of luck and competition in winner-take-all tournaments," *Theoretical Economics*, 1587–1626.

Scarsini, M., 1988, "Multivariate stochastic dominance with fixed dependence structure," *Operations Research Letters*, 7: 237–240.

Shaked, M. and J.G. Shanthikumar, 2007, "*Stochastic Orders*," Springer.

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

Tullock, G., 1975, "On the Efficient Organization of Trials," *Kyklos*, 28: 745–762.

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.

Vojnović, M., 2016, "*Contest Theory: Incentive Mechanisms and Ranking Methods*," Cambridge University Press.

# Appendix A: Omitted proofs

**Proof of Proposition 1.** Existence follows from Brouwer's fixed-point theorem. Given a solution $(\overline{p}_1, \overline{p}_2, ..., \overline{p}_n) \in \times_{i=1}^{n} \left[\overline{p}_i^{\min}, \overline{p}_i^{\max}\right]$ to (5), the best response to $\overline{\mathbf{p}}_{-i}$ is $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$. This implies that agent $i$'s expected impact is $\overline{p}_i = E_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right]$. Thus, agents are mutually best responding. ■

**Proof of Proposition 2.** The main body of the text contains a proof which aims to bring out the intuition. A more direct proof is provided here.

In two-player contests, integration by parts yields

$$\begin{aligned}
\frac{\partial k_i(\overline{p}_j)}{\partial \overline{p}_j} &= \int_{\underline{x}}^{\overline{x}} (H(x) - G(x))\,(h(x) - g(x))\,dx \\
&= \left[\frac{1}{2}\,(H(x) - G(x))^2\right]_{\underline{x}}^{\overline{x}} \\
&= 0,
\end{aligned}$$

which proves the result. For larger contests,

$$\frac{\partial k_i\left(\overline{\mathbf{p}}_{-i}\right)}{\partial p_j} = \int \left(\prod\nolimits_{l\neq i,j} \left(\overline{p}_l H(x) + (1-\overline{p}_l)\,G(x)\right)\right)(H(x) - G(x))\,(h(x) - g(x))\,dx.$$

Let $\omega(x)$ denote the product in the first parenthesis. This is strictly increasing in $x$, with $\omega'(x) > 0$. Thus, the proposition follows from

$$\begin{aligned}
\frac{\partial k_i\left(\overline{\mathbf{p}}_{-i}\right)}{\partial \overline{p}_j} &= \int \omega(x)\,(H(x) - G(x))\,(h(x) - g(x))\,dx \\
&= -\int \omega'(x)\frac{1}{2}\,(H(x) - G(x))^2\,dx < 0,
\end{aligned}$$

where integration by parts was used once again. ■

**Proof of Proposition 3.** The result is trivial in the case of two-player contests, since equilibrium is then in strictly dominant strategies. Thus, consider contests with $n \geq 3$ agents. In a symmetric equilibrium, $\overline{p}_j = \overline{p}^*$ for all $j$. At such a symmetric impact profile,

$$k_i\left(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*\right) = \int \left(\overline{p}^* H(x) + (1-\overline{p}^*)\,G(x)\right)^{n-1}(h(x) - g(x))\,dx.$$

Following the same argument as in the last part of the proof of Proposition 2, the derivative with respect to $\overline{p}^*$,

$$\int (n-1)\left(\overline{p}^* H(x) + (1-\overline{p}^*)\, G(x)\right)^{n-2} (H(x) - G(x))\,(h(x) - g(x))\,dx,$$

is strictly negative. Hence, taking the possibility of corner solutions into account, $B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)$ is weakly decreasing in $\overline{p}^*$. It follows that $p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly decreasing in $\overline{p}^*$ for all $\boldsymbol{\theta}_i$, and therefore that $\mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right]$ is weakly decreasing in $\overline{p}^*$. However, a symmetric equilibrium must satisfy

$$\overline{p}^* = \mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right].$$

The left hand side is strictly increasing in $\overline{p}^*$, whereas the right hand side is weakly decreasing. Thus, one and only one solution exists. ■

**Proof of Proposition 4.** The two-player case is trivial because the equilibrium strategy does not change, but higher types are more likely. Thus, consider contests with more than two agents. Given how the type space is ordered, $\mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right]$ weakly increases for any $\overline{p}^*$ when the distribution of types improves in the sense of the usual stochastic order. Thus, using the argument following (6), the equilibrium value of $\overline{p}^*$ weakly increases and equilibrium actions weakly decrease for any given type. ■

**Proof of Proposition 5.** First, $\mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right]$ increases for any $\overline{p}^*$ if $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly increasing and supermodular (submodular) in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$ and the joint distribution becomes stronger in the upper (lower) orthant order. The argument following (6) then implies that the $\overline{p}^*$ weakly increases and actions weakly decrease for any given type. ■

**Proof of Proposition 6.** $\mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right]$ increases for any $\overline{p}^*$ if $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is supermodular in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$ and the joint distribution becomes greater in the supermodular order. Using the argument after (6) then implies that the $\overline{p}^*$ weakly increases and actions weakly decrease for any given type.

If $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is submodular in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$ then $-p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is supermodular in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$. Hence, $\mathbb{E}_{\boldsymbol{\theta}_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right]$ weakly decreases for any $\overline{p}^*$ if the joint distribution becomes greater in the supermodular order. The right hand side of (6) further decreases if $\overline{p}^*$ increases, while the left hand side increases. This is a contradiction, which thus implies that $\overline{p}^*$ must weakly decrease. Equilibrium actions in turn weakly increase for any given type. ■

**Proof of Corollary 1.** Since $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is increasing and supermodular in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$, the corollary follows from Propositions 5 and 6. ■

**Proof of Corollary 2.** Since $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is increasing and submodular in $\boldsymbol{\theta}_i$ for all $\overline{\mathbf{p}}_{-i}$, the corollary follows from Propositions 5 and 6. ■

**Proof of Corollary 3.** From the first-order condition $\theta_i^1 + B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)\theta_i^2$ is a function only of $\theta_i^2$ and $\overline{\mathbf{p}}_{-i}$. Thus, it is possible to write

$$\begin{aligned} \theta_i^1 + B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)\theta_i^2 &= b(\theta_i^2, \overline{\mathbf{p}}_{-i}) \\ p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i) &= p\left(b(\theta_i^2, \overline{\mathbf{p}}_{-i})\right). \end{aligned}$$

Hence, the right-hand side of (6) depends only on $\overline{p}^*$ and the marginal distribution of $\theta_i^2$. However, this is unchanged when the type distribution becomes greater in the supermodular order, and it follows that the solution, $\overline{p}^*$, to (6) is unchanged. This proves the first part of the corollary.

For the second part, $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i) = \frac{b(\theta_i^2, \overline{\mathbf{p}}_{-i}) - \theta_i^1}{\theta_i^2}$ is supermodular in $\boldsymbol{\theta}_i = \left(\theta_i^1, \theta_i^2\right)$ since $\frac{\partial^2 B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial\theta_i^1 \partial\theta_i^2} = 1/\left(\theta_i^2\right)^2 > 0$. Since $\overline{p}^*$ does not change when the type distribution becomes greater in the supermodular order, $\mathbb{E}_{\boldsymbol{\theta}_i}\left[B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)\right]$ increases. Indeed,

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}_i}\left[B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\boldsymbol{\theta}_i)\right] &= \mathbb{E}_{\boldsymbol{\theta}_i}\left[\frac{b(\theta_i^2, \overline{p}^*, \overline{p}^*, ..., \overline{p}^*) - \theta_i^1}{\theta_i^2}\right] \\ &= \mathbb{E}_{\boldsymbol{\theta}_i}\left[\frac{b(\theta_i^2, \overline{p}^*, \overline{p}^*, ..., \overline{p}^*)}{\theta_i^2}\right] + \mathbb{E}_{\boldsymbol{\theta}_i}\left[-\frac{\theta_i^1}{\theta_i^2}\right]. \end{aligned}$$

The first term is unaffected since the marginal distribution of $\theta_i^2$ does not change. The second term is the expectation of a supermodular function, which increases. ■

**Proof of Corollary 4.** Since the (independence) copula is held fixed, it follows from Scarsini (1988) that the improvement in the marginal distribution of $\theta_i^2$ causes the joint distribution to improve in the usual stochastic order. By Proposition 4 and (9), the equilibrium value of $\overline{p}^*$ therefore increases and each type takes a weakly lower action. Fixing $\theta_i^1$ and taking the expectation over $\theta_i^2$, the latter effect means that the expected action decreases for a (counterfactually) fixed marginal distribution of $\theta_i^2$. Since $p_{a\theta_i^2} < 0$ implies that the equilibrium action is decreasing in $\theta_i^2$, the expected action further decreases as the marginal distribution of $\theta_i^2$ improves in the sense of first order stochastic dominance. Thus, conditional on $\theta_i^1$, the expected action decreases. Since the marginal distribution of $\theta_i^1$ has not changed, it follows that the equilibrium expected action decreases. ■

**Proof of Corollary 5.** For the first part, a marginal increase in $\theta_i^2$ must be met with a marginal change in $a_i$ of $\frac{-p_{\theta_i^2}}{p_a}$ if $p(a_i|\theta_i^2)$ is to be held constant. As a result, $p_a(a_i|\theta_i^2)$ changes at a rate of

$$p_{aa}\frac{-p_{\theta_i^2}}{p_a} + p_{a\theta_i^2} = \frac{p_{a\theta_i^2}p_a - p_{aa}p_{\theta_i^2}}{p_a},$$

which is positive, by (9). Hence, $p_a(a_i|\theta_i^2)$ increases at the action that holds $p(a_i|\theta_i^2)$ constant, meaning that the agent is incentivized to increase $a_i$ and thus $p(a_i|\theta_i^2)$. This proves that $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly increasing in $\theta_i^2$. It is trivial to see that $p_i(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is weakly increasing in $\theta_i^1$, as higher $\theta_i^1$ means that the agent is more interested in winning the contest.

For the second part, the agent's first-order condition must be satisfied since $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is interior. Hence, $\theta_i^1 p_a(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\theta_i^2)$ is constant, which implies that

$$\begin{aligned}
\frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial \theta_i^1} &= -\frac{p_a}{\theta_i^1 p_{aa}} > 0 \\
\frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial \theta_i^2} &= -\frac{p_{a\theta_i^2}}{p_{aa}},
\end{aligned}$$

where the inequality utilizes $p_a > 0 > p_{aa}$. From the second expression, the sign of

$$\frac{\partial^2 B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial \theta_i^1 \partial \theta_i^2} = \frac{\partial}{\partial \theta_i^2}\left(-\frac{p_{a\theta_i^2}}{p_{aa}}\right) \times \frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial \theta_i^1}$$

is thus determined by (10). Finally,

$$\begin{aligned}
\frac{\partial^2 p(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\theta_i^2)}{\partial \theta_i^1 \partial \theta_i^2} &= \frac{\partial}{\partial \theta_i^2}\left(p_a(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\theta_i^2)\frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial \theta_i^1}\right) \\
&= p_a(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\theta_i^2)\frac{\partial^2 B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)}{\partial \theta_i^1 \partial \theta_i^2},
\end{aligned}$$

where the last equality follows from the fact that $p_a(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\theta_i^2)$ does not depend on $\theta_i^2$ because $\theta_i^1 p_a(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\theta_i^2)$ is the same for all $\boldsymbol{\theta}_i$ for which the optimal action is interior. In short, $p(B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)|\theta_i^2)$ is submodular (supermodular) in $\boldsymbol{\theta}_i$ if and only if the action $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is submodular (supermodular) in $\boldsymbol{\theta}_i$. ■

**Proof of Proposition 7.** The two-player case is trivial once again. In contests with more than two agents, if $p_i(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)|\theta_i)$ is convex (concave) in $\theta_i$ for all $\overline{\mathbf{p}}_{-i}$, then the mean-preserving spread leads $\mathbb{E}_{\theta_i}\left[p_i(B_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*|\theta_i)|\theta_i)\right]$ to increase (decrease). The argument following (6) then completes the proof. ■

**Proof of Lemma 1.** From the agent's first order condition it follows that

$$p'(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)) = \frac{1}{\theta_i k_i\left(\overline{\mathbf{p}}_{-i}\right)}$$

and therefore that

$$p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))\frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\theta_i)}{\partial \theta_i} = -\frac{1}{\theta_i^2 k_i\left(\overline{\mathbf{p}}_{-i}\right)}$$

and

$$p'''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))\left(\frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\theta_i)}{\partial \theta_i}\right)^2 + p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))\frac{\partial^2 B_i(\overline{\mathbf{p}}_{-i}|\theta_i)}{\partial \theta_i^2} = \frac{2}{\theta_i^3 k_i\left(\overline{\mathbf{p}}_{-i}\right)}.$$

Simplifying the last expression,

$$\begin{aligned}
\frac{\partial^2 B_i(\overline{\mathbf{p}}_{-i}|\theta_i)}{\partial \theta_i^2} &= \frac{-1}{p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))}\left(p'''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))\left(\frac{\partial B_i(\overline{\mathbf{p}}_{-i}|\theta_i)}{\partial \theta_i}\right)^2 - \frac{2}{\theta_i^3 k_i\left(\overline{\mathbf{p}}_{-i}\right)}\right) \\
&= \frac{-1}{\theta_i^3 k_i\left(\overline{\mathbf{p}}_{-i}\right) p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))^3}\left(p'''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))\frac{1}{\theta_i k_i\left(\overline{\mathbf{p}}_{-i}\right)} - 2p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))^2\right) \\
&= \frac{p'''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))p'(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)) - 2p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))^2}{-\theta_i^3 k_i\left(\overline{\mathbf{p}}_{-i}\right) p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))^3},
\end{aligned}$$

which proves that $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ is concave in $\theta_i$ if $p'(a)$ is $\rho$-concave of order $\rho = -1$. The proof that $p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$ is concave in $\theta_i$ if $p'(a)$ is $\rho$-concave of order $\rho = -2$ is similarly based on the fact that

$$\frac{\partial^2 p(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))}{\partial \theta_i^2} \propto p'''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))p'(B_i(\overline{\mathbf{p}}_{-i}|\theta_i)) - 3p''(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))^2.$$

The last part of the lemma follows immediately. ■

**Proof of Corollary 6.** By Proposition 7, the expected equilibrium impact weakly decreases. Since $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ is convex in $\theta_i$, its expected value increases due to the mean-preserving spread, for any given $\overline{\mathbf{p}}_{-i}$. This on its own drives the expected action up, and Proposition 7 in addition reveals that the action increases type-for-type. The two forces work in the same direction, implying that the expected equilibrium action weakly increases. ■

**Proof of Proposition 8.** In the two-player contest, $k_i(\overline{p}_{-i})$ is constant and identity-dependent. Hence equilibrium strategies are independent of who the competitor is and they are weakly increasing in valuations. Thus, agent 3's expected action and expected total impact if he is a participant in a two-player contest is lower than that of agents 1 and 2 if they participate. Hence, replacing agent 3 with whichever agent was excluded in the two-player contest weakly increases total expected effort and total expected impact. ■

# Appendix B: Homogeneous technologies

Table of contents:



## B.1 Properties of the polynomial $k_i\left(\overline{p}_{-i}\right)$

With homogenous technologies, $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ is a symmetric polynomial in the $n-1$ variables $\overline{p}_1, ..., \overline{p}_{i-1}, \overline{p}_{i+1}, ...\overline{p}_n$. Expanding $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ gives $\binom{n-1}{1}$ terms involving a single element of $\overline{\mathbf{p}}_{-i}$, $\binom{n-1}{2}$ terms involving the product of two elements of $\overline{\mathbf{p}}_{-i}$, $\binom{n-1}{3}$ terms involving the product of three elements, and so on. By symmetry, the coefficient to each term depends only on the number of elements in the product. Thus, $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ can alternatively be written as

$$k_i\left(\overline{\mathbf{p}}_{-i}\right)=\kappa_0^n+\kappa_1^n\sum\nolimits_{j\neq i}\overline{p}_j+\kappa_2^n\sum\nolimits_{j\neq i}\sum\nolimits_{j'\neq i,,j'>j}\overline{p}_j\overline{p}_{j'}+...+\kappa_{n-1}^n\overline{p}_1\times...\times\overline{p}_{i-1}\times\overline{p}_{i+1}\times...\times\overline{p}_n,$$

where $\kappa_m^n$ is the coefficient to the terms of degree $m$, $m=0,1,...,n-1$. The qualifier that $j'>j$ in the summation after $\kappa_2^n$ prevents double-counting. For notational simplicity, the superscript in $\kappa_m^n$ will be omitted when no confusion arises as a result.

**Lemma 2** *Any contest with homogeneous technologies has the following properties:*

1. $k_i\left(\overline{\mathbf{p}}_{-i}\right)$ *is a polynomial of degree* $n-2$, *or* $\kappa_{n-1}=0$.
2. *The coefficients alternate in sign, with* $\kappa_m$ *positive if* $m$ *is even and negative if* $m$ *is odd, and* $\kappa_m\in(-1,1)$.
3. *The coefficients diminish in magnitude, or* $\kappa_0>|\kappa_1|>\kappa_2>|\kappa_3|>...>\kappa_{n-1}=0$.

**Proof.** It is easy to see that

$$\kappa_m = \int G(x)^{n-m-1} \left(H(x) - G(x)\right)^m \left(h(x) - g(x)\right) dx, \quad m = 0, 1, ..., n-1. \tag{13}$$

Thus, $\kappa_m$ is the difference between the expectation of $G(x)^{n-m-1} (H(x) - G(x))^m$ under $H$ and $G$, respectively. It follows that $\kappa_m \in (-1, 1)$. Using integration by parts,

$$\begin{aligned} \kappa_{n-1} &= \int \left(H(x) - G(x)\right)^{n-1} \left(h(x) - g(x)\right) dx \\ &= \left[\frac{1}{n} \left(H(x) - G(x)\right)^n\right]_{\underline{x}}^{\overline{x}} = 0 \end{aligned}$$

and

$$\kappa_m = -\int \frac{n-m-1}{m+1} \left(H(x) - G(x)\right)^{m+1} G(x)^{n-m-2} g(x) dx, \qquad m = 0, 1, ..., n-2. \tag{14}$$

First, the fact that $\kappa_{n-1} = 0$ implies that $k_i\left(\overline{p}_{-i}\right)$ is a polynomial of only degree $n-2$. This is consistent with the conclusion in Proposition 2 that $k_i\left(\overline{p}_{-i}\right)$ is constant when $n = 2$. Second, recall that since $H$ first order stochastically dominates $G$, $H(x) - G(x) \leq 0$. Hence, it follows from (14) that $\kappa_m$ is positive if $m$ is even and negative if $m$ is odd. Thus, $\kappa_0, \kappa_2, \kappa_4...$ are positive while $\kappa_1, \kappa_3, \kappa_5...$ are negative. Third, using (13) to calculate $\kappa_m + \kappa_{m+1}$ and then using integration by parts as in (14) reveals that $\kappa_m + \kappa_{m+1} > 0$ when $m$ is even and that $\kappa_m + \kappa_{m+1} < 0$ when $m$ is odd. This in turn implies that the coefficients diminish in magnitude, or $\kappa_0 > |\kappa_1| > \kappa_2 > |\kappa_3| > ... > \kappa_{n-1} = 0.$[13] ■

The first part of the lemma provides one possible way to quantify how the complexity of the problem increases as the number of agents grows. With homogenous technologies, any equilibrium "just" involves handling polynomials of degree $n-2$. The next example further illustrates and discusses the implications of the lemma.

Example 3: Assume that agents have *homogeneous power technologies*: $H$ and $G$ are generated from a parent distribution $F_0$, with $H(x_i) = F_0(x_i)^\alpha$ and $G(x_i) = F_0(x_i)^\beta$,

[13] From (14), note that $\kappa_m$ can be bounded above or below (depending on the sign of $m$) by letting $H(x) \to 0$. The economic interpretation of $H(x) \to 0$ is that $H$ approaches a degenerate distribution in which the top performance is guaranteed. Thus, letting $H(x) \to 0$ yields the conclusion that $0 < \kappa_m < \frac{n-m-1}{n(m+1)}$ if $m$ is even and $0 > \kappa_m > -\frac{n-m-1}{n(m+1)}$ if $m$ is odd, $m = 0, 1, ..., n-2$.

where $\alpha > \beta > 0$. Equivalently, $H(x_i) = G(x_i)^\gamma$, where $\gamma = \frac{\alpha}{\beta} > 1$. Then, from (14),

$$\kappa_m^n = -\int \frac{n-m-1}{m+1} \left(G(x)^\gamma - G(x)\right)^{m+1} G(x)^{n-m-2} g(x) dx.$$

Letting $z = G(x)$ and thus $dz = g(x)dx$, integration by substitution yields

$$\kappa_m^n = -\int_0^1 \frac{n-m-1}{m+1} \left(z^\gamma - z\right)^{m+1} z^{n-m-2} dz.$$

Expanding $\left(z^\gamma - z\right)^{m+1}$ using the Binomial Theorem produces

$$\begin{aligned}
\kappa_m^n &= -\int_0^1 \frac{n-m-1}{m+1} \sum_{j=0}^{m+1} \binom{m+1}{j} \left(z^\gamma\right)^{m+1-j} (-z)^j z^{n-m-2} dz \\
&= -\int_0^1 \frac{n-m-1}{m+1} \sum_{j=0}^{m+1} \binom{m+1}{j} z^{\gamma(m+1-j)+j+n-m-2} (-1)^j dz \\
&= -\frac{n-m-1}{m+1} \sum_{j=0}^{m+1} \binom{m+1}{j} (-1)^j \frac{1}{\gamma(m+1-j)+j+n-m-1} \\
&= -\frac{n-m-1}{m+1} \sum_{j=0}^{m+1} \binom{m+1}{j} (-1)^j \frac{1}{(\gamma-1)(m+1-j)+n}.
\end{aligned}$$

Using this formula, it is easy to see that

$$k_i(\overline{p}_j) = \frac{(\gamma-1)}{2(\gamma+1)}$$

when $n = 2$. In comparison, when $n = 3$,

$$k_i(\overline{p}_j, \overline{p}_{j'}) = \frac{2(\gamma-1)}{3(\gamma+2)} - \frac{(\gamma-1)^2}{3(\gamma+2)(2\gamma+1)} (\overline{p}_j + \overline{p}_{j'}),$$

and when $n = 4$,

$$\begin{aligned}
k_i(\overline{p}_j, \overline{p}_{j'}, \overline{p}_{j''}) = \frac{3(\gamma-1)}{4(\gamma+3)} &- \frac{(\gamma-1)^2}{4(\gamma+1)(\gamma+3)} (\overline{p}_j + \overline{p}_{j'} + \overline{p}_{j''}) \\
&+ \frac{(\gamma-1)^3}{4(\gamma+1)(\gamma+3)(3\gamma+1)} \left(\overline{p}_j\overline{p}_{j'} + \overline{p}_j\overline{p}_{j''} + \overline{p}_{j'}\overline{p}_{j''}\right),
\end{aligned}$$

where $j$, $j'$, and $j''$ are agent $i$'s competitors. If $\gamma = 2$ then $\kappa_0^2 = \frac{1}{6}$, $(\kappa_0^3, \kappa_1^3) = \left(\frac{1}{6}, -\frac{1}{60}\right)$, and $(\kappa_0^4, \kappa_1^4, \kappa_2^4) = \left(\frac{3}{20}, -\frac{1}{60}, \frac{1}{420}\right)$. In line with the previous result, note that the sequence of coefficients alternate in sign and rapidly diminish in numerical magnitude.

The latter property suggests the idea that a good "initial guess" for a solution might be obtained by ignoring the terms of higher degree. This may prove to be a useful starting point in complicated environments where numerical methods are required to obtain a solution. For an analytical demonstration of this point, assume that $\gamma = 2$ and that there are $n = 4$ symmetric agents with known valuation $v$, impact function $p_i(a_i) = \sqrt{a_i}$, and cost function, $c_i(a_i) = a_i$, $a_i \in [0, 1]$. A symmetric and interior equilibrium exists as long as $v \in (0, \frac{56}{3})$, in which case it is characterized by

$$\overline{p}^{4^*} = \frac{1}{2v}\left(7v - \sqrt{35\left(112v - v^2 + 2240\right) + 280}\right), \quad i = 1, 2, 3, 4.$$

A guess based on $\kappa_2^4 = 0$ instead yields $\overline{p}^* = \frac{3v}{v+40}$, which is no more than 5% off the true value (depending on $v$). In this case, the guess underestimates the true value of $\overline{p}_i$ because the guess underestimates the magnitude of $k_i\left(\overline{p}_{-i}\right)$. For instance, if $v = 10$ then the true $\overline{p}_i$ is 0.611 whereas the underestimate is 0.6. In the same setting but with $n = 5$ agents, $\overline{p}^{5^*}$ is 0.533. By setting $\kappa_2^5 = \kappa_3^5 = 0$ the underestimate 0.519 is obtained. Setting only $\kappa_3^5 = 0$ but keeping the correct $\kappa_2^5$ instead gives the overestimate 0.534. These examples demonstrate that it is often possible to bound the true solution by examining simpler but artificial problems where the higher-degree terms are eliminated from $k_i\left(\overline{\mathbf{p}}_{-i}\right)$. ▲

## B.2 Asymmetric equilibria of symmetric contests

Example 4 illustrates that symmetric contests may have asymmetric equilibria.

Example 4: There are three symmetric agents and information is complete. Hence, types are suppressed from the notation. Assume that $v_i = 132$, $p_i(a_i) = a_i$, $c_i(a_i) = 20a_i + a_i^2$, $a_i \in [0, 1]$, $i = 1, 2, 3$. Assume that the mixture components satisfy $H(x_i) = G(x_i)^2$. It is then easy to verify that there is a symmetric equilibrium in which $(a_1, a_2, a_3) = (\overline{p}_1, \overline{p}_2, \overline{p}_3) = (\frac{5}{16}, \frac{5}{16}, \frac{5}{16})$. However, there are also asymmetric equilibria. For instance, $(a_1, a_2, a_3) = (\overline{p}_1, \overline{p}_2, \overline{p}_3) = (0, 0, 1)$ is also an equilibrium. In the symmetric equilibrium, any agent faces a total rival impact of $\frac{10}{16} = \frac{5}{8}$, whereas agents 1 and 2 in the asymmetric equilibrium faces a total rival impact of 1. Given the high rival impact in the latter case, agents 1 and 2 are deterred from taking a positive action.

The total action and the total impact coincide in this setting, or $\sum a_i = \sum p_i(a_i)$. If the contest organizer only cares about the aggregate values, then she is better off in the asymmetric equilibria than the symmetric equilibrium. ▲

## B.3 A valuation/impact contest

The next example describes a bivariate valuation/impact contest. The example is used to illustrate Propositions 3–6 and Corollaries 4 and 5 in a more direct and concrete way.

Example 5: Assume that agents are ex ante symmetric with $v_i(\boldsymbol{\theta}_i) = \theta_i^1$ and $p_i(a_i|\boldsymbol{\theta}_i) = p(a|\theta_i^2) = 1 - e^{-a\theta_i^2}$, where $\theta_i^2 > 0$ measures how productive the agent's action is. Assume also that $c_i(a_i|\boldsymbol{\theta}_i) = a_i$ and $A_i(\boldsymbol{\theta}_i) = [0, \overline{a}_i]$, with $\overline{a}_i > 0$. Note that (9) is satisfied and that $-\frac{p_{a\theta_i^2}}{p_{aa}}$ is strictly decreasing in $a$. Hence, Corollary 5 implies that $p_i(B_i(\overline{p}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is increasing and submodular in $\boldsymbol{\theta}_i$, as will be verified more directly in the following. However, with Corollary 4 in mind, note that $p_{a\theta_i^2} = (1 - a\theta_i^2)e^{-a\theta_i^2}$ can be positive or negative.

Assume that $\overline{a}_i \geq \max \theta_i^1$, meaning that it is never profitable to take the highest possible action. Similarly, assume that $k_i(\overline{\mathbf{p}}_{-i})\theta_i^1\theta_i^2 - 1 > 0$ for all $\boldsymbol{\theta}_i$ when $\overline{\mathbf{p}}_{-i} = (1, 1, ..., 1)$. This rules out low valuations and unproductive impact functions. It implies that $a_i = 0$ is never a best response, because it is not a best-response even in the worst-case scenario where all rivals have the highest possible impacts. Thus, the best response is always interior. For type $\boldsymbol{\theta}_i$, the first-order condition reveals that

$$\left(1 - p_i(B_i(\overline{p}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)\right) k_i(\overline{\mathbf{p}}_{-i}) = \frac{1}{\theta_i^1\theta_i^2}.$$

Taking the expectation over $\boldsymbol{\theta}_i$ yields $(1-\overline{p}_i)k_i(\overline{\mathbf{p}}_{-i}) = \rho$, where $\rho = \mathbb{E}_{\boldsymbol{\theta}_i}[\left(\theta_i^1\theta_i^2\right)^{-1}] > 0$. A symmetric equilibrium impact, $\overline{p}^*$, solves

$$(1 - \overline{p}^*)k_i(\overline{p}^*, \overline{p}^*, ..., \overline{p}^*) = \rho. \tag{15}$$

The left-hand side is decreasing in $\overline{p}^*$. Thus, the solution is unique, as implied by Proposition 3. It also follows that $\overline{p}^*$ is decreasing in $\rho$. In this sense, $\rho$ is a sufficient statistics of the multivariate uncertainty faced by agents. Now, $-\rho$ is increasing and submodular. Hence, $-\rho$ increases when the type distribution becomes stronger in the lower orthant order. In other words, $\rho$ decreases and $\overline{p}^*$ increases. This illustrates Proposition 4, as the usual stochastic order is stronger than the lower orthant order.

Another way to understand Proposition 4 in this context begins by noting that

$$p_i(B_i(\overline{p}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i) = 1 - \frac{1}{\theta_i^1\theta_i^2 k_i(\overline{\mathbf{p}}_{-i})} \tag{16}$$

is increasing in $\theta_i^1\theta_i^2$. Hence, even though types are multivariate, what matters in this particular example is the distribution of the univariate random variable $\theta_i^1\theta_i^2$. The set of $\boldsymbol{\theta}_i$ for which $\theta_i^1\theta_i^2$ exceeds some threshold is an increasing set. As the type distribution improves in the usual stochastic order, this increasing set is assigned more mass. Hence, the distribution of $\theta_i^1\theta_i^2$ improves in the sense of (univariate) first-order stochastic dominance, implying that its expected value increases.[14]

To apply Propositions 5 and 6, note from (16) that $p_i(B_i(\overline{p}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is increasing and submodular in $\boldsymbol{\theta}_i$. By Proposition 5, $\overline{p}^*$ increases when the type distribution becomes stronger in the lower orthant order, as already mentioned above. By Proposition 6, $\overline{p}^*$ decreases when the type distribution becomes greater in the supermodular order.

Next, the left-hand side of (15) is a polynomial of degree $n-1$. Polynomials can be solved analytically up to quartic equations, which means that (15) can be solved analytically for $n \leq 5$. The case with $n = 3$ is considered next.

Given that $n = 3$, each $k_i(\overline{\mathbf{p}}_{-i})$ is a polynomial of degree 1. By symmetry, it can therefore be written $k_i(\overline{\mathbf{p}}_{-i}) = \kappa_0 + \sum_{j \neq i} \kappa_1 \overline{p}_j$. Thus, $k_i(\overline{p}^*, \overline{p}^*) = \kappa_0 + 2\kappa_1\overline{p}^*$. Note that $\kappa_0 > 0$ since it equals $k_i(\overline{\mathbf{p}}_{-i})$ when $\overline{\mathbf{p}}_{-i} = (0, 0)$. By Proposition 2, $\kappa_1 < 0$. The assumption that $k_i(\overline{\mathbf{p}}_{-i})\theta_i^1\theta_i^2 - 1 > 0$ for all $\boldsymbol{\theta}_i$ when $\overline{\mathbf{p}}_{-i} = (1, 1)$ implies that $\kappa_0 + 2\kappa_1 > \rho$, which in turn implies that $\kappa_0 > \rho$. The correct root to (15) can now be identified. The expected equilibrium impact is

$$\overline{p}^* = \frac{-\kappa_0 + 2\kappa_1 + \sqrt{(\kappa_0 + 2\kappa_1)^2 - 8\rho\kappa_1}}{4\kappa_1}. \tag{17}$$

[14] If the distribution of $\theta_i^1\theta_i^2$ undergoes a mean-preserving spread, then $\overline{p}^*$ decreases because $p_i(B_i(\overline{p}_{-i}|\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is concave in $\theta_i^1\theta_i^2$. See also Proposition 7 in Section 7.

Combining the first-order condition with (15) and (17), the equilibrium strategy is

$$\begin{aligned} s_i(\boldsymbol{\theta}_i) &= \frac{1}{\theta_i^2} \ln\left(\theta_i^1 \theta_i^2 k_i(\overline{p}^*, \overline{p}^*)\right) \\ &= \frac{1}{\theta_i^2} \ln\left(\theta_i^1 \theta_i^2 \frac{1}{2}\left(\kappa_0 + 2\kappa_1 + \sqrt{(\kappa_0 + 2\kappa_1)^2 - 8\rho\kappa_1}\right)\right), \; i = 1, 2, 3. \end{aligned}$$

As mentioned, the equilibrium impact $p_i(s_i(\boldsymbol{\theta}_i)|\boldsymbol{\theta}_i)$ is monotonic in $\theta_i^1\theta_i^2$. Thus, two types with the same $\theta_i^1\theta_i^2$ value have the same equilibrium impacts but not necessarily the same equilibrium actions. In fact, $s_i(\boldsymbol{\theta}_i)$ is generally hump-shaped in $\theta_i^2$. The underlying reason is that $p_{a\theta_i^2}$ may change sign. However, if $\theta_i^2$ is large for all $\boldsymbol{\theta}_i$, then the equilibrium action is decreasing in $\theta_i^2$ because it then holds that $p_{a\theta_i^2} < 0$. The same is true if the action set and type space is such that $a\theta_i^2 > 1$ for all actions and types. In the latter case, Corollary 4 applies. ▲

## B.4 A contest with three heterogeneous agents

This subsection presents a contest with three heterogenous agents in which equilibrium is unique and strategies can be characterized in closed form. Types are univariate and written as $\theta_i$ rather than $\boldsymbol{\theta}_i$. Technologies are homogeneous. Impact and cost functions are also symmetric across agents, with

$$p_i(a_i|\theta_i) = \sqrt{a_i}, \; c_i(a_i|\theta_i) = a_i, \; A_i(\theta_i) = [0,1] \text{ for all } \theta_i \in \Theta_i \text{ and all } i = 1, 2, 3. \quad (18)$$

This specification is isomorphic to a setting with a linear impact function but quadratic cost function. Types capture private information about valuations, $v_i(\theta_i)$. Let $v_i$ denote the expected value of agent $i$'s valuation, $i = 1, 2, 3$. Thus, agents are allowed to have different type distributions. Given that $k_i\left(\overline{\mathbf{p}}_{-i}\right) \leq \kappa_0$ for all $\overline{\mathbf{p}}_{-i}$, the best response for all types is strictly below 1 regardless of $\overline{p}_{-i}$ if

$$\max_{\theta_i \in \Theta_i} \left(\frac{1}{2}\kappa_0 v_i(\theta_i) - 1\right) < 0, \quad (19)$$

which is henceforth assumed. For future reference, note that since $\kappa_0 > |\kappa_1|$, it holds that $2 + \kappa_1 v_i(\theta_i) > 0$ for all types, and therefore it holds that $2 + \kappa_1 v_i > 0$.

**Proposition 9** *Consider a three-player contest with symmetric mixture components* *and which satisfies (18) and (19). In the unique equilibrium,*

$$\overline{p}_i = \frac{\kappa_0 \upsilon_i \left(\kappa_1 \upsilon_j + 2\right)\left(\kappa_1 \upsilon_{j'} + 2\right)}{8 - 2\left(\kappa_1\right)^2 \left(\upsilon_1 \upsilon_2 + \upsilon_1 \upsilon_3 + \upsilon_2 \upsilon_3\right) - 2\left(\kappa_1\right)^3 \upsilon_1 \upsilon_2 \upsilon_3}, \quad i = 1, 2, 3 \tag{20}$$

*and the equilibrium strategies are given by*

$$s_i(\theta_i) = \left(\frac{\upsilon_i(\theta_i)}{\upsilon_i}\right)^2 \left(\frac{\kappa_0 \upsilon_i \left(\kappa_1 \upsilon_j + 2\right)\left(\kappa_1 \upsilon_{j'} + 2\right)}{8 - 2\left(\kappa_1\right)^2 \left(\upsilon_1 \upsilon_2 + \upsilon_1 \upsilon_3 + \upsilon_2 \upsilon_3\right) - 2\left(\kappa_1\right)^3 \upsilon_1 \upsilon_2 \upsilon_3}\right)^2, \quad i = 1, 2, 3 \tag{21}$$

*where agents $j$ and $j'$ are agent $i$'s competitors.*

**Proof.** If $\upsilon_i(\theta_i) = 0$, then the best response of agent $i$ with type $\theta_i$ is $B_i(\overline{\mathbf{p}}_{-i}|\theta_i) = 0$. More generally, $B_i(\overline{\mathbf{p}}_{-i}|\theta_i)$ can be derived from the first-order condition as in Example 1. This satisfies

$$\sqrt{B_i(\overline{\mathbf{p}}_{-i}|\theta_i)} = \frac{1}{2}\upsilon_i(\theta_i) k_i\left(\overline{\mathbf{p}}_{-i}\right),$$

where the left-hand side is $p_i(B_i(\overline{\mathbf{p}}_{-i}|\theta_i))$. Thus, in equilibrium,

$$\overline{p}_i = \frac{1}{2}\upsilon_i\left(\kappa_0 + \kappa_1(\overline{p}_j + \overline{p}_{j'})\right), \quad i = 1, 2, 3, \tag{22}$$

where agents $j$ and $j'$ are agent $i$'s competitors. The unique solution to this system of equations is (20). With this in hand, the equilibrium strategy is

$$\begin{aligned} s_i(\theta_i) &= \left(\frac{1}{2}\upsilon_i(\theta_i) k_i\left(\overline{\mathbf{p}}_{-i}\right)\right)^2 \\ &= \left(\upsilon_i(\theta_i)\frac{\overline{p}_i}{\upsilon_i}\right)^2 \end{aligned}$$

by (22),which thus gives (21). Using (19), it can be readily confirmed that $s_i(\theta_i) \in [0, 1)$. ■

Using (20), it is easy to verify that $\overline{p}_i$ is increasing in $\upsilon_i$ but decreasing in $\upsilon_j$ and $\upsilon_{j'}$. Thus, agent $i$ is discouraged and lowers his effort if agent $j$'s mean valuation increases. The reason is that actions are strategic substitutes. When agent $j$'s value increases, his incentives to work harder increases, which in turn lowers agent $i$'s incentives. However, it is clear from Example 1 that $\upsilon_j$ is a good measure of how agent $i$ perceives of agent $j$'s expected impact only because $p_i(a_i|\theta_i) = \sqrt{a_i}$.

The last observation in Section 7.2 implies that when $p_i(a_i|\theta_i) = \sqrt{a_i}$, total output in two-player contests does not depend on how the "total valuation" is distributed across agents. Does this result extend to three-player contests? The complication is that adding a third agent opens the door to strategic interactions that are absent in the two-player contest with strictly dominant strategies.

Total output in the three-player contest is proportional to $\sum_{i=1}^{3} \overline{p}_i$, which in turn depends on $\upsilon_1, \upsilon_2$ and $\upsilon_3$. Thus, write $\sum_{i=1}^{3} \overline{p}_i$ as $w(\upsilon_1, \upsilon_2, \upsilon_3)$ and note that

$$w(\upsilon_1, \upsilon_2, \upsilon_3) = \frac{4(\upsilon_1+\upsilon_2+\upsilon_3) + 4\kappa_1(\upsilon_1\upsilon_2+\upsilon_1\upsilon_3+\upsilon_2\upsilon_3) + 3(\kappa_1)^2 \upsilon_1\upsilon_2\upsilon_3}{8 - 2(\kappa_1)^2(\upsilon_1\upsilon_2+\upsilon_1\upsilon_3+\upsilon_2\upsilon_3) - 2(\kappa_1)^3 \upsilon_1\upsilon_2\upsilon_3}\kappa_0.$$

Next, compare a set of contests with the same fixed "total strength", $\sum_{i=1}^{3} \upsilon_i = 3\upsilon > 0$. Assume that in all the contests that are being compared, the profile of mean valuations can be written $(\upsilon_1, \upsilon_2, \upsilon_3) = (\upsilon - z, \upsilon, \upsilon + z)$ for some $z \in [-\upsilon, \upsilon]$.

**Corollary 7** *$w(\upsilon - z, \upsilon, \upsilon + z)$ is u-shaped in $z$ and minimized at $z = 0$. Thus, the lowest total expected performance among all contests in which $(\upsilon_1, \upsilon_2, \upsilon_3) = (\upsilon - z, \upsilon, \upsilon + z)$ and (19) is satisfied occurs when agents have the same expected valuation, or $z = 0$. Moreover, in complete-information contests in which (19) is satisfied, the sum of actions is also u-shaped in $z$ and minimized at $z = 0$.*

**Proof.** Simple differentiation yields

$$\frac{\partial w(\upsilon - z, \upsilon, \upsilon + z)}{\partial z} = \frac{-2\kappa_0\kappa_1(\upsilon\kappa_1 + 2)^3}{\left(4 + (z^2 - 3\upsilon^2)(\kappa_1)^2 + \upsilon(z^2 - \upsilon^2)(\kappa_1)^3\right)^2} z.$$

Since the first term is positive whenever (19) holds, the entire expression has the same sign as $z$. Thus, $w(\upsilon - z, \upsilon, \upsilon + z)$ is first decreasing in $z$ and then increasing in $z$ on $[-\upsilon, \upsilon]$. Consequently, it is minimized at $z = 0$, i.e. when agents have symmetric strengths.

In complete-information contests, agent $i$'s action is simply $\overline{p}_i^2$. Thus, the sum of actions is $\sum_{i=1}^{3} \overline{p}_i^2$. Differentiating this sum with respect to $z$ yields

$$\frac{2\kappa_0^2(\upsilon\kappa_1 + 2)^2\left(z^2\kappa_1^2(\upsilon^2\kappa_1^2 + 2) - (\upsilon^2\kappa_1^2 + 4\upsilon\kappa_1 - 2)(\upsilon\kappa_1 + 2)^2\right)}{\left(4 + (z^2 - 3\upsilon^2)(\kappa_1)^2 + \upsilon(z^2 - \upsilon^2)(\kappa_1)^3\right)^3} z,$$

where $(v^2\kappa_1^2 + 4v\kappa_1 - 2) < 0$ since $v\kappa_1 \in (-2, 0)$ as a consequence of (19). Hence, both numerator and denominator are positive and the derivative therefore has the same sign as $z$. It now follows that the sum of actions is u-shaped in $z$ as long as (19) is satisfied. ■

If $x > 0$ then $v_3 > v_2 > v_1$ and $\overline{p}_3 > \overline{p}_2 > \overline{p}_1$ as a consequence. Hence, agent 1 faces a greater aggregate expected impact from his competitors than agent 3 does, or $\overline{p}_2 + \overline{p}_3 > \overline{p}_1 + \overline{p}_2$. This means that agent 1 has less of an incentive to provide effort. Thus, $\overline{p}_1$ is less sensitive to a small change in $z$ than $\overline{p}_3$ is. In other words, a further increase in $z$ causes $\overline{p}_1$ to decrease less than $\overline{p}_3$ increases. Thus, $\overline{p}_1 + \overline{p}_3$ increases when $z > 0$ increases. This is the direct effect of a change in $z$. A secondary and indirect effect comes from the fact that the increase in $\overline{p}_1 + \overline{p}_3$ causes agent 2 to work less hard, which in turn tends to increase the incentives for both agents 1 and 3 to work hard. The indirect effects are smaller and overall the fact that agent 2 works less hard cannot overcome the fact that agents 1 and 3 jointly work harder than before.

Fang (2002) characterizes total effort in complete information lottery contests with linear impact functions and an arbitrary number of agents. Using his results, it can be verified that total effort in a three-player contest is maximized when $z = 0$.

The exclusion result for mixture contests in Example 2 in Section 7.3 can be verified by comparing $w(120, 9, 6)$ with the dominant strategy equilibrium when agent 3 is excluded, while using $\gamma = \frac{17}{16}$ and the characterization in Example 3. This subsection concludes with another example of the exclusion principle.

Example 6: Assume that (18) applies, and that technologies are power technologies with $\gamma \in (1, 2)$. Assume also that information is complete and that $(v_1, v_2, v_3) = (v, v, 0)$, with $v > 0$ and $v < \frac{3(\gamma+2)}{\gamma-1}$. The last part ensures that (19) is satisfied. The total impact is $w(v, v, 0)$. It is straightforward to compute the total impact if agent 3 is excluded. Such exclusion increases the total impact if

$$v < \min\left\{\frac{2\,(2\gamma+1)\,(2-\gamma)}{(\gamma-1)^2}, \frac{3\,(\gamma+2)}{\gamma-1}\right\}.$$

Since $(i)$ agents 1 and 2 are symmetric and information is complete and $(ii)$ agent 3's action is zero even if he is included, the total action moves in the same direction as the total impact. Hence, total effort also increases when agent 3 is excluded. ▲

## B.5 Non-concave mixture contests

This section considers an extension to a more general mixture contests in which $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is not necessarily single-valued. That is, $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is allowed to be a correspondence.

Fixing $\boldsymbol{\theta}_i$ and recalling that $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is a closed set, the impact of any action in $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ is somewhere between $\min_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)$ and $\max_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)$. Since the agent is indifferent between any action in $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$, he is willing to randomize between $\arg\min_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)$ and $\arg\max_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)$. By varying the mixed strategy, any expected impact for type $\boldsymbol{\theta}_i$ between $\min_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)$ and $\max_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)$ can therefore be obtained. Note that there is no loss of generality in assuming that the agent randomizes between only these two actions. However, there may be multiple equilibria because the same expected impact may be achieved by assigning positive probability to other best responses as well. Nevertheless, these equilibria are payoff equivalent. A fully solved example is given momentarily.

Repeating this argument for every $\boldsymbol{\theta}_i \in \Theta_i$, the ex ante expected impact of agent $i$ can be made to be any number in the interval

$$\left[\mathbb{E}_{\boldsymbol{\theta}_i}\left[\min_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)\right], \mathbb{E}_{\boldsymbol{\theta}_i}\left[\max_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)\right]\right].$$

Thus, starting from the correspondence $B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)$ that maps $\overline{\mathbf{p}}_{-i}$ into actions, then allowing randomization, and finally integrating over types gives another correspondence that maps $\overline{\mathbf{p}}_{-i}$ into the uncertainty measure $\overline{p}_i$. Since there is one such correspondence for each agent, the aim is now to find a fixed-point $\overline{\mathbf{p}}$ where

$$\overline{p}_i \in \left[\mathbb{E}_{\boldsymbol{\theta}_i}\left[\min_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)\right], \mathbb{E}_{\boldsymbol{\theta}_i}\left[\max_{a_i \in B_i(\overline{\mathbf{p}}_{-i}|\boldsymbol{\theta}_i)} p_i(a_i|\boldsymbol{\theta}_i)\right]\right] \text{ for all } i = 1, ...n. \quad (23)$$

**Proposition 10** *Any solution* $\overline{\mathbf{p}}$ *to (23) has one or more associated Bayesian Nash Equilibria. These equilibria have the same interim expected payoffs and ex ante winning probabilities.*

**Proof.** The first part is in the text. Interim payoff equivalence holds because the only difference between equilibria with the same $\overline{\mathbf{p}} = (\overline{p}_1, \overline{p}_2, ..., \overline{p}_n)$ is that agents are varying the randomization probabilities between actions they are indifferent between in the first place. Ex ante winning probabilities are the same because $\overline{\mathbf{p}}$ is the same across the equilibria in question. ■

The following example illustrates Proposition 10.

Example 7: There are two agents. Assume that agents have heterogenous power technologies, with $H_i(x_i) = F_0(x_i)^{\alpha_i}$ and $G_i(x_i) = F_0(x_i)^{\beta_i}$, where $F_0$ is a parent distribution common to both agents and $\alpha_i > \beta_i > 0$, $i = 1, 2$. For concreteness, let $\alpha_1 = 10$, $\beta_1 = 1$, $\alpha_2 = 5$, $\beta_2 = 1$. Since technologies are heterogenous, equilibrium is not in strictly dominant strategies even though there are only two agents. In fact,

$$k_1(\overline{p}_2) = \frac{9}{22} + \frac{1}{11}\overline{p}_2 \text{ and } k_2(\overline{p}_1) = \frac{1}{3} - \frac{1}{11}\overline{p}_1.$$

Agent 1 has one of two types, $\theta_1'$ and $\theta_1''$, both of which are equally likely. The two types are characterized by

$$\begin{aligned} v_1(\theta_1') &= v_1',\ p_1(a_1|\theta_1') = a_1,\ c_1(a_1|\theta_1') = a_1,\ A_1(\theta_1') = \{0,1\} \\ v_1(\theta_1'') &= v_1'',\ p_1(a_1|\theta_1'') = a_1 + \frac{1}{2}\left(a_1^2 - a_1^3\right),\ c_1(a_1|\theta_1'') = a_1,\ A_1(\theta_1'') = [0,1]. \end{aligned}$$

Here, type $\theta_1'$ has a discrete action set. Type $\theta_1''$ has a continuous action set but $p_1(a_1|\theta_1'')$ is not globally concave. As a result, the utility maximization problem is not concave and some actions can never be best responses. This impact function is taken from Kirkegaard (2017), who (in a moral hazard context) explains why it can never be optimal to take an action in $(0, \frac{1}{2})$. It can be verified that

$$B_1(\overline{p}_2|\theta_1') = \begin{cases} 0 & \text{if } v_1'k_1(\overline{p}_2) < 1 \\ \{0,1\} & \text{if } v_1'k_1(\overline{p}_2) = 1 \\ 1 & \text{if } v_1'k_1(\overline{p}_2) > 1 \end{cases},$$

and

$$B_1(\overline{p}_2|\theta_1'') = \begin{cases} 0 & \text{if } v_1''k_1(\overline{p}_2) < \frac{8}{9} \\ \{0, \frac{1}{2}\} & \text{if } v_1''k_1(\overline{p}_2) = \frac{8}{9} \\ \frac{1}{3} + \frac{1}{3}\sqrt{\frac{7v_1''k_1(\overline{p}_2) - 6}{v_1''k_1(\overline{p}_2)}} & \text{if } v_1''k_1(\overline{p}_2) \in (\frac{8}{9}, 2) \\ 1 & \text{if } v_1''k_1(\overline{p}_2) \geq 2 \end{cases}.$$

Note that there are values of $v_1'k_1(\overline{p}_2)$ and $v_1''k_1(\overline{p}_2)$ for which each of the two types would be willing to randomize. Indeed, if $v_1'' = \frac{8}{9}v_1'$ then there may even be $k_1(\overline{p}_2)$

values for which both types are willing to randomize at the same time.

Agent 2 has only one type (complete information) with $v_2(\theta_2) = v_2$, $p_2(a_2|\theta_2) = \sqrt{a_2}$, $c_2(a_2|\theta_2) = a_2$, and $A_2(\theta_2) = [0,1]$. Thus, $B_2(\overline{p}_1|\theta_2) = \left(\frac{1}{2}v_2 k_2\left(\overline{p}_1\right)\right)^2$.

It remains to specify $v_1'$, $v_1''$, and $v_2$. These will now be chosen in such a way that an equilibrium is generated in which $\overline{p}_1 = \overline{p}_2 = \frac{1}{2}$ and both types of agent 1 randomize. First, note that $k_1(\frac{1}{2}) = \frac{5}{11}$. Thus, if $v_1' = \frac{11}{5}$ and $v_1'' = \frac{8}{9}v_1' = \frac{88}{45}$ then both types are willing to randomize if $\overline{p}_2 = \frac{1}{2}$. Second, note that $k_2(\frac{1}{2}) = \frac{19}{66}$. Thus, if $v_2 = \frac{66}{19}$, the best response to $\overline{p}_1 = \frac{1}{2}$ is $a_2 = \frac{1}{4}$, with impact $p_2(a_2|\theta_2) = \frac{1}{2}$. This in turns means that agent 1 is willing to randomize for both his types.

Let $q'$ and $q''$ be the probability that type $\theta_1'$ and $\theta_1''$ assign to $a_1 = 1$ and $a_1 = \frac{1}{2}$, respectively (the remaining probability is assigned to $a_1 = 0$). Since each type is equally likely,

$$\overline{p}_1 = \frac{1}{2}q' p_1(1|\theta_1') + \frac{1}{2}q'' p_1(\frac{1}{2}|\theta_1'') = \frac{1}{2}\left(q' + q''\frac{9}{16}\right).$$

Since $\overline{p}_1 = \frac{1}{2}$ in equilibrium, this means that any $(q', q'')$ for which $q' + q''\frac{9}{16} = 1$ forms an equilibrium strategy for agent 1. Thus, there is a continuum of equilibria, but they are all interim payoff equivalent. Note that agent 1's expected action is $\frac{1}{2}q' + \frac{1}{4}q''$, which differs from equilibrium to equilibrium.

Next, assume instead that $v_1' = 2$ and $v_1'' = \frac{16}{9}$ and observe that $k_1(1) = \frac{1}{2}$. This means that both types of agent 1 are willing to randomize if $\overline{p}_2 = 1$. A range of $\overline{p}_1$ values can thus be sustained as long as $\overline{p}_2 = 1$. Conversely, $\overline{p}_2 = 1$ is optimal for agent 2 for low enough $\overline{p}_1$ values if $v_2 > 6$. Thus, if $v_1' = 2$, $v_1'' = \frac{16}{9}$, and $v_2 > 6$, then there are multiple equilibria that have varying $\overline{p}_1$ but all with $\overline{p}_2 = 1$. These equilibria are not payoff equivalent to agent 2 and agent 1's expected performance differs from equilibrium to equilibrium as well. ▲

## B.6 Varying the number of agents

Assume agents are ex ante symmetric. Consider the problem from agent $i$'s point of view, assuming that all his competitors have identical expected impacts. Write this scalar as $\overline{p}^n$ and define $K^n(\overline{p}^n) = k_i(\overline{p}^n, \overline{p}^n, ..., \overline{p}^n)$ as the value of $k_i(\overline{\mathbf{p}}_{-i})$ when all the $n-1$ competing agents have the same expected impact, $\overline{p}^n$. Consider the extreme case in which the agent's competitors has zero impact, or $\overline{p}^n = 0$. Note that this does not mean that the competitors necessarily perform poorly, since there is a chance that even a draw from the bad mixture component may be high. In particular, an agent with zero impact wins the contest with positive probability regardless of other agents' strategies. Hence, the strategic considerations in a mixture contest depends on the number of rivals, even if some of them are inactive. Incidentally, recall that it is also the case that an inactive agent contributes expected performance of $\mu_G$. For the contest organizer, excluding an inactive agent may therefore come at a cost if she is interested in total performance.

Given that all rivals have zero impacts, agent $i$'s incentives are determined by

$$K^n(0) = \int G(x)^{n-1} \left(h(x) - g(x)\right) dx = \int G(x)^{n-1} h(x) dx - \frac{1}{n},$$

which is the increase in winning probability when agent $i$'s performance is drawn from the good mixture component rather than the bad mixture component, given all rivals' draws are known to be from the bad components. Imagine now that $G$ is a very bad distribution in which outcomes close to $\underline{x}$ are extremely likely, while $H$ is a very good distribution in which outcomes close to $\overline{x}$ are extremely likely. Then, the probability that agent $i$ wins is near one if his draw comes from the good component, and this does not change much when the number of rivals changes. However, the probability that he wins if his draw comes from $G$ is $\frac{1}{n}$, which decreases rapidly in the number of rivals when $n$ is small. Thus, the difference between the two probabilities increases when the contest is small to start with and another agent joins the fray. In other words, agent $i$'s incentives actually *increase* when another competitor shows up.

To be more specific, fix any bad component $G$ and any two integers $m$ and $n$, with $m > n \geq 2$. For *any* such combination, there exists a good component $H$ such that $K^m(0) > K^n(0)$. The proof is by construction. Assume agents have power technologies, with $\gamma > 1$. Then, integration by substitution yields

$$K^n(0) = \frac{(\gamma - 1)(n-1)}{n(\gamma + n - 1)}$$

and

$$K^m(0) - K^n(0) = \frac{(m-n)(\gamma - 1)(\gamma - (n-1)(m-1))}{mn(\gamma + m - 1)(\gamma + n - 1)} \tag{24}$$

is positive when $\gamma$ is large enough, i.e. when $H$ assigns enough mass to very high realizations. This construction works for more values of $\gamma$ the smaller $n$ or $m$ are. For a fixed $\gamma$, $K^n(0)$ is either decreasing in $n$ or hump-shaped in $n$. Thus, incentives may be non-monotonic in the number of agents.

Let $\overline{p}^{m^*}$ and $\overline{p}^{n^*}$ denote the individual expected equilibrium impacts of the two contests. If *some* type takes a strictly higher action in the $m$ player contest, then *all* types must take a weakly higher action. The reason is that the incentives of all types are determined by the same number in the two settings, $K^m(\overline{p}^{m^*})$ and $K^n(\overline{p}^{n^*})$.

Of course, it is endogenous whether the competitors work so little that equilibrium impacts are near zero. Propositions 4–6 and Corollaries 1–5 reveal when the expected impact may be expected to be low. For instance, this is the case if valuations are low or if valuations and budgets are very negatively dependent. An additional competitor may then increase incentives and cause all types to take higher actions.[15]

**Proposition 11** *For any bad component $G$ and any integers $m$ and $n$, with $m > n \geq 2$, there exists a good component $H$ for which $K^m(0) > K^n(0)$. Whenever $K^m(0) > K^n(0)$, there exists preferences and abilities such that individual actions and impacts are strictly higher in a contest with $m$ than with $n$ players, or $\overline{p}^{m^*} > \overline{p}^{n^*}$ and $K^m(\overline{p}^{m^*}) > K^n(\overline{p}^{n^*})$.*

**Proof.** The first part is proven in the text. The intuition for the second part is also explained in the text. A more formal proof follows. In equilibrium, $K^m(\overline{p}^{m^*}) > K^n(\overline{p}^{n^*})$ is necessary for individual actions to be higher in the $m$ player contest, and sufficient if actions are in the interior of the action set. By continuity, and given $K^m(0) > K^n(0)$, there exists a number $\widehat{p} \leq 1$ such that $K^m(\overline{p}) > K^n(\overline{p})$ for all $\overline{p} \in [0, \widehat{p}]$.

Now construct preferences and abilities such that actions are interior for all types in the $n$ player contest and the equilibrium impact satisfies $\overline{p}^{n^*} < \widehat{p}$. For example, this can be achieved with univariate types by letting $p_i(a|\theta_i) = \sqrt{a_i}$, $c_i(a_i|\theta_i) = a_i$, $A_i(\theta_i) = [0, 1]$, and $v_i(\theta_i)$ be strictly positive but small for all $\theta_i \in \Theta_i$. It follows that best responses are strictly higher type-for-type in the $m$ player contest than in the $n$ player contest, given that the expected impact by the competitors in the two settings is the same and below $\widehat{p}$. In other words, the right hand side of (6) is now strictly

[15] With power technologies, $K^n(1)$ always decreases in $n$, meaning that if equilibrium impacts are high to start with, then the addition of more competitors lowers incentives. However, it is possible to find examples that do not use power technologies and where $K^n(1)$ may increase in $n$.

greater in the $m$ player contest than in the $n$ player contest for impacts below $\widehat{p}$. The fixed point thus occurs at a higher value, or $\overline{p}^{m^*} > \overline{p}^{n^*}$. This in turn requires that actions are higher in the $m$ player contest, or $K^m(\overline{p}^{m^*}) > K^n(\overline{p}^{n^*})$. ■

EXAMPLE 8: Building on Example 1, assume that agents are symmetric and that $v_i(\theta_i) = \theta_i$, with $c_i(a_i|\theta_i) = a_i$, $p_i(a_i|\theta_i) = \sqrt{a_i}$, and $A_i(\theta_i) = [0, 1]$ for all $\theta_i \in \Theta_i$. Let $\upsilon$ denote the expected valuation.

Assume that agents have power technologies and consider an increase in $n$ from 2 to 3. With (24) in mind, assume that $\gamma > 2$. This means that any agent's incentives are higher if he faces two rather than one competitor, and highest if his competitors have zero impact. His best response in this case is characterized by the first-order condition as long as $\theta_i \frac{1}{2} K^3(0) - 1 \leq 0$, or $\theta_i \leq \frac{3(\gamma+2)}{(\gamma-1)} \equiv \overline{\theta}$. Thus, assume that $\Theta_i \subseteq [0, \overline{\theta}]$. Then,

$$\overline{p}^{2^*} = \frac{1}{4}\frac{\upsilon(\gamma - 1)}{\gamma + 1}$$

and

$$\overline{p}^{3^*} = \frac{\upsilon(\gamma - 1)(2\gamma + 1)}{3(2\gamma + 1)(\gamma + 2) + (\gamma - 1)^2 \upsilon}.$$

It is now easy to verify that the expected equilibrium impact strictly increases when the third agent joins the contest if and only if

$$0 < \upsilon < \frac{(\gamma - 2)(2\gamma + 1)}{(\gamma - 1)^2}, \tag{25}$$

where the term on the right hand side can be verified to be less than $\overline{\theta}$. In either case, the equilibrium strategy is $s_i(\theta_i) = \left(\frac{\theta_i}{\upsilon}\overline{p}^{n^*}\right)^2$. Hence, actions are higher type-for-type in the three-player contest than in the two-player contest when (25) holds. ▲

EXAMPLE 9: Consider the environment in the previous example, but allow for an arbitrary number of agents. Assuming best responses are interior, the symmetric equilibrium impact is found by solving

$$\overline{p}^{n^*} = \frac{1}{2}\upsilon k_i\left(\overline{p}^{n^*}, \overline{p}^{n^*}, ..., \overline{p}^{n^*}\right),$$

which requires solving a polynomial of degree $n-2$. As noted in Example 5, polynomials can be solved analytically up to quartic equations, which means that an analytical solution can in principle be obtained for $n = 2, 3, ..., 6$. Once the correct root has been identified, the strategy is again given by $s_i(\theta_i) = \left(\frac{\theta_i}{v}\overline{p}^{n^*}\right)^2$, as in Example 8. Thus, there is not the same curse of dimensionality as Ryvkin (2010) identifies in Tullock contests with private information. Of course, this example is made easier by the fact that an analytical solution to each type's maximization problem can be obtained for any value of $k_i$. However, even if numerical solutions for each type must be relied upon, it is still the case that the complexity of each type's maximization problem is independent of the number of competitors (and the cardinality of the type space), for any given $\overline{p}^n$. Thus, with ex ante symmetric agents, the only minor headache of increasing $n$ is that the mapping from $\overline{p}^{n^*}$ to $k_i\left(\overline{p}^{n^*}, \overline{p}^{n^*}, ..., \overline{p}^{n^*}\right)$ is a higher-degree polynomial.

Assume next that agents have power technologies with $\gamma = 2$ and valuations are commonly known to equal 10 (there is no incomplete information). Table 2 reports equilibrium values of $\overline{p}^{n^*}$ and other comparative statics. The impact and action of an individual agent decreases when he faces more competitors. Nevertheless, the total impact, $n \times \overline{p}^{n^*}$, increases, which in turns implies that total expected output increases as long as $\mu_G \geq 0$. What is perhaps more surprising is that total effort, $n \times a^* = n \times (\overline{p}^{n^*})^2$, decreases. Thus, the total effort and the total performance can go in opposite directions. Whether it is in the contest organizer's interest to grow the contest therefore depends on what exactly it is she is trying to maximize. ▲

| | $n=2$ | $n=3$ | $n=4$ | $n=5$ | $n=6$ | $n=10$ |
|---|---|---|---|---|---|---|
| $\overline{p}^{n^*}$ | 0.8333 | 0.7143 | 0.6107 | 0.5334 | 0.4744 | 0.3326 |
| $n \times \overline{p}^{n^*}$ | 1.6667 | 2.1429 | 2.4426 | 2.6670 | 2.8462 | 3.325 |
| $a^*$ | 0.6944 | 0.5102 | 0.3729 | 0.2845 | 0.2250 | 0.1106 |
| $n \times a^*$ | 1.3889 | 1.5306 | 1.4916 | 1.4225 | 1.3501 | 1.1060 |

**Table 2:** Equilibrium properties as a function of $n$.

## B.7 Endogenous entry with private outside options

Consider a contest in which types are univariate and capture the value of the outside option. Assume that the ensuing reservation utility, $\overline{u}(\theta_i)$ is continuous and strictly monotonic. For expositional simplicity, assume the type distribution is continuous and has no mass points. These assumptions rule out that there is a mass of types that are indifferent between entry and exit. Write the impact and cost functions as $p(a)$ and $c(a)$, respectively, and let $v$ denote the valuation.

Assume agents are symmetric ex ante and search for a symmetric equilibrium. Let $n$ be the commonly known number of potential entrants. For any $(\overline{p}^{out}, \overline{p}^{in})$ candidate, define $Q_i(a_i, \overline{p}^{out}, \overline{p}^{in})$ as the (type independent) value of (2) when $\overline{p}_j = (\overline{p}^{out}, \overline{p}^{in})$ for all $j \neq i$. Let

$$\begin{aligned} a^*(\overline{p}^{out}, \overline{p}^{in}) &= \arg\max_{a_i} vQ_i(a_i, \overline{p}^{out}, \overline{p}^{in}) - c(a) \\ \Theta^{out} &= \{\theta_i \in \Theta | vQ_i(a^*(\overline{p}^{out}, \overline{p}^{in}), \overline{p}^{out}, \overline{p}^{in}) - c(a^*(\overline{p}^{out}, \overline{p}^{in})) \leq \overline{u}(\theta_i)\}, \end{aligned}$$

both of which are unique given $(\overline{p}^{out}, \overline{p}^{in})$. Then, any symmetric equilibrium $(\overline{p}^{out}, \overline{p}^{in})$ solves

$$\begin{aligned} \overline{p}^{out} &= \Pr\left(\theta_i \in \Theta^{out}\right) \\ \overline{p}^{in} &= \left(1 - \overline{p}^{out}\right) \times p(a^*(\overline{p}^{out}, \overline{p}^{in})). \end{aligned}$$

The second part clarifies that $\overline{p}^{in}$ is the expected impact before knowing whether the agent entered or not. Thus, there are two clues for finding equilibrium: $(i)$ $\overline{p}^{out}$ is determined by an indifference condition (when interior) and $(ii)$ $\overline{p}^{in}$ is determined by the best response upon entry.

EXAMPLE 10: Assume $p(a) = \sqrt{a}$, $c(a) = a$, $a \in [0, 1]$ and that $H(x) = x^2$, $G(x) = x$, $x \in [0, 1]$. Assume that $v = 8$ and that reservation utility is uniformly distributed on $[4, 5]$. These values are calibrated to produce a nice solution when $n = 2$ in the sense that $\overline{p}^{out}$ and $\overline{p}^{in}$ are rational numbers. For $n > 2$, the system described above can be solved numerically. Table 3 reports equilibrium values for various values of $n$ (details are available upon request). In this contest, the more potential entrants there are the more likely each individual agent is to stay out, and any agent who does enter works less hard. While $\overline{p}^{out}$ increases rapidly in $n$, the expected number of entrants decreases only slowly from 1.5 to 1.45. The action and impact upon entry also decline at a slow pace. Since both the expected number of entrants and the impact of entrants decrease in $n$, the expected aggregate performance in the contest decreases in $n$. More formally, the expected total performance is

$$n \times \left(\overline{p}^{in}\mu_H + \left(1 - \overline{p}^{out} - \overline{p}^{in}\right)\mu_G\right) = n \times \left(1 - \overline{p}^{out}\right)\left(p(a^*)\mu_H + \left(1 - p(a^*)\right)\mu_G\right)$$

and both $n \times (1 - \overline{p}^{out})$ and $p(a^*)$ decrease with $n$. Thus, the contest organizer may have an incentive to limit the set of potential participants. ▲

| | $n=2$ | $n=3$ | $n=4$ | $n=5$ | $n=6$ | $n=10$ |
|---|---|---|---|---|---|---|
| $\overline{p}^{out}$ | 0.25 | 0.509 74 | 0.634 96 | 0.708 85 | 0.757 71 | 0.854 81 |
| $\overline{p}^{in}$ | 0.375 | 0.234 41 | 0.172 47 | 0.136 81 | 0.113 48 | 0.067 48 |
| $p(a^*)$ | 0.5 | 0.478 13 | 0.472 47 | 0.469 89 | 0.468 35 | 0.465 93 |

**Table 3:** Equilibrium of a contest with endogenous entry.