Determinants of the Group-Size Paradox*

Martin Kolmar[†], Hendrik Rommeswinkel[‡]

August 6, 2014

Abstract

This paper analyzes the occurrence of the group-size paradox in situations in which groups compete for rents, allowing for degrees of rivalry of the rent among group members. We provide two intuitive criteria for the group-impact function which for groups with homogeneous valuations of the rent determine whether there are advantages or disadvantages for larger groups: social-interactions effects and returns to scale. For groups with heterogeneous valuations, the complementarity of group members' efforts and the composition of valuations are shown to play a role as further factors.

JEL classification: D74, H41

Keywords: Contests, Public Goods, Private Goods, Group-Size Paradox

---

*We thank Stefan Bühler, Philipp Denter, Jörg Franke, Reto Föllmi, Magnus Hoffmann, Nick Netzer, Marco Runkel, Dana Sisak, Felix Várdy, and three anonymous referees for very helpful comments.

†Institute of Economics, University of St. Gallen, Varnbüelstrasse 19, CH-9000 St. Gallen, Email: martin.kolmar@unisg.ch

‡California Institute of Technology, Division of the Humanities and Social Sciences, MC 228-77, 1200 E. California Blvd., Pasadena, CA 91125. Email: hendrik@caltech.edu. Financial support by the Swiss National Science Foundation through grant P1SGP1_148727 is gratefully acknowledged.

1 Introduction

The group-size paradox is perceived as being a result of an unresolved free-rider problem between group members that becomes the more accentuated the larger the group is. Olson (1965) already discussed the alleged advantage of small interest groups over larger ones. His arguments gave rise to a debate about the so-called group-size paradox, which Esteban and Ray (2001) define as: *"larger groups may be less successful than smaller groups in furthering their interests"* (p.663).

The starting point of our paper is to ask which properties of a conflict environment between groups explain the relative advantage or disadvantage of larger compared to smaller groups. We focus on three properties of the group impact function:[1] social-interactions effects, returns to scale, and complementarities between group-members' efforts. As will become clear throughout the paper, all three technological factors are independent. Since we also allow group members to differ in valuations within the group, a fourth crucial property will be the heterogeneity of the valuations of a group. In order to analyze the impact of group size on group performance we use a comparative-static approach where we ask for the effect of adding an additional set of group members to a given group.[2] The main contribution of the paper is the complete characterization of the influence of the above factors on the group-size paradox.

Whereas returns to scale, complementarities in efforts and heterogeneity of valuations are standard concepts, the use of the term social-interactions effects has to be clarified.[3] We say that (positive or negative) symmetric social-interactions effects exist if group impact changes in group size while holding the total group effort constant and distributing it evenly among the group members. There are diverse causes for social-interactions effects in contests such as returns to the division of labor, network effects among group members, or learning between group members.

If group members have equal valuations of winning the contest (which may still differ *between* groups), returns to scale and symmetric social-interactions effects completely deter-

[1] Impact functions are defined as the functions with which individuals transform effort into relative chance of success in a contest (Wärneryd, 2001). Group impact functions correspondingly play the role of production functions with which group members jointly "produce" a higher relative chance of their group winning the contest.

[2] In Appendix K we show that this approach yields the same results as a comparison between groups and argue that it is slightly more general.

[3] The term "social-interactions effect" has a number of different meanings in the literature. Definitions reach from the very narrow concept of direct interdependencies between preferences (Scheinkman, 2008; Bernheim, 1994; Akerlof, 1997) to the very wide concept of aggregative games (e.g. Manski, 2000).

mine whether the group-size paradox occurs or not.[4] Social-interactions effects work in the predictable way: positive symmetric social-interactions effects *ceteris paribus* make it less likely that the group-size paradox occurs. Returns to scale play the role of the discriminatory power of the contest and may thus favor either smaller or larger groups, depending on whether the valuation of winning the contest decreases or increases with group size.

Despite the fact that there is a growing interest in the influence of heterogeneity within and between groups[5], with only a few exceptions the literature on group contests[6] has either focused on homogenous individuals or on situations where the effort levels of group members are perfect substitutes, i.e. are aggregated by summation. For the analysis of the case of heterogeneous individuals, this paper employs CES-impact functions with varying degrees of complementarity. In order to analyze this case, we were able to characterize two useful technical properties that help to simplify future research on group conflicts and comparative statics for CES production functions in general. First, the generalized-mean structure of the impact functions maps onto a generalized mean structure of the valuations of group members that explains equilibrium behavior. This is a quite useful technical property because it allows to analyze the impact of the composition of valuations within a group on group performance using properties of generalized means. Second, we derive a theorem for comparative statics of the elasticity of substitution for a ratio of two generalized means over vectors that differ in heterogeneity. This theorem helps us perform comparative statics in the present model, but is applicable in any other setting where such ratios occur, for example New Keynesian models of inflation, where the inflation rate is a ratio of two CES aggregates.

The effect of adding additional individuals to a group ("new" group members as opposed to "old" ones) depends on the relationship between social-interactions effects and returns to scale on the one hand, and the ratio of power means of the valuations of the group with smaller and larger group size on the other hand. The latter effect is new, and we further explore how adding new group members influences this power mean. As a general conjecture that follows from the above results one would expect that the group-size paradox becomes more likely for higher levels of complementarity between group-members' efforts if adding new group members to an additional group makes the extended group weakly more heterogeneous.

Our paper is most closely related to Esteban and Ray (2001) who argue that in a contest

[4]This claim may appear to be at odds with Esteban and Ray (2001) who focus on convexities in the cost-of effort functions. However, their model is isomorphic to a model with linear costs and nonlinear impact functions that is a special case of our model.

[5]See Esteban and Ray (2008, 2010).

[6]The literature on contests between groups has recently been surveyed by Corchón (2007, Section 4.2), Garfinkel and Skaperdas (2007, Section 7), and Konrad (2009, Chapters 5.5 and 7).

between groups of different sizes, larger groups may profit from cost advantages if the costs of effort are sufficiently convex. In this case, *ceteris paribus*, members of larger groups face sufficiently lower marginal costs that reverse the group-size paradox. This is a very important insight that helps to explain the prevalence of groups in conflicts.[7] Our model differs from the model by Esteban and Ray (2001) in several ways. First, we take a comparative-statics view on the group-size paradox instead of a comparison between groups. As we show in Appendix K, this approach is slightly more general. Moreover, it allows us to transfer our results and methods to other collective action problems as we show in Appendix L. Second, we allow for heterogeneous valuations within a group. Third, given that the model by Esteban and Ray (2001) is isomorphic to a specific contest model with linear costs and impact functions which are sums of concave functions of efforts (Siegel, 2009), their model is a special case of the model analyzed in this paper. In addition, our results are directly relevant for models of Cournot-competition in oligopolistic markets with hyperbolic demand if firms consist of teams (Raab & Schipper, 2009) and team output is some (in general nonadditive) function of team-members' efforts.

The paper is organized as follows. We introduce the model in Section 2 and cover the case of homogeneous group members in Section 3. In Section 4 we allow for heterogeneity of agents and use a CES type impact function to aggregate group members' efforts. We characterize the simultaneous Nash equilibrium and we show the effect of complementarity on the group-size paradox for heterogeneous agents. Section 5 concludes.

2 The model

Assume that n groups compete for a given rent. Let $\overline{m} \in \mathbb{N}$ be the maximum possible number of group members and let $m_i \in 2, \ldots, \overline{m}$ be the number of individuals in group i where k is the index of a generic member of this group. We refer to the set of group members by $M_i = \{1, \ldots, m_i\}$. The rent can be completely rival or completely non-rival between group members, and every intermediate case where additional group members dilute the value of the rent for the remaining group are also taken into consideration. To cover these cases it suffices to assume that the valuation of the rent for each individual k of group i is a function of the size of the group, $v_i^k(m_i) > 0$. If $v_i^k(\hat{m}_i) < v_i^k(m_i)$, whenever $\hat{m}_i > \tilde{m}_i$, then the rent is partly rival among group members as some degree of crowding is involved as group size is

[7] See also Marwell and Oliver (1993); Pecorino and Temini (2008); Nitzan and Ueda (2009, 2010).

increased. If $v_i^k(\hat{m}_i) = v_i^k(m_i)$ for all $\hat{m}_i, m_i$ the rent is a group-specific public good[8]. In the following it will be assumed that $v_i^k(\hat{m}_i) \leq v_i^k(m_i)$ whenever $\hat{m}_i > m_i$. [9]

Sometimes it will be necessary to refer to vectors of valuations of (subsets of) the group members: $\vec{v}_{i,M}(m_i) \equiv (v_i^{k_1}(m_i), \ldots, v_i^{k_{\sharp M}}(m_i))$ where $M \subset M_i$ and $v_i^{k_1}(m_i)$ refers to the valuation of the first element of M.[10] This somewhat elaborate notation is necessary since later on we will analyze comparative statics if sets of individuals are added to a group. It is the easiest to think of $\vec{v}_{i,M}(m_i)$ as the vector of valuations of a subset of group members M of group i if the total group size is m_i.

p_i represents the probability of group $i = 1, ..., n$ to win the contest. Individuals can influence the winning probability by contributing effort x_i^k. The group members' efforts are then aggregated by a function $q_i(x_i^1, ..., x_i^{m_i}) = q_i(\vec{x}_i) \geq 0$. Following the literature, it will be called *impact function*. Since we are most of the time interested in comparative statics with respect to the size of a single group, we define a class of impact functions for this group to specify the impact functions which are used at different sizes of the group $\{q_{i,m_i}(x_i^1, ..., x_i^{m_i})\}_{m_i=2}^{\overline{m}}$. The winning probability p_i is a function of these impacts. $p_i(.)$ is called a *contest-success function*. We focus on Tullock-form contest-success functions where the winning probability of a group i is defined as:[11]

Assumption 1.

$$p_i(q_1, ..., q_n) = \begin{cases} \dfrac{q_i}{\sum_{j=1}^{n} q_j}, i = 1, ...n, & \exists j : q_j > 0 \\ \dfrac{1}{n}, & \forall j : q_j = 0 \end{cases}.$$

Further, we impose the following assumptions on the individuals:

Assumption 2. Individuals are risk neutral, face linear costs, and maximize their net rent.

[8]See Cornes and Sandler (1996) for a precise and ample discussion of different types of public goods with crowding. Note that the linear case (as for example in Esteban & Ray, 2001) $v_i^k = \alpha \frac{w}{m_i} + (1-\alpha)w'$, where w is the utility from the rival dimensions (with an equal-sharing rule being applied) of the rent and w' the utility from the non-rival part, is a special case of our formulation.

[9]It is also possible to consider other cases, but for reasons of space these will only shortly be discussed.

[10]To illustrate this notation assume that group i has three members, $m_i = 3$ and $M_i = (1, 2, 3)$ with valuations $v_i^1(3) = 5, v_i^2(3) = 10, v_i^1(3) = 15$. Let $M' = (1, 2)$ and $M'' = (2, 3)$ be two subsets of group members. Then, $\vec{v}_{i,M'}(3) = (5, 10)$ and $\vec{v}_{i,M''}(3) = (10, 15)$.

[11]An axiomatic foundation for the Tullock function for group contests can be found in Münster (2009). An interpretation of the Tullock contest as a perfectly discriminatory noisy ranking contest can be found in Fu and Lu (2008).

Assumptions 1 and 2 imply that we can write expected utility as:

$$\pi_i^k(x_i^k, \vec{x}_{-x_i^k}) = \frac{q_i(x_i^1, ..., x_i^{m_i})}{\sum_{j=1}^n q_j(x_j^1, ..., x_j^{m_j})} v_i^k(m_i) - x_i^k. \tag{1}$$

We are looking for Nash equilibria of this game where individuals choose their effort x_i^k simultaneously to maximize their expected utility,

$$x_i^{k*} \in \arg\max_{x_i^k} \pi_i^k(x_i^k, \vec{x}^*_{-x_i^k}) \quad \forall i, k. \tag{2}$$

where "$*$" refers to equilibrium values and $\vec{x}^*_{-x_i^k}$ to the vector of efforts by all individuals except k in group i. In order to facilitate the analysis, we will focus on situations where a unique Nash equilibrium with respect to the total effort produced by each group exists. Formally,

Assumption 3. $q_i(.)$ is at least twice continuously differentiable,
$\forall k, \vec{x}_i$: $\frac{\partial q(\vec{x}_i)}{\partial x_i^k} > 0$,
$\forall k, \vec{x}_i$: $\frac{\partial^2 q(\vec{x}_i)}{\partial (x_i^k)^2} \leq 0$, and
$\forall \lambda \geq 1, k, \vec{x}_i$: $q_i(\lambda \vec{x}_i) \leq \lambda q_i(\vec{x}_i)$.[12]

Assumption 4. $q_i(.)$ has symmetric partial derivatives at $\{x, ..., x\}$, i.e. $\partial q_i(x, ..., x)/\partial x_i^k = \partial q_i(x, ..., x)/\partial x_i^l \; \forall x \; \forall \; k, l \; \forall \; i$

Assumption 5. If $\vec{x}_i$ is such that $x_i^k > x_i^l$, then $\frac{\partial q_i(\vec{x}_i)}{\partial x_i^k} < \frac{\partial q_i(\vec{x}_i)}{\partial x_i^l}$.

In some of the below results we also need the assumption that the impact functions are homogeneous.

Assumption 6. $q_i(.)$ is homogeneous of degree r_i, i.e. $\forall \lambda \geq 0, \vec{x}_i$: $q_i(\lambda \vec{x}_i) = \lambda^{r_i} \cdot q_i(\vec{x}_i)$.

[12] These assumptions rule out impact functions with for example hyperbolic (Cobb-Douglas) or L-shaped (perfect complements) indifference curves. Impact functions with $\partial q_i(0, ..., 0)/\partial x_i^k = 0$ usually lead to multiple equilibria because $\{0, ..., 0\}$ at the group as well as as the aggregate level is always a Nash equilibrium. See Skaperdas (1992) for an extensive discussion in a somewhat different context. This would cause additional and merely technical problems that would divert attention from the main focus of the paper.

3 Homogeneous valuations within groups

Before we turn to the analysis of the effects of group size on winning probabilities, we establish that a unique equilibrium exists. A proof of existence of a Nash equilibrium cannot rely on standard fixed-point arguments because with a Tullock lottery contest the best-response function for an individual k of group i are not well defined if all other groups exert zero effort, and the alternative approach to make use of the aggregative nature of contests does not work because group contests lack such an aggregative structure.[13] The proof of this and all of the following results can be found in the appendix.

Theorem 1. *Suppose a contest fulfills Assumptions 1, 2, 3, and 4 for all groups i. Then, a Nash equilibrium exists where the equilibrium efforts are symmetric such that $\forall i, k : x_i^{k*} = x_i^*$. There exists only one symmetric equilibrium given $\forall i, k : x_i^{k*} = x_i^*$.*

Therefore, under the given assumptions there may exist other, nonsymmetric equilibria. Using a stronger assumption instead of Assumption 4 we obtain a unique equilibrium:

Theorem 2. *Suppose a contest fulfills Assumptions 1, 2, 3, and 5 for all groups i. Then, there exists a unique Nash equilibrium.*

Since Assumption 4 is weaker than 5, some cases are of course not covered by the latter theorem. Most prominent is the case of additive impact functions where infinitely many equilibria exist in which only the level of total effort of each group is fixed.

In some cases, a group may decide to exhibit zero effort, which implies that it makes sense to distinguish between active and inactive groups:

Definition 1. (**Participation**) An individual k of group i is said to *participate* if $x_i^{k*} > 0$. A group i is said to *participate* if there exists some k such that $x_i^{k*} > 0$. A group is said to *fully participate* if $\forall k : x_i^{k*} > 0$.

The group-size paradox was first discussed by Olson (1965), who stated that *"the larger the group, the farther it will fall short of providing an optimal amount of a collective good"* (p. 35). One particular interpretation of the statement has been given by Esteban and Ray (2001): In a contest environment in which different groups compete for a rent, larger groups should win with lower probability if the group-size paradox was true. We take a comparative-static perspective on the group-size paradox:

[13] See Acemoglu and Jensen (2009) for a definition of aggregative games.

Definition 2. (**Group-size paradox**) Suppose there are n groups competing for a prize and each group $j \neq i$ consists of a set $M_j = \{1, \ldots, m_j\}$ of individuals with equal valuations v_j. Let group i have either members $M_i = \{1, \ldots, m_i\}$ or $\hat{M}_i = \{1, \ldots, \hat{m}_i\}$ with $m_i < \hat{m}_i$ with valuations $v_i(m_i)$ and $v_i(\hat{m}_i)$, respectively. Let the corresponding equilibrium winning probabilities be p_i^* and $\hat{p}_i^*$. Then the group-size paradox holds strictly (weakly) for group i at sizes m_i and $\hat{m}_i$ if and only if $p_i^* > (\geq) \hat{p}_i^*$.

In order to have a simple language we will refer to M_i as the "old" group members and to $\Xi_i = \hat{M}_i / M_i$ as the "new" group members in the following. The definition of the group size paradox is therefore local with respect to the original group size m_i and the size of the group after the increase in group members, $\hat{m}_i$. We will give precise conditions under which the group-size paradox occurs but it may well be the case that an impact function is such that the group size paradox only occurs for small group sizes but not for large ones or vice versa. One can naturally also take the perspective of a comparison across groups of different size in the same contest. In Appendix K, we show for all propositions in this paper that the comparative static perspective on the group size paradox yields the same results as a comparison of winning probabilities across groups. However, for some cases that can be analyzed via the comparative-static perspective, no corresponding contest exists which can be analyzed by comparing groups of different size in the same contest. Moreover, using our approach one can also analyze the group-size paradox in other collective action problems, as we show in Appendix L.

We will also consider welfare effects and their relation to the group size paradox.

Definition 3. (**Group Welfare**) The total group welfare is defined as the sum of expected utilities $\pi_i^T = \sum_{k \in M_i} \pi_i^k (x_i^{k*}, \vec{x}^*_{-x_i^k})$ and the average group welfare is defined as $\pi_i^A = \frac{1}{m_i} \pi_i^T$.

Next we formulate two intuitive criteria that will turn out to be able to explain the occurrence of the group-size paradox if individuals of a group have identical valuations of the rent. The first one defines the concept of social-interactions effects for within-group symmetric effort contributions.

Definition 4. (**Symmetric Social-interactions effects (SSIE)**) A class of impact functions $\{q_{m_i}(x_i^1, \ldots, x_i^{m_i})\}_{m_i=2}^{\overline{m}}$ with m_i being the number of group members is said to have *absent (positive, negative) symmetric social-interactions effects* at effort level x_i for an increase in group size from m_i to $\hat{m}_i$ if it holds that $q_{\hat{m}_i}(\frac{x_i m_i}{\hat{m}_i} \ldots \frac{x_i m_i}{\hat{m}_i}) = (>, <)\, q_{m_i}(x_i, \ldots, x_i)$.

Definition 4 can be used to define a measure of SSIE:

Definition 5. For a class of impact functions, $\{q_{m_i}(x_i^1, \ldots, x_i^m)\}_{m_i=2}^{\overline{m}}$, SSIE are measured by $s_i(x_i, m_i, \hat{m}_i) = q_{\hat{m}_i}(x_i m_i/\hat{m}_i, \ldots, x_i m_i/\hat{m}_i)/q_{m_i}(x_i, \ldots, x_i)$. SSIE are absent (positive, negative), if and only if $s_i(x_i, m_i, \hat{m}_i) = (>, <)\ 1$.

Note that SSIE are defined as a local measure and may change for different values of x_i, m_i, and $\hat{m}_i$. If impact functions are homogeneous, then $s_i(\ldots)$ does not depend on x_i and can be written as $s_i(m_i, \hat{m}_i)$.

To gain intuition it is instructive to look at an impact function that is the sum of efforts of all group members, $\sum_{k=0}^{m_i} x_i^k$. This function has absent SSIE: If all group members k exert the same effort $x_i^k = x_i$, then $\sum_{k=1}^{m_i} x_i = m_i \cdot x_i = \sum_{k=1}^{\hat{m}_i} x_i m_i/\hat{m}_i$. In this case, adding additional group members has no influence on the productivity of the group, social-interactions effects are absent.

Another property of an impact function is its returns to scale:

Definition 6. (**Returns to scale (RTS)**) A class of impact functions $\{q_{m_i}(\vec{x}_i)\}_{m_i=2}^{\overline{m}}$ is said to have *constant (increasing, decreasing) returns to scale* if $\forall m_i : q_{m_i}(\lambda \vec{x}_i) = (>, <)\ \lambda \cdot q_{m_i}(\vec{x}_i)$ where $\lambda > 0$.

Based on this definition, it is plausible to measure returns to scale in the following way:

Definition 7. For a class of homogeneous impact functions, $\{q_{m_i}(\vec{x}_i)\}_{m_i=2}^{\overline{m}}$, returns to scale are measured by the degree of homogeneity r_i, such that for all m_i: $q_{m_i}(\lambda \vec{x}_i) = \lambda^{r_i} \cdot q_{m_i}(\vec{x}_i)$

Note that Definition 7 immediately implies that if we speak of a class of impact functions having certain returns to scale, each of the impact functions of this class has the same returns to scale. Further, since we focus on concave impact functions (see Assumption 3), the results will only be stated for decreasing or constant returns to scale. However, our results also hold in those cases where even with increasing returns to scale there still exists a unique interior equilibrium. One example would be the case of two groups with symmetric valuations $v_i = v_j$ and $r_i = r_j < 2$.

Both properties, SSIE and RTS are independent: Assume that the impact functions have the generalized CES-form

$$q_{m_i}(\vec{x}_i) = m_i^{s_i + r_i} \cdot \left(\frac{1}{m_i} \sum \left(x_i^k \right)^{\gamma_i} \right)^{r_i/\gamma_i}. \tag{3}$$

In this case, $q_{\hat{m}_i}(\frac{x_i m_i}{\hat{m}_i}, \ldots, \frac{x_i m_i}{\hat{m}_i}) = (\hat{m}_i/m_i)^{s_i} \cdot q_{m_i}(x_i, \ldots, x_i)$ and $q_{m_i}(\lambda x_i, \ldots, \lambda x_i) = \lambda^{r_i} q_{m_i}(x_i, \ldots, x_i)$, which shows that RTS and SSIE can be chosen independently.

Before presenting the main results of this section it makes sense to discuss them informally. In the graphs in Figure 1, RTS are measured along the ordinate and SSIE, which are independent of efforts under Assumption 6, are measured along the abscissa. At the point $\{1, 1\}$, SSIE are absent and RTS are constant. Moving right from this point creates positive and moving left creates negative SSIE, moving downwards reduces RTS.

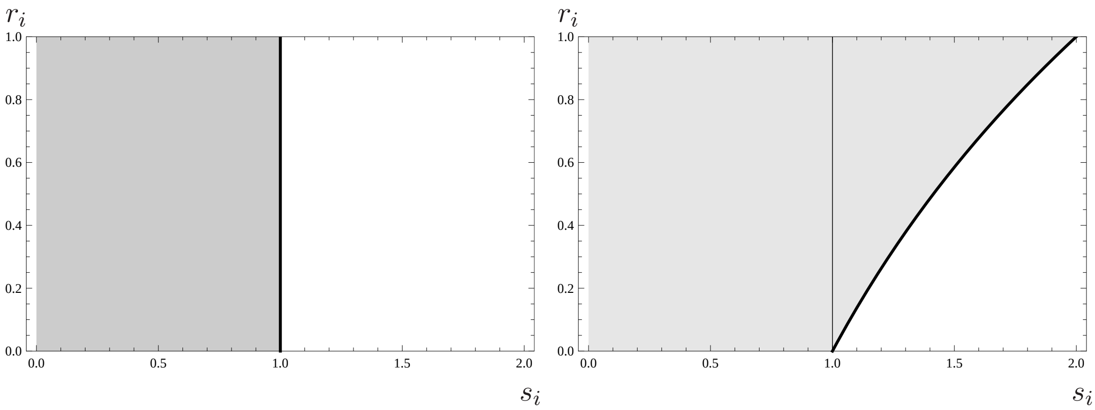

Figure 1: Group-size paradox for the case of a non-rival rent (left) and for the case of crowding (right).

The left panel of Figure 1 focuses on the special case that the rent is a pure public good among group members. In the case of a perfectly non-rival rent, only SSIE turn out to be relevant for the occurrence of the group-size paradox: the winning probability decreases in group-size in the left quadrant (shaded gray) whereas it increases in group size in the right quadrant (shaded white).

Allowing for crowding makes the group-size paradox more likely, due to the dilution of per-capita rents that follows from increases in group-size. The separating line moves to the right compared to the case of a pure public good and is given by the upward-sloping line in the right panel of Figure 1. The group-size paradox again holds in the gray shaded areas of the figure, which means the RTS now also play a role. The adverse effect of crowding must be compensated by an increase in SSIE, and the increase has to be the stronger, the larger the RTS. This is due to the fact that the RTS of the impact function control the discriminatory power of the contest with respect to the average valuation of that group. If due to crowding there is an inherent disadvantage from larger group size, then this disadvantage is amplified by higher RTS.

Though unlikely, we can also imagine cases where an increase in group size leads to an increase in the valuation. In this case, the RTS will favor the larger group, as evident from the dotted line in Figure 2.

This is again due to the role of the RTS as the discriminatory power, which amplifies the

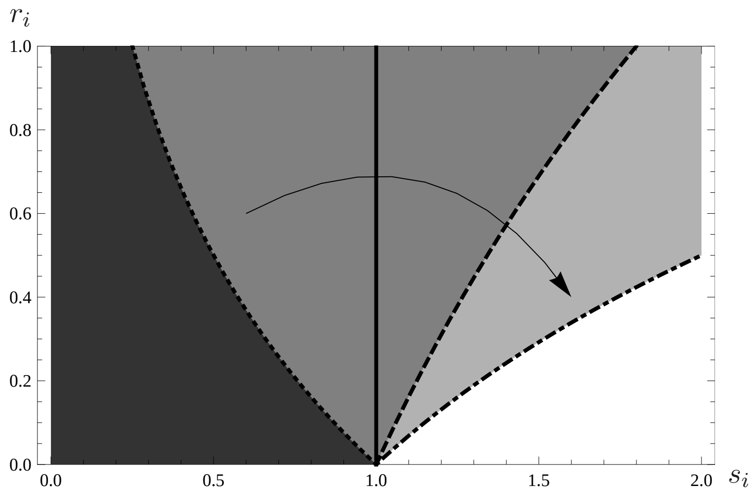

Figure 2: Group-size paradox for increasing rivalry of the rent.

effect of differences in valuation on the winning probabilities. Since larger groups have higher valuations than smaller ones, they are favored by large RTS and therefore the larger the RTS, the lower the SSIE must be in order for the group-size paradox to occur. Figure 2 also shows the effect of an increase in rivalry on the occurrence of the group-size paradox. The dividing line pivots clockwise around the point of zero RTS and absent SSIE. The more rival the rent becomes, the higher the level of SSIE that is necessary to compensate for the increase in the dilution of the rent.

We now turn to the formal presentation of the results.

Proposition 1. *Consider two contests fulfilling Assumptions 1, 2, 3, and 4 for all groups, which differ only in the group size of group i, m_i and $\hat{m}_i > m_i$. For all j, k: $v_j^k = v_j$ and let the equilibrium winning probabilities in the symmetric equilibrium be p_i^* and $\hat{p}_i^*$, respectively. Group i participates at group size m_i with effort level x_i^*. The class of impact functions of group i has $s_i(x_i^*, m_i, \hat{m}_i) = 1$ and constant or decreasing RTS. If $v_i(m_i) = v_i(\hat{m}_i)$, then $p_i^* = \hat{p}_i^*$, $\pi_i^A < \hat{\pi}_i^A$, and $\pi_i^T < \hat{\pi}_i^T$. If $v_i(m_i) > v_i(\hat{m}_i)$, then $p_i^* > \hat{p}_i^*$ and $\exists \underline{v_i(\hat{m}_i)} : \pi_i^T \geq \hat{\pi}_i^T$.*

The case of a pure public good establishes a link between our model and the special case of additively linear impact functions which have been standard in the literature so far (e.g. Baik, 2008; Konrad, 2009, Chapters 5.5 and 7). For the case of non-rival rents, the equilibrium group impact and the winning probability are independent of group size as long as the valuation remains unchanged. This leads to a welfare advantage for larger groups. If rents are rival, the increasing dilution of rents (and therefore lower marginal returns) for larger group sizes, bring larger groups into a worse position. If the rent is sufficiently rival, both total and average welfare will decrease after an increase in group size.

The results on the group-size paradox can be strengthened if we assume that the impact functions are homogeneous and allow for SIE. However, welfare effects will be less clear in this case:

Proposition 2. *Consider two contests fulfilling Assumptions 1, 2, 3, and 4 for all groups, which differ only in the group size of group i, m_i and $\hat{m}_i > m_i$. For all j, k: $v_j^k = v_j$ and let the equilibrium winning probabilities in the symmetric equilibrium be p_i^* and $\hat{p}_i^*$, respectively. The class of impact functions $\{q_{m_i}(.)\}_{m_i=2}^{\overline{m}}$ fulfills Assumption 6 with $s_i(m_i, \hat{m}_i)$ as the measure of SSIE. Suppose group i participates at group size m_i. Then:*

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad \frac{v_i(m_i)}{v_i(\hat{m}_i)} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i} \tag{4}$$

$$\pi_i^A \gtreqless \hat{\pi}_i^A \quad \Leftrightarrow \quad p_i^* v_i(m_i)\left(1 - (1 - p_i^*)\frac{r_i}{m_i}\right) \gtreqless \hat{p}_i^* v_i(\hat{m}_i)\left(1 - (1 - \hat{p}_i^*)\frac{r_i}{\hat{m}_i}\right) \tag{5}$$

$$\pi_i^T \gtreqless \hat{\pi}_i^T \quad \Leftrightarrow \quad p_i^* m_i v_i(m_i)\left(1 - (1 - p_i^*)\frac{r_i}{m_i}\right) \gtreqless \hat{p}_i^* \hat{m}_i v_i(\hat{m}_i)\left(1 - (1 - \hat{p}_i^*)\frac{r_i}{\hat{m}_i}\right). \tag{6}$$

In other words, if for some class of impact functions the group-size paradox holds, then increasing the RTS or decreasing the SSIE further will imply that the group-size paradox still holds if the rent is partly rival. The reverse holds for classes of impact functions for which the group-size paradox does not hold: With crowding, decreasing the RTS or increasing the SSIE will imply that for the new class of impact functions, the group-size paradox also does not hold. It also follows from the proposition as a corollary that RTS play no role in the case of non-rival rents, since in that case the LHS of (4) equals one. In the case of $v_i(m_i) < v_i(\hat{m}_i)$, the effect of the SSIE remains the same, but the effect of the RTS is opposite: If for some class of impact functions the group-size paradox holds, then it will *ceteris paribus* continue to hold under lower RTS, but not necessarily under higher RTS. An example where $v_i(m_i) < v_i(\hat{m}_i)$ is meaningful is the case when groups can rely on mechanisms to internalize within-group externalities.[14]

Solving (4) for r_i further reveals that (a) the locus of RTS–SSIE pairs that constitute the dividing line between group-size paradox and no group-size paradox has a positive slope in Figures 1 and 2 and (b) an increase in the privateness of the rent shifts this dividing line in the direction of either more increasing SSIE and/or lower returns to scale as seen in Figure 2.

[14] Suppose a group has access to a mechanism solving its collective action problem. In this case, agents fully internalize their effect on the payoff of others and thus optimize as if their valuation of the rent were $v_i(m_i) = \sum_{k=1}^{m_i} v_i^k(m_i)$. Therefore, equilibrium efforts (and thus winning probabilities) will be equal to those obtained in a contest with a homogeneous group and valuations $v_i(m_i)$. If the rent is not too rival and the new group members' valuations are not too low, we will further have $v_i(m_i) < v_i(\hat{m}_i)$ and thus the described case. For details see Kolmar and Wagener (2011).

The above analysis shows that SSIE and RTS fully explain the occurrence of the group-size paradox if individuals of the same group have the same valuation of the rent. They enable us to understand how the technological and cultural determinants of group impact influence the relative success of larger or smaller groups.[15] For the non-rival case, the case of absent SSIE is the watershed for the existence of the group-size paradox so that this very simple rule is easy to check empirically. In case that dilution is important, empirical tests are more difficult because the quantitative extent of SSIE becomes important, but it nevertheless gives a clear guideline.

The welfare effects do not follow such a clear pattern, since we cannot solve for them explicitly. While it is obvious that an increase in winning probability ceteris paribus increases average and total group welfare, it becomes clear from setting $p_i^* = \hat{p}_i^*$ in (6), that even for equal winning probabilities it is not clear whether $\pi_i^T > \hat{\pi}_i^T$ or the opposite holds. Also, the public good case where in Proposition 1 larger groups still held an advantage may have $\pi_i^T > \hat{\pi}_i^T$ if SIE are sufficiently low.

4 Heterogeneous valuations within groups

While the literature on the group-size paradox has focused on the case of homogeneous groups, we will now proceed to examine the heterogeneous case. Naturally then, the above mentioned connection between an analysis relying on cost functions and one allowing for more general impact functions with SSIE and RTS as the main properties no longer holds. Since individuals may have different valuations, they may end up with different marginal returns on impact to effort. The following analysis however establishes that SSIE and RTS continue to play the same role, thus generalizing the results from the previous section. We also introduce a further parameter that will gain importance, the complementarity between members' efforts, whose effect depends on the heterogeneity of the new and old group members.

To simplify the analysis, we concentrate on a CES impact function with SSIE given by $s_i(m_i, \hat{m}_i) = (\frac{\hat{m}_i}{m_i})^{s_i}$ and RTS r_i. These properties are fulfilled by the following CES-type impact function:[16]

[15] In Appendix L we show that this also holds for voluntary contriution games as Bergstrom, Blume, and Varian (1986).

[16] It could be argued that a class of impact functions should fulfill a condition such as $\forall m_i, \hat{m}_i$: $q_{m_i}(x_i^1, \ldots, x_i^{m_i}) = q_{\hat{m}_i}(x_i^1, \ldots, x_i^{m_i}, 0, \ldots, 0)$. We do not require (and indeed violate) this for the following reason: Consider 100 individuals participating in a demonstration. It may matter for their impact on policies a lot whether they belong to a group of 100 or 1000 individuals affected by that policy. The notion of complementarity captures this: High complementarity means that an interest group of 1000 individuals will only

Assumption 7. $q_i(x_i^1, ..., x_i^{m_i}) = m_i^{s_i+r_i} \cdot \left(\sum_{l=1}^{m_i} \frac{1}{m_i}(x_i^l)^{\gamma_i}\right)^{r_i/\gamma_i}$, $\gamma_i \in (0,1)$, $r_i \in (0,1]$, $s_i \in \mathbb{R}$ $i = 1, ...n$.

γ_i accounts for different elasticities of substitution of the group members' efforts.[17] Since the CES-type impact function is essentially a power mean of the contributions, and power means will play an important role in the following, it is useful to introduce them formally:

Definition 8. (**Power Mean**) If $\vec{a}$ is a vector with s elements $a_1, a_2, \ldots, a_s$, then the θ-power mean of $\vec{a}$ is defined as:

$$\mathcal{M}(\vec{a}, \theta) \equiv \left(\sum_{i=1}^{s} \frac{a_i^\theta}{s}\right)^{1/\theta}. \tag{7}$$

Therefore, we can express the CES-type impact function as: $q_i(\vec{x_i}) = m_i^{s_i+r_i} \cdot \mathcal{M}(\vec{x_i}, \gamma_i)^{r_i}$. To analyze the interplay of γ_i and the heterogeneity of a group, one needs a tractable definition of heterogeneity. The most common idea associated with higher heterogeneity is that of a mean-preserving spread:

Definition 9. A vector $\vec{v} = (v^{(1)}, \ldots, v^{(m)})$ is a θ-power mean preserving spread of a vector $\vec{v'} = (v'^{(1)}, \ldots, v'^{(m)})$ if there exist i, j such that $v^{(i)} > v'^{(i)} \geq v'^{(j)} > v^{(j)}$ with $\mathcal{M}(\vec{v}) = \mathcal{M}(\vec{v'})$ and $v^{(k)} = v'^{(k)}$ for all $k \neq i, j$.

The definition of a power mean preserving spread differs from mean preserving spreads by Rothschild and Stiglitz (1970) in two important ways: First, it is generalized to power means since – as discussed before – an arithmetic mean preserving spread of valuations is not always neutral to the winning probability of a group. Second, it is restricted by the assumption of equal weights of each element, since in the CES impact function employed here, all individual efforts have equal weights. We may want to compare groups with different average effort levels and thus employ a slightly more general definition of heterogeneity than power mean preserving spreads:

Definition 10. $\vec{v'}$ is more heterogeneous than $\vec{v}$ at mean parameter θ if $\vec{v'}$ is a permutation of $\vec{v''} \cdot \omega$ where $\omega \in \mathbb{R}^+$ and $\vec{v''}$ can be obtained from a sequence of θ-power mean preserving spreads of $\vec{v}$.

have an impact if all individuals demonstrate and not only a subset. The above condition however violates this intuition.

[17] We restrict attention to $\gamma_i \in (0,1)$ to guarantee uniqueness of the equilibrium. If $\gamma_i \leq 0$ multiple equilibria can occur because of a within-group coordination failure: If at least one group member exerts zero effort, group impact is zero and it is rational for the other group members to also exert zero effort. However, Propositions 3 and 4 continue to hold for $\gamma_i \leq 0$ in all but the extreme equilibrium where all members of all groups exert zero effort.

According to this definition, a vector is more heterogeneous if it can be obtained from another vector via the application of power mean preserving spreads and multiplying it with a positive constant. From the definition of heterogeneity, the following theorem follows.

Theorem 3. *Suppose $\vec{v}'$ is more heterogeneous than $\vec{v}$ at power mean parameter θ, then:*

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \frac{\mathcal{M}(\vec{v}',\theta)}{\mathcal{M}(\vec{v},\theta)} \gtreqless \frac{\mathcal{M}(\vec{v}',\phi)}{\mathcal{M}(\vec{v},\phi)}$$

While this theorem will be applied in the context of contests in this paper, it is applicable in many other settings with heterogeneity and CES aggregates. For example, it also applies to ratios of price indices P_{t+1}/P_t in models with monopolistic competition and heterogeneous producers (e.g. New Keynesian models such as Yun, 1996). Theorem 3 can be used in these contexts to analyze the effects of changes in the elasticity of substitution on the inflation measure if heterogeneity differs across periods. Similarly, growth rates of consumption in models with CES production functions can be analyzed.

It follows from Assumptions 1, 2, and 7 that the individual expected utility functions are as follows:

$$\pi_i^k(x_1^1,...,x_n^{m_n}) = v_i^k(m_i)\frac{m_i^{s_i+r_i} \cdot \mathcal{M}(\vec{x_i},\gamma_i)^{r_i}}{\sum_j m_j^{s_j+r_j} \cdot \mathcal{M}(\vec{x_j},\gamma_j)^{r_j}} - x_i^k. \tag{8}$$

The Nash equilibrium of this model can only be obtained explicitly for the case $r_i = 1$.[18] For $r_i < 1$ it turns out that comparative statics results can still be derived. We proceed as follows: First, existence and uniqueness of the Nash equilibrium will be proven. Second, it will be shown that the winning probability of a group is strictly increasing in an aggregate valuation V_i of the group (to be determined). This reduces the question of whether the group-size paradox holds to the question whether V_i increases or decreases after adding a set of individuals to the group. Third, we will examine how various combinations of heterogeneity and complementarity affect V_i.

Theorem 4. *Suppose a contest fulfills Assumptions 1, 2, 7 for all groups. Then, a unique Nash equilibrium exists in which $\forall r_i < 1$ all groups fully participate and $\forall r_i = 1$, $n^* \geq 2$ groups fully participate.*

Having established existence and uniqueness of the Nash equilibrium, we can now turn to the comparative-static analysis. It follows from the proof of Theorem 4 that if Q^* is the equilibrium total impact, the following equilibrium relation must hold for all participating

[18] These results are given in Kolmar and Rommeswinkel (2013). For our purposes, explicit results are not necessary.

groups i with members M_i:

$$V_i \cdot (1 - p_i^*) = (Q^*)^{1/r_i} \cdot (p_i^*)^{1/r_i - 1}, \tag{9}$$

where $V_i \equiv r_i m_i^{s_i/r_i} \cdot \mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma_i}{1-\gamma_i})$. Notice that (as discussed in Section 2) $\vec{v}_{i,M_i}(m_i)$ is a vector valued function of total group size m_i. For example, if there are three members in M_i, then $\vec{v}_{i,M_i}(5)$ would give the vector of valuations which these three members would have if the actual group size was five. $\mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma_i}{1-\gamma_i})$ is then a power mean of these valuations.

In the following the term "average valuation" will refer to this power mean (which does not have to coincide with the arithmetic mean). There is a tight relation between changes in p_i^* and V_i which we can use for comparative statics of our model:

Theorem 5. *Consider two contests fulfilling Assumptions 1, 2, 7 for all groups, which differ only by the set of group members of group i, M_i and $\hat{M}_i$ and their valuations $\vec{v}_{M_i}$, $\vec{v}_{\hat{M}_i}$,. Let the equilibrium winning probabilities in each equilibrium be p_i^* and $\hat{p}_i^*$, respectively. Then, $p_i^* \geq \hat{p}_i^*$ if and only if $V_i \geq \hat{V}_i$.*

The above theorem holds for changes in valuations and group size from M_i to $\hat{M}_i$. Hence, we can obtain comparative-static results on p_i^* by only examining the effect of a change in group size on V_i. The question whether a change in group size increases or decreases the winning probability of that group reduces to whether the change in group size increases or decreases V_i. This is a noteworthy result because it implies that the strategic interaction *between* groups has no *qualitative* influence on the comparative-static properties of the model. We show this in Appendix L by extending our results to voluntary contribution games with linear costs of effort.

The following proposition summarizes the effect of adding a set of individuals Ξ_i to group i on its winning probability:

Proposition 3. *Consider two contests fulfilling Assumptions 1, 2, 7 for all groups, which differ only by the set of group members M_i and $\hat{M}_i = M_i \cup \Xi_i$. Let the equilibrium winning probabilities in each equilibrium be p_i^* and $\hat{p}_i^*$, respectively. Suppose group i participates at group size m_i. Then:*

$$p_i^* \gtreqless \hat{p}_i^* \Leftrightarrow \frac{\mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma_i}{1-\gamma_i})}{\left(\frac{m_i}{\hat{m}_i} \cdot \mathcal{M}(\vec{v}_{i,M_i}(\hat{m}_i), \frac{\gamma_i}{1-\gamma_i})^{\frac{\gamma_i}{1-\gamma_i}} + (1 - \frac{m_i}{\hat{m}_i}) \cdot \mathcal{M}(\vec{v}_{i,\Xi_i}(\hat{m}_i), \frac{\gamma_i}{1-\gamma_i})^{\frac{\gamma_i}{1-\gamma_i}}\right)^{\frac{1-\gamma_i}{\gamma_i}}} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i}. \tag{10}$$

The proposition shows that the results for heterogeneous groups are very similar to the ones derived for the case of homogeneous groups given in Proposition 2. However, because of the heterogeneity of the group, there is no longer an obvious choice for the valuation of the new group members. The LHS of the above expression may therefore be smaller than one even with crowding if high-valuation individuals join the group and the crowding effect is not too strong. In this case, the effect of the RTS is opposite to the effect we have observed for homogeneous groups: The higher the RTS, the lower the minimal SSIE such that the group-size paradox does not occur. The addition of new group members with high valuations may therefore compensate for average valuation losses due to crowding. With this exception, all other results from the homogeneous valuation analysis carry over to the heterogeneous case.

The relaxation of the assumption of homogeneous valuations introduces another important property which has an influence on the performance of large and small groups and that is somewhat hidden in Proposition 3. The complementarity of efforts influences the degree to which the new group members influence group effort.

To better understand the interplay of group heterogeneity and complementarity, we impose a further assumption to simplify the LHS of (10):

Assumption 8. The valuation of an individual k in group i is given by $v_i^k(m_i) = w_i^k \alpha_i(m_i)$, where $\alpha_i : \mathbb{N}_+ \to \mathbb{R}$ is a weakly decreasing function.

This assumption encompasses the often used functional form $v_i^k = \alpha w_i^k + (1-\alpha) w_i^k / m_i$ by setting $\alpha_i(m_i) = \alpha + (1-\alpha)/m_i$, where α denotes the fraction of the rent that is a public good. Other functional forms are also possible, for example $\alpha_i(m_i) = 1/m_i^{1-\alpha}$ corresponds to Cobb-Douglas preferences of the form $v_i^k = (w_i^k)^\alpha (w_i^k/m_i)^{1-\alpha}$. It is however restrictive in the sense that it does not allow for heterogeneity in the way group members' valuations respond to additional group members which join the group: All valuations in the group are reduced by a common factor.

Assumption 8 yields a natural measure for the degree of rivalry in the rent:

Definition 11. The degree of rivalry of the rent is measured by:

$$R_i(m_i, \hat{m}_i) = \frac{\alpha_i(m_i)}{\alpha_i(\hat{m}_i)}. \tag{11}$$

Notice that $R_i(m_i, \hat{m}_i) \geq 1$ if we focus the analysis on public goods and rents that are partly rival. Also, under homogeneous valuations, this ratio is equivalent to the LHS ratio in (4), which neatly extends the homogeneous case.

Assumption 8 allows us to simplify (10) and obtain comparative statics on γ_i for cases where new and old group members can clearly be ranked in their heterogeneity:

Proposition 4. *Consider two contests fulfilling Assumptions 1, 2, 7 for all groups, which differ only by the set of group members M_i and $\hat{M}_i = M_i \cup \Xi_i$. The valuations of M_i fulfill Assumption 8. Let the equilibrium winning probabilities in each equilibrium be p_i^* and $\hat{p}_i^*$, respectively. Suppose group i participates for the set of group members M_i.*
a) Then:

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad \Gamma(\gamma_i, m_i, \hat{m}_i, R_i(m_i, \hat{m}_i), s_i, r_i) \gtreqless \frac{\mathcal{M}(\vec{w}_{i,\Xi_i}, \frac{\gamma_i}{1-\gamma_i})}{\mathcal{M}(\vec{w}_{i,M_i}, \frac{\gamma_i}{1-\gamma_i})}. \tag{12}$$

where $\Gamma(\dots) \equiv \left(\frac{\hat{m}_i}{\hat{m}_i - m_i} \left(\frac{R_i(m_i,\hat{m}_i)}{s_i(\hat{m}_i/m_i)^{1/r_i}} \right)^{\frac{\gamma_i}{1-\gamma_i}} - \frac{m_i}{\hat{m}_i - m_i} \right)^{\frac{1-\gamma_i}{\gamma_i}}.$
b) Suppose $\vec{w}_{\Xi_i}$ is more heterogeneous than $\vec{w}_{M_i}$ at mean parameter $\frac{\gamma_i}{1-\gamma_i}$. Then:

$$\gamma_i \gtreqless \gamma_i' \quad \Leftrightarrow \quad \frac{\mathcal{M}(\vec{w}_{i,\Xi_i}, \frac{\gamma_i}{1-\gamma_i})}{\mathcal{M}(\vec{w}_{i,M_i}, \frac{\gamma_i}{1-\gamma_i})} \gtreqless \frac{\mathcal{M}(\vec{w}_{i,\Xi_i}, \frac{\gamma_i'}{1-\gamma_i'})}{\mathcal{M}(\vec{w}_{i,M_i}, \frac{\gamma_i'}{1-\gamma_i'})}$$

c) Then $\Gamma(\gamma_i, m_i, \hat{m}_i, R_i, s_i, r_i)$ is weakly decreasing in γ_i, increasing in R_i, and decreasing in s_i. It is strictly decreasing in γ_i if $R_i \neq s_i^{1/r_i}$.

The occurence of the group-size paradox is therefore dependent on the behavior of Γ_i and the ratio of power means on the RHS of (12). $\Gamma(.)$ captures the complex interplay between the rivalry of the rent R_i, social interaction effects s_i, returns to scale r_i, and the complementarity in efforts γ_i. Obviously, the SSIE and the rivalry of the rent have opposite effects, with high SSIE making the group-size paradox less likely and higher rivalry making it more likely. This is in line with our results from the first part. $\Gamma(.)$ is decreasing in γ_i, which makes the group-size paradox more likely under lower γ_i when not considering the effect on the RHS.

The effects on the RHS can easily be derived from Theorem 3. If Ξ_i is more heterogeneous than M_i, the RHS of (12) will increase for a discrete increase in γ_i. We therefore know that in case the new group members are less heterogeneous than the old group members, the effects of γ_i the RHS of (12) and Γ_i will work in opposite directions and the total effect remains undetermined. However, if the new group members are equally or more heterogeneous than the former group members, a higher level of complementarity will make the group-size paradox more likely to occur.

5 Concluding remarks

Some empirical findings support the existence of a group-size paradox, but as noted by Marwell and Oliver (1993), it also stands in contrast to a significant body of empirical findings pointing to a positive relationship between group size and group performance in conflicts. Oliver (1993) has complained that in most theoretical studies, the results on the group-size paradox depend on some implicit assumptions that drive the result and that the theoretical understanding of the problem is too weak to permit confident conclusions, especially in light of the fact that empirical results reveal complex interactions that prevent simple generalizations. According to our model one can expect that four crucial factors determine the effect of group size on the outcome of a group contest: symmetric social-interactions effects, returns to scale, complementarity between group members' efforts, and the composition of their valuations in case of heterogeneous valuations within groups. We are confident that our analysis helps to clarify the different dimensions that contribute to the logic of collective action.

Appendix A: Proof of Theorem 1

For ease of notation, define $Q_{-i} = \sum_{j \neq i} q_j(\vec{x}_j)$ and $Q = \sum_j q_j(\vec{x}_j)$. The first order conditions (FOCs) are:

$$\frac{Q_{-i}}{Q^2} \cdot \frac{\partial q_i}{\partial x_i^k} \cdot v_i(m_i) - 1 \leq 0 \quad \wedge \quad x_i^k \geq 0 \quad \wedge \quad \left(\frac{Q_{-i}}{Q^2} \cdot \frac{\partial q_i}{\partial x_i^k} \cdot v_i(m_i) - 1 \right) \cdot x_i^k = 0 \; \forall \, i, k. \tag{A.1}$$

We start by showing that the first order conditions are sufficient conditions for an equilibrium. The second order conditions for a local maximum are:

$$\frac{Q_{-i} \cdot v_i(m_i)}{Q^2} \left(\frac{\partial^2 q_i(\vec{x}_i)}{\partial (x_i^k)^2} - \frac{\partial q_i(\vec{x}_i)}{\partial x_i^k} \frac{2}{Q} \right) < 0$$

which holds for $Q_{-i} \geq 0$ and $\frac{\partial^2 q_i(\vec{x}_i)}{\partial (x_i^k)^2} \leq 0$, which holds by Assumption 3. Since the above concavity condition holds for all $x_i^k \in [0, \infty)$ we only need to verify that $\pi(\infty, \vec{x}_{/x_i^k}) \leq \pi(x_i^{k*}, \vec{x}_{/x_i^k})$ and $\pi(0, \vec{x}_{/x_i^k}) \leq \pi(x_i^{k*}, \vec{x}_{/x_i^k})$. Since the FOC is strictly decreasing in x_i^k, we must have for all $x_i^k \in [0, x_i^*)$: $\frac{\partial \pi_i^k(x_i^k, \vec{x}_{/x_i^k})}{\partial x_i^k} > \frac{\partial \pi_i^k(x_i^*, \vec{x}_{/x_i^k})}{\partial x_i^k}$. This means profits are strictly increasing in x_i^k over the interval $[0, x_i^*)$ and thus $\pi(0, \vec{x}_{/x_i^k}) < \pi(x_i^{k*}, \vec{x}_{/x_i^k})$. Further, since $\pi(\infty, \vec{x}_{/x_i^k}) = -\infty < 0 \leq \pi(0, \vec{x}_{/x_i^k})$ the solution to the FOCs indeed yields a global maximum of the expected payoff for each player.

What is left to show is that there exists a unique solution to the system of FOCs given that $\forall i, k : x_i^{k*} = x_i^*$.[19] By Assumption 4 we have for all k, l: $\frac{\partial q_i(x_i, \ldots, x_i)}{\partial x_i^k} = \frac{\partial q_i(x_i, \ldots, x_i)}{\partial x_i^l}$. Therefore, if $\forall i, k : x_i^{k*} = x_i^*$, then the system of FOCs can be reduced to:

$$\frac{Q_{-i}}{Q^2} \cdot \frac{\partial q_i(x_i^*, \ldots, x_i^*)}{\partial x_i^1} \cdot v_i(m_i) - 1 \leq 0 \quad \wedge \quad x_i^* \geq 0 \quad \wedge \quad \left(\frac{Q_{-i}}{Q^2} \cdot \frac{\partial q_i(x_i^*, \ldots, x_i^*)}{\partial x_i^1} \cdot v_i(m_i) - 1 \right) \cdot x_i^* = 0 \; \forall \, i. \tag{A.2}$$

We furthermore have the following relation between Q, p_i and x_i:

$$p_i = \frac{q_i(x_i, \ldots, x_i)}{Q} \quad \wedge \quad \underline{p}_i(Q) = \frac{q_i(0, \ldots, 0)}{Q}$$

[19] In a contest with $q_i(0, \ldots, 0) = 0$ at least two groups participate. Since we do no make this assumption, it may be that all groups contribute zero effort because the starting advantage of one group is too large. However, then at least one group will have $q_i(0, \ldots, 0) > 0$.

where $\underline{p}_i(Q)$ is the lower bound on the winning probability given a specific Q. Since $q_i(x_i, \ldots, x_i)$ is strictly increasing in x_i, we can solve this for $x_i(Q, p_i)$ as long as $Q > 0$ and $p_i \geq \underline{p}_i(Q)$. Finally, we can rewrite

$$\frac{\partial q_i(x_i(Q,p_i), \ldots, x_i(Q,p_i))}{\partial x_i^k} = \rho(Q, p_i), \quad \forall p_i \geq \underline{p}_i(Q)$$

which by Assumption 3 is weakly decreasing in Q and p_i. The left equation of the Kuhn-Tucker conditions then becomes:

$$\frac{1-p_i}{Q} \cdot v_i(m_i) \cdot \rho(Q, p_i) - 1 \leq 0, \quad \forall i. \tag{A.3}$$

Since ρ is weakly decreasing in Q and p_i, the LHS of (A.3) is strictly decreasing in p_i and Q. Further, for $p_i \to \infty$, the LHS is negative while for $p_i = q_i(0, \ldots, 0)/Q$ it can be negative or positive. Therefore, for each strictly positive Q there exists a unique $p_i \in [0, \infty)$ which solves the Kuhn-Tucker conditions where $p_i = q_i(0, \ldots, 0)/Q$ if $v_i(m_i)/Q \cdot \rho(Q, 0) \leq 1$ and $p_i = 1$ if $Q = 0$. We can therefore form the function $p_i(Q)$ as the solution to the FOC of each group i.

What remains to be shown is that a unique strictly positive Q^* exists such that the winning probabilities $p_i(Q^*)$ sum to one. Notice that $p_i(Q)$ has the following properties: It is continuous, $\lim_{Q\to 0} p_i(Q) = 1$ and $\lim_{Q\to\infty} p_i(Q) = 0$ and it is strictly decreasing. Therefore, $\sum_i p_i(Q)$ is also strictly decreasing, continuous and has $\lim_{Q\to 0} \sum_i p_i(Q) > 1$ as well as $\lim_{Q\to\infty} \sum_i p_i(Q) = 0$. It follows by the intermediate value theorem that a $Q^* \in (0, \infty)$ exists such that $\sum_i p_i(Q^*) = 1$. Since $\sum_i p_i(Q)$ is strictly decreasing, this Q^* is unique. Given a unique Q^*, we can obtain unique solutions for $p_i(Q^*)$ and thus x_i^* and via $\forall i, k : x_i^{k*} = x_i^*$ also for all x_i^{k*}.

Appendix B: Proof of Theorem 2

First notice that Assumption 5 implies Assumption 4, i.e. we are only considering a subset of the impact functions, therefore the results from the proof of Theorem 1 carry over. Since we thus know that the equilibrium is unique given $\forall i, k : x_i^{k*}$, we only need to show that under the more strict Assumption 5, any equilibrium must fulfill $\forall i, k : x_i^{k*}$.

From Assumption 5 we have that

$$x_i^k > x_i^l \quad \Leftrightarrow \quad \frac{\partial q_i(\vec{x}_i)}{\partial x_i^k} < \frac{\partial q_i(\vec{x}_i)}{\partial x_i^l}.$$

Therefore, in equilibrium it can never be that case that $x_i^{l*} = 0$ if $x_i^{k*} > 0$ since then the above FOC (A.1) does not hold for at least one group member. Thus, either $\forall k : x_i^{k*} = 0$ or $\forall k : x_i^{k*} > 0$. If $x_i^{k*} > 0$ then inserting the FOC for player k into the FOC for player l yields $x_i^{k*} = x_i^{l*}$ and thus the desired condition.

Appendix C: A useful Lemma

Lemma 1. *Suppose a contest fulfills Assumptions 1, 2, 3, and 4 for all groups. Further, $\forall j, k : v_j^k(m_j) = v_j(m_j)$. Consider two within-group symmetric equilibria, which only differ by the group sizes $m_i \neq \hat{m}_i$ and/or the impact functions, $q_{m_i}(\ldots) \neq q_{\hat{m}_i}(\ldots)$. Suppose group i participates under group size m_i and impact function $q_{m_i}(\ldots)$ with winning probability p_i^*. Let the winning probability under group size $\hat{m}_i$ and impact function $q_{\hat{m}_i}(\ldots)$ be $\hat{p}_i^*$. Then the following equivalence holds:*

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad v_i(m_i) \frac{\partial q_{m_i}(x_i^*, \ldots, x_i^*)}{\partial x_i^k} \gtreqless v_i(\hat{m}_i) \frac{\partial q_{\hat{m}_i}(\hat{x}_i, \ldots, \hat{x}_i)}{\partial x_i^k} \tag{C.1}$$

where $\hat{x}_i$ is defined such that

$$q_{m_i}(x_i^*,\ldots,x_i^*) = q_{\hat{m}_i}(\hat{x}_i,\ldots,\hat{x}_i).$$

This Lemma has a very intuitive explanation: Iff for a switch from m_i to $\hat{m}_i$ while holding the winning probability constant the LHS of the FOC is too low, the group will respond by increasing the effort from which a higher winning probability results. The only complication in the proof is that one has to address the possibility of a response by other groups which overcompensates this effect.

Proof. The first-order condition for an interior solution, evaluated at the solution, becomes after rearranging terms:

$$\forall i,k: \quad v_i(m_i)\frac{\partial q_{m_i}(x_i^*,\ldots,x_i^*)}{\partial x_i^k} = \frac{Q^*}{(1-p_i^*)}. \tag{C.2}$$

We first show that

$$p_i^* \gtreqless \hat{p}_i^* \quad \Rightarrow \quad v_i(m_i)\frac{\partial q_{m_i}(x_i^*,\ldots,x_i^*)}{\partial x_i^k} \gtreqless v_i(\hat{m}_i)\frac{\partial q_{\hat{m}_i}(\hat{x}_i,\ldots,\hat{x}_i)}{\partial x_i^k} \tag{C.3}$$

and since the cases are exhaustive, the reverse implication is then automatically proven.

$p_i^* > \hat{p}_i^*$: This implies that either $Q^* < \hat{Q}^*$ or $q_{m_i}(x_i^*,\ldots,x_i^*) > q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)$. We first show that $Q^* < \hat{Q}^*$ yields a contradiction: If $p_i^* > \hat{p}_i^*$ then there exists a group j: $p_j^* < \hat{p}_j^*$. Together with $Q^* < \hat{Q}^*$, this implies that

$$\frac{Q^*}{1-p_j^*} < \frac{\hat{Q}^*}{1-\hat{p}_j^*}.$$

By (C.2) this is equivalent with:

$$v_j\frac{\partial q_j(x_j^*,\ldots,x_j^*)}{\partial x_j^k} < v_j\frac{\partial q_j(\hat{x}_j^*,\ldots,\hat{x}_j^*)}{\partial x_j^k} \tag{C.4}$$

Since $q_j(\ldots)$ has constant or decreasing RTS, this implies $q_j(x_j^*,\ldots,x_j^*) \geq q_j(\hat{x}_j^*,\ldots,\hat{x}_j^*)$. But since $Q^* < \hat{Q}^*$, we have $p_j^* > \hat{p}_j^*$ and thus a contradiction. From this follows that if $p_i^* > \hat{p}_i^*$, then $Q^* \geq \hat{Q}^*$ and $q_{m_i}(x_i^*,\ldots,x_i^*) > q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)$. The latter implies $\hat{x}_i > \hat{x}_i^*$ via the definition of $\hat{x}_i$. Since q_i has constant or decreasing RTS, we have:

$$\frac{\partial q_{\hat{m}_i}(\hat{x}_i,\ldots,\hat{x}_i)}{\partial x_i^k} \leq \frac{\partial q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)}{\partial x_i^k} \tag{C.5}$$

From $p_i^* > \hat{p}_i^*$ and $Q^* \geq \hat{Q}^*$ follows $Q^*/(1-p_i^*) > \hat{Q}^*/(1-\hat{p}_i^*)$. Using the FOCs, we have:

$$v_i(m_i)\frac{\partial q_{m_i}(x_i,\ldots,x_i)}{\partial x_i^k} > v_i(\hat{m}_i)\frac{\partial q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)}{\partial x_i^k}.$$

Combining this equation with C.5, immediately yields the $p_i^* > \hat{p}_i^*$ part of (C.3).

The proof for $p_i^* < \hat{p}_i^*$ follows the same steps with reverse inequalities and is therefore omitted.

$p_i^* = \hat{p}_i^*$: This implies that either $Q^* = \hat{Q}^*$ and $q_{m_i}(x_i^*,\ldots,x_i^*) = q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)$ or $Q^* \lesseqgtr \hat{Q}^*$ $q_{m_i}(x_i^*,\ldots,x_i^*) \lesseqgtr q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)$.

Suppose $Q^* = \hat{Q}^*$ and $q_{m_i}(x_i^*,\ldots,x_i^*) = q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)$ hold. Then it immediately follows from the FOCs that

$$v_i(m_i)\frac{\partial q_{m_i}(x_i^*,\ldots,x_i^*)}{\partial x_i^k} = v_i(m_i)\frac{\partial q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)}{\partial x_i^k}.$$

By definition of $\hat{x}_i$ we then have the symmetric part of (C.1).

For the case $Q^* \lesseqgtr \hat{Q}^*$ $q_{m_i}(x_i^*,\ldots,x_i^*) \lesseqgtr q_{\hat{m}_i}(\hat{x}_i^*,\ldots,\hat{x}_i^*)$ we can show that this yields a contradiction. It follows from these assumptions that there exists a group j with $p_j^* \gtreqless \hat{p}_j^*$ such that:

$$q_j(x_j^*,\ldots,x_j^*) \gtrless q_j(\hat{x}_j^*,\ldots,\hat{x}_j^*). \tag{C.6}$$

Furthermore,

$$\frac{Q^*}{1-p_j^*} \gtreqless \frac{\hat{Q}^*}{1-\hat{p}_j^*}.$$

Applying the FOCs gives us:

$$\frac{\partial q_j(x_j^*,\ldots,x_j^*)}{\partial x_j^k} \gtreqless \frac{\partial q_j(\hat{x}_j^*,\ldots,\hat{x}_j^*)}{\partial x_j^k}.$$

But from this follows $x_j^* < \hat{x}_j^*$ and thus

$$q_{m_i}(x_j^*,\ldots,x_j^*) \lesseqgtr q_{\hat{m}_i}(\hat{x}_j^*,\ldots,\hat{x}_j^*) \tag{C.7}$$

which contradicts (C.6).

Since the cases considered are exhaustive, it follows that the implication holds in both directions. □

Appendix D: Proof of Proposition 1

Proof. We employ the total differential:

$$\Delta q_{m_i}(\vec{x_i}) = \sum_k \Delta x_i^k \frac{\partial q(\vec{x_i})}{\partial x_i^k} \tag{D.1}$$

For equal inputs we can write $q_{m_i}(x_i,\ldots,x_i) = g(x_i,m_i)$. The total differential then becomes for symmetric efforts:

$$\Delta g(x_i,m_i) = m_i \cdot \Delta x_i \frac{\partial q(x_i,\ldots,x_i)}{\partial x_i^k}, \tag{D.2}$$

and thus

$$\frac{\partial q(x_i,\ldots,x_i)}{\partial x_i^k} = \frac{g(x_i,m_i) - g(x_i,m_i)}{m_i \cdot (x_i - x_i')} \tag{D.3}$$

for $x_i - x_i' \to 0$. Similarly, we have:

$$\frac{\partial q_{\hat{m}_i}(x_i\frac{m_i}{\hat{m}_i},\ldots,x_i\frac{m_i}{\hat{m}_i})}{\partial x_i^k} = \frac{g(x_i\frac{m_i}{\hat{m}_i},\hat{m}_i) - g(x_i'\frac{m_i}{\hat{m}_i},\hat{m}_i)}{m_i \cdot (x_i - x_i') \cdot \frac{m_i}{\hat{m}_i}} \tag{D.4}$$

for $x_i - x_i' \to 0$. It follows from (D.3) and (D.4) that:

$$\frac{\partial q(x_i,\ldots,x_i)}{\partial x_i^k} = \frac{\partial q_{\hat{m}_i}(x_i\frac{m_i}{\hat{m}_i},\ldots,x_i\frac{m_i}{\hat{m}_i})}{\partial x_i^k} \tag{D.5}$$

since by absent SSIE it holds that $g(x_i,m_i) = g(x_i\frac{m_i}{\hat{m}_i},\hat{m}_i)$ and $g(x_i,m_i) = g(x_i'\frac{m_i}{\hat{m}_i},\hat{m}_i)$.

To apply Lemma 1, we need to know what the symmetric effort level $\hat{x}_i$ of the group after the increase in size would need to be in order to obtain $q_{m_i}(x_i^*,\ldots,x_i^*) = q_{\hat{m}_i}(\hat{x}_i,\ldots,\hat{x}_i)$. With absent SSIE we have:

$$\hat{x}_i = \frac{x_i^* m_i}{\hat{m}_i}.$$

Lemma 1 then yields:

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad v_i(m_i)\frac{\partial q_{m_i}(x_i^*,\ldots,x_i^*)}{\partial x_i^k} \gtreqless v_i(\hat{m}_i)\frac{\partial q_{\hat{m}_i}(\frac{x_i^* m_i}{\hat{m}_i},\ldots,\frac{x_i^* m_i}{\hat{m}_i})}{\partial x_i^k}$$

which given (D.5) reduces to the desired condition:

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad v_i(m_i) \gtreqless v_i(\hat{m}_i)$$

Therefore, if the rent is a public good, the winning probability is independent of group size for an impact function with absent SSIE at the equilibrium effort x_i^* for group size m_i. If the rent is partly private, it is strictly decreasing in group size.

For the welfare effects, we have:

$v_i(m_i) = v_i(\hat{m}_i)$: It follows from $p_i^* = \hat{p}_i^*$ and $s_i(x_i^*, m_i, \hat{m}_i)$ that $x_i^* > \hat{x}_i^*$. Inserting this into $\pi_i^* - \hat{\pi}_i^*$ we get: $(p_i^* - \hat{p}_i^*)(v_i(m_i)) - (x_i^* - \hat{x}_i^*)$. Since the first term is equal to zero and the second term negative, we get that $\pi_i^{k*} < \hat{\pi}_i^{k*}$ from which the statements for average and total utility follow for $v_i(m_i) = v_i(\hat{m}_i)$.

$v_i(m_i) > v_i(\hat{m}_i)$: Let $v_i(\hat{m}_i) \to 0$. Then $\hat{\pi}_i^{k*} \to 0$. Since $\pi_i^{k*} > 0$, and $\hat{\pi}_i^{k*}$ is continuous in $\hat{v}_i(m_i)$ the existence of $\underline{v_i(\hat{m}_i)}$ follows.

We cannot specify $\underline{v_i(\hat{m}_i)}$ more precisely under the very general assumptions. Especially, we cannot know, whether $\underline{v_i(\hat{m}_i)} \lesseqgtr m_i v_i(m_i)/\hat{m}_i$ which is the private good case. □

Appendix E: Proof of Proposition 2

Proof. We assume throughout that we are in a symmetric, interior equilibrium. By homogeneity of degree r_i, we have from Euler's theorem

$$r_i \cdot q_{m_i}(x_i, \ldots, x_i) = m_i \cdot x_i \cdot \frac{\partial q_{m_i}(x_i, \ldots, x_i)}{\partial x_i^k}, \tag{E.1}$$

and further

$$\frac{\partial q_{m_i}(x_i, \ldots, x_i)}{\partial x_i^k} = \frac{r_i \cdot q_{m_i}(x_i, \ldots, x_i)}{m_i \cdot x_i} = \frac{r_i \cdot q_{m_i}(1, \ldots, 1)}{m_i \cdot (x_i)^{1-r_i}}. \tag{E.2}$$

Using homogeneity of the impact function and the above expression for the partial derivative, we get for the measure of SSIE:

$$s_i(m_i, \hat{m}_i) = \frac{q_{\hat{m}_i}(1, \ldots, 1) \cdot m_i^{r_i}}{q_{m_i}(1, \ldots, 1) \cdot \hat{m}_i^{r_i}}. \tag{E.3}$$

Lemma 1 now tells us that

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad v_i(m_i) \frac{\partial q_{m_i}(x_i^*, \ldots, x_i^*)}{\partial x_i^k} \gtreqless v_i(\hat{m}_i) \frac{\partial q_{\hat{m}_i}(\hat{x}_i, \ldots, \hat{x}_i)}{\partial x_i^k} \tag{E.4}$$

where $\hat{x}_i$ is defined such that

$$q_{m_i}(x_i^*, \ldots, x_i^*) = q_{m_i}(\hat{x}_i, \ldots, \hat{x}_i).$$

We can make use of homogeneity of degree r_i and solve for $\hat{x}_i$:

$$\hat{x}_i = \frac{x_i^* m_i}{\hat{m}_i} \left(\frac{q_{m_i}(1, \ldots, 1) \cdot \hat{m}_i^{r_i}}{q_{\hat{m}_i}(1, \ldots, 1) \cdot m_i^{r_i}} \right)^{1/r_i} = \frac{x_i^* \cdot m_i}{s_i(m_i, \hat{m}_i)^{1/r_i} \cdot \hat{m}_i} \tag{E.5}$$

where the last step follows from (E.3). Plugging this definition back into (E.4), we get using (E.2):

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad v_i(m_i) \frac{r_i \cdot q_{m_i}(x_i^*, \ldots, x_i^*)}{m_i \cdot x_i^*} \gtreqless v_i(\hat{m}_i) \frac{r_i \cdot q_{\hat{m}_i}(\hat{x}_i, \ldots, \hat{x}_i) s_i(m_i, \hat{m}_i)^{1/r_i}}{m_i x_i^*} \tag{E.6}$$

By canceling terms, this simplifies to:

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad \frac{v_i(m_i)}{v_i(\hat{m}_i)} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i} \tag{E.7}$$

For the welfare effects, we have:

$\pi_i^A \gtreqless \hat{\pi}_i^A \Leftrightarrow p_i^* v_i(m_i) - x_i^* \gtreqless \hat{p}_i^* v_i(\hat{m}_i) - \hat{x}_i^*$. We furthermore have:

$$x_i^* = \frac{r_i v_i(m_i)}{m_i} p_i^*(1 - p_i^*)$$

from inserting (E.2) into (A.1). Inserting this into the above equation for x_i^* and $\hat{x}_i^*$ and rearranging terms yields the result. The total welfare effects follow by multiplying with m_i and $\hat{m}_i$ on the LHS and RHS, respectively.

□

Appendix F: Proof of Theorem 3

A useful result will be the following:

Lemma 2. *If $a_1 \geq a_2 > a_3$ or $a_3 > a_2 \geq a_1$ and f is a convex function, then*

$$f(a_1 + a_2 - a_3) > f(a_1) + f(a_2) - f(a_3) \tag{F.1}$$

Proof. Case $a_3 < a_1$: From convexity of f, we have for all h

$$f'(a_1 - a_3 + h) > f'(h) \quad \Leftrightarrow \quad a_3 < a_1 \tag{F.2}$$

Integrating both sides gives:

$$\int_{a_3}^{a_2} f'(a_1 - a_3 + h)dh > \int_{a_3}^{a_2} f'(h)dh \tag{F.3}$$

which yields the desired condition:

$$f(a_1 + a_2 - a_3) - f(a_1) > f(a_2) - f(a_3). \tag{F.4}$$

Case $a_3 > a_1$: From convexity of f, we have for all h

$$f'(a_1 - a_3 + h) < f'(h) \quad \Leftrightarrow \quad a_3 > a_1 \tag{F.5}$$

Integrating both sides gives:

$$\int_{a_2}^{a_3} f'(a_1 - a_3 + h)dh > \int_{a_2}^{a_3} f'(h)dh \tag{F.6}$$

which yields the desired condition:

$$f(a_1) - f(a_1 + a_2 - a_3) < f(a_3) - f(a_2). \tag{F.7}$$

□

The above Lemma can be used to derive the following result:

Lemma 3. *Suppose $\vec{v}''$ is obtained from a sequence of θ-power mean preserving spreads of $\vec{v}$. Then*

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \mathcal{M}(\vec{v}, \phi) \gtreqless \mathcal{M}(\vec{v}'', \phi) \tag{F.8}$$

Proof. Suppose $\vec{v}^{(1)}, \ldots \vec{v}^{(n)}$ is a sequence of vectors generated by a sequence of θ-power mean preserving spreads. If for all i it holds that

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \mathcal{M}(\vec{v}^{(i)}, \phi) \gtreqless \mathcal{M}(\vec{v}^{(i+1)}, \phi) \tag{F.9}$$

then it clearly also holds that:

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \mathcal{M}(\vec{v}^{(1)}, \phi) \gtreqless \mathcal{M}(\vec{v}^{(n)}, \phi) \tag{F.10}$$

We therefore only need to show this property for vectors which differ by a single power mean preserving spread. Notice that for any ϕ:

$$\mathcal{M}(\vec{v}^{(i)}, \phi) \gtreqless \mathcal{M}(\vec{v}^{(i+1)}, \phi) \quad \Leftrightarrow \quad \mathcal{M}((v_H^{(i)}, v_L^{(i)}), \phi) \gtreqless \mathcal{M}((v_H^{(i+1)}, v_L^{(i+1)}), \phi) \tag{F.11}$$

where $(v_H^{(i)}, v_L^{(i)})$ refers to the vector of the two elements that are changed by the spreading operation and $(v_H^{(i+1)}, v_L^{(i+1)})$ to the vector of these two elements after application of the spreading operation. Let w.l.o.g. $v_H^{(i)} \geq v_L^{(i)}$ from which immediately follows $v_H^{(i+1)} > v_H^{(i)} \geq v_L^{(i)} > v_L^{(i+1)}$ by the properties of the power mean preserving spread. That is, $v_H^{(i)}$ refers to the element of $\vec{v}^{(i)}$ which is increased to $v_H^{(i+1)}$ and $v_L^{(i)}$ to the decreased element of $\vec{v}^{(i)}$.

From the power mean preserving spread also follows via evaluating (F.11) at equality:

$$\left(\frac{1}{2}(v_H^{(i+1)})^\theta + \frac{1}{2}(v_L^{(i+1)})^\theta\right)^{1/\theta} = \left(\frac{1}{2}(v_H^{(i)})^\theta + \frac{1}{2}(v_L^{(i)})^\theta\right)^{1/\theta}, \tag{F.12}$$

which – after solving for $v_H^{(i+1)}$ – yields:

$$(v_H^{(i+1)})^\theta = \left((v_H^{(i)})^\theta + (v_L^{(i)})^\theta - (v_L^{(i+1)})^\theta\right)^{1/\theta}. \tag{F.13}$$

Combining this condition with (F.11) and (F.9), what is left to show is:

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \left((v_H^{(i)})^\phi + (v_L^{(i)})^\phi\right)^{1/\phi} \gtreqless \left(\left((v_H^{(i)})^\theta + (v_L^{(i)})^\theta - (v_L^{(i+1)})^\theta\right)^{\phi/\theta} + (v_L^{(i+1)})^\phi\right)^{1/\phi} \tag{F.14}$$

which is implied by the following, more general condition:

$$\forall\, \psi > \eta: \quad \left((v_H^{(i)})^\eta + (v_L^{(i)})^\eta - (v_L^{(i+1)})^\eta\right)^{1/\eta} > \left((v_H^{(i)})^\psi + (v_L^{(i)})^\psi - (v_L^{(i+1)})^\psi\right)^{1/\psi} \tag{F.15}$$

Notice that standard mean inequalities or the reverse Jensen inequality from the previous appendix do not apply to prove (F.15). This would also be counterintuitive as then the proof would not rely on $v_H^{(i)} \geq v_L^{(i)} > v_L^{(i+1)}$. We have to distinguish the cases $\psi > 0$ and $\psi < 0$.

$\psi > 0$: Define $f(a) = a^{\psi/\phi}$, which is strictly convex and $a_1 = (v_H^{(i)})^\phi$, $a_2 = (v_L^{(i)})^\phi$, and $a_3 = (v_L^{(i+1)})^\phi$. If $\phi > 0$, we have $a_1 \geq a_2 > a_3$, while if $\phi < 0$, we have $a_3 > a_2 \geq a_1$. In both cases Lemma 2 applies. Employing these definitions in Lemma 2 gives:

$$((v_H^{(i)})^\phi + (v_L^{(i)})^\phi - (v_L^{(i+1)})^\phi)^{\psi/\phi} > ((v_H^{(i)})^\phi)^{\psi/\phi} + ((v_L^{(i)})^\phi)^{\psi/\phi} - ((v_L^{(i+1)})^\phi)^{\psi/\phi}, \tag{F.16}$$

which simplifies to (F.15).

$\psi < 0$: Define $f(a) = a^{\phi/\psi}$, which is strictly convex and $a_1 = (v_H^{(i)})^\psi$, $a_2 = (v_L^{(i)})^\psi$, and $a_3 = (v_L^{(i+1)})^\psi$. Since $\phi < 0$, we have $a_3 > a_2 \geq a_1$. Employing these definitions in Lemma 2 gives us:

$$((v_H^{(i)})^\psi + (v_L^{(i)})^\psi - (v_L^{(i+1)})^\psi)^{\phi/\psi} > ((v_H^{(i)})^\psi)^{\phi/\psi} + ((v_L^{(i)})^\psi)^{\phi/\psi} - ((v_L^{(i+1)})^\psi)^{\phi/\psi} \tag{F.17}$$

which is equivalent with (F.15) since ϕ is negative and the inequality sign thus changes direction once we exponentiate both sides with ϕ. □

We now turn to the main proof of the theorem.

Proof. Suppose $\vec{v}'$ is more heterogeneous than $\vec{v}$ at mean parameter θ. Then $\exists\omega : \omega \cdot \vec{v}'' = \vec{v}'$ and $\vec{v}''$ is obtained from a sequence of θ-power mean preserving spreads of $\vec{v}$. Thus by Lemma 2,

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \mathcal{M}(\vec{v}, \phi) \gtreqless \mathcal{M}(\vec{v}'', \phi) \tag{F.18}$$

By definition of a θ-power mean preserving spread we have $\mathcal{M}(\vec{v}, \theta) = \mathcal{M}(\vec{v}'', \theta)$. We therefore obtain:

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \frac{\omega\mathcal{M}(\vec{v}'', \theta)}{\mathcal{M}(\vec{v}, \theta)} \gtreqless \frac{\omega\mathcal{M}(\vec{v}'', \phi)}{\mathcal{M}(\vec{v}, \phi)} \tag{F.19}$$

Making use of the homogeneity of degree 0 of $\mathcal{M}$ and the definition of $\omega\vec{v}''$ we have:

$$\theta \gtreqless \phi \quad \Leftrightarrow \quad \frac{\mathcal{M}(\vec{v}', \theta)}{\mathcal{M}(\vec{v}, \theta)} \gtreqless \frac{\mathcal{M}(\vec{v}', \phi)}{\mathcal{M}(\vec{v}, \phi)} \tag{F.20}$$

□

Appendix G: Proof of Theorem 4

Proof. The proof proceeds similarly to the one for single player contests Cornes and Hartley (2005), with the main difference that first one has to obtain equilibrium conditions that fix relative efforts within each group. The optimality condition for individual k in group i yields:

$$\frac{\partial q_i}{\partial x_i^k} \frac{Q_{-i} \cdot v_i^k}{Q^2} \leq 1 \tag{G.1}$$

with equality if $x_i^k > 0$. Notice first that in equilibrium it can never be the case that $Q_{-i} = 0$, since then some individual will have an incentive to provide effort $x_i^k = \epsilon$ with $\epsilon \to 0$ to win the rent with probability 1. The expression for the partial derivative of the impact function becomes after rearranging terms:

$$\frac{\partial q_i}{\partial x_i^k} = \frac{r_i \cdot Q_i^{1-\gamma/r_i} \cdot (x_i^k)^{\gamma-1}}{m_i^{1-(s_i+r_i)\gamma/r_i}} \tag{G.2}$$

From this expression can already be derived that if one group member participates in equilibrium, all group members do: Notice that if some group member l of group i participates in equilibrium, $Q_i > 0$ and therefore at $x_i^k = 0$, $\frac{\partial q_i}{\partial x_i^k} = \infty$. But then the first order condition cannot hold for individual k in that equilibrium, since we have $Q > 0$ and $Q_{-i} > 0$ and thus the RHS of the optimality condition is infinite, which is greater than the RHS.

Since either all group members participate or none, we can express the following relationship among efforts within a group:

$$(x_i^k)^{\gamma-1} \cdot v_i^k = (x_i^l)^{\gamma-1} \cdot v_i^l \quad \forall l, k \tag{G.3}$$

Notice that this relation trivially also holds for groups that do not participate. Rearranging and summing over all l yields:

$$\left(\frac{1}{m_i}\sum (x_i^l)^\gamma\right)^{1/\gamma} = \frac{x_i^k}{(v_i^k)^{\frac{1}{1-\gamma}}}\left(\frac{1}{m_i}\sum (v_i^l)^{\frac{\gamma}{1-\gamma}}\right)^{1/\gamma} \tag{G.4}$$

Substituting this relation into the optimality condition yields:

$$Q_i^{1-\frac{1}{r_i}} \cdot Q_{-i} \cdot V_i \leq Q^2 \tag{G.5}$$

where $V_i = r_i m_i^{s_i/r_i} \cdot \left(\frac{1}{m_i}\sum_l (v_i^l)^{\frac{\gamma}{1-\gamma}}\right)^{\frac{1-\gamma}{\gamma}}$. We now differentiate our analysis between the cases $\forall i : r_i < 1$ and $\forall i : r_i = 1$.

Notice that for $\forall i : r_i < 1$ we have that $Q_i = 0$ can never be a best response to any positive Q_{-i}, since then the LHS of the above optimality condition is infinite. Therefore, if there exists a Nash equilibrium, it must be such that all groups fully participate. We rewrite the optimality condition therefore in terms of winning probabilities $p_i = Q_i/Q$:

$$V_i \cdot (1 - p_i) = Q^{1/r_i} \cdot p_i^{1/r_i - 1} \tag{G.6}$$

It is now easy to see that for all Q there exists a $p_i(Q)$ such that the optimality condition is fulfilled: Notice that for $p_i = 1$, the LHS is strictly smaller than the RHS, while for $p_i = 0$, the RHS is strictly smaller than the LHS. Since both are continuous functions of p_i, by the intermediate value theorem there then exists at least one $p_i(Q)$ such that the optimality condition holds with equality. Further, this point is unique, since the LHS is strictly decreasing in p_i, while the RHS is strictly increasing in p_i. This proves that there exists a unique best response $p_i(Q)$ to any level of Q. $p_i(Q)$ corresponds to a share function of a single player contest (Cornes & Hartley, 2005), only with the change of interpretation that it is the share of the whole group on which the within-group equilibrium condition (G.3) has been imposed.

Naturally, the remainder of the proof proceeds similarly. $p_i(Q)$ is decreasing in Q as can be verified from the following argument: Suppose Q increases, then the RHS is larger than the LHS of the optimality condition. Since the RHS is strictly increasing in p_i and the LHS strictly decreasing, p_i must decrease in order to maintain equality. A Nash equilibrium is now given by a Q^* such that $\sum_i p_i(Q^*) = 1$. Notice that for $Q = 0$, the solution to the optimality condition is $p_i(0) = 1$, while for $Q \to \infty$, we have that $p_i(\infty) \to 0$. Therefore, $\sum_i p_i(0) > 1 > \sum_i p_i(\infty)$. Since $\sum_i p_i(Q)$ is strictly decreasing in Q and continuous, there must then exist exactly one Q^* such that the equilibrium condition is fulfilled. Thus, there exists a unique Nash equilibrium, where all groups fully participate.

For the case of $r_i = 1$, we instead have the simplified optimality condition:

$$Q_{-i} \cdot V_i \leq Q^2 \tag{G.7}$$

We can therefore directly solve for the best response winning probability:

$$p_i(Q) = \max[0, 1 - Q/V_i] \tag{G.8}$$

Which has the properties $p_i(0) = 1$ and $p_i(V_i) = 0$. Noticing that the best response $p_i(Q)$ is weakly decreasing in Q, we can repeat a similar argument as above to prove that there exists a unique Nash equilibrium: Without loss of generality reorder the groups such that $V_1 > V_2 > \cdots > V_n$. We have

$$\sum_i p_i(0) = n > 1 > 0 = \sum_i p_i(V_1) \tag{G.9}$$

Since $p_1(Q)$ is strictly decreasing in Q for $Q \in [0, V_1]$ and strictly decreasing if , we have that $\sum_i p_i(Q)$ is also strictly decreasing in Q, since it is the sum of a strictly decreasing function and weakly decreasing functions in Q. From this then readily follows existence and uniqueness of a Q^* such that $\sum_i p_i(Q^*) = 1$. Depending on the level of this Q^*, it may very well be for some low enough V_i, that $0 \geq 1 - Q^*/V_i$, such that group i does not participate. Define n^* as the index of the group with the lowest V_i such that $0 > 1 - Q^*/V_i$, which completes the proof. □

Appendix H: Proof of Theorem 5

Proof. The proof goes by contradiction. Suppose we have that $V_i \geq \hat{V}_i$ and $p_i^* < \hat{p}_i^*$. Then it follows that $V_i \cdot (1 - p_i^*) > \hat{V}_i \cdot (1 - \hat{p}_i^*)$. By (9) this is equivalent to:

$$(Q^*)^{1/r_i} \cdot (p_i^*)^{1/r_i - 1} > (\hat{Q}^*)^{1/r_i} \cdot (\hat{p}_i^*)^{1/r_i - 1}$$

and thus $Q^* > \hat{Q}^*$.

Since $p_i^* < \hat{p}_i^*$ there must exist at least one group j such that: $p_j^* > \hat{p}_j^*$. Therefore, $V_j \cdot (1-p_j^*) > V_j \cdot (1-\hat{p}_j^*)$, since V_j does not differ between both equilibria. Using (9) for group j gives us:

$$(Q^*)^{1/r_j} \cdot (p_j^*)^{1/r_j - 1} < (\hat{Q}^*)^{1/r_j} \cdot (\hat{p}_j^*)^{1/r_j - 1}.$$

and thus $Q^* < \hat{Q}^*$ which yields a contradiction. By an analogous proof for $V_i \leq \hat{V}_i$ and $p_i^* > \hat{p}_i^*$ then follows the theorem.

□

Appendix I: Proof of Proposition 3

Proof. From Theorem 5, we have that:

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad r_i m_i^{s_i/r_i} \cdot \mathcal{M}(\vec{v}_{i,M_i}(m_i), \tfrac{\gamma_i}{1-\gamma_i}) \gtreqless r_i(\hat{m}_i)^{s_i/r_i} \cdot \mathcal{M}(\vec{v}_{i,M_i \cup \Xi_i}(\hat{m}_i), \tfrac{\gamma_i}{1-\gamma_i}). \tag{I.1}$$

Rearranging terms, using the definition of $s_i(m_i, \hat{m}_i)$ and writing out $\mathcal{M}(\vec{v}_{i,M_i \cup \Xi_i}(\hat{m}_i), \frac{\gamma_i}{1-\gamma_i})$ gives us:

$$p_i^* \gtreqless \hat{p}_i^* \quad \Leftrightarrow \quad \frac{\mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma_i}{1-\gamma_i})}{\left(\frac{\sum_{k=1}^{m_i} \left(v_i^k(\hat{m}_i)\right)^{\frac{\gamma_i}{1-\gamma_i}} + \sum_{k=m_i+1}^{\hat{m}_i} \left(v_i^k(\hat{m}_i)\right)^{\frac{\gamma_i}{1-\gamma_i}}}{\hat{m}_i} \right)^{\frac{1-\gamma_i}{\gamma_i}}} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i}, \tag{I.2}$$

which yields:

$$p_i^* \gtreqless \hat{p}_i^* \Leftrightarrow \frac{\mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma_i}{1-\gamma_i})}{\left(\frac{m_i}{\hat{m}_i} \frac{1}{m_i} \sum\limits_{k \in M_i} \left(v_i^k(\hat{m}_i)\right)^{\frac{\gamma_i}{1-\gamma_i}} + (1 - \frac{m_i}{\hat{m}_i}) \frac{1}{(\hat{m}_i - m_i)} \sum\limits_{k \in \Xi_i} \left(v_i^k(\hat{m}_i)\right)^{\frac{\gamma_i}{1-\gamma_i}} \right)^{\frac{1-\gamma_i}{\gamma_i}}} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i} \tag{I.3}$$

from which follows (10), since $m_i = |M_i|$ and $\hat{m}_i - m_i = |\Xi_i|$. □

Appendix J: Proof of Proposition 4

Proof. a) Inserting $R_i(m_i, \hat{m}_i)$ into (10) and simplifying gives the desired result.
b) Follows directly from Theorem 3 and the assumption that $\vec{w}_{i,\Xi_i}$ is more heterogeneous than $\vec{w}_{i,M_i}$ at $\gamma_i/(1-\gamma_i)$. Note that we only consider cases with $\gamma_i < 1$ and thus $\gamma_i/(1-\gamma_i) < \infty$.
c) It is immediately obvious that $\Gamma(\gamma_i, m_i, \hat{m}_i, R_i, s_i, r_i)$ is increasing in R_i and decreasing in s_i. The only difficulty is thus the proof of the behavior of $\Gamma(\gamma_i, m_i, \hat{m}_i, R_i, s_i, r_i)$ with changes in γ_i. A useful result on which the proof is based is the reverse Jensen inequality (for a more general version and its proof, see Bullen, 2003, p. 43):

Lemma 4. *If f is convex, $\varrho_1 > 0$ and $\varrho_i < 0$ for all $2 \geq i \geq n$ and $\sum_{j=1}^n \varrho_j = 1$, then $f(\sum_{j=1}^n \varrho_j a_j) \geq \sum_{j=1}^n \varrho_j f(a_j)$ for all $\sum_{j=1}^n \varrho_j a_j \in I$, and the inequality holds strictly, if f is strictly convex and $\exists i, j : a_i \neq a_j$.*

We now show that the following term is decreasing in γ_i:

$$\Gamma(\gamma_i, m_i, \hat{m}_i, R_i, s_i, r_i) = \left(\frac{\hat{m}_i}{\hat{m}_i - m_i} \cdot \left(\frac{R_i(m_i, \hat{m}_i)}{s_i(\hat{m}_i/m_i)^{1/r_i}} \right)^{\frac{\gamma_i}{1-\gamma_i}} - \frac{m_i}{\hat{m}_i - m_i} \right)^{\frac{1-\gamma_i}{\gamma_i}}. \tag{J.1}$$

Define $\theta_i = \frac{\gamma_i}{1-\gamma_i}$, which is increasing in γ_i. What is to be shown is that the above term is decreasing in θ_i, thus:

$$\left(\frac{\hat{m}_i}{\hat{m}_i - m_i}\Psi^{\theta_i} - \frac{m_i}{\hat{m}_i - m_i}\right)^{1/\theta_i} > \left(\frac{\hat{m}_i}{\hat{m}_i - m_i}\Psi^{\phi_i} - \frac{m_i}{\hat{m}_i - m_i}\right)^{1/\phi_i} \tag{J.2}$$

whenever $\phi_i > \theta_i$ and where $\Psi = \frac{R_i(m_i,\hat{m}_i)}{s_i(\hat{m}_i/m_i)^{1/r_i}}$. Note that $\phi_i, \theta_i \in (-1, \infty)$ and therefore we will distinguish the cases $0 < \phi_i$, and $\phi_i < 0$. We will furthermore assume that θ_i and ϕ_i are not zero (which is equivalent to assuming that $\gamma_i \neq 0$.

$0 < \phi_i$: Since ϕ_i is positive, we can rewrite condition (J.2) to:

$$\left(\frac{\hat{m}_i}{\hat{m}_i - m_i}\Psi^{\theta_i} - \frac{m_i}{\hat{m}_i - m_i}\left(1^{\theta_i}\right)\right)^{\phi_i/\theta_i} > \left(\frac{\hat{m}_i}{\hat{m}_i - m_i}\left(\Psi^{\theta_i}\right)^{\phi_i/\theta_i} - \frac{m_i}{\hat{m}_i - m_i}\left(1^{\theta_i}\right)^{\phi_i/\theta_i}\right). \tag{J.3}$$

Setting $f(a) = a^{\phi_i/\theta_i}$ (which is strictly convex also for negative θ_i) and $\varrho_1 = \frac{\hat{m}_i}{\hat{m}_i - m_i}$, $\varrho_2 = -\frac{m_i}{\hat{m}_i - 1}$, and $a_1 = (\Psi)^{\theta_i}$, $a_2 = 1^{\theta_i}$ in the above reverse Jensen inequality directly yields equation (J.3).

$\phi_i < 0$: Since θ_i is negative, we can rewrite condition (J.2) to:

$$\left(\frac{\hat{m}_i}{\hat{m}_i - m_i}\Psi^{\phi_i} - \frac{m_i}{\hat{m}_i - m_i}\left(1^{\phi_i}\right)\right)^{\theta_i/\phi_i} > \left(\frac{\hat{m}_i}{\hat{m}_i - m_i}\left(\Psi^{\phi_i}\right)^{\theta_i/\phi_i} - \frac{m_i}{\hat{m}_i - m_i}\left(1^{\phi_i}\right)^{\theta_i/\phi_i}\right). \tag{J.4}$$

Setting $f(a) = a^{\theta_i/\phi_i}$ (which is strictly convex since the absolute value of ϕ_i is smaller than that of θ_i) and $\varrho_1 = \frac{\hat{m}_i}{\hat{m}_i - m_i}$, $\varrho_2 = -\frac{m_i}{\hat{m}_i - m_i}$, and $a_1 = (\Psi)^{\phi_i}$, $a_2 = 1^{\phi_i}$ in the above reverse Jensen inequality directly yields equation (J.4). Therefore, condition (J.2) follows, which concludes the proof of part c) of the proposition. □

Appendix K: Relation between comparative statics analysis and inter-group comparisons.

Theorem 6. *Consider two contests fulfilling Assumptions 1, 2, 3, and 4 for all groups, which differ only in the group size of group i, m_i and $\hat{m}_i$. Moreover, let $m_j = \hat{m}_i$, $q_j = q_{i,\hat{m}_i}$, and $v_i(\hat{m}_i) = v_j(m_j)$. For all h, k: $v_h^k = v_h$ and let the equilibrium winning probabilities in the symmetric equilibria be p_i^*, p_j^* and $\hat{p}_i^*$, $\hat{p}_j^*$ respectively. Group i participates at group size m_i with effort level x_i^*.*

Then:

$$p_i^* \gtreqless p_j^* \quad \Leftrightarrow \quad p_i^* \gtreqless \hat{p}_i^*$$

The theorem shows that the comparative static interpretation of the group-size paradox and the interpretation of inter-group comparisons yield the same results for equal valuations within groups.

Proof. Define first $\hat{x}$ as the solution to $q_{m_i}(x_i^*, \ldots, x_i^*) = q_{\hat{m}_i}(\hat{x}, \ldots, \hat{x})$. By the first order conditions (A.1) evaluated at the equilibrium effort of group i, we have for arbitrary group members k and l of groups i and j, respectively: $p_i^* > p_j^*$ iff

$$v_i(m_i)\frac{\partial q_{i,m_i}(x_i^*, \ldots, x_i^*)}{\partial x_i^k} < v_j(m_j)\frac{\partial q_j(\hat{x}, \ldots, \hat{x})}{\partial x_j^l} \tag{K.1}$$

since $\partial q_j(x_j, \ldots, x_j)/\partial x_j^l$ is weakly decreasing in x_j. (K.1) is equivalent with:

$$v_i(m_i)\frac{\partial q_{i,m_i}(x_i^*, \ldots, x_i^*)}{\partial x_i^k} < v_i(\hat{m}_i)\frac{\partial q_{i,\hat{m}_i}(\hat{x}, \ldots, \hat{x})}{\partial x_i^k}. \tag{K.2}$$

This is by Lemma 1 equivalent with $p_i^* > \hat{p}_i^*$. The proof for $p_i^* = \hat{p}_i^*$ and $p_i^* < \hat{p}_i^*$ is analogous. □

For the heterogeneous case, a similar statement can be made:

Theorem 7. *Consider two contests fulfilling Assumptions 1, 2, 7 for all groups, which differ only by the set of group members M_i and $\hat{M}_i = M_i \cup \Xi_i$. Moreover, let $M_j = \hat{M}_i$, $r_i = r_j$, $\gamma_i = \gamma_j$. Let the equilibrium winning probabilities in each equilibrium be p_i^*, p_j^* and $\hat{p}_i^*$, $\hat{p}_j^*$ respectively. Suppose group i participates with group members M_i. Then,*

$$p_i^* \gtreqless p_j^* \quad \Leftrightarrow \quad p_i^* \gtreqless \hat{p}_i^*.$$

To show this, note the following Lemma:

Lemma 5. *Consider a contest fulfilling Assumptions 1, 2, 7 for all groups. Suppose $r_i = r_j$. Let the equilibrium winning probabilities in equilibrium be p_i^*, p_j^*. Then,*

$$p_i^* \gtreqless p_j^* \quad \Leftrightarrow \quad V_i \gtreqless V_j.$$

Proof. Suppose $V_i > V_j$, then from (9) of groups i and j and $r_i = r_j$ we have:

$$\frac{(p_i^*)^{1/r_i - 1}}{(1 - p_i^*)} > \frac{(p_j^*)^{1/r_i - 1}}{(1 - p_j^*)} \tag{K.3}$$

Since the RHS is increasing in p_i (note that $1/r_i - 1 \geq 0$ by assumption) and the LHS in p_j, it follows that $p_i > p_j$. By the symmetry of the problem, for $V_i < V_j$ it follows that $p_i > p_j$. Next suppose $V_i = V_j$, then by (9) of groups i and j and $r_i = r_j$ we directly have:

$$\frac{(p_i^*)^{1/r_i - 1}}{(1 - p_i^*)} = \frac{(p_j^*)^{1/r_i - 1}}{(1 - p_j^*)} \tag{K.4}$$

which only holds for $p_i^* = p_j^*$. Since the considered cases are exhaustive, it directly follows that: $p_i^* \gtreqless p_j^*$ if and only if $V_i \gtreqless V_j$. □

From here the proof of Theorem 7 directly follows from Theorem 5 and Lemma 5 and the fact that $V_j = \hat{V}_i$.

Appendix L: Extensions of Propositions 2 and 3 to voluntary contributions games.

It turns out that the key properties which have been examined for the group-size paradox in a contest setting are also at work in collective action problems without the contest environment. In this appendix we show that for two collective action problems without the contest environment, our methods and to some extent even the results can be transfered.

We use the model by (Bergstrom et al., 1986) with the simplification of identical preferences across players and the generalization of allowing $v_i(m_i)$ to depend on group size.

Assumption 9. Individuals k maximize:

$$\pi_i^k(x_i^k, \vec{x}_{-x_i^k})) = u(w - x_i^k, v_i(m_i) q_{m_i}(\vec{x}_i)) \tag{L.1}$$

where u is a binormal utility function increasing in both arguments.

(We could drop the group index i here since there is only one group, but leave it for cross-referencing to our results on group contests.) As discussed in (Bergstrom et al., 1986; Cornes & Hartley, 2007), binormality implies that the marginal rate of substitution $MRS_i^k(x_i^k, v_i(m_i)q_{m_i}(\vec{x}_i)) = \frac{\frac{\partial u(\ldots)}{\partial v_i(m_i)q_{m_i}(\vec{x}_i)}}{\frac{\partial u(\ldots)}{\partial w_i - x_i^k}}$ is decreasing in $v_i(m_i)q_{m_i}(\vec{x}_i)$ and non-increasing in x_i^k. Equilibrium existence has been proven by (Cornes & Hartley, 2007). Symmetry of the equilibrium follows from Assumption 4.

We can now obtain similar results to Proposition 2:

Proposition 5. *Consider two voluntary contribution games fulfilling Assumptions 9, 3, and 4, which differ only in the group size of group i, m_i and $\hat{m}_i > m_i$. For all k: $v_i^k = v_i$ and the class of impact functions $\{q_{m_i}(.)\}_{m_i=2}^{\overline{m}}$ fulfills Assumption 6 with $s_i(m_i, \hat{m}_i)$ as the measure of SSIE. Suppose group i contributes $0 < x_i^* < w_i$ at group size m_i and $0 < \hat{x}_i^* < w_i$ at group size $\hat{m}_i$. Then:*

$$(d_i)^{r_i} v_i(m_i) q_i(x_i^*, \ldots, x_i^*) \gtreqless v_i(\hat{m}_i) q_i(\hat{x}_i^*, \ldots, \hat{x}_i^*) \quad \Leftrightarrow \quad 1 \gtreqless s_i(m_i, \hat{m}_i)^{1/ri} \cdot d_i^{r_i} \tag{L.2}$$

where

$$d_i \gtreqless 1 \quad \Leftrightarrow \quad 1 \gtreqless \frac{v_i(m_i)}{v_i(\hat{m}_i)} \frac{1}{s_i(m_i, \hat{m}_i)} \left(\frac{m_i}{\hat{m}_i}\right)^{r_i} \tag{L.3}$$

In this proposition we used the value-adjusted consumption of the public good as the criterion for the group-size paradox. Results on group welfare are naturally even more difficult to obtain than in the contest case, since they strongly depend on the shape of $u(\ldots)$. Since we assumed a very general form of preferences, we also do not obtain a closed form solution for d_i. However, both SSIE and the rivalry in the rent still work in the predictable manner of making the group-size paradox less and more likely, respectively.[20] We also see that the term $m_i/\hat{m}_i$ (in the expression determining the orientation of d_i) provides a starting advantage for larger groups. The larger the returns to scale, the more pronounced this starting advantage is.

Proof. We will prove the result only for the case of $MRS(\ldots)$ being strictly decreasing in the first argument. The extension to the case where the utility function can be locally linear in the first argument is trivial, but would require many case distinctions.[21] The method of the proof is similar to the contest case. We first find an effort level $\breve{x}_i$ of the group with size $\hat{m}_i$ such that the RHS of the FOC is identical to the RHS of the FOC under the equilibrium efforts x_i^* and then compare the LHS of the FOC to determine whether $\breve{x}_i \gtreqless \hat{x}_i^*$.

The first order condition yields for all k:

$$\frac{\partial q(\vec{x}_i)}{\partial x_i^k} v_i(m_i) = \Big(MRS(x_i^k, v_i(m_i)q(\vec{x}_i))\Big)^{-1} \tag{L.4}$$

Evaluated in a symmetric equilibrium, we have

$$\frac{\partial q(x_i^*, \ldots, x_i^*)}{\partial x_i^k} v_i(m_i) = \left(MRS(x_i^*, v_i(m_i)q(x_i^*, \ldots, x_i^*))\right)^{-1}. \tag{L.5}$$

Define $\hat{x}_i$ such that $q_{m_i}(x_i^*, \ldots, x_i^*) = q_{\hat{m}_i}(\hat{x}_i, \ldots, \hat{x}_i)$. A simplification in comparison to the contest model is that we do not need to consider the responses of other groups to a change in efforts after a change in group size. Instead, we face the difficulty that the RHS of the FOC under group size m_i given efforts x_i^* is not identical

[20] An increase in d_i is unanimously good for a larger group: It helps fulfilling the $1 \gtreqless s_i(m_i, \hat{m}_i)^{1/ri} \cdot d_i^{r_i}$ condition and at the same time increases the critical level $(d_i)^{r_i} v_i(m_i) q_i(x_i^*)$ which will be surpassed if the former condition is met.

[21] u being linear in $w_i - x_i^k$ implies $d_i = 0$.

to the RHS of the FOC under group size $\hat{m}_i$ given efforts $\hat{x}_i$. To obtain an identical RHS we define $\breve{x}_i$ such that $MRS(x_i^*, v_i(m_i)q(x_i^*, \ldots, x_i^*)) = MRS(\breve{x}_i, v_i(m_i)q(\breve{x}_i, \ldots, \breve{x}_i))$. We need to determine how $\breve{x}$ compares with $\hat{x}$. For this, define d_i such that $\breve{x}_i = d_i \left(\frac{v_i(m_i)}{v_i(\hat{m}_i)}\right)^{1/r_i} \hat{x}_i$. Note that due to homogeneity of q_i, we have that: $v_i(m_i)q_{m_i}(x_i^*, \ldots, x_i^*) = v_i(\hat{m}_i)q_{\hat{m}_i}\left(\frac{\breve{x}_i}{d_i}, \ldots, \frac{\breve{x}_i}{d_i}\right)$.

$$\frac{\breve{x}_i}{d_i} \gtreqless x_i^* \quad \Leftrightarrow \quad MRS\left(\frac{\breve{x}_i}{d_i}, v_i(\hat{m}_i)q_{\hat{m}_i}\left(\frac{\breve{x}_i}{d_i}, \ldots, \frac{\breve{x}_i}{d_i}\right)\right) \gtreqless MRS(x_i^*, q_{m_i}(x_i^*, \ldots, x_i^*)) \tag{L.6}$$

since the second argument of the two MRS is identical and the MRS is strictly decreasing in the first argument. Since $MRS(x, v(\hat{m}_i)q_{\hat{m}_i}(x, \ldots, x))$ is strictly decreasing in x and $MRS(x_i^*, v(m_i)q_{m_i}(x_i^*, \ldots, x_i^*)) = MRS(\breve{x}_i, v(m_i)q_{\hat{m}_i}(\breve{x}_i, \ldots, \breve{x}_i))$, we have that:

$$\frac{\breve{x}_i}{d_i} \gtreqless x_i^* \quad \Leftrightarrow \quad 1 \gtreqless d_i. \tag{L.7}$$

Using (E.5), which holds in virtue of Assumption 6, and our definition of d_i we can solve for the left condition as:

$$\frac{v_i(m_i)}{v_i(\hat{m}_i)} \frac{1}{s_i(m_i, \hat{m}_i)} \left(\frac{m_i}{\hat{m}_i}\right)^{r_i} \gtreqless 1 \quad \Leftrightarrow \quad 1 \gtreqless d_i. \tag{L.8}$$

Switching gears, we can now look at what determines whether $\breve{x}_i \gtreqless \hat{x}_i^*$. Since the RHS of the FOC is strictly increasing in x_i and the LHS is strictly decreasing in x_i, we have that:

$$\breve{x}_i \gtreqless \hat{x}_i^* \quad \Leftrightarrow \quad \frac{\partial q_{\hat{m}_i}(\breve{x}_i, \ldots, \breve{x}_i)}{\partial x_i^k} v_i(\hat{m}_i) \gtreqless \left(MRS(\breve{x}_i, q_{m_i}(\breve{x}_i, \ldots, \breve{x}_i))\right)^{-1} \tag{L.9}$$

Substituting the $MRS(\ldots)$ term:

$$\breve{x}_i \gtreqless \hat{x}_i^* \Leftrightarrow \frac{\partial q_{\hat{m}_i}(\breve{x}_i, \ldots, \breve{x}_i)}{\partial x_i^k} v_i(\hat{m}_i) \gtreqless \frac{\partial q_{m_i}(x_i^*, \ldots, x_i^*)}{\partial x_i^k} v_i(m_i) \tag{L.10}$$

Making use of the results up to (E.7), we get via homogeneity of q_{m_i}:

$$\breve{x}_i \gtreqless \hat{x}_i^* \quad \Leftrightarrow \quad \frac{v_i(m_i)}{v_i(\hat{m}_i)} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i} \left(\frac{\breve{x}_i}{\hat{x}_i}\right)^{r_i} \tag{L.11}$$

Finally, cancelling terms:

$$d_i\hat{x}_i \left(\frac{v_i(m_i)}{v_i(\hat{m}_i)}\right)^{1/r_i} \gtreqless \hat{x}_i^* \quad \Leftrightarrow \quad 1 \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i} (d_i)^{r_i} \tag{L.12}$$

By using (E.5) and homogeneity of q_i, we can simplify the left condition:

$$(d_i)^{r_i} v_i(m_i)q_{m_i}(x_i^*) \gtreqless v_i(\hat{m}_i)q_{\hat{m}_i}(\hat{x}_i^*) \quad \Leftrightarrow \quad 1 \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i} (d_i)^{r_i} \tag{L.13}$$

□

Using a utility function u which is homogeneous in each argument, we could also obtain similar results for the case of heterogeneous valuations. Instead, for a variant of Proposition 3 as a voluntary contributions game we assume individuals maximize the following utility function:

Assumption 10. Individuals k maximize:

$$\pi_i^k(x_i^k, \vec{x}_{-x_i^k}) = g(q_{m_i}(\vec{x}_i))v_i^k(m_i) - x_i^k \tag{L.14}$$

where g is twice continuously differentiable and $\frac{\partial g(x)}{\partial x} > 0$ and $\frac{\partial^2 g(x)}{\partial x^2} < 0$.

Note that this model is not covered by Assumptions 9 and 6, since $g(q(\vec{x}_i))$ is allowed to be non-homogeneous in efforts. However, it assumes linear costs instead. The model can be understood as a voluntary contributions game to some intermediate impact q_i, from which some final good $g(\dots)$ with value $v_i(m_i)$ is produced. A characterization of the group-size paradox just by the properties of q_i is helpful in case we have a clear idea how the intermediate good is produced (e.g. media impact of demonstrations), but not how the final good is produced (e.g. political influence).

Proposition 6. *Consider two voluntary contribution games fulfilling Assumptions 10, 2, 7 which differ only by the set of group members M_i and $\hat{M}_i = M_i \cup \Xi_i$. Let the equilibrium efforts in each equilibrium be $\vec{x}_i^* \geq 0$ and $\vec{\hat{x}}_i^*$, respectively. Then:*

$$g(q_{m_i}(\vec{x}_i^*)) \gtreqless g(q_{\hat{m}_i}(\vec{\hat{x}}_i^*)) \Leftrightarrow \frac{\mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma_i}{1-\gamma_i})}{\left(\frac{m_i}{\hat{m}_i} \cdot \mathcal{M}(\vec{v}_{i,M_i}(\hat{m}_i), \frac{\gamma_i}{1-\gamma_i})^{\frac{\gamma_i}{1-\gamma_i}} + (1 - \frac{m_i}{\hat{m}_i}) \cdot \mathcal{M}(\vec{v}_{i,\Xi_i}(\hat{m}_i), \frac{\gamma_i}{1-\gamma_i})^{\frac{\gamma_i}{1-\gamma_i}}\right)^{\frac{1-\gamma_i}{\gamma_i}}} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i}. \tag{L.15}$$

It can easily be verified that the game has a unique Nash equilibrium. For ease of comparison, we again keep the index i for the group even though there is just one group in this game.

Proof. The first order conditions are for all k:

$$\left(g'(q_{m_i}(\vec{x}_i^*)) m_i^{s_i+r_i-1} r_i \mathcal{M}(\vec{x}_i, \gamma)^{r_i-\gamma} v_i^k\right) = (x_i^{k*})^{1-\gamma} \tag{L.16}$$

Taking the γ mean over all x_i^k gives us:

$$\left(\sum_k \frac{1}{m_i} \left(g'(q_{m_i}(\vec{x}_i^*)) m_i^{s_i+r_i-1} r_i \mathcal{M}(\vec{x}_i^*, \gamma)^{r_i-\gamma} v_i^k\right)^{\frac{\gamma}{1-\gamma}}\right)^{1/\gamma} = \mathcal{M}(\vec{x}_i^*, \gamma) \tag{L.17}$$

Cancelling terms and rearranging yields:

$$m_i^{s_i/r_i} r_i \mathcal{M}(\vec{v}_i, \frac{\gamma}{1-\gamma}) = q_{m_i}(\vec{x}_i^*)^{1/r_i - 1} \left(g'(q_{m_i}(\vec{x}_i^*))\right)^{-1} \tag{L.18}$$

We now compare the production of the good between two groups consisting of members M_i and $M_i \cup \Xi_i$. Noting that the RHS of the above equation is strictly increasing in $q_{m_i}(\vec{x}_i^*)$ and $g(q_{m_i}(\vec{x}_i^*))$ is also strictly increasing in $q_{m_i}(\vec{x}_i^*)$, we have:

$$g(q_{m_i}(\vec{x}_i^*)) \gtreqless g(q_{\hat{m}_i}(\vec{\hat{x}}_i^*)) \Leftrightarrow m_i^{s_i/r_i} \mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma}{1-\gamma}) \gtreqless \hat{m}_i^{s_i/r_i} \mathcal{M}(\vec{v}_{i,M_i \cup \Xi_i}(\hat{m}_i), \frac{\gamma}{1-\gamma}) \tag{L.19}$$

Which using the definition of $s_i(m_i, \hat{m}_i)$ can be rewritten as:

$$g(q_{m_i}(\vec{x}_i^*)) \gtreqless g(q_{\hat{m}_i}(\vec{\hat{x}}_i^*)) \Leftrightarrow \frac{\mathcal{M}(\vec{v}_{i,M_i}(m_i), \frac{\gamma}{1-\gamma})}{\mathcal{M}(\vec{v}_{i,M_i \cup \Xi_i}(\hat{m}_i), \frac{\gamma}{1-\gamma})} \gtreqless s_i(m_i, \hat{m}_i)^{1/r_i} \tag{L.20}$$

□

Since the condition for the occurrence of the group-size paradox in this voluntary contribution game is identical to the one from the contest, it follows that also Proposition 4 continues to hold in the voluntary contribution game.

It should be noted that the interpretation of the RTS is not as straightforward as in the contest model, where it represented the discriminatory power of the contest. For example, the function g may be of the form

$g(x) = x_i^t$ in which case the RTS of the overall model are $r_i + t_i$ instead of r_i. This result must therefore be understood as a decomposition property. If one can rewrite the production function of the collective good as a concave function g applied to the impact produced via a CES aggregate, then equation L.15 determines whether the group-size paradox occurs. As mentioned above, such a decomposition may be helpful in many cases where we observe the impact of groups (e.g. media attention) but not final outcomes (e.g. political influence).

References

Acemoglu, D., & Jensen, M. K. (2009). *Aggregate comparative statics.* (Working Paper, MIT)

Akerlof, G. A. (1997). Social distance and social decisions. *Econometrica*, *65*(5), 1005-1028.

Baik, K. H. (2008). Contests with group-specific public goods prizes. *Social Choice and Welfare*, *30*(1), 103-117.

Bergstrom, T., Blume, L., & Varian, H. (1986). On the private provision of public goods. *Journal of public economics*, *29*(1), 25–49.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, *102*(5), 841-77.

Bullen, P. (2003). *Handbook of means and their inequalities.* Dordrecht: Kluwer.

Corchón, L. C. (2007). The theory of contests: A survey. *Review of Economic Design*, *11*(2), 69-100.

Cornes, R., & Hartley, R. (2005). Asymmetric contests with general technologies. *Economic Theory*, *26*(4), 923-946.

Cornes, R., & Hartley, R. (2007). Weak links, good shots and other public good games: Building on BBV. *Journal of Public Economics*, *91*(9), 1684-1707.

Cornes, R., & Sandler, T. (1996). *The theory of externalities, public goods, and club goods.* Cambridge, MA: Cambridge University Press.

Esteban, J., & Ray, D. (2001). Collective action and the group size paradox. *American Political Science Review*, *95*(3), 663-672.

Esteban, J., & Ray, D. (2008). On the salience of ethnic conflict. *The American Economic Review*, *98*(5), 21852202.

Esteban, J., & Ray, D. (2010). *A model of ethnic conflict.* (forthcoming in: The Journal of the European Economic Association.)

Fu, Q., & Lu, J. (2008). *A micro foundation of generalized multi-prize lottery contests: A noisy ranking perspective.* (Unpublished working paper)

Garfinkel, M. R., & Skaperdas, S. (2007). Economics of conflict: An overview. In K. Hartley & T. Sandler (Eds.), *Handbook of defense economics* (1st ed., Vol. 2, p. 649-709). Elsevier.

Kolmar, M., & Rommeswinkel, H. (2013). Contests with group-specific public goods and complementarities in efforts. *Journal of economic behavior and organization*, *89*, 9-22.

Kolmar, M., & Wagener, A. (2011). *Inefficiency as a strategic device in group contests against dominant opponents.* (Working Paper, University of St. Gallen)

Konrad, K. (2009). *Strategy and dynamics in contests.* Oxford, UK: Oxford University Press.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, *14*(3), 115-136.

Marwell, G., & Oliver, P. (1993). *The critical mass in collective action.* New York: Cambridge University Press.

Münster, J. (2009). Group contest success functions. *Economic Theory*, *41*(2), 345-357.

Nitzan, S., & Ueda, K. (2009). Collective contests for commons and club goods. *Journal of Public Economics*, *93*(1), 48-55.

Nitzan, S., & Ueda, K. (2010). *Prize sharing in collective contests.* (forthcoming in: The European Economic Review)

Oliver, P. (1993). Formal models of collective action. *Annual Review of Sociology*, *19*, 271-300.

Olson, M. (1965). *The logic of collective action.* Cambridge, MA: Harvard University Press.

Pecorino, P., & Temini, A. (2008). The group size paradox revisited. *Journal of Public Economic Theory*, *10*, 785-799.

Raab, P., & Schipper, B. C. (2009). Cournot competition between teams: An experimental study. *Journal of Economic Behavior & Organization*, *72*(2), 691-702.

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk i: A definition. *Journal of Economic Theory*, *2*, 225-243.

Scheinkman, J. A. (2008). Social interactions (theory). In S. N. Durlauf & L. E. Blume (Eds.), *The new palgrave dictionary of economics, 2nd edition.* Basingstoke: Palgrave Macmillan.

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

Skaperdas, S. (1992). Cooperation, conflict, and power in the absence of property rights. *American Economic Review*, *82*, 720-739.

Wärneryd, K. (2001). Replicating contests. *Economics Letters*, *71*(3), 323 - 327.

Yun, T. (1996). Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics*, *37*, 345-370.