ASHOKA UNIVERSITY ECONOMICS
DISCUSSION PAPER NO. 40

Affirmative Actions in a Large Population Contests

September 2020

Ratul Lahkar, Ashoka University
Rezina Sultana, Indian Institute of Management, Udaipur

https://ashoka.edu.in/economics-discussionpapers

Affirmative Action in Large Population Contests

Ratul Lahkar* Rezina Sultana†

July 30, 2020

Abstract

We consider affirmative action in large population Tullock contests. The standard Tullock contest is an equal treatment contest in which agents who exert equal effort have an equal probability of success. In contrast, under affirmative action, agents with equal cost of effort have equal probability of success. We analyze such contests as generalized aggregative potential games and characterize their Nash equilibria. We show that affirmative action equalizes equilibrium payoffs without causing any loss of aggregate welfare. It enhances the welfare and effort levels of agents facing high effort cost. Thus, affirmative action engenders equality without having any detrimental effects on efficiency, at least when the number of agents involved are large. It does, however, reduce aggregate effort in society.

Keywords: Affirmative Action; Tullock Contests; Aggregative Games; Potential Games.

JEL Classification: C72; C73; D63; I38; J78.

*(corresponding author) Department of Economics, Ashoka University, Rajiv Gandhi Education City, Sonipat, Haryana 131029, India, email: ratul.lahkar@ashoka.edu.in.

†Economics area, Indian Institute of Management Udaipur, Udaipur, Rajasthan, 313001, India, email: rezina.sultana@iimu.ac.in

1 Introduction

Affirmative action policies are usually applied in competitive situations, like college and university admissions, or filling job vacancies, where the credentials of applicants from historically discriminated groups are weighed differently in order to ameliorate the effects of such discrimination. It is well recognized that affirmative action contributes towards greater equality by creating a level playing field for historically disadvantaged groups (Holzer and Neumark [24]). There are, however, debates about the effect of such policies on efficiency. Both efficient and inefficient outcomes are considered as possibilities under affirmative action, depending on how such policies affect effort incentives (Coate and Loury [6], Sultana [45]). Opponents of affirmative action point at possible distortion in effort incentives under such policies (Sowell [43]). Proponents, on the contrary, cite the possible greater incentives for effort created by a level playing field (Bowen and Bok [3]). The vigorousness of this debate notwithstanding, there isn't much unanimity among economists about such efficiency implications. For example, as Fryer and Loury [18] note, "Confident a priori assertions about how affirmative action affects incentives are unfounded. Indeed, economic theory provides little guidance".

Our paper is a contribution in this direction. Due to the competitive nature of situations in which affirmative action policies are usually applied, contests provide a convenient theoretical tool to model the effects of such policies.[1] Several papers, therefore, have analyzed affirmative action using the three different forms of contests—Tullock contests (Franke [15], Dahm and Esteve-González [11], Franke et al. [16]), all–pay auctions (Fu [21], Franke et al. [17]) and rank–order tournaments (Schotter and Weigelt [42], Fryer and Loury [19], Fain [14]). Following this literature, we model affirmative action in the environment of a Tullock contest (Tullock [49]), which is probably the most well known model of a contest. Within this environment, we assess affirmative action on the basis of two important criteria, namely the welfare of agents in society and the incentives faced by agents to exert effort.

Affirmative actions models in Tullock contests introduce heterogeneity among agents through asymmetries in the cost of effort function of different agents (Franke [15], Dahm and Esteve-González [11], Franke et al. [16]). Such heterogeneity is then interpreted as the effect of historical discrimination. Thus, agents who have faced higher levels of discrimination have a higher cost of effort function. An important feature of such models is that effort is taken to be inherently valuable. This is unlike usual applications of Tullock contests to understand rent–seeking in environments where effort is wasteful. With this premise, these models introduce affirmative action in different ways with the primary objective of assessing the implications of such a policy on the total equi-

[1] Contests are used to model situations where a resource is sought to be allocated among agents who expend effort to increase their share or chance of obtaining the resource. The literature on contests started with the seminal contributions of Tullock [48, 49] who applied them to study rent seeking. Apart from Tullock contests, two other canonical forms of contests are all–pay auctions (Hillman and Riley [23]) and rank–order tournaments (Lazear and Rosen [30]). Such contests have been applied to analyze a variety of areas like litigation, lobbying, labor market tournaments, awarding a contract etc. See Corchón [8], Konrad [25] or Corchón and Serena [9] for a review of this literature.

librium effort by agents. For example, Franke [15] modifies the standard *equal treatment* Tullock contest by introducing an *affirmative action* policy that biases the success function of agents in the Tullock contest so that equal *cost or disutility of effort* generates equal chances of success.[2] The primary conclusion of that paper then is that affirmative action can increase total effort if the number of agents is low but unlikely to do so when that number is large (Proposition 3, Franke [15]). Dahm and Esteve–Gonzáles [11] model affirmative action differently by introducing an extra prize over which only disadvantaged agents can compete. Their conclusion is different from Franke's [15]; affirmative action increases total effort. Franke et al. [16] seek to find the optimal way to bias the success function of disadvantaged agents so as to maximize total effort.

In this paper, we focus closely on the model of affirmative action by Franke [15]. The importance of this paper is that it is perhaps the first model of affirmative action in a general $n-$player Tullock contest environment. Given the seminal importance of Tullock contests as a model of competitive behavior, exploring the implications of affirmative action in such a framework is an worthwhile exercise. Hence, the present paper also assumes that effort is valuable and considers the two policies of equal treatment and affirmative action in a Tullock contest environment. The crucial difference is that while Franke [15] analyzes Tullock contests with a finite number of players, our paper is in the context of large population Tullock contests, i.e. contests where the set of agents is a continuum. Thus, the society in our model is divided into a finite number of types (or populations), each with a common effort cost function. Each such type, therefore, represents a group which has faced similar levels of historical discrimination. The large population model is an approximation for a situation when the number of agents are finite but large, a condition that is nearly always satisfied in real world applications of affirmative action.

The motivation behind this change is not just technical. The more important reason is that using a large population framework opens up the possibility of generalizing results obtained in a finite player framework. It is well–known that it is not feasible to derive closed form solutions in asymmetric Tullock contests in the presence of non–linearities in the success and cost functions. Hence, for example, Franke's [15] result on total effort under affirmative action in the general $n-$player case is only for the case where these functions are linear. On the other hand, as we show, our large population framework is general enough that it can accommodate such asymmetries and non–linearities. We are, therefore, able to show under both linear and non–linear conditions, affirmative action reduces total effort as compared to equal treatment in a Tullock contest environment.[3] But even as total effort declines, we also show that the effort of agents belonging to highly disadvantaged groups increase under affirmative action. This finding is intuitive but of interest because one reason why policy makers may favour affirmative action is to diversify certain sectors of the economy or the government (for example, the civil service) by enhancing representation of socially disadvantaged

[2]The standard Tullock contest is interpreted as an equal treatment contest because agents who exert equal effort have equal chances of success in that contest. Other papers which have also compared equal treatment with affirmative action include Schotter and Weigelt [42] and Sultana [46].

[3]In fact, showing this result under non–linear conditions is simpler in our framework as we are then assured of a unique Nash equilibrium in our Tullock contest models. This is not the case under linear conditions.

groups (Fryer and Loury [20]).

The existing literature on affirmative action in contests focus almost exclusively on the effort implications of affirmative action (for example, the papers cited in the second paragraph of this section). The tractability of our large population approach, in contrast, also allows us to arrive at certain strong results on welfare. We find that aggregate welfare in equilibrium under both equal treatment and affirmative action are equal. This result is of significance because a common criticism of affirmative action policies, starting with Becker [1], is that such policies reduce total welfare. Our model shows that this is not necessarily the case.[4]. Of course, the caveat remains that ours is large population result whereas real world populations are finite. Therefore, in real world situations, this result will be relevant (even if approximately) only when the number of agents involved are sufficiently many. In that case, even if affirmative action does have any deleterious effect of aggregate welfare, it will not be significant. Instead, by increasing effort by disadvantaged agents, it also enhances their welfare thereby contributing to a more equal society. In fact, affirmative action in our model leads to perfect equality.

Further, much of this existing literature on affirmative action in contests takes an instrumental approach to this policy. In this approach, the policy maker seeks to design a contest that maximizes some prior objective such as aggregate effort. In contrast, we, like Franke [15] and Dahm and Esteve–Gonzáles [11], take a normative approach to affirmative action. Thus, the policy maker is inspired not by any prior policy objective but by the norm of levelling the playing field.

In addition to the literature on affirmative action, this paper also contributes to the theory of contests and, in particular, large population contests. As far as we know, this is the first paper that models large population Tullock contests. Our formal analysis of these contests rely on the machinery of potential games (Monderer and Shapley [32], Sandholm [40], Cheung and Lahkar [5]) and aggregative games (Corchón [7]), which together define the class of generalized aggregative potential games. The tractability of the large population approach suggests that if interest is in characterizing equilibrium with a large number of contestants, then it is a more feasible methodology than directly analyzing the finite player contest.[5,6] Through the use of potential games, we also contribute to the literature on applications of evolutionary game theory (Newton [34]). It is well known that a variety of evolutionary dynamics converge to Nash equilibria in large population potential games (Sandholm [40], Cheung and Lahkar [5]). Therefore, Tullock contests and applications of such contests emerge as a new area of application of evolutionary game theory through this paper.

There is a literature on large population contests (Olszewski and Siegel [38], Bodoh–Creed and

[4]Of course, ours is not the only model of affirmative action in which this result arises. For example, Fryer and Loury [20] and Harel and Segal [22] also establish conditions under which diversity enhancing policies need not reduce social welfare. Details of these models differ from ours. In particular, they are not in a contest environment.

[5]We note that the strategy set in our large population contests is continuous. This greatly simplifies equilibrium characterization even though it introduces certain measure theoretic complications.

[6]Ewerhart [13] analyzes a finite player Tullock contest with symmetric and linear impact and cost functions using the best response potential, which is an extension of Monderer and Shapley's [32] original notion of a potential function.

Hickman [2]). These papers also explore the idea that large population contests provide a more tractable way of approximating equilibria of more realistic contests with a finite but large number of players. But their details vary considerably from ours. Olszewski and Siegel [38] study large contests in quasi–linear settings while we focus on Tullock contests. Bodoh–Creed and Hickman [2] use large contests to compare two types of affirmative action in college admission—admission preference and quotas. They do not use Tullock contests so that their definitions of affirmative action differ from ours. The analytical methods of these papers also vary from ours. They consider finite player contests and show that the equilibrium of such games is approximated by the large population limit as the number of agents increase. In contrast, we analyze the large population model directly using the theory of large population potential games.

The rest of the paper is as follows. In Section 2, we introduce large population Tullock contests with equal treatment and affirmative action. In Section 3, we define the category of generalized aggregative potential games and interpret Tullock contests as such games. Section 4 characterizes Nash equilibrium of such games and applies them to Tullock contests. In Section 5, we do a comparative static analysis of Nash equilibria under equal treatment and affirmative action and establish our main conclusions about the implications of affirmative action on welfare and effort incentives. In all these sections, we assume certain non–linearities in our Tullock contests which ensure a unique Nash equilibrium. Section 6 considers the fully linear case. Section 7 concludes.

2 Tullock Contests and Affirmative Action

We consider a continuum of agents of mass 1, which we call a *society*. We divide the society into a finite set of populations or *types* $\mathcal{P} = \{1, 2, \cdots, n\}$. We denote the mass of population p as $m_p \in (0, 1)$ and call the distribution $(m_1, \cdots, m_n)$ the type distribution. Since the total mass of agents is 1, it must be that $\sum_{p \in \mathcal{P}} m_p = 1$.

Every agent in the society has a common strategy set $\mathcal{S} = [\underline{x}, \bar{x}] \subset \mathbf{R}_+$, with $\underline{x} > 0$.[7] We interpret $x \in \mathcal{S}$ as the effort exerted by an agent in the contest. We denote by $\mu_p \in \mathcal{M}^+_{m_p}(\mathcal{S})$ the distribution of strategies in population p, where $\mathcal{M}^+_\nu(\mathcal{S})$ is the set of finite positive measures that impose a mass ν on $\mathcal{S}$. Equivalently, μ_p is the state of population p with $\mu_p(A) \in [0, m_p]$ denoting the mass of type$-p$ agents playing strategies in $A \subseteq \mathcal{S}$.[8] If all agents in a population p play the same strategy x, then that is a monomorphic population state which we denote as $m_p \delta_x$. If a population state is not monomorphic, then we call it polymorphic.

We denote the set of states in the entire society as $\Delta = \prod_{p=1}^n \mathcal{M}^+_{m_p}(\mathcal{S})$. A social state is, therefore, $\mu = (\mu_1, \cdots, \mu_n) \in \Delta$, where $\mu_p \in \mathcal{M}^+_{m_p}(\mathcal{S})$. We then define a population game as a mapping

$$F : \mathcal{S} \times \mathcal{P} \times \Delta \to \mathbf{R} \tag{1}$$

[7]In certain applications, it may be possible for $\underline{x}$ to be zero. However, in our main application to contests in this paper, we require $\underline{x} > 0$ for payoffs to be well defined. This is not a very serious restriction as we can take $\underline{x}$ to be arbitrarily close to zero.

[8]Note that $\mu_p(\mathcal{S}) = m_p$ for all $p \in \mathcal{P}$.

such that $F_{x,p}(\mu)$ is the payoff of an agent in population p who uses strategy $x \in \mathcal{S}$ at the social state μ. We assume that such a payoff function is bounded and weakly continuous with respect to μ. We define a Nash equilibrium of such a population game as follows.

Definition 2.1 *A Nash equilibrium of a multipopulation game F as defined in (1) is a social state $\mu^* = (\mu_1^*, \cdots, \mu_n^*) \in \Delta$ such that for all $x \in \mathcal{S}$, all $p \in \mathcal{P}$, if x lies in the support of μ_p^*, then $F_{x,p}(\mu^*) \geq F_{y,p}(\mu^*)$, for all $y \in \mathcal{S}$.*

To introduce Tullock contests as a large population game, we assume that the set of agents are contesting over a resource of value $V > 0$. Further, we assume that if an agent of type p exerts effort x, then that agent incurs an effort cost of $c_p(x) = k_p x^\gamma$, $k_p > 0$ and $\gamma \geq 1$. Thus, the cost functions are asymmetric across types according to the parameter k_p but within a type, the cost of effort is identical. The common parameter γ ensures that the cost function is convex. We then define a large population Tullock contest to be a population game F in which the payoff of a type$-p$ agent who exerts effort x is

$$F_{x,p}(\mu) = \frac{x^r}{\sum_{q \in \mathcal{P}} \int_{\mathcal{S}} x^r \mu_q(dx)} V - k_p x^\gamma. \tag{2}$$

To interpret (2), we follow Myerson and Wärneryd [33] and refer to the mapping $x \mapsto x^r$ as a player's strategy *impact function*. The fraction $\frac{x^r}{\sum_{q \in \mathcal{P}} \int_{\mathcal{S}} x^r \mu_q(dx)}$ is then the contest success function (CSF) in F. It measures, as a density function, the share of an agent who exerts effort x in the prize V (if V is divisible) or the probability of success of that agent in the contest (if V is indivisible). Thus, exerting effort x in the contest gives benefit $\frac{x^r}{\sum_{q \in \mathcal{P}} \int_{\mathcal{S}} x^r \mu_q(dx)} V$. Subtracting the cost of effort, we obtain the payoff (2).[9,10]

We assume that $r \in (0, 1]$ so that the impact function x^r is concave.[11] In addition, for most of the paper, we also assume that if $r = 1$, then the cost function parameter $\gamma > 1$. Together, these assumptions will allow us to establish uniqueness of Nash equilibrium in our models. We will consider the specific case of $r = \gamma = 1$ separately in Section 6.

[9]The terminology used here follows from finite player Tullock contests (see, for example, Konrad [25]). Consider such a contest with N players. Player i exerts effort x_i and obtains payoff $\frac{x_i^r}{\sum_{j=1}^N x_j^r} V - c_i(x_i)$, where $c_i(x_i)$ is the player's cost of effort x_i. The fraction $\frac{x_i^r}{\sum_{j=1}^N x_j^r}$ is then the CSF in such a contest.

[10]Due to the differences in the cost functions, (2) is strategically equivalent to a model where the valuation of the resource differs according to type but costs are symmetric. Stein [44] consider a finite–player contest model with such asymmetric valuations among players.

[11]Concavity of the impact function is a usual assumption in the finite player contest literature (see, for example, Szidarovszky and Okuguchi [47] and Yamazaki [50]) It ensures a the existence of a pure strategy Nash equilibrium. Specifically for finite $N-$ player Tullock contests, a weaker assumption that $r \leq \frac{N}{N-1}$ is required for the existence of a pure strategy equilibrium (for example, Nitzan [35]). Observe though that as $N \to \infty$, $\frac{N}{N-1} \to 1$. This provides another justification for our assumption about r in our large population model.

2.1 Affirmative Action

In the literature on affirmative action, the standard Tullock contest (2) is called a contest with *equal treatment* (Franke [15]). This is because in this contest, if two participants exert equal effort x, then the CSF for the two agents are equal. Under equal treatment, two contestants are treated equally disregarding the underlying differences in effort costs across types.

However, as per the standard interpretation in the affirmative action literature, agents face different costs of effort due to past discrimination. Thus, in our model where $c_p(x) = k_p x^\gamma$, we assume that k_p is higher for agents who have faced higher levels of historical discriminations.[12] But historical discrimination is a factor that is beyond the control of individual agents. Therefore, it may be argued that on grounds of fairness, there is a need for policy intervention that levels the playing field in favor of agents who have a high cost of effort. *Affirmative action* is premised on such a notion of fairness. A fair policy that seeks to redress the effect of such discrimination should ensure that agents who have equal cost of effort should have the same success level in the contest even if the effort levels are not the same. Formally, if two agents in populations p and q exert effort levels x_p and x_q respectively such that $c_p(x_p) = c_q(x_q)$, then affirmative action would require that the two agents have the same CSF. Kranich [26] justifies such a policy in terms of the "moral intuition that two people incurring equal disutility deserve equal rewards".

Recall that the impact function in the equal treatment contest (2) is x^r. This function is symmetric across populations. Following Franke [15], we now model affirmative action in terms of an asymmetric impact function $x \mapsto R_p x^r$, where we interpret R_p as a *bias* parameter. This function, therefore, differs between populations but is the same for all agents within a population. For a type$-p$ agent playing strategy x, the CSF would then be $\frac{R_p x^r}{\sum_{q \in \mathcal{P}} \int_{\mathcal{S}} R_q x^r \mu_q(dx)}$. The condition $c_p(x_p) = c_q(x_q)$ reduces to $k_p x_p^\gamma = k_q x_q^\gamma$, which implies $x_p = \left(\frac{k_q}{k_p}\right)^{\frac{1}{\gamma}} x_q$. We then want, by the definition of affirmative action, that the CSFs of the two agents are also equal, which is equivalent to $R_p x_p^r = R_q x_q^r$, or $\frac{R_p}{k_p^{\frac{r}{\gamma}}} = \frac{R_q}{k_q^{\frac{r}{\gamma}}}$. Since the CSF is homogeneous of degree zero, it suffices to make $R_p = k_p^{\frac{r}{\gamma}}$ for all p.[13] We thus obtain the CSF $\frac{k_p^{\frac{r}{\gamma}} x^r}{\sum_{q \in \mathcal{P}} \int_{\mathcal{S}} k_q^{\frac{r}{\gamma}} x^r \mu_q(dx)}$.

We can now formally define a Tullock contest with affirmative action. In order to distinguish it from the contest F with equal treatment as defined in (2), we denote this contest as $\hat{F}$. Thus, a contest $\hat{F}$ is a Tullock contest with affirmative action if the payoff to an agent who uses $x \in \mathcal{S}$ at the social state μ is

$$\hat{F}_{x,p}(\mu) = \frac{k_p^{\frac{r}{\gamma}} x^r}{\sum_{q \in \mathcal{P}} \int_{\mathcal{S}} k_q^{\frac{r}{\gamma}} x^r \mu_q(dx)} V - k_p x^\gamma. \tag{3}$$

[12] Observe that we have groups (populations) with different levels of k_p. Thus, agents in our society have faced different levels of historical discrimination. Within a group, however, the level of discrimination faced is the same. Hence, k_p is the same for all agents in population p.

[13] Notice that, with the affirmative action mandate of equal representation from different types, agents with higher cost parameter, k_p, also obtain a higher bias parameter, $k_p^{\frac{r}{\gamma}}$. For instance, reservation policy in India allows the different social groups based on *'caste'*, (e.g., Scheduled Castes and Scheduled Tribes) to get admission in colleges with different cut-off marks.

We retain all the assumptions on k_p, r and γ in (3) that we made for (2). Thus, $k_p > 0$, $r \in (0,1]$, $\gamma \geq 0$ and if $r = 1$, then $\gamma > 1$, except in Section 6 where we will consider $r = \gamma = 1$. Note that real–world affirmative action policies seek to alleviate discrimination through different means like quotas and preferential treatment. Such measures bias the playing field in favor of the disadvantaged agents. The contest (3) captures this essential aspect of affirmative action policies without going into the details of any such policy. It is the large population analogue of the finite player contest with affirmative action considered in Franke [15].

Before analyzing the two contests F and $\hat{F}$, we need to clarify the interpretation of effort in these models. The usual application of a Tullock contest is to model rent–seeking behavior (Konrad [25]). In such applications, effort is interpreted as being wasteful since exertion of effort does not have any impact on the value of the rent V. The Pareto efficient outcome is then the one where agents exert the lowest possible effort and simply divide up the rent amongst themselves. This may also be seen from the fact that the aggregate payoff in the Tullock contest (2) is $\bar{F}(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} F_{x,p}(\mu)\mu_p(dx) = V - \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p x^{\gamma} \mu_p(dx)$, which is maximized at the state where all agents exert effort $\underline{x}$.[14]

Such an interpretation of effort is, however not tenable is our model. Because then, the policy maker can simply distribute the prize equally among the agents without requiring anything more than the minimal effort. This would achieve Pareto efficiency. Hence, to justify the policy maker designing the affirmative action contest (3), we have to assume that effort is socially valuable.[15] For example, the policy maker may seek to democratize or diversify the civil service by increasing representation of socially disadvantaged groups. In that case, a policy that enhances effort by such agents will be desired. Or while implementing affirmative action, the policy maker may also view employment as a labor market tournament and will be concerned with total effort.

But even though effort is valuable, the affirmative action contest is not designed for instrumental objectives like maximizing total effort. Instead, we assume that the goal of the policy maker is the normative one of alleviating the effects of historical discrimination according to Kranich's [26] principle of equal reward for equal disutility. This is the perspective we adopt in the following analysis as we assess the implications of affirmative action on welfare and effort incentives.

3 Generalized Aggregative Potential Games

In order to solve our large population contest models, we now introduce a class of population games called generalized aggregative potential games. As we will see, the two contests we introduced in Section 2 belong to this class of games. These properties will be helpful in characterizing Nash equilibria of these contests. In particular, once we establish these properties, we will be able to

[14]Formally, this state is $\underline{\mu} = (m_1\delta_{\underline{x}}, m_2\delta_{\underline{x}}, \cdots, m_n\delta_{\underline{x}})$. Note that the aggregate payoff maximizing state in the affirmative action contest $\hat{F}$ is also $\underline{\mu}$. The CSF in both contest add up to 1.

[15]Interpretation of effort as being socially valuable is common in the affirmative action literature and such policies are often evaluated on the basis of their incentive effect on effort (see, for example, Bowen and Bok [3], Fryer and Loury [18, 19], Sowell [43]). In the Introduction, we have also cited papers in the contest literature where maximizing effort is the main objective of the contest designer.

characterize equilibria by maximizing the underlying potential function.

To define such games, we retain the modeling structure introduced in Section 2, with the set of populations $\mathcal{P}$, the strategy set $\mathcal{S}$ and the mass distribution $(m_1, \cdots, m_n)$. We introduce the impact function $\phi_p : \mathcal{S} \to \mathbf{R}_+$ and define the generalized aggregate strategy level in population p as

$$a(\mu_p) = \int_{\mathcal{S}} \phi_p(x)\mu_p(dx), \tag{4}$$

where μ_p is the state of population p as described in Section 2. Since $\mu_p \in \mathcal{M}^+_{m_p}(\mathcal{S})$, $a(\mu_p) \in [m_p\phi(\underline{x}), m_p\phi(\bar{x})]$. The generalized aggregate strategy level in the whole society at the social state $\mu = (\mu_1, \cdots, \mu_n)$ is then

$$A(\mu) = \sum_{p\in\mathcal{P}} a(\mu_p) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} \phi_p(x)\mu_p(dx). \tag{5}$$

Therefore, $A(\mu) \in \left[\sum_{p\in\mathcal{P}} m_p\phi_p(\underline{x}), \sum_{p\in\mathcal{P}} m_p\phi_p(\bar{x})\right]$. If $\phi_p(x) = x$, then (5) is simply aggregate strategy. This is the special form of aggregation considered in Lahkar [28] and Cheung and Lahkar [5]. Therefore, (5) broadens the notion of aggregate strategy by taking aggregation with respect to a more general strategy impact function.

We first describe a generalized aggregative game. Consider a population game F in which the payoff of an agent in population p who plays strategy $x \in \mathcal{S}$ is

$$F_{x,p}(\mu) = \phi_p(x)\beta(A(\mu)) - c_p(x), \tag{6}$$

where ϕ_p is the impact function, $\beta : \left[\sum_{p\in\mathcal{P}} m_p\phi_p(\underline{x}), \sum_{p\in\mathcal{P}} m_p\phi_p(\bar{x})\right] \to \mathbf{R}_+$ is the benefit function for the agent that depends upon the generalized aggregate strategy level $A(\mu)$ and $c_p : \mathcal{S} \to \mathbf{R}_+$ is the cost function of the agent. This payoff function clearly depends upon only the agent's own strategy and the generalized aggregate strategy level $A(\mu)$. This particular form of F, therefore, generalizes Corchón's [7] notion of an aggregative game to the generalized notion of aggregate strategy (5). It is also a generalization of the large population aggregative games considered in Lahkar [28] and Cheung and Lahkar [5] who only considered a single population and aggregation with respect to $\phi_p(x) = x$.[16] Due to such extensions of the notion of aggregative games, we refer to the population game F defined by (6) as a generalized aggregative game.

Of course, (6) is not the only possible form that an aggregative game can take. This particular form is motivated by the two contests in Section 2. In the standard Tullock contest (2), $\phi_p(x) = x^r$, $A(\mu) = \sum_{q\in\mathcal{P}} \int_{\mathcal{S}} x^r \mu_q(dx)$ and $\beta(A(\mu)) = \frac{V}{\sum_{q\in\mathcal{P}} \int_{\mathcal{S}} x^r \mu_q(dx)}$. In the affirmative action contest (3), $\phi_p(x) = k_p^{\frac{r}{\gamma}} x^r$, $A(\mu) = \sum_{q\in\mathcal{P}} \int_{\mathcal{S}} k_p^{\frac{r}{\gamma}} x^r \mu_q(dx)$ and $\beta(A(\mu)) = \frac{V}{\sum_{q\in\mathcal{P}} \int_{\mathcal{S}} k_p^{\frac{r}{\gamma}} x^r \mu_q(dx)}$. Thus, both contests are generalized aggregative games. Therefore, motivated by the assumptions in these models, we also make the following assumptions about (6).

[16] Lahkar [28] considers finite strategy single population aggregative games while Cheung and Lahkar [5] consider such games with a continuous strategy set. These papers apply this particular form of aggregation to important economic examples like Cournot competition and tragedy of the commons.

Assumption 3.1 *Consider the large population contest F defined by (6). We assume that*

1. *The benefit function β is strictly decreasing on $\left[\sum_{p\in\mathcal{P}} m_p\phi_p(\underline{x}), \sum_{p\in\mathcal{P}} m_p\phi_p(\bar{x})\right]$.*

2. *The impact function ϕ_p is strictly increasing and concave on $\mathcal{S}$ for every $p \in \mathcal{P}$.*

3. *For every $p \in \mathcal{P}$, the cost function c_p is strictly increasing and convex.*

4. *For every $p \in \mathcal{P}$, if ϕ_p is linear, then c_p is strictly convex.*

The fact that $\beta(A(\mu)) = \frac{V}{A(\mu)}$ in our contest models imply that these models satisfy part 1 of Assumption 3.1. The conditions that $r \in (0, 1]$, $\gamma \geq 1$ and that if $r = 1$, then $\gamma > 1$ ensure that the remaining parts of this assumption are also satisfied by the contest models. Assumption 3.1 will allow us to show that (6) has a unique Nash equilibrium.

We follow Lahkar and Mukherjee [28] to define potential games in our context of multipopulation games with a continuous strategy set.[17] Certain technical details involved in this definition, namely Fréchet differentiability and the gradient of a Fréchet differentiable function, are in Appendix A.1.1. These more general notions of differentiability and gradient are required because of the measure theoretic nature of our model. We also introduce the notation $\mathcal{M}$ to denote the extension of Δ to the space of all finite signed measures. Further details are in Appendix A.1.1.

Definition 3.2 *A population game F as defined in (1) is a potential game if there exists a Fréchet differentiable function $f : \mathcal{M} \to \mathbf{R}$ such that*

$$\nabla f(\mu) = F(\mu) \text{ for all } \mu = (\mu_1, \cdots, \mu_n) \in \Delta.$$

The function f is called the potential function of the game F.

Therefore, according to Definition 3.2, the large population contest F defined by (6) is a potential game if there exists a real–valued function f, called the potential function, such that $\nabla f(\mu)(x, p) = F_{x,p}(\mu)$, for all $x \in \mathcal{S}$, $p \in \mathcal{P}$ and $\mu \in \Delta$.[18] We now show that such a potential function does exist for (6). First, we define the aggregate cost at social state μ as

$$C(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} c_p(x)\mu_p(dx). \tag{7}$$

The following proposition states the result. It also shows that the relevant potential function is concave. The proof is in Appendix A.1.2.

[17]The original concept of potential games is due to Monderer and Shapley [32], who defined it for finite strategy games. Sandholm [40] adapts that definition to large population games with finite strategy sets. This has been further extended to games with continuous strategy sets by Cheung [4], Lahkar and Riedel [27] and Cheung and Lahkar [5] for single population games, and by Lahkar and Mukherjee [29] for multipopulation games.

[18]Here, $\nabla f(\mu)(x, p)$ is the value of the gradient of f at strategy x and type p.

Proposition 3.3 *The population game F defined by (6) is a potential game with potential function*

$$f(\mu) = \int_{\sum_{p\in\mathcal{P}} m_p \phi_p(\underline{x})}^{A(\mu)} \beta(z)dz - C(\mu), \tag{8}$$

where $A(\mu)$ is the generalized aggregate strategy level as defined in (5). Further, f is concave but not strictly concave on Δ.

The concavity of f is important because, as shown by Sandholm [40], maximizing the potential function is sufficient in that case to characterize the set of Nash equilibrium of the underlying potential game.[19] We note that we do not need Assumption 3.1 to establish that F is a potential game. However, showing the concavity of f does require us to use Assumption 3.1(1).

Thus, F defined by (6), in addition to being a generalized aggregative game is also a potential game. Hence, we call such games *generalized aggregative potential games*. The two contests introduced in Section 2 are such games. The following corollary characterizes their potential functions. The main difference between these functions lie in the definition of the generalized aggregate strategy level $A(\mu)$.

Corollary 3.4 *Consider the two contests F and $\hat{F}$ defined by (2) and (3) respectively.*

1. *The contest with equal treatment F is a generalized aggregative potential game with potential function*
$$f(\mu) = \int_{\underline{x}^r}^{A(\mu)} \frac{V}{z} dz - \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p x^\gamma \mu_p(dx), \tag{9}$$
for $\mu \in \Delta$, where $A(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} x^r \mu_p(dx)$.

2. *The contest with affirmative action $\hat{F}$ is a generalized aggregative potential game with potential function*
$$\hat{f}(\mu) = \int_{\sum_{p\in\mathcal{P}} m_p k_p^{\frac{r}{\gamma}} \underline{x}^r}^{A(\mu)} \frac{V}{z} dz - \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p x^\gamma \mu_p(dx), \tag{10}$$
for $\mu \in \Delta$, where $A(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p^{\frac{r}{\gamma}} x^r \mu_p(dx)$.

Hence, both F and $\hat{F}$ are generalized aggregative potential games. Further, the potential functions (9) and (10) are concave in Δ.

Proof. Follows from Proposition 3.3. ■

[19] Indeed, the strict concavity of f is sufficient to establish that the potential game F has a unique Nash equilibrium. In our case, though, we only obtain f. Hence, establishing uniqueness will require additional results.

4 Nash Equilibrium of Generalized Aggregative Potential Games

Proposition 3.3 implies that we can characterize Nash equilibria of a generalized aggregative potential game (6) by maximizing the potential function (8) on Δ. However, maximizing (8) directly is difficult as it is defined on an abstract measure space. Cheung and Lahkar [5] resolve this problem in their special context of aggregative games by introducing a simpler function called the *quasi–potential function*, which serves as a proxy for the original potential function and which is amenable to direct maximization. We now extend that method to our generalized aggregative games. We denote the quasi–potential function as $g : \prod_{p=1}^{n} [\underline{x}, \bar{x}] \to \mathbf{R}$ and define as

$$g(\alpha_1, \alpha_2, \cdots, \alpha_n) = \int_{\sum_{p \in \mathcal{P}} m_p \phi_p(\underline{x})}^{\sum_{p \in \mathcal{P}} m_p \phi_p(\alpha_p)} \beta(z) dz - \sum_{p \in \mathcal{P}} m_p c_p(\alpha_p). \tag{11}$$

As with the extension from aggregative games to generalized aggregative games described in Section 3, (11) generalizes the quasi–potential function in Cheung and Lahkar [5] in two ways. First, it allows for multiple populations. Second, it allows for a more general notion of aggregate strategy level $A(\mu)$.[20]

Lemma A.2 in Appendix A.2 establishes the relationship between the potential function and the quasi–potential function. An important aspect of this relationship which highlights the proxy nature of the quasi–potential function is that at social states consisting entirely of monomorphic population states $\mu_p = m_p \delta_{\alpha_p}$, the two functions have the same value. Otherwise, the quasi–potential function always has a strictly higher value. Moreover, because the quasi–potential function is defined with respect to the real vector $(\alpha_1, \cdots, \alpha_n)$, maximizing it is straightforward. Indeed, Lemma A.3 in Appendix A.2 shows that this function is strictly concave and, therefore, has a unique maximizer. These two lemmas follow from Assumption 3.1 and the aggregative nature of the game. Moreover, the two lemmas imply the following result on the Nash equilibrium of generalized aggregative potential games. The proof of this theorem, along with the proofs of Lemmas A.2 and A.3, are in Appendix A.2.

Theorem 4.1 *Let Assumption 3.1 hold. Then, the generalized aggregative potential game F defined by (6) has a unique Nash equilibrium $\mu^* = \left(m_1 \delta_{\alpha_1^*}, m_2 \delta_{\alpha_2^*}, \cdots, m_n \delta_{\alpha_n^*}\right)$, where $(\alpha_1^*, \cdots, \alpha_n^*) \in \prod_{p \in \mathcal{P}} [\underline{x}, \bar{x}]$ is the unique maximizer of the quasi–potential function (11), as established in Lemma A.3.*

Thus, at the unique Nash equilibrium of F, every agent in population p plays α_p^*, where α_p^* is obtained from the maximizer of the quasi–potential function (11). In this sense, maximizing the quasi–potential function provides a complete characterization of the Nash equilibrium of F. In order to prove this theorem, we use Lemma A.2 to show that $f(\mu^*) > f(\mu)$ for all $\mu \in \Delta \setminus \{\mu^*\}$,

[20] Lahkar [28] introduces quasi–potential functions in the context of single population aggregative potential games with a finite strategy set. Recall from footnote 16 that an aggregative game is a special case of a generalized aggregative game with $\phi_p(x) = x$. Cheung and Lahkar [5] extend that technique to single population aggregative potential games with a continuous strategy set..

where f is the potential function (8). Therefore, μ^* is the unique maximizer of f in Δ. Since, by Proposition 3.3, f is concave on Δ, we conclude that μ^* is the unique Nash equilibrium of F.[21]

How important is it to use the potential game method to obtain Theorem 4.1? In large population games, direct methods of computing Nash equilibria using best responses are typically not available when payoffs are non–linear. Having certain additional structure then helps find equilibria. The feature of being a potential game grants that additional structure to F. It allows us to characterize equilibria and, in fact, establish uniqueness of equilibrium by maximizing the concave potential function. In addition, of course, as discussed in the Introduction, the potential game approach makes our equilibrium prediction more robust as such games have attractive convergence properties under evolutionary dynamics (Sandholm [40]).[22] The aggregative game characteristic, in turn, allows us to replace the task of maximizing the potential game with the much simpler task of maximizing the quasi–potential function.

4.1 Nash Equilibrium in Tullock Contests

Recall from Corollary 3.4 that the two contests introduced in Section 2 are generalized aggregative potential games. We can, therefore, apply Theorem 4.1 to characterize their Nash equilibria. First, we consider the Tullock contest F with equal treatment as defined in (2), where $\phi_p(x) = x^r$, $\beta(z) = \frac{V}{z}$ and $c_p(x) = k_p x^\gamma$. Hence, using (11), we write the quasi–potential function for F as

$$\begin{aligned} g(\alpha_1, \alpha_2, \cdots, \alpha_n) &= \int_{\underline{x}^r}^{\sum_{p\in\mathcal{P}} m_p \alpha_p^r} \frac{V}{z} dz - \sum_{p\in\mathcal{P}} m_p k_p \alpha_p^\gamma \\ &= V \log\left(\sum_{p\in\mathcal{P}} m_p \alpha_p^r\right) - V\log(\underline{x}^r) - \sum_{p\in\mathcal{P}} m_p k_p \alpha_p^\gamma. \end{aligned} \tag{12}$$

We now maximize g to obtain the unique Nash equilibrium of F. In establishing this result, we assume that (12) has an interior maximizer in $\prod_{p=1}^n [\underline{x}, \bar{x}]$. The proof of the proposition is in Appendix A.2.1.

Proposition 4.2 *Consider the Tullock contest with equal treatment F as defined in (2). Assume that $r \in (0, 1]$ and $\gamma \geq 1$, with the further restriction that if $r = 1$, then $\gamma > 1$. Further assume that $\underline{x}$ is sufficiently small and $\bar{x}$ is sufficiently large that (12) has an interior maximizer in $\prod_{p=1}^n [\underline{x}, \bar{x}]$. Then, F has a unique Nash equilibrium $\mu^* = \left(m_1 \delta_{\alpha_1^*}, m_2 \delta_{\alpha_2^*}, \cdots, m_n \delta_{\alpha_n^*}\right)$ such that for every $p \in \mathcal{P}$,*

$$\alpha_p^* = \left(\frac{rV k_p^{\frac{\gamma}{r-\gamma}}}{\gamma \sum_{q\in\mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}}\right)^{\frac{1}{\gamma}}. \tag{13}$$

[21] Theorem 4.1 is a generalization of Proposition 4.2 in Lahkar [28] and Proposition 3 in Cheung and Lahkar [5] to the case of multiple populations and a more general aggregate strategy level.

[22] Well–known dynamics that converge in potential games are the replicator dynamic, the Brown–von Neumann–Nash dynamic, the pairwise comparison dynamic and the logit dynamic. See, for example, Cheung and Lahkar [5] for a discussion of such convergence in potential games with a continuous set of strategies.

Thus, at the unique Nash equilibrium of F, every agent in population p plays $\alpha_p^* \in \mathcal{S}$ as defined in (13).

Next, we characterize the Nash equilibrium of the Tullock contest $\hat{F}$ with affirmative action as defined in (3). In order to distinguish it form (12), we denote the quasi–potential function for $\hat{F}$ as $\hat{g}$. Recalling that $\hat{F}$ is characterized by $\phi_p(x) = k_p^{\frac{r}{\gamma}} x^r$, $\beta(z) = \frac{V}{z}$ and $c_p(x) = k_p x^\gamma$, we use (11) to write $\hat{g}$ as

$$\begin{aligned}\hat{g}(\alpha_1, \alpha_2, \cdots, \alpha_n) &= \int_{\underline{x}^r \sum_p m_p k_p^{\frac{r}{\gamma}}}^{\sum_{p\in\mathcal{P}} m_p k_p^{\frac{r}{\gamma}} \alpha_p^r} \frac{V}{z} dz - \sum_{p\in\mathcal{P}} m_p k_p \alpha_p^\gamma \\ &= V \log\left(\sum_{p\in\mathcal{P}} m_p k_p^{\frac{r}{\gamma}} \alpha_p^r\right) - V \log\left(\underline{x}^r \sum_p m_p k_p^{\frac{r}{\gamma}}\right) - \sum_{p\in\mathcal{P}} m_p k_p \alpha_p^\gamma. \qquad (14)\end{aligned}$$

We now maximize $\hat{g}$ to characterize the Nash equilibrium of $\hat{F}$. As in Proposition 4.2, we assume that $\hat{g}$ has an interior maximizer in $\prod_{n=1}^n [\underline{x}, \bar{x}]$. We state the result in the following proposition. The proof of the result is in Appendix A.2.1.

Proposition 4.3 *Consider the Tullock contest with affirmative action $\hat{F}$ as defined in (3). Assume that $r \in (0, 1]$ and $\gamma \geq 1$, with the further restriction that if $r = 1$, then $\gamma > 1$. Further assume that $\underline{x}$ is sufficiently small and $\bar{x}$ is sufficiently large that (14) has an interior maximizer in $\prod_{p=1}^n [\underline{x}, \bar{x}]$. Then, $\hat{F}$ has a unique Nash equilibrium $\hat{\mu}^* = \left(m_1\delta_{\hat{\alpha}_1^*}, m_2\delta_{\hat{\alpha}_2^*}, \cdots, m_n\delta_{\hat{\alpha}_n^*}\right)$ such that for every $p \in \mathcal{P}$,*

$$\hat{\alpha}_p^* = k_p^{\frac{-1}{\gamma}} \left(\frac{rV}{\gamma}\right)^{\frac{1}{\gamma}}. \qquad (15)$$

Thus, at the unique Nash equilibrium of $\hat{F}$, every agent in population p plays $\hat{\alpha}_p^* \in \mathcal{S}$ as defined in (15).

Finite player Tullock contests with more than two agents can be solved in closed form only when the strategy impact and cost functions are symmetric or linear (Franke [15]).[23] Propositions 4.2 and 4.3, however, show that the large population approach is tractable enough to yield closed form solutions even under asymmetric and non–linear conditions. In fact, as our application to the affirmative action contest shows, our technique is general enough to accommodate such non–linearities or asymmetries in the impact and cost functions without any modification.[24]

[23]Results on existence and uniqueness of equilibrium are, however, possible even with such asymmetries or non–linearities (for example, Perez-Castrillo and Verdier [39], Nti [36], Szidarovszky and Okuguchi [47], Cornes and Hartley [10], Yamazaki [50]).

[24]In contrast, in finite player contests, the generalization from symmetric to asymmetric CSFs and cost functions is not so straightforward. The finite player contest literature has, therefore, proceeded sequentially by first considering symmetric CSFs and cost functions (Perez-Castrillo and Verdier [39], Nti [36]) and then introducing asymmetries in the CSF (Szidarovszky and Okuguchi [47], Yamakazi [50]) and the cost functions (Cornes and Hartley [10]). Also see Corchón [8] for a review of existence and uniqueness results in both the symmetric and asymmetric cases.

Further, our approach need not be restricted to Tullock contests. Tullock contests belong to the more general class of logit–CSF contests in which the strategy impact function take the form $\phi_p(x)$ (Dixit [12]).[25] Payoffs in such a large population contest will take the form $F_{x,p}(\mu) = \frac{\phi_p(x)}{A(\mu)} V - c_p(x)$, where $A(\mu) = \sum_p \int_{\mathcal{S}} \phi_p(x)\mu_p(dx)$. As long as the impact functions ϕ_p and cost functions c_p satisfy Assumption 3.1, our equilibrium characterization method will apply.

5 Equal Treatment and Affirmative Action: Comparative Statics

We now consider the implications of equal treatment and affirmative action on aggregate welfare and effort incentives. We measure individual welfare using equilibrium payoffs in the two contests and aggregate welfare by using aggregate payoffs. It is also relevant to note that in both contests, the fact that the shares of agents must add up to 1 implies that the aggregate payoff at any social state μ must be identical. Thus, from (2) and (3), we obtain that the aggregate payoffs in the two contests F and $\hat{F}$ are

$$\bar{F}(\mu) = \bar{\hat{F}}(\mu) = V \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} \frac{\phi_p(x)}{A(\mu)} \mu_p(dx) - \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} k_p x^\gamma \mu_p(dx) = V - C(\mu), \tag{16}$$

with $\phi_p(x) = x^r$ in F, $\phi_p(x) = k_p^{\frac{r}{\gamma}} x^r$ in $\hat{F}$, $A(\mu)$ as defined in (5) and aggregate cost $C(\mu)$ as defined in (7). We now state the following proposition on equilibrium payoffs. The proof is in Appendix A.3

Proposition 5.1 *Consider the two contests F and $\hat{F}$ as defined in (2) and (3) with their respective Nash equilibrium μ^* and $\hat{\mu}^*$ as characterized in Propositions 4.2 and 4.3. Recall that $r < \gamma$ in both contests.*

1. *At the Nash equilibrium μ^* of the equal treatment contest F, every agent in population $p \in \mathcal{P}$ obtains payoff*

$$F_{\alpha_p^*,p}(\mu^*) = \frac{(\alpha_p^*)^r}{A(\mu^*)} V - k_p(\alpha_p^*)^\gamma = \frac{k_p^{\frac{r}{r-\gamma}}}{\sum_{q \in \mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}} V \left(1 - \frac{r}{\gamma}\right), \tag{17}$$

where $\alpha_p^* \in \mathcal{S}$ is as defined in (13) and $A(\mu^*) = \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} x^r \mu_p^*(dx)$ in F. Hence, aggregate payoff at μ^* in F is

$$\bar{F}(\mu^*) = V - C(\mu^*) = V\left(1 - \frac{r}{\gamma}\right). \tag{18}$$

2. *At the Nash equilibrium $\hat{\mu}^*$ of the affirmative action contest $\hat{F}$, every agent in population*

[25] Dixit [12] defines a logit–CSF contest with $N-$players as one in which the CSF for player i playing strategy x_i is $\frac{\phi_i(x_i)}{\sum_{j=1}^N \phi_j(x_j)}$. Thus, in a Tullock contest, $\phi_i(x_i) = x_i^r$.

$p \in \mathcal{P}$ obtains payoff

$$\hat{F}_{\hat{\alpha}_p^*,p}(\hat{\mu}^*) = \frac{k_p^{\frac{r}{\gamma}}(\hat{\alpha}_p^*)^r}{A(\hat{\mu}^*)}V - k_p(\hat{\alpha}_p^*)^\gamma = V\left(1 - \frac{r}{\gamma}\right), \tag{19}$$

where $\hat{\alpha}_p^* \in \mathcal{S}$ is as defined in (15) and $A(\hat{\mu}^*) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p^{\frac{r}{\gamma}} x^r \hat{\mu}_p^*(dx)$ in $\hat{F}$. Hence, aggregate payoff at $\hat{\mu}^*$ in $\hat{F}$ is

$$\bar{\hat{F}}(\hat{\mu}^*) = V - C(\hat{\mu}^*) = V\left(1 - \frac{r}{\gamma}\right). \tag{20}$$

We obtain two important conclusions from Proposition 5.1. First, as can be seen from (19), affirmative action in our model ensures perfect equality in equilibrium. Agents of every type, irrespective of their cost parameter k_p, obtain the same equilibrium payoff. The intuition behind this result is that at the Nash equilibrium under affirmative action, the cost of effort of every agent is equal. Thus, $c_p(\hat{\alpha}_p^*) = k_p(\hat{\alpha}_p^*)^\gamma = \frac{rV}{\gamma}$, for all $p \in \mathcal{P}$, as can be seen in (57). By the design of our affirmative action policy, this implies that the CSF for every agent at the affirmative action equilibrium is 1. The equilibrium payoff then follows from the definition of the affirmative action contest (3). Equal treatment does not ensure equality. Instead, as is to be expected, the lower is the cost parameter k_p, the higher is equilibrium payoff, as is evident from (17).

The second important conclusion is that the aggregate equilibrium payoff in both contests is identical, as is shown by (18) and (20). This does not immediately follow from (16) because the two equilibria μ^* and $\hat{\mu}^*$ are different. Instead, it requires some calculation to show, as we have done in (55) and (58) respectively, that the aggregate cost of effort at both equilibria are equal. Thus, $C(\mu^*) = C(\hat{\mu}^*) = \frac{rV}{\gamma}$. The equilibrium values of aggregate payoff then follows from (16). Together, these two conclusions encapsulate the welfare implications of affirmative action in our model. It contributes towards equality while being as efficient (measured in terms of aggregate payoff) as the equal treatment contest (2). It is worth observing that these strong conclusions about welfare arise from the closed form solutions we have been able to derive for our Tullock contest models under a wider variety of conditions; something that would not have been feasible in finite player models.

Our next objective is to assess the effect of affirmative action on effort incentives, which we do in two ways. First, we consider whether affirmative action increases effort of certain types of agents as compared to equal treatment. Second, we compare aggregate effort level in equilibrium under the two contests. The comparison of equilibrium effort levels follows immediately from Propositions 4.2 and 4.3.

Corollary 5.2 *Consider the Nash equilibrium effort levels α_p^* and $\hat{\alpha}_p^*$ under the two contests F and $\hat{F}$ as established in Propositions 4.2 and 4.3 respectively. Define*

$$\bar{k} = \left(\frac{1}{\sum_{q\in\mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}}\right)^{\frac{\gamma-r}{r}}. \tag{21}$$

Then, $\hat{\alpha}_p^* > \alpha_p^*$ if and only if $k_p > \bar{k}$. Further, $\hat{F}_{\hat{\alpha}_p^*,p}(\hat{\mu}^*) > F_{\alpha_p^*,p}(\mu^*)$ if and only if $k_p > \bar{k}$, where $F_{\alpha_p^*,p}(\mu^*)$ and $\hat{F}_{\hat{\alpha}_p^*,p}(\hat{\mu}^*)$ are the equilibrium payoffs (17) and (19) respectively.

Proof. The relationship between α_p^* and $\hat{\alpha}_p^*$ follows from (13) and (15). The relationship between $F_{\alpha_p^*,p}(\mu^*)$ and $\hat{F}_{\hat{\alpha}_p^*,p}(\hat{\mu}^*)$ follows from (17) and (19). ■

Corollary 5.2 shows that there exists a threshold (21) such that agents with cost parameter k_p above that threshold exert higher effort under affirmative action while agents with cost parameter below the threshold exert lower effort. This conclusion is along expected lines as affirmative action favors agents with a higher cost of effort. We do note, however, that this comparison is between the two contests of equal treatment and affirmative action. Within each contest, low–cost agents exert higher effort than high–cost agents in equilibrium, as is evident from (13) and (15). The same conclusions hold with respect to equilibrium payoffs.

But does affirmative action also increase total effort? To consider this question, recall the Nash equilibrium effort levels α_p^* and $\hat{\alpha}_p^*$ in the two contests as characterized in (13) and (15) respectively. Since these are the effort levels by every agent in population p and the mass of population p is m_p, the equilibrium aggregate effort levels in F and $\hat{F}$ are $\sum_p m_p \alpha_p^*$ and $\sum_p m_p \hat{\alpha}_p^*$ respectively. The following proposition shows that affirmative action reduces aggregate effort in equilibrium. The proof of the result is in Appendix A.3.

Proposition 5.3 *Consider the two contests F and $\hat{F}$ with their respective Nash equilibria μ_p^* and $\hat{\mu}_p^*$ as characterized in Propositions 4.2 and 4.3 respectively. Recall our assumptions that $r \in (0,1]$ and $\gamma \geq 1$ with the further condition that if $r = 1$, then $\gamma > 1$. Then,*

$$\sum_p m_p \hat{\alpha}_p^* < \sum_p m_p \alpha_p^*, \tag{22}$$

where α_p^* and $\hat{\alpha}_p^*$ are the equilibrium effort levels as characterized in (13) and (15) respectively. Hence, equilibrium aggregate effort in the affirmative action contest $\hat{F}$ is strictly less than that in the equal treatment contest F.

The proof of this result involves some tedious algebra. But we can obtain the intuition behind it from two of our previous results. Corollary 5.2 shows that affirmative action increases the effort level of high cost agents while reducing that of low cost agents. On the other hand, we have shown in the proof of Proposition 5.1 that $C(\mu^*) = C(\hat{\mu}^*) = \frac{rV}{\gamma}$. Therefore, the aggregate cost in equilibrium doesn't go up under affirmative action even though it is high cost agents who exert more effort. It is then reasonable to conclude that total effort goes down under affirmative action.

Thus, the broad consequences of affirmative action in our model are as follows. The effort of agents who have been subjected to high levels of discrimination goes up without any adverse effect on aggregate welfare.[26] In fact, social equality improves. But aggregate effort goes down.

[26] We qualify this statement by noting that we identify aggregate welfare entirely with the aggregate payoff of agents. We do not consider the welfare of the planner in such calculations. If the welfare of the planner is measured by the aggregate effort of agents, then that welfare does go down under affirmative action.

6 The Linear Case

We now drop Assumption 3.1(4) and consider Tullock contests (2) and (3) in which $r = \gamma = 1$. Franke's [15] analysis of affirmative action in finite player Tullock contests is restricted to this case of $r = \gamma = 1$, except when the number of player is two in which case $r \leq 1$. Therefore, the analysis in this section will also allow us to compare our results with that of Franke [15].

Even with $r = \gamma = 1$, Proposition 3.3 continues to apply to the two contests F and $\hat{F}$.[27] Thus, both contests are potential games with concave potential functions f and $\hat{f}$ defined by (9) and (10) respectively.[28] Hence, both F and $\hat{F}$, under this particular condition, have a convex set of Nash equilibria that can be characterized by maximizing their respective potential functions. The reason we had to leave out this case from our equilibrium characterization in Section 4.1 is that if both ϕ_p and c_p are linear, we cannot apply Lemma A.3 to show that the quasi–potential functions g and $\hat{g}$ defined by (12) and (14) respectively are strictly convex and, therefore, have a unique maximizer. However, the arguments in Lemma A.3 still suffice to show that both g and $\hat{g}$ are concave. This allows us to apply alternative arguments to characterize their set of maximizers.[29]

For this purpose, we now order the populations such that $k_1 < k_2 < \cdots < k_n$. We also assume that $\underline{x}$ is sufficiently close to zero and $\bar{x}$ is sufficiently high so that we don't have to deal with tedious boundary issues. We first characterize Nash equilibria and payoffs of F using its quasi–potential function g. The proof of the result is in Appendix A.4.

Proposition 6.1 *Consider the Tullock contest F with equal treatment as defined in (2) with $r = \gamma = 1$. Hence, $\phi_p(x) = x$ and $c_p(x) = k_p x$ in F. Assume that $k_1 < k_2 < \cdots < k_n$. Then, F has a convex set of Nash equilibria*

$$NE(F) = \left\{\mu^* \in \Delta : a(\mu_1^*) = m_1\alpha_1^*, \mu_p^* = m_p\delta_{\underline{x}} \text{ for all } p \neq 1\right\}, \tag{23}$$

where $a(\mu_1^*) = \int_{\mathcal{S}} x\mu_1^*(dx)$ as defined in (4) and $\alpha_1^* = \frac{V-(1-m_1)k_1\underline{x}}{m_1k_1}$ as characterized in Lemma A.4.

Further, the payoff of every agent of every type at any $\mu^* \in NE(F)$ is 0. Therefore, the aggregate payoff at any such Nash equilibrium is also 0.

We may relate Proposition 6.1 to the analysis in Franke [15]. When the number of players is more than two in that paper, then the unique Nash equilibrium under equal treatment may involve only a subset of the players exerting positive effort. In our case, there is a convex set of Nash equilibria. But in any such equilibrium, all agents except the ones with the lowest cost parameter exert the lowest possible effort $\underline{x}$. In that sense, only those agents whose cost of effort is the lowest participate actively in the contest. Their payoff, though, in equilibrium is zero as is the case with everybody else.

[27] Recall that the potential game property in Proposition 3.3 does not depend upon Assumption 3.1. The concavity of the potential function only requires that $\beta(z) = \frac{V}{z}$ be strictly declining (Assumption 3.1(1)).

[28] A concave potential function implies that the underlying potential game has a convex set of Nash equilibria, which are maximizers of the potential function (Sandholm [40], Cheung and Lahkar [5]).

[29] As we will see in Lemma A.4, g does have a unique maximizer. But that does not follow from Lemma A.3.

We now characterize Nash equilibria and payoffs in the affirmative action contest $\hat{F}$ as defined in (3). The proof of the result is in Appendix A.4.

Proposition 6.2 *Consider the Tullock contest $\hat{F}$ with affirmative action as defined in (3) with $r = \gamma = 1$. Hence, $\phi_p(x) = c_p(x) = k_p x$ in $\hat{F}$. Then, $\hat{F}$ has a convex set of Nash equilibria*

$$NE(\hat{F}) = \{\hat{\mu}^* \in \Delta : A(\hat{\mu}^*) = V\}, \tag{24}$$

where $A(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} \phi_p(x)\mu_p(dx) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p x \mu_p(dx)$ in $\hat{F}$.

Further, the payoff of every agent of every type at any $\hat{\mu}^* \in NE(\hat{F})$ is 0. Therefore, the aggregate payoff at any such Nash equilibrium is also 0.

As in Proposition 6.2, the affirmative action contest with linear parameters also has a convex set of equilibria with zero payoff for all agents in any such equilibrium.[30] Therefore, in a trivial sense, there is no distinction between equal treatment and affirmative action on grounds of aggregate payoff and equality when $r = \gamma = 1$.

Nevertheless, the policy maker may still favor the affirmative action policy (3) because it at least creates the possibility of active participation beyond the group of agents in the lowest cost group. This is because the set of Nash equilibria in $\hat{F}$ differs from that of F significantly in one respect. In any Nash equilibrium of F, as shown in Proposition 6.1, only agents of population 1, i.e. agents with the lowest cost parameter, participate actively in the contest in the sense of playing $x > \underline{x}$. However, in $\hat{F}$, there are Nash equilibria in which all agents participate actively. All agents at such an equilibrium play a strategy $x > \underline{x}$.[31] This observation is also consistent with Franke's [15] finding that once affirmative action is introduced in finite player contests with more than two players, all players participate in the unique Nash equilibrium of that contest. Of course, in our case, due to the convexity of $NE(\hat{F})$, there may also be equilibria in which all agents in some population play $\underline{x}$.

Thus, as in Section 5, affirmative action gives rise to equilibria where high cost agents participate more. But what about aggregate effort? The following result shows that it goes down under affirmative action. Thus, even when $r = \gamma = 1$, which is the one case left out in Proposition 5.3, affirmative action is detrimental to aggregate effort. The proof of the result is in Appendix A.4.

Proposition 6.3 *Consider the two contests F and $\hat{F}$ defined by (2) and (3) respectively under the condition $r = \gamma = 1$. Then, at every Nash equilibrium of F as characterized in (23), the total effort in the society is $\frac{V}{k_1}$. Further, at any Nash equilibrium $\hat{\mu}^*$ of $\hat{F}$ as characterized in (24), the aggregate effort $\sum_p \int_{\mathcal{S}} x\hat{\mu}_p^*(dx) < \frac{V}{k_1}$. Hence, affirmative action reduces aggregate effort as compared to equal treatment.*

[30] Note that if $r = \gamma = 1$, then $V\left(1 - \frac{r}{\gamma}\right) = 0$. Hence, the zero equilibrium payoff conclusion is along expected lines if we mechanically apply the equilibrium payoffs calculated in Proposition 5.1 to this particular case.

[31] For example, one such Nash equilibrium $\hat{\mu}^*$ would be where all agents in population p play strategy $\frac{V}{k_p}$. This equilibrium would satisfy $A(\hat{\mu}^*) = V$ as required by (24).

We can relate Proposition 6.3 to Proposition 3 in Franke [15]. There, it is shown affirmative action may increase aggregate effort. In contrast, the above result shows that this is not possible in a large population Tullock contest. Intuitively, this occurs because satisfying the condition in that paper that determines whether affirmative action raises aggregate effort becomes increasingly difficult as the number of agents increases. Hence, in the continuum limit, it should become impossible to satisfy that condition, which means affirmative action cannot raise aggregate effort in the large population case.[32]

To conclude, we find that in both the linear and non–linear cases, affirmative action induces participation by disadvantaged agents even though aggregate effort goes down. But unlike in the non–linear case, we do not find any effect of affirmative action on agents' payoffs under linear conditions.

7 Conclusion

We have considered affirmative action in large population Tullock contests. The standard Tullock contest is a contest with equal treatment policy wherein agents who exert equal effort have an equal contest success function (CSF). In contrast, under affirmative action policy, the standard contest is modified so that agents who have an equal cost of effort have an equal CSF. The underlying justification for this policy is the normative ideal that differences in costs have their roots in historical discrimination which agents are not responsible for. Hence, social policy should aim to redress the effects of such discrimination.

We model the two contests as generalized aggregative potential games. Using the method of quasi–potential functions, we characterize the Nash equilibria of such contests under fairly general conditions of asymmetries and non–linearities in the CSFs and cost functions. Our results show that affirmative action increases effort by more disadvantaged agents. It promotes equality without reducing aggregate payoff in a significant way. On the other hand, affirmative action reduces aggregate effort in society. Therefore, our results suggest that affirmative action succeeds in promoting diversity and equality in society without significantly sacrificing aggregate welfare. But it does so at the cost of distorting effort incentives which reduces aggregate effort. In establishing these results on affirmative action, the paper also contributes to the theory of large population contests by introducing the potential game methodology to analyze such contests.

Our exercise raises certain interesting questions for further research. First, this paper has primarily focused on large population Tullock contests. Is it possible to extend this approach to

[32] Condition (12) in Franke [15] decides whether affirmative action increases aggregate effort. We can translate that condition to our large population context. Under equal treatment, only type 1 agents participate (Proposition 6.1). Hence, the arithmetic mean of the cost parameter of the participants is k_1. Under affirmative action, there are equilibria where all agents participate (Proposition 6.2). The harmonic mean of the cost parameters of all agents in the society is $\left(\sum_p \frac{m_p}{k_p}\right)^{-1}$. The appropriate translation of Franke's [15] condition (12) to our large population context is that affirmative action will increase effort if $\frac{k_1}{\left(\sum_p \frac{m_p}{k_p}\right)^{-1}} > 1$. But this condition cannot be satisfied because $k_1 < k_p$, for all $p > 1$.

large population versions of other canonical contest models like all–pay auctions and rank–order tournaments? This would require further generalization of the theory of generalized aggregative games. For example, it is well known that all–pay auctions, where the winner gets the prize with probability one and all participants pay, is equivalent to a Tullock contest with $r = \infty$. Hence, the strategy impact function in such a contest would no longer be concave. This would require us to relax some of the conditions in Assumption 3.1. Another interesting research question, alluded to in the Introduction, is taking a mechanism design approach and designing the CSF in large population contests so as to maximize aggregate effort. Yet another possible extension is to analyze other forms of affirmative action, for example, extra prizes in the manner of Dahm and Esteve–Gonzáles [11], in large population Tullock contests.

One issue we have not addressed in this paper is the informational requirement to coordinate on the Nash equilibrium of the large population contest models or more generally, in generalized aggregative potential games. Finite player contests have been analyzed in the classical framework of both complete and incomplete information. The informational requirement in large population models are, however, best assessed in the framework of evolutionary game theory. This is because in such models, it is unrealistic to expect agents to immediately coordinate on a Nash equilibrium. Instead, as per the evolutionary viewpoint, agents adjust their strategies myopically on the basis of their current payoff. In general, this strategy adjustment process may or may not lead society to a Nash equilibrium. But in potential games such as our contest models, it will converge to a Nash equilibrium (Sandholm [40], Cheung and Lahkar [5]).

From this evolutionary perspective, the relevance of information arises in the context of what an agent needs to know to revise strategies. Sandholm [41] addresses this question in detail. The answer depends upon which evolutionary model we choose. For example, under the replicator dynamic, which is the most well known evolutionary dynamic, the informational requirements are fairly minimal. This dynamic is generated when an agent imitates some other member of his type or population, either because the other agent is more successful or because the agent is dissatisfied with his own payoff. All this model calls for is that an agent should be able to recognize another agent of her own type. In applications like contests or affirmative action, this is a reasonable assumption since we are envisaging each type as representing a particular social group with some specific demographic characteristics. In particular, the key question that arises in classical complete or incomplete information models—whether agents know the type of other agents or just the type distribution—is irrelevant in the replicator dynamic. Agents do not need to know the type of any agent outside their own population or even the type distribution. In other evolutionary dynamics, the informational requirement may be more demanding. For example, the logit dynamic, which is another well known model of evolution, depends upon a perturbed notion of best response where agents choose the best response with probability nearly but not completely one. This would require agents to compute the payoff of all their strategies. In generalized aggregative games such as (6), this would require the agent to know $\beta(A(\mu))$ which, in turn, may require the agent to know the generalized aggregate strategy level $A(\mu)$. Depending upon context, this may be easy or difficult.

But knowing it does not require the agent to know types or the type distribution. Therefore, even under this dynamic, the core issue in the complete versus incomplete information debate about knowledge of types or type distribution does not arise.

A Appendix

A.1 Appendix to Section 3

A.1.1 Fréchet Derivative

The definition of the Fréchet derivative is as follows (Zeidler [51], Chapter 4).

Definition A.1 *Let X and Y are Banach spaces. We say that $g : X \to Y$ is Fréchet-differentiable at x if there exists a continuous linear map $T : X \to Y$ such that $g(x + \vartheta) = g(x) + T\vartheta + o(\|\vartheta\|)$ for all ϑ in some neighborhood of zero in X. If it exists, this T is called the Fréchet-derivative of g at x, and is written as $Dg(x)$.*

To apply the Fréchet derivative in our context, we denote by $\mathcal{M}(\mathcal{S})$ the space of finite signed measures on $(\mathcal{S}, \mathcal{B})$, where $\mathcal{B}$ is the Borel $\sigma-$algebra on $\mathcal{S}$. We then impose the strong topology on $\mathcal{M}(\mathcal{S})$. The strong topology is the topology induced by the variational norm on $\mathcal{M}(\mathcal{S})$. For $\mu \in \mathcal{M}(\mathcal{S})$, the variational norm is given by $\|\mu\| = \sup_g |\int_{\mathcal{S}} gd\mu|$ where g is a measurable function $g : \mathcal{S} \to \mathbf{R}$ such that $\sup_{x\in S} |g(x)| \leq 1$. We define the variational norm on $\mathscr{M} = \prod_{p=1}^{n} \mathcal{M}(S)$ as (Lahkar and Mukherjee [29])

$$\|\nu\| = \max\{\|\nu^1\|, \|\nu^2\|, \cdots, \|\nu^n\|\} \text{ for } \nu = (\nu^1, \nu^2, \cdots, \nu^n) \in \mathscr{M}. \tag{25}$$

We seek to calculate Fréchet derivatives on the Banach space $(\mathscr{M}, \|\cdot\|)$, where $\|\cdot\|$ is the variational norm on $\mathscr{M}$ as defined in (25). Consider a function $f : \mathscr{M} \to \mathbf{R}$ that is Fréchet differentiable when $\mathscr{M}$ is endowed with the variational norm. The Fréchet derivative of f at $\mu \in \mathscr{M}$ is a continuous linear map

$$Df(\mu) : \mathscr{M} \to \mathbf{R}$$

that maps the direction $\zeta = (\zeta^1, \cdots, \zeta^n) \in \mathscr{M}$ to rates of change in f when μ changes in the direction ζ. We write this linear transformation as $Df(\mu)\zeta$.

Consider now a function $f : \mathscr{M} \to \mathbf{R}$ that is Fréchet differentiable when $\mathscr{M}$ is endowed with the variational norm. We denote the Fréchet derivative of f at $\mu \in \mathscr{M}$ in the direction $\zeta = (\zeta_1, \cdots, \zeta_n) \in \mathscr{M}$ as $Df(\mu)\zeta$. Intuitively, $Df(\mu)\zeta$ is the change in f when μ changes in the direction ζ.

Suppose now that there exists an element $\nabla f(\mu) : \mathcal{S} \times \mathcal{P} \to \mathfrak{M}_b(\mathcal{S} \times \mathcal{P})$ such that

$$Df(\mu)\zeta = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} \nabla f(\mu)(x, p)\zeta_p(dx) = \langle \nabla f(\mu), \zeta \rangle, \text{ for all } \zeta = (\zeta_1, \cdots, \zeta_n) \in \mathscr{M}, \tag{26}$$

where we use the "inner product" notation $\langle\cdot,\cdot\rangle : \mathfrak{M}_b(\mathcal{S}\times\mathcal{P})\times\mathscr{M}\to\mathbf{R}$ to denote $\langle g,\nu\rangle = \sum_{p=1}^{n}\int_{\mathcal{S}} g(x,p)\nu_p(dx)$, for $g\in\mathfrak{M}_b(\mathcal{S}\times\mathcal{P})$ and $\nu\in\mathscr{M}$. We then call $\nabla f(\mu)$ the *gradient* of the function f.

A.1.2 Proofs in Section 3

Proof of Proposition 3.3: We first extend the domain of the generalized aggregate strategy function $A(\cdot)$ defined in (5) from Δ to $\mathscr{M}$ (and, therefore, the domain of the $a(\cdot)$ defined in (4) from $\mathcal{M}^{+}_{m_p}(\mathcal{S})$ to $\mathcal{M}(\mathcal{S})$). We also extend the domain of the benefit function β from $\left[\sum_p m_p\phi_p(\underline{x}),\sum_p m_p\phi_p(\bar{x})\right]$ to $\mathbf{R}$ such that this extension is bounded and differentiable on $\mathbf{R}$.[33]

Consider f as defined as in (8) with appropriate extensions of $A(\cdot)$, $C(\cdot)$ and $\beta(\cdot)$. We show that for all $\mu\in\Delta$ and all $(x,p)\in\mathcal{S}\times\mathcal{P}$,

$$\nabla f(\mu)(x,p) = F_{x,p}(\mu) = \phi_p(x)\beta(A(\mu)) - c_p(x), \tag{27}$$

where $A(\mu)$ is as defined in (5).

Let $\zeta = (\zeta_1,\zeta_2,\cdots,\zeta_n)\in\mathscr{M}$, where ζ_p represents a direction of change in μ_p. Then,

$$Df(\mu)\zeta = \beta(A(\mu))DA(\mu)\zeta - DC(\mu)\zeta, \tag{28}$$

where $C(\mu)$ is as defined in (7). Note that

$$A(\mu+\zeta) = \sum_{p\in\mathcal{P}} a(\mu_p+\zeta_p) = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}}\phi_p(\tilde{x})(\mu_p+\zeta_p)(d\tilde{x}) = A(\mu) + \sum_{p\in\mathcal{P}}\int_{\mathcal{S}}\phi_p(\tilde{x})\zeta_p(d\tilde{\tilde{x}}).$$

Therefore,

$$DA(\mu)\zeta = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}}\phi_p(\tilde{x})\zeta_p(d\tilde{x}). \tag{29}$$

Further,

$$\begin{aligned} C(\mu+\zeta) &= \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} c_p(x)(\mu_p+\zeta_p)(dx) \\ &= \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} c_p(x)\mu_p(dx) + \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} c_p(x)\zeta_p(dx) \\ &= C(\mu) + \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} c_p(x)\zeta_p(dx). \end{aligned}$$

Hence,

$$DC(\mu)\zeta = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} c_p(x)\zeta_p(dx). \tag{30}$$

[33] This extension is required because we have extended the domain of $A(\cdot)$ from Δ to $\mathscr{M}$. Hence, $\beta(A(\mu))$ may take any value in $\mathbf{R}$, including negative ones.

Inserting (29) and (30) into (28) and using (26), we obtain

$$\sum_{p\in\mathcal{P}}\int_{\mathcal{S}}\nabla f(\mu)(x,p)\zeta_p(dx)=\beta(A(\mu))\sum_{p\in\mathcal{P}}\int_{\mathcal{S}}\phi_p(\tilde{x})\zeta_p(d\tilde{x})-\sum_{p\in\mathcal{P}}\int_{\mathcal{S}}c_p(\tilde{x})\zeta_p(d\tilde{x})$$

This equation holds for all $\zeta\in\mathscr{M}$. In particular, it holds for ζ such that $\zeta_p=\delta_x$ and $\zeta_k=0$ for all $k\neq p$. With this ζ, we obtain

$$\nabla f(\mu)(x,p)=\phi_p(x)\beta(A(\mu))-c_p(x),$$

which gives us (27). This establishes the result.

To establish concavity of f, let $\mu,\nu\in\Delta$, $\mu\neq\nu$, be two social states. We need to show that

$$\begin{aligned}&\int_{\sum_p m_p\phi_p(\underline{x})}^{A(\lambda\mu+(1-\lambda)\nu)}\beta(z)dz-C(\lambda\mu+(1-\lambda)\nu)\\&\quad\geq\lambda\int_{\sum_p m_p\phi_p(\underline{x})}^{A(\mu)}\beta(z)dz+(1-\lambda)\int_{\sum_p m_p\phi_p(\underline{x})}^{A(\nu)}\beta(z)dz-\lambda C(\mu)-(1-\lambda)C(\nu).\end{aligned}\tag{31}$$

Fix $\lambda\in(0,1)$. Clearly,

$$\sum_{p\in\mathcal{P}}\int_{\mathcal{S}}c_p(x)(\lambda\mu_p+(1-\lambda)\nu_p)(dx)=\lambda\sum_{p\in\mathcal{P}}\int_{\mathcal{S}}c_p(x)\mu_p(dx)+(1-\lambda)\sum_{p\in\mathcal{P}}\int_{\mathcal{S}}c_p(x)\nu_p(dx).$$

Hence,

$$C(\lambda\mu+(1-\lambda)\nu)=\lambda C(\mu)+(1-\lambda)C(\nu).\tag{32}$$

Since β is strictly decreasing on $\left[\sum_{p\in\mathcal{P}}m_p\phi_p(\underline{x}),\sum_{p\in\mathcal{P}}m_p\phi_p(\bar{x})\right]$, $\int_{\sum_p m_p\phi_p(\underline{x})}^{\alpha}\beta(z)dz$ is strictly concave for all $\alpha\in\left[\sum_{p\in\mathcal{P}}m_p\phi_p(\underline{x}),\sum_{p\in\mathcal{P}}m_p\phi_p(\bar{x})\right]$. Moreover, since $\mu,\nu\in\Delta$, $A(\lambda\mu+(1-\lambda)\nu)\in\left[\sum_{p\in\mathcal{P}}m_p\phi_p(\underline{x}),\sum_{p\in\mathcal{P}}m_p\phi_p(\bar{x})\right]$ for all $\lambda\in[0,1]$. Combining these facts with the linearity of $A(\cdot)$, we obtain

$$\begin{aligned}\int_{\sum_p m_p\phi_p(\underline{x})}^{A(\lambda\mu+(1-\lambda)\nu)}\beta(z)dz&=\int_{\sum_p m_p\phi_p(\underline{x})}^{\lambda A(\mu)+(1-\lambda)A(\nu)}\beta(z)dz\\&\geq\lambda\int_{\sum_p m_p\phi_p(\underline{x})}^{A(\mu)}\beta(z)dz+(1-\lambda)\int_{\sum_p m_p\phi_p(\underline{x})}^{A(\nu)}\beta(z)dz,\end{aligned}\tag{33}$$

with $\geq$ holding with equality only if μ,ν such that $A(\mu)=A(\nu)$.

Combining (32) and (33), we obtain (31). This establishes that f is concave but not strictly concave on Δ, strict concavity failing because there are $\mu\neq\nu$ such that $A(\mu)=A(\nu)$. The convexity of the set of Nash equilibria of F then follows from Corollary 2 in Cheung and Lahkar [5]. ■

A.2 Appendix to Section 4

First, we establish Lemmas A.2 and A.3 required for the proof of Theorem 4.1.

Lemma A.2 *Consider the generalized aggregative potential game F defined by (6) with potential function f defined by (8) and quasi–potential function g defined by (11). Let $\mu \in \Delta$ and denote $\frac{a(\mu_p)}{m_p} = \phi_p(\alpha_p)$, for some $\alpha_p \in [\underline{x}, \bar{x}]$, for all $p \in \mathcal{P}$, where $a(\mu_p)$ is as defined in (4). Let Assumption 3.1 hold.*

1. *Suppose μ_p is monomorphic for all p. Then, $f(\mu) = g(\alpha_1, \cdots, \alpha_n)$.*
2. *Suppose there exists at least one p such that μ_p is polymorphic. Then, $g(\alpha_1, \alpha_2, \cdots, \alpha_n) > f(\mu)$.*

Proof. Consider the potential function f defined by (8) and the quasi–potential function g defined by (11). Note that since $\frac{a(\mu_p)}{m_p} = \phi_p(\alpha_p)$, $A(\mu) = \sum_{p\in\mathcal{P}} a(\mu_p) = \sum_{p\in\mathcal{P}} m_p\phi_p(\alpha_p)$. Therefore, any difference between f and g can only arise due to the difference between $C(\mu)$ and $\sum_{p\in\mathcal{P}} m_p c_p(\alpha_p)$.

1. Let μ be monomorphic for every p. In that case, the fact that $\frac{a(\mu_p)}{m_p} = \phi_p(\alpha_p)$ implies $\mu_p = m_p\delta_{\alpha_p}$ for all p. This implies $\int_{\mathcal{S}} c_p(x)\mu_p(dx) = m_p c_p(\alpha_p)$ so that $C(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} c_p(x)\mu_p(dx) = \sum_{p\in\mathcal{P}} m_p c_p(\alpha_p)$. Hence, $f(\mu) = g(\alpha_1, \cdots, \alpha_n)$.
2. Let μ_p be polymorphic for some $p \in \mathcal{P}$, with $\frac{a(\mu_p)}{m_p} = \phi_p(\alpha_p)$. For any population q, ϕ_q is concave. Moreover, it is given that $\frac{a(\mu_q)}{m_q} = \phi_q(\alpha_q)$. Hence,

$$\phi_q\left(\int_{\mathcal{S}} x\frac{\mu_q}{m_q}(dx)\right) \geq \int_{\mathcal{S}} \phi_q(x)\frac{\mu_q}{m_q}(dx) = \frac{a(\mu_q)}{m_q} = \phi(\alpha_q), \tag{34}$$

with the weak inequality holding with equality if $\mu_q = m_q\delta_{\alpha_q}$ or ϕ_q is linear. Along with (34), the fact that ϕ_q is strictly increasing implies

$$\int_{\mathcal{S}} x\frac{\mu_q}{m_q}(dx) \geq \alpha_q, \tag{35}$$

with strict inequality only if μ_q is polymorphic and ϕ_q is strictly concave.

Further, the cost function c_q is convex and strictly increasing (part 3 of Assumption 3.1). Hence,

$$\int_{\mathcal{S}} c_q(x)\frac{\mu_q}{m_q}(dx) \geq c_q\left(\int_{\mathcal{S}} x\frac{\mu_q}{m_q}(dx)\right) \geq c_q(\alpha_q) \tag{36}$$

$$\Rightarrow \int_{\mathcal{S}} c_q(x)\mu_q(dx) \geq m_q c_q(\alpha_q), \tag{37}$$

with the second weak inequality in (36) following from (35). From (37), it follows that $C(\mu) > \sum_q m_q c_q(\alpha_q)$ if (37) holds with strict inequality for at least one $p \in \mathcal{P}$. We show that this is the case if μ_p is polymorphic and satisfies $\frac{a(\mu_p)}{m_p} = \phi_p(\alpha_p)$. We divide the proof into two cases.

(a) Suppose ϕ_p is strictly concave. In that case, the fact that μ_p is polymorphic implies (35) holds with strict inequality. Hence, so does the second inequality of (36) and, therefore, (37).

(b) Suppose ϕ_p is linear. In that case, (35) and, therefore, second weak inequality of (36) holds with equality. But if ϕ_p is linear, then, by part 4 of Assumption 3.1, c_p is strictly convex. Hence, the first inequality of (36) holds strictly. Therefore, so does (37).

In both cases, $C(\mu) > \sum_q m_q c_q(\alpha_q)$. Therefore, $g(\alpha_1, \alpha_2, \cdots, \alpha_n) > f(\mu)$. ■

Lemma A.3 *Let Assumption 3.1 hold. Then, the quasi–potential function* $g : \prod_{p=1}^n [\underline{x}, \bar{x}] \to \mathbf{R}$ *defined in (11) is a strictly concave function. Hence, it has a unique maximizer* $\alpha^* = (\alpha_1^*, \cdots, \alpha_n^*) \in \prod_{p=1}^n [\underline{x}, \bar{x}]$.

Proof. Consider the quasi–potential function (11). Take two points $(\hat{\alpha}_1, \cdots, \hat{\alpha}_n) \neq (\tilde{\alpha}_1, \cdots, \tilde{\alpha}_n) \in [\underline{x}, \bar{x}]$. Denote $\sum_p m_p \phi_p(\underline{x}) = \underline{\phi}$. We need to show that for all $\lambda \in (0,1)$,

$$\int_{\underline{\phi}}^{\sum_q m_q \phi_q(\lambda\hat{\alpha}_q + (1-\lambda)\tilde{\alpha}_q)} \beta(z)dz - \sum_{q \in \mathcal{P}} m_q c_q(\lambda\hat{\alpha}_q + (1-\lambda)\tilde{\alpha}_q)$$
$$> \lambda \left(\int_{\underline{\phi}}^{\sum_q m_q \phi_q(\hat{\alpha}_q)} \beta(z)dz - \sum_{q \in \mathcal{P}} m_q c_q(\hat{\alpha}_q) \right) + (1-\lambda) \left(\int_{\underline{\phi}}^{\sum_q m_q \phi_q(\tilde{\alpha}_q)} \beta(z)dz - \sum_{q \in \mathcal{P}} m_q c_q(\tilde{\alpha}_q \right). \tag{38}$$

We first note that since β is a strictly decreasing function, $\int_{\underline{\phi}}^{y} \beta(z)dz$ is a strictly concave function, for $y \in \mathbf{R}$. Hence,

$$\int_{\underline{\phi}}^{\sum_q m_q \phi_q(\lambda\hat{\alpha}_q + (1-\lambda)\tilde{\alpha}_q)} \beta(z)dz \geq \int_{\underline{\phi}}^{\lambda \sum_q m_q \phi_q(\hat{\alpha}_q) + (1-\lambda) \sum_q m_q \phi_q(\tilde{\alpha}_q)} \beta(z)dz \tag{39}$$
$$\geq \lambda \int_{\underline{\phi}}^{\sum_q m_q \phi_q(\hat{\alpha}_q)} \beta(z)dz + (1-\lambda) \int_{\underline{\phi}}^{\sum_q m_q \phi_q(\tilde{\alpha}_q)} \beta(z)dz, \tag{40}$$

where the weak inequality in (39) arises due to the concavity of ϕ_q, for all $q \in \mathcal{P}$, and the weak inequality in (40) arises due to the strict concavity of $\int_{\underline{\phi}}^{y} \beta(z)dz$, with the equality holding only if $\sum_q m_q \phi_q(\hat{\alpha}_q) = \sum_q m_q \phi_q(\tilde{\alpha}_q)$.

Further, due to the convexity of c_q, for all $q \in \mathcal{P}$, we have

$$c_q(\lambda\hat{\alpha}_q + (1-\lambda)\tilde{\alpha}_q) \leq \lambda c_q(\hat{\alpha}_q) + (1-\lambda) c_q(\tilde{\alpha}_q). \tag{41}$$

Hence,

$$\sum_{q\in\mathcal{P}} m_q c_q(\lambda\hat{\alpha}_q + (1-\lambda)\tilde{\alpha}_q) \leq \lambda \sum_{q\in\mathcal{P}} m_q c_q(\hat{\alpha}_q) + (1-\lambda)\sum_{q\in\mathcal{P}} m_q c_q(\tilde{\alpha}_q). \tag{42}$$

We now divide the proof into three cases. First, we consider the case where for at least one $p \in \mathcal{P}$, ϕ_p is strictly concave. Further, for at least one $p \in \mathcal{P}$ such that ϕ_p is strictly concave, $\hat{\alpha}_p \neq \tilde{\alpha}_p$. Note that this case covers the possibility that ϕ_q is strictly concave, for all $q \in \mathcal{P}$. Then, for such p for which $\hat{\alpha}_p \neq \tilde{\alpha}_p$, $\phi_p(\lambda\hat{\alpha}_p + (1-\lambda)\tilde{\alpha}_p) > \lambda\phi_p(\hat{\alpha}_p) + (1-\lambda)\phi_p(\tilde{\alpha}_p)$. In that case, (39) holds with strict inequality. Then, (39)–(42) imply (38).

Second, for all $q \in \mathcal{P}$ such that ϕ_q is strictly concave, $\hat{\alpha}_q = \tilde{\alpha}_q$. In that case, there must exist at least one $p \in \mathcal{P}$ such that ϕ_p is linear and $\hat{\alpha}_p \neq \tilde{\alpha}_p$. Further, by Assumption 3.1(4), c_p is strictly convex for such p. Then, (41) holds with strict inequality for such p. Hence, so does (42). Then, (39)–(42) imply (38).

Third, ϕ_q is linear for all q. Therefore, by Assumption 3.1(4), c_q is strictly convex for all $q \in \mathcal{P}$ and there must exist at least one $p \in \mathcal{P}$ such that $\hat{\alpha}_p \neq \tilde{\alpha}_p$. Thus, (41) and (42) hold with strict inequality so that (39)–(42) imply (38).

These three cases exhaust all possibilities. Thus, the quasi–potential function (11) is strictly concave. The existence of a unique maximizer then follows. ■

Using Lemmas A.2 and A.3, we now establish Theorem 4.1.

Proof of Theorem 4.1: Consider the potential function on Δ as defined in (8) and the quasi–potential function g as defined in (11).

First, note that $\mu^* = (m_1\delta_{\alpha_1^*}, \cdots, m_n\delta_{\alpha_n^*})$, where $(\alpha_1^*, \cdots, \alpha_n^*) \in \prod_{p=1}^n [\underline{x}, \bar{x}]$ is as characterized in Lemma A.3. Hence, μ_p^* is monomorphic for all $p \in \mathcal{P}$ and $\frac{a(\mu_p^*)}{m_p} = \phi_p(\alpha_p^*)$. Therefore, by part 1 of Lemma A.2, $f(\mu^*) = g(\alpha_1^*, \cdots, \alpha_n^*)$.

Now consider $\mu \neq \mu^*$ such that μ_p is monomorphic for every p. Hence, $\mu = (m_1\delta_{\alpha_1}, \cdots, m_n\delta_{\alpha_n})$ for some $(\alpha_1, \cdots, \alpha_n) \in \prod_{p=1}^n [\underline{x}, \bar{x}]$, with $\alpha_p \neq \alpha_p^*$ for at least one $p \in \mathcal{P}$. Note that in this case, $\frac{a(\mu_p)}{m_p} = \phi(\alpha_p)$, $\alpha_p \in [\underline{x}, \bar{x}]$. Therefore, by Lemma A.2(1), $f(\mu) = g(\alpha_1, \cdots, \alpha_n)$.

Since $\mu \neq \mu^*$ and μ_p is monomorphic for all p, it must be that $\alpha_p \neq \alpha_p^*$ for at least one population p so that $(\alpha_1, \cdots, \alpha_n) \neq (\alpha_1^*, \cdots, \alpha_n^*)$. Hence, by the definition of α^*, $g(\alpha_1^*, \cdots, \alpha_n^*) > g(\alpha_1, \cdots, \alpha_n)$. Therefore, we obtain

$$f(\mu^*) = g(\alpha_1^*, \cdots, \alpha_n^*) > g(\alpha_1, \cdots, \alpha_n) = f(\mu), \tag{43}$$

for any $\mu \neq \mu^*$ such that μ_p is monomorphic for every $p \in \mathcal{P}$.

Next, consider μ such that μ_p is polymorphic for at least one $p \in \mathcal{P}$. For any $p \in \mathcal{P}$, define $\alpha_p \in [\underline{x}, \bar{x}]$ such that $\frac{a(\mu_p)}{m_p} = \phi(\alpha_p)$. Then,

$$f(\mu^*) = g(\alpha_1^*, \cdots, \alpha_n^*) \geq g(\alpha_1, \cdots, \alpha_n) > f(\mu), \tag{44}$$

where the weak inequality holds if $(\alpha_1^*, \cdots, \alpha_n^*) = (\alpha_1, \cdots, \alpha_n)$ and the strict inequality follows from Lemma A.2(2).

Combining (43) and (44), we conclude $f(\mu^*) > f(\mu)$ for every $\mu \in \Delta \setminus \{\mu^*\}$. Hence, μ^* is the unique maximizer of f in Δ. Since f is concave on Δ (Lemma **??**), it follows that μ^* is the unique Nash equilibrium of F. ■

A.2.1 Appendix to Section 4.1

Proof of Proposition 4.2: Given the assumption that (12) has an interior maximizer, the FOC for α_p^* is

$$\frac{Vr\alpha_p^{r-1}}{\sum_q m_q \alpha_q^r} = k_p \gamma \alpha_p^{\gamma-1}. \tag{45}$$

Using (45), we obtain

$$\alpha_q^* = \left(\frac{k_p}{k_q}\right)^{\frac{1}{\gamma-1}} \alpha_p^*, \text{ for } q \neq p. \tag{46}$$

Inserting (46) into (45) and simplifying gives us (13). The result follows from Corollary 3.4 and Theorem 4.1.

For equilibrium payoff, note that since $\phi_p(x) = x^r$ in (2), $A(\mu^*) = \sum_p \int_{\mathcal{S}} x^r \mu_p^*(dx)$ follows from (5). Hence, the generalized aggregate strategy at the Nash equilibrium $\mu^* = (m_1\delta_{\alpha_1^*}, \cdots, m_n\delta_{\alpha_n^*})$ of F (see Proposition 4.2), where α_p^* is given by (13), is

$$A(\mu^*) = \sum_{p\in\mathcal{P}} m_p(\alpha_p^*)^r = \left(\frac{rV}{\gamma}\right)^{\frac{r}{\gamma}} \left(\sum_{p\in\mathcal{P}} m_p k_p^{\frac{r}{r-\gamma}}\right)^{1-\frac{r}{\gamma}}. \tag{47}$$

Therefore, the CSF at μ^* for a type p agent is $\frac{(\alpha_p^*)^r}{A(\mu^*)} = \frac{k_p^{\frac{r}{r-\gamma}}}{\sum_{q\in\mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}}$. The equilibrium cost of effort is

$$c_p(\alpha_p^*) = k_p(\alpha_p^*)^\gamma = \frac{rV}{\gamma} \frac{k_p^{\frac{r}{r-\gamma}}}{\sum_{q\in\mathcal{P}} k_q^{\frac{r}{r-\gamma}}}. \tag{48}$$

The equilibrium payoff $F_{\alpha_p^*,p}(\mu_p^*)$ then follows from the definition of F in (2). ■

Proof of Proposition 4.3: Given the assumption that (14) has an interior maximizer, the FOC for $\hat{\alpha}_p^*$ is

$$\frac{Vrk_p^{\frac{r}{\gamma}}\alpha_p^{r-1}}{\sum_q m_q k_q^{\frac{r}{\gamma}} \alpha_q^r} = k_p \gamma \alpha_p^{\gamma-1}. \tag{49}$$

Using (45), we obtain

$$\hat{\alpha}_q^* = \left(\frac{k_p}{k_q}\right)^{\frac{1}{\gamma}} \hat{\alpha}_p^*, \text{ for } q \neq p. \tag{50}$$

Inserting (50) into (49) and simplifying gives us (15). The result follows from Corollary 3.4 and Theorem 4.1.

For equilibrium payoff, note that $\phi_p(x) = k_p^{\frac{r}{\gamma}} x^r$ in $\hat{F}$. Hence, $A(\hat{\mu}^*) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p^{\frac{r}{\gamma}} x^r \hat{\mu}_p^*(dx)$ in $\hat{F}$ follows from (5). Therefore, the generalized aggregate strategy at the Nash equilibrium $\hat{\mu}^* = (m_1\delta_{\hat{\alpha}_1^*}, \cdots, m_n\delta_{\hat{\alpha}_n^*})$ of $\hat{F}$ (see Proposition 4.3), where $\hat{\alpha}_p^*$ is given by (15), is

$$A(\hat{\mu}^*) = \sum_p m_p k_p^{\frac{r}{\gamma}} k_p^{\frac{-r}{\gamma}} \left(\frac{rV}{\gamma}\right)^{\frac{r}{\gamma}} = \left(\frac{rV}{\gamma}\right)^{\frac{r}{\gamma}} \tag{51}$$

Hence, the CSF at μ^* for a type p agent is $\frac{k_p^{\frac{r}{\gamma}}(\hat{\alpha}_p^*)^r}{A(\hat{\mu}^*)} = 1$. The equilibrium cost of effort is

$$c_p(\hat{\alpha}_p^*) = k_p(\hat{\alpha}_p^*)^\gamma = \frac{rV}{\gamma}. \tag{52}$$

The equilibrium payoff $\hat{F}_{\hat{\alpha}_p^*,p}(\hat{\mu}_p^*)$ then follows from the definition of $\hat{F}$ in (3). ■

A.3 Appendix to Section 5

Proof of Proposition 5.1:

1. Since $\phi_p(x) = x^r$ in F, $A(\mu^*) = \sum_p \int_{\mathcal{S}} x^r \mu_p^*(dx)$ follows from (5). Hence, the generalized aggregate strategy at the Nash equilibrium $\mu^* = (m_1\delta_{\alpha_1^*}, \cdots, m_n\delta_{\alpha_n^*})$ of F (see Proposition 4.2), where α_p^* is given by (13), is

$$A(\mu^*) = \sum_{p\in\mathcal{P}} m_p(\alpha_p^*)^r = \left(\frac{rV}{\gamma}\right)^{\frac{r}{\gamma}} \left(\sum_{p\in\mathcal{P}} m_p k_p^{\frac{r}{r-\gamma}}\right)^{1-\frac{r}{\gamma}}. \tag{53}$$

   Therefore, the CSF at μ^* for a type p agent is $\frac{(\alpha_p^*)^r}{A(\mu^*)} = \frac{k_p^{\frac{r}{r-\gamma}}}{\sum_{q\in\mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}}$. The equilibrium cost of effort is

$$c_p(\alpha_p^*) = k_p(\alpha_p^*)^\gamma = \frac{rV}{\gamma} \frac{k_p^{\frac{r}{r-\gamma}}}{\sum_{q\in\mathcal{P}} k_q^{\frac{r}{r-\gamma}}}. \tag{54}$$

   The equilibrium payoff $F_{\alpha_p^*,p}(\mu_p^*)$ then follows from the definition of F in (2).

   For aggregate payoff, we apply (16) to μ^*. From (54), we obtain

$$C(\mu^*) = \sum_p m_p c_p(\alpha_p^*) = \frac{rV}{\gamma}. \tag{55}$$

   Hence, aggregate payoff at μ^* is

$$\bar{F}(\mu^*) = V - C(\mu^*) = V - \frac{rV}{\gamma} = V\left(1 - \frac{r}{\gamma}\right).$$

2. Similar to part 1, $\phi_p(x) = k_p^{\frac{r}{\gamma}} x^r$ in $\hat{F}$. Hence, $A(\hat{\mu}^*) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p^{\frac{r}{\gamma}} x^r \hat{\mu}_p^*(dx)$ in $\hat{F}$ follows from (5). Therefore, the generalized aggregate strategy at the Nash equilibrium $\hat{\mu}^* = (m_1\delta_{\hat{\alpha}_1^*}, \cdots, m_n\delta_{\hat{\alpha}_n^*})$ of $\hat{F}$ (see Proposition 4.3), where $\hat{\alpha}_p^*$ is given by (15), is

$$A(\hat{\mu}^*) = \sum_p m_p k_p^{\frac{r}{\gamma}} k_p^{\frac{-r}{\gamma}} \left(\frac{rV}{\gamma}\right)^{\frac{r}{\gamma}} = \left(\frac{rV}{\gamma}\right)^{\frac{r}{\gamma}} \tag{56}$$

Hence, the CSF at μ^* for a type p agent is $\frac{k_p^{\frac{r}{\gamma}} (\hat{\alpha}_p^*)^r}{A(\hat{\mu}^*)} = 1$. The equilibrium cost of effort is

$$c_p(\hat{\alpha}_p^*) = k_p(\hat{\alpha}_p^*)^\gamma = \frac{rV}{\gamma}. \tag{57}$$

The equilibrium payoff $\hat{F}_{\hat{\alpha}_p^*,p}(\hat{\mu}_p^*)$ then follows from the definition of $\hat{F}$ in (3)

For aggregate payoff, we apply (16) to $\hat{\mu}^*$. From (57), we obtain

$$C(\hat{\mu}^*) = \sum_p m_p c_p(\hat{\alpha}_p^*) = \frac{rV}{\gamma}. \tag{58}$$

Hence, aggregate payoff at $\hat{\mu}^*$ is

$$\bar{F}(\hat{\mu}^*) = V - C(\hat{\mu}^*) = V - \frac{rV}{\gamma} = V\left(1 - \frac{r}{\gamma}\right). \blacksquare$$

Proof of Proposition 5.3: Using (13) and (15), we have

$$\sum_p m_p \alpha_p^* = \left(\frac{rV}{\gamma}\right)^{\frac{1}{\gamma}} \frac{\sum_{p\in\mathcal{P}} m_p k_p^{\frac{1}{r-\gamma}}}{\left(\sum_{q\in\mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}\right)^{\frac{1}{\gamma}}},$$

$$\sum_p m_p \hat{\alpha}_p^* = \left(\frac{rV}{\gamma}\right)^{\frac{1}{\gamma}} \left(\sum_{p\in\mathcal{P}} \frac{m_p}{k_p^{\frac{1}{\gamma}}}\right).$$

Hence, the result will be proven if we can show that

$$\frac{\sum_{p\in\mathcal{P}} m_p k_p^{\frac{1}{r-\gamma}}}{\left(\sum_{q\in\mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}\right)^{\frac{1}{\gamma}}} > \left(\sum_{p\in\mathcal{P}} \frac{m_p}{k_p^{\frac{1}{\gamma}}}\right). \tag{59}$$

Note that if $r = 0$, the two sides of (59) are equal. Hence, to establish this inequality, it suffices to show that the LHS of (59) is strictly increasing in r.

For this purpose, we assume, without loss of generality, that $k_1 < k_2 < \cdots < k_n$. Through some tedious algebra, we then calculate

$$\frac{\partial}{\partial r}\left(\frac{\sum_{p\in\mathcal{P}} m_p k_p^{\frac{1}{r-\gamma}}}{\left(\sum_{q\in\mathcal{P}} m_q k_q^{\frac{r}{r-\gamma}}\right)^{\frac{1}{\gamma}}}\right)$$

$$=\frac{\left(\sum_p m_p k_p^{\frac{r}{r-\gamma}}\right)^{\frac{-1}{\gamma}} \sum_p \sum_{q>p}\left(m_p m_q (\log k_q - \log k_p) k_p^{\frac{r}{r-\gamma}} k_q^{\frac{r}{r-\gamma}}\right)\left(k_p^{\frac{1-r}{r-\gamma}} - k_q^{\frac{1-r}{r-\gamma}}\right)}{(\gamma-r)^2 \sum_p m_p k_p^{\frac{r}{r-\gamma}}} > 0,$$

given our assumptions that $k_p < k_q$ if $p < q$ and $r < \gamma$. Thus, (59) holds for all $r \in (0,1]$ and $\gamma \geq 1$ with the additional restriction that both r and γ are not equal to 1. ■

A.4 Appendix to Section 6

The proof of Proposition 6.1 requires us to first characterize the maximizer of the quasi–potential function g. We state the result in the following lemma.

Lemma A.4 *Consider the quasi–potential function g defined by (12) with $r = \gamma = 1$. Assume that $k_1 < k_2 < \cdots < k_n$. Then, g has a unique maximizer $\alpha^* = (\alpha_1^*, \cdots, \alpha_n^*)$ such that $\alpha_1^* = \frac{V-(1-m_1)k_1\underline{x}}{m_1 k_1}$ and $\alpha_p^* = \underline{x}$ for all $p \neq 1$.*

Proof. Since $\underline{x}$ is sufficiently low and $\bar{x}$ is sufficiently high, the FOC for maximizing (12) is satisfied with equality for $p = 1$. With $r = \gamma = 1$, this FOC is

$$\frac{V}{\sum_q m_q \alpha_q^*} = k_1. \tag{60}$$

But because $k_1 < k_2 < \cdots < k_n$, (60) implies that the FOC for all $p \neq 1$ must be

$$\frac{V}{\sum_q m_q \alpha_q^*} < k_p. \tag{61}$$

From (61), we conclude that $\alpha_p^* = \underline{x}$, for all $p \neq 1$. Using this in (60), we obtain

$$\frac{V}{m_p\alpha_p^* + \sum_{q\neq p} m_q \alpha_q^*} = \frac{V}{m_p \alpha_p^* + (1-m_1)\underline{x}} = k_1 \Rightarrow \alpha_1^* = \frac{V-(1-m_1)k_1\underline{x}}{m_1 k_1}. \ ■$$

Proof of Proposition 6.1: Consider the potential function f defined in (9) and the quasi–potential function g defined in (12) with $\alpha = \gamma = 1$. Recall from (5) that $A(\mu) = \sum_p a(\mu_p)$. Further, because $\gamma = 1$, $\sum_p \int_{\mathcal{S}} k_p x \mu_p(dx) = \sum_p k_p a(\mu_p)$ in (9). It, therefore, follows from (9) and (12) that if $\mu = (\mu_1, \cdots, \mu_n) \in \Delta$ and $(\alpha_1, \cdots, \alpha_n) \in \prod_p [\underline{x}, \bar{x}]$ are such that $\frac{a(\mu_p)}{m_p} = \alpha_p$ for every $p \in \mathcal{P}$, then $f(\mu) = g(\alpha_1, \cdots, \alpha_n)$.

Consider $\mu^* \in NE(F)$ as defined in (23). By the definition of $NE(F)$, $\frac{a(\mu_1^*)}{m_1} = \alpha_1^*$ and $\frac{a(\mu_p^*)}{m_p} = \underline{x}$ for every $p \neq 1$. Hence, by the preceding argument,

$$f(\mu^*) = g(\alpha_1^*, \cdots, \alpha_n^*) \tag{62}$$

where, as characterized in Lemma A.4, $\alpha_1^* = \frac{V-(1-m_1)k_1\underline{x}}{m_1 k_1}$ and $\alpha_n^* = \underline{x}$ for every $p \neq 1$.

Next, for every $\mu \in \Delta$, we can find $(\alpha_1, \cdots, \alpha_n) \in \prod_p [\underline{x}, \bar{x}]$ such that $\frac{a(\mu_p)}{m_p} = \alpha_p$ for every $p \in \mathcal{P}$. This is because $a(\mu_p) \in [m_p\underline{x}, m_p\bar{x}]$ for every $p \in \mathcal{P}$.

Hence, if we now consider $\mu \notin NE(F)$, then there exists such an element $(\alpha_1, \cdots, \alpha_n) \in \prod_p [\underline{x}, \bar{x}]$ such that, by the earlier argument, $f(\mu) = g(\alpha_1, \cdots, \alpha_n)$. Moreover, because $\mu \notin NE(F)$, it must be that $(\alpha_1, \cdots, \alpha_n) \neq (\alpha_1^*, \cdots, \alpha_n^*)$. But because by Lemma A.4, $(\alpha_1^*, \cdots, \alpha_n^*)$ is the unique maximizer of g, it must then be that $g(\alpha_1, \cdots, \alpha_n) < g(\alpha_1^*, \cdots, \alpha_n^*)$. Combining this argument with (62), we conclude

$$f(\mu) = g(\alpha_1, \cdots, \alpha_n) < g(\alpha_1^*, \cdots, \alpha_n^*) = f(\mu^*).$$

Therefore, $NE(F)$ is the set of maximizers of f. Hence, due to the concavity of f, it must also be the set of Nash equilibria of F. Due to the linearity of $a(\cdot)$, $NE(F)$ is clearly a convex set.

We now calculate equilibrium payoffs. If $\mu^* \in NE(F)$, then by (23),

$$A(\mu^*) = \sum_p a(\mu_p^*) = m_1 \frac{V-(1-m_1)k_1\underline{x}}{m_1 k_1} + \sum_{p\neq 1} m_p \underline{x} = \frac{V-(1-m_1)k_1\underline{x}}{k_1} + (1-m_1)\underline{x} = \frac{V}{k_1}. \tag{63}$$

Note that μ_1^* is possibly polymorphic. If an agent in population 1 plays $x \in \mathcal{S}$ at μ^*, then by (2), his payoff is $\frac{x}{A(\mu^*)}V - k_1 x = 0$ by (63). All agents other than in population 1 play $\underline{x}$. Their payoff is $\frac{\underline{x}}{A(\mu^*)}V - k_1\underline{x} = 0$ by (63).

Hence, at every Nash equilibrium in $NE(F)$, all agents have payoff 0. The aggregate payoff at any such Nash equilibrium must also be 0. ■

The proof of Proposition 6.2 requires us to characterize the set of maximizers of the quasi–potential function $\hat{g}$. We state the result in the following lemma.

Lemma A.5 *If $r = \gamma = 1$, then the quasi–potential function $\hat{g}$ defined in (14) has a convex set of maximizers*

$$M(\hat{g}) = \left\{ \{a_1, \cdots, \alpha_n\} \in \prod_{p\in\mathcal{P}} [\underline{x}, \bar{x}] : \sum_{p\in\mathcal{P}} m_p k_p \alpha_p = V \right\}. \tag{64}$$

Proof. If $r = \gamma = 1$, then the FOC for maximizing (14) is $\frac{V}{\sum_{p\in\mathcal{P}} m_p k_p \alpha_p} = 1$. This establishes the result. ■

Proof of Proposition 6.2: Consider the potential function $\hat{f}$ defined in (10) and the quasi–potential function $\hat{g}$ defined in (14) with $r = \gamma = 1$. Clearly if μ and $(\alpha_1, \cdots, \alpha_n)$ are such that $A(\mu) = \sum_p m_p k_p \alpha_p$, then $\hat{f}(\mu) = \hat{g}(\alpha_1, \cdots, \alpha_n)$.

Note that in $\hat{F}$, $a(\mu_p) = \int_{\mathcal{S}} k_p x \mu_p(dx)$. This follows from (4) and the fact that $\phi_p(x) = k_p x$. Since $a(\mu_p) \in [m_p k_p \underline{x}, m_p k_p \bar{x}]$, it follows that for any μ_p, there exists a unique $\alpha_p \in [\underline{x}, \bar{x}]$ such that $a(\mu_p) = k_p \alpha_p$. Therefore, for any $\mu = (\mu_1, \cdots, \mu_p)$, we can find a unique $(\alpha_1, \cdots, \alpha_n) \in \prod_p [\underline{x}, \bar{x}]$ such that $A(\mu) = \sum_p a(\mu_p) = \sum_p m_p k_p \alpha_p$.

Let $\hat{\mu}^* \in NE(\hat{F})$ as defined in (24). Hence, $A(\hat{\mu}^*) = V$. Further, by the argument in the preceding paragraph, there exists $(\hat{\alpha}_1^*, \cdots, \hat{\alpha}_n^*) \in \prod_p [\underline{x}, \bar{x}]$ such that $A(\hat{\mu}^*) = \sum_p m_p k_p \hat{\alpha}_p^*$. Hence,

$$\hat{f}(\hat{\mu}^*) = g(\hat{\alpha}_1^*, \cdots, \hat{\alpha}_n^*). \tag{65}$$

Moreover, because $A(\hat{\mu}^*) = V$, $\sum_p m_p k_p \hat{\alpha}_p^* = V$. Hence, $(\hat{\alpha}_1^*, \cdots, \hat{\alpha}_n^*) \in M(\hat{g})$ as defined in (64).

Now consider $\mu \notin NE(\hat{F})$. So $A(\mu) \neq V$. There exists $(\alpha_1, \cdots, \alpha_n) \in \prod_p [\underline{x}, \bar{x}]$ such that $A(\mu) = \sum_p m_p k_p \alpha_p$. Hence, $\hat{f}(\mu) = \hat{g}(\alpha_1, \cdots, \alpha_n)$. Moreover, as $A(\mu) \neq V$, $\sum_p m_p k_p \alpha_p \neq V$ so that $\hat{g}(\alpha_1, \cdots, \alpha_n) \notin M(\hat{g})$. It then follows from (65) and Lemma A.5 that

$$\hat{f}(\mu) = \hat{g}(\alpha_1, \cdots, \alpha_n) < g(\hat{\alpha}_1^*, \cdots, \hat{\alpha}_n^*) = \hat{f}(\hat{\mu}^*).$$

Hence, $NE(\hat{F})$ as defined in (24) is the set of maximizers of $\hat{f}$. The concavity of $\hat{f}$ then implies that $NE(\hat{F})$ is also the set of Nash equilibria of $\hat{F}$. Clearly, $NE(\hat{F})$ is a convex set.

We now calculate equilibrium payoffs at a Nash equilibrium $\hat{\mu}^*$. From (3), and due to $r = \gamma = 1$, $F_{x,p}(\hat{\mu}^*) = \frac{k_p x}{A(\hat{\mu}^*)} V - k_p x = 0$ because by (24), $A(\hat{\mu}^*) = V$. Since all agents in all populations have zero payoff, aggregate payoff at a Nash equilibrium $\hat{\mu}^*$ is also 0. ■

Proof of Proposition 6.3: In the equal treatment contest, the set of Nash equilibria is given by (23). We calculate aggregate effort at any such Nash equilibrium. Since $r = 1$, the aggregate strategy level $A(\mu)$ is equivalent to aggregate effort in this case. Hence, from (23), the aggregate effort at any Nash equilibrium $\mu^* \in NE(F)$ is

$$A(\mu^*) = a(\mu_1^*) + \sum_{p>1} a(\mu_p^*). \tag{66}$$

From (23), we know that at any $\mu^* \in NE(F)$, $a(\mu_1^*) = m_1 \frac{V - (1-m_1) k_1 \underline{x}}{m_1 k_1} = \frac{V - (1-m_1) k_1 \underline{x}}{k_1}$ and, for $p > 1$, $a(\mu_p^*) = \int_{\mathcal{S}} x \mu_p^*(dx) = m_p \underline{x}$. Hence, using (66), we obtain

$$\begin{aligned} A(\mu^*) &= a(\mu_1^*) + \sum_{p>1} a(\mu_p^*). \\ &= \frac{V - (1-m_1) k_1 \underline{x}}{k_1} + \sum_{p>1} m_p \underline{x} \\ &= \frac{V - (1-m_1) k_1 \underline{x}}{k_1} + (1-m_1) \underline{x} \\ &= \frac{V}{k_1}. \end{aligned} \tag{67}$$

Under affirmative action, the set of Nash equilibria is characterized by (24). Note that in this case, $A(\mu) = \sum_p \int_{\mathcal{S}} k_p x \mu_p(dx)$ is not aggregate effort. Instead, aggregate effort is $\sum_p \int_{\mathcal{S}} x \mu_p(dx)$. To calculate aggregate effort, note from (24) that at any Nash equilibrium $\hat{\mu}^*$,

$$A(\hat{\mu}^*) = \sum_p k_p \int_{\mathcal{S}} x \hat{\mu}_p^*(dx) = V. \tag{68}$$

But $\sum_p k_p \int_{\mathcal{S}} x \hat{\mu}_p^*(dx) > k_1 \sum_p \int_{\mathcal{S}} x \hat{\mu}_p^*(dx)$ as $k_1 < \cdots < k_n$. Hence, from (68), we obtain

$$\begin{aligned} & k_1 \sum_p \int_{\mathcal{S}} x \hat{\mu}_p^*(dx) < V \\ \Rightarrow & \sum_p \int_{\mathcal{S}} x \hat{\mu}_p^*(dx) < \frac{V}{k_1}. \end{aligned} \tag{69}$$

But $\sum_p \int_{\mathcal{S}} x \hat{\mu}_p^*(dx)$ is aggregate effort under affirmative action at $\hat{\mu}^*$. Hence, comparing (67) and (69), we conclude that aggregate effort at any equilibrium under affirmative action is lower than equilibrium aggregate effort under equal treatment. ■

References

[1] Becker GS (1971) The economics of discrimination, 2nd edn. University of Chicago Press, Chicago

[2] Bodoh–Creed AL, Hickman BR (2018) College assignment as a large contest. J. Econ. Theory 175:344–375.

[3] Bowen WG, Bok D (2000) The Shape of the River. Princeton University Press, Princeton.

[4] Cheung MW (2014) Pairwise comparison dynamics for games with continuous strategy space. J. Econ. Theory 153:344–375.

[5] Cheung MW, Lahkar R (2018) Nonatomic Potential Games: the Continuous Strategy Case, Games Econ. Behav. 108:341–362.

[6] Coate S, Loury G C (1993) Will affirmative action policies eliminate negative stereotypes? Am Econ Rev 83:1220–1240.

[7] Corchón L (1994) Comparative statics for aggregative games the strong concavity case. Math. Soc. Sci. 28:151–165.

[8] Corchón L (2007) The theory of contests: a survey. Rev. Econ. Design. 11:69–100.

[9] Corchón L, Serena M (2018) Contest theory. Handbook of game theory and industrial organization 2:125–136.

[10] Cornes R, Hartley R (2005) Asymmetric contests with general technologies. Economic Theory 26:923–946.

[11] Dahm M, Esteve-González P (2018) Affirmative action through extra prizes. J. Econ. Behav. Organ. 153:123–142.

[12] Dixit A (1987) Strategic Behavior in Contests. Am. Econ. Rev. 77:891–898.

[13] Ewerhart C (2017) The lottery contest is a best-response potential game. Econ. Lett. 155:168–171.

[14] Fain JR (2009) Affirmative action can increase effort. J. Labor Res. 30:168–175.

[15] Franke J (2012) Affirmative action in contest games. Eur. J. Political Econ. 28:105–118.

[16] Franke J, Kanzow C, Leininger W, Schwartz A (2013) Effort maximization in asymmetric contest games with heterogeneous contestants. Econ. Theory 52:589–630.

[17] Franke J, Kanzow C, Leininger W, Schwartz A (2014) Lottery versus all-pay auction contests: A revenue dominance theorem. Games Econ. Behav. 83:116–126.

[18] Fryer RG, Loury GC (2005) Affirmative action and its mythology. J. Econ. Perspect. 19:147–162.

[19] Fryer RG, Loury GC (2005) Affirmative action in winner take-all markets. J. Econ. Inequal. 3:263-–280.

[20] Fryer RG, Loury GC (2013) Valuing Diversity. J. Political Econ. 121:747-774.

[21] Fu Q (2006) A theory of affirmative action in college admissions. Econ. Inq. 44:420–428.

[22] Harel A, Segal U (2014) Utilitarianism and Discrimination. Soc. Choice Welf. 42:367—380.

[23] Hillman A, Riley JG (1989) Politically contestable rents and transfers. Economics and Politics 1:17–40.

[24] Holzer H, Neumark D (2000) Assessing affirmative action. J. Econ. Lit. 38:483–568.

[25] Konrad KA (2009). Strategy and Dynamics in Contests. (OUP Catalogue).

[26] Kranich L (1994). Equal division, efficiency, and the sovereign supply of labor. Am. Econ. Rev. 84:178-–189.

[27] Lahkar R, Riedel F (2015) The logit dynamic for games with continuous strategy sets. Games Econ. Behav. 91:268–282.

[28] Lahkar R (2017) Large population aggregative potential games. Dyn Games Appl. 7:443–467.

[29] Lahkar R, Mukherkee, S. (2019) Evolutionary Implementation in a Public Goods Game. J. Econ. Theory 181:423–460.

[30] Lazear EP, Rosen S (1981) Rank–order tournaments as optimum labor contracts. J. Political Econ. 89:841-–864.

[31] Lee, SH (2013) The incentive effect of a handicap. Econ. Lett. 118:42–45.

[32] Monderer D, Shapley L (1996) Potential games. Games Econ. Behav. 14:124–143.

[33] Myerson RB, Wärneryd K (2006) Population uncertainty in contests. Econ. Theory 27:469-–474.

[34] Newton J (2018) Evolutionary Game Theory: A Renaissance. Games 9:31.

[35] Nitzan S (1994) Modelling Rent Seeking Contest. Eur. J. Political Econ. 10:41–60.

[36] Nti KO (1997) Comparative statics of contests and rent-seeking games. Int Econ Rev 38:43–59.

[37] Nti KO (2004) Maximum efforts in contests with asymmetric valuations. Eur. J. Political Econ. 20: 1059-–1066.

[38] Olszewski W, Siegel R (2014) Large Contests. Econometrica 84:835–854.

[39] Pérez-Castrillo D, Verdier T (1992) A general analysis of rent-seeking games. Public Choice 71:351–361.

[40] Sandholm WH (2001) Potential games with continuous player sets. J. Econ. Theory 97:81–108.

[41] Sandholm WH (2010) Population Games and Evolutionary Dynamics. MIT Press, Cambridge, MA.

[42] Schotter A, Weigelt K (1992) Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. Q.J. Econ. 107:511–539.

[43] Sowell T (2004) Affirmative Action Around the World. Yale University Press, New Haven, London.

[44] Stein WE (2002) Asymmetric rent-seeking with more than two contestants. Public Choice 113:325–336.

[45] Sultana R (2017) Affirmative action and dynamics of work-ethic preferences. Econ. Inq. 55:1350–1369.

[46] Sultana R (2019) The incentive and efficiency effects of affrmative action: does envy matter? forthcoming, Oxf. Econ. Pap.

[47] Szidarovsky F, Okuguchi K (1997) On the existence and uniqueness of pure Nash equilibrium in rent–seeking games. Games Econ. Behav. 18:135-–140.

[48] Tullock G (1967) The welfare cost of tariffs, monopolies and theft. West Econ J 5:224—232

[49] Tullock G (1980) Efficient rent-seeking. In: Buchanan JM, Tollison RD, Tullock G (eds) Towards a theory 994 of a rent-seeking society. Texas A&M University Press, College Station, 97—112.

[50] Yamazaki T (2008) On the existence and uniqueness of pure-strategy Nash equilibrium in asymmetric rent–seeking contest. J. Public Econ. Theory 10:317—327.

[51] Zeidler E (1986) Nonlinear Functional Analysis and Its Applications, vol.I. Springer-Verlag, New York.