Ashoka University Economics
Discussion Paper 82

Optimal Large Population Tullock Contests

July 2022

Ratul Lahkar, Ashoka University
Saptarshi Mukherjee, IIT Delhi

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

Plot No. 2, Rajiv Gandhi Education City,
National Capital Region,
P.O.Rai, Sonepat, Haryana 131029

Optimal Large Population Tullock Contests

Ratul Lahkar* Saptarshi Mukherjee†

July 20, 2022

Abstract

We consider large population Tullock contests in which agents are divided into different types according to their strategy cost function. A planner assigns type specific bias parameters to affect the likelihood of success with the objective of maximizing the Nash equilibrium level of aggregate strategy. We characterize such optimal bias parameters and identify conditions under which those parameters are increasing or decreasing according to the cost parameters. The parameters are biased in favor of high cost agents if the cost functions are strictly convex and the likelihood of success is sufficiently responsive to strategy. We also identify conditions under which a planner can truthfully implement the optimal parameters under incomplete information. In fact, under such conditions, dominant strategy implementation is equivalent to Nash implementation in our model. Hence, our mechanism double implements the optimal bias parameters.

Keywords: Tullock Contests; Contest Design; Implementation.

JEL Classification: C72; D72; D82.

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

†(corresponding author) Department of Humanities and Social Sciences, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India, email: saptarshi.isi@gmail.com

1 Introduction

Contests are used to model a variety of competitive situations where contestants exert effort (or make a payment) to obtain a fixed prize. The rules of a contest are usually set by a contest administrator or a planner who may, therefore, exercise discretion and set rules that favor or handicap certain types of contestants. The effects of such contest rules have been explored in a variety of settings like labour and promotion tournaments (Lazear and Rosen [23], Tsoulouhas et al. [35]), lobbying and public procurement (Baye et al. [3], Fang [12], Epstein et al. [11]), affirmative action (Fu [15], Franke [13]), sports tournaments (Runkel [31]) and in more general abstract models (Fu and Lu [16], Franke et al. [14], Fu and Wu [17]). Studying optimal contest design rules to meet specific objectives is, therefore, an interesting research goal given the variety of contexts in which they may be applied.

Perhaps the most well known and most widely used model of a contest is that of the Tullock contest (Tullock [36]). Tullock contests are characterized by a contest success function (CSF) that determines the probability or share of players winning the prize as a functions of their strategy. Such contests are mathematically tractable, which accounts for their popularity in applications.[1] There are multiple ways in which a designer may seek to influence the rules of a Tullock contest. These include applying different bias parameters in the CSF of different players, modifying players' valuation for the reward or giving some players a head start by adding some extra effort to what they are already exerting (Mealem and Nitzan [25]). Some other means of contest design are influencing the structure of prizes (Moldovanu and Sela [24]), the choice of contestants (Baye et al. [3]) and the number of stages in a contest (Fu and Lu [16]). Typically, the objective of the contest designer is to maximize the aggregate strategy of the contestants. This is relevant in contexts like employment tournaments or lobbying and public procurement where the planner wishes to maximize rent–seeking behavior.[2] Or, alternatively, the planner may regard the aggregate effort by contestants as being of intrinsic importance as in certain models of affirmative action (Franke [13]).

In this paper, we consider one particular way of designing Tullock contests in order to maximize the aggregate strategy by agents; that of finding the optimal way to bias the CSFs of different players. Dasgupta and Nti [10] and Nti [27] are two early papers who examined this problem; the former considered a contest with n homogeneous players while the latter considered two–player contests with possible heterogeneity. Franke et al. [14] and Fu and Wu [17] have considered more general versions of this question by considering n player Tullock contests with heterogeneity among contestants. A key difference between these two later papers is that Franke et al. [14] consider contests with linearities in both the CSF and effort cost functions while Fu and Wu [17] allow non–linearities in the CSF, although cost functions are still linear. An important insight of such

[1] See, for example, Corchón [7] and Congleton et al. [5] for reviews of such applications. See also Skarpedas [33] and Clark and Riis [9] for axiomatic foundations of Tullock contests.

[2] In rent seeking models, the fixed prize is interpreted as the rent in the sense that exertion of effort by contestants cannot increase the size of the prize but only affect the likelihood of success.

models is that symmetric CSFs allow effort maximization only when players are homogeneous. With heterogeneity in cost functions, or equivalently, in players' valuation of the prize, CSFs need to be biased in order to maximize total effort. The important question then is whether the bias should be in favor of stronger players or weaker players. Our paper also addresses these questions but with two important differences. First, our setting is that of a large population of agents or, technically, when each contestant is of measure zero. Second, we also analyze the possibility of implementing the optimal bias parameters in an environment of incomplete information. We provide more careful justifications of both these features of our approach after the following brief description of our model.

Models of optimal contest design, including Franke et al. [14] and Fu and Wu [17], are usually under the restriction that the contestants continue to play the Nash equilibrium of the contest.[3] Following this approach, we divide our large population of agents into a finite number of types, with each type having a distinct strategy cost function. Agents with lower cost parameters are interpreted as stronger contestants. The planner assigns type specific bias parameters and for each such allocation of parameters, there exists a Nash equilibrium and an associated level of aggregate strategy. The objective of the planner, and the problem we solve, is to assign that vector of bias parameters that maximizes this equilibrium level of aggregate strategy. As we show, the solution depends upon two important parameters of the model; r and γ. The former captures how responsive the CSF is to the strategy choice of an agent while the latter determines the convexity of the cost function. We then examine whether the optimal bias parameters favor or discriminate against weaker contestants. The answer, as we discuss in detail in Section 4, again depends upon r and γ. We note that it is also possible to have more general objective functions in which the planner cares not only about the aggregate strategy but also, for example, the variance of the strategy choices (Fu and Wu [17]). In this paper, though, we focus only on maximizing aggregate strategy.

But why analyze large population Tullock contests? Such large population contests have been introduced by Lahkar and Sultana [21] in the context of a model of affirmative action. As described in that paper, and as we also discuss in more detail later in this paper, it is technically more convenient to analyze a large population contest than with a finite but large number of players. Solving the model becomes easier and arguably more elegant in the large population setting. This applies both to characterizing the Nash equilibrium of the contest as well as deriving the optimal bias parameters that maximize the aggregate strategy. In fact, if there are non–linearities in the CSF or strategy cost functions, and the cost functions are heterogeneous (or asymmetric), a finite player Tullock contest cannot be analytically solved (Franke [13]). Hence, characterizing such equilibrium or optimal bias parameters in n heterogeneous player models become a fairly complex mathematical exercise (Franke et al. [14], Fu and Wu [17]). On the other hand, our approach allows us to characterize Nash equilibrium directly by computing best responses and then finding the optimal bias parameters through a straightforward maximization exercise. Intuitively, the large population approach depends upon the fact that with players being of measure zero, changes

[3]See Section 6, Corchón [7] for a wider survey of this approach in the literature.

in individual behavior do not affect aggregate behavior in society. This simplifies our analysis considerably which allows us to go beyond fully linear Tullock contests, as in Franke et al. [14], or Tullock contests with non–linearities only in the CSF. Instead, we can allow for asymmetries and non–linearities in both the CSF and the cost function without any additional complications.[4]

Of course, in reality, no agent is ever of measure zero. Therefore, as in an idealized model of a perfectly competitive market, our results will be valid only when the number of agents is sufficiently large so that agents behave as if their individual actions will not affect aggregate behavior. Even with this caveat, a large population model can provide important insights, particularly in the contexts where number of participants involved are naturally large. This is true, for example, in contexts like affirmative action where Tullock contests have been applied (Franke [13], Lahkar and Sultana [21]). Optimal contest design can be applied to such situations to understand whether leveling the playing field in favor of weaker contestants can maximize total effort. This is the conventional wisdom but Fu and Wu [17] show that this wisdom holds only if the CSF is sufficiently responsive to effort.[5] We also obtain similar results but under a broader set of conditions since we allow not just linear cost functions but also strictly convex ones. In fact, we find that if the cost function is linear, which is the case that Fu and Wu [17] examines, then maximizing total effort involves the complete reversal of leveling the field as all but the strongest contestants are eliminated from the contest. This remains true when, along with linearity in cost functions, we also have a linearity in the CSF, as in Franke et al. [14]. Franke et al.'s [14] and Fu and Wu's [17] results are not as extreme as these but our large population approach suggests that even in their finite player models, as the number of participants become very large, the CSFs of weaker players will approach zero. This, for all practical purposes, removes them from effective participation in the contest. In fact, in our approach, leveling the field maximizes total effort only when the cost function is strictly convex. Even then, the CSF needs to be responsive enough to strategy in the sense that $r > 1$ (see footnote 5).

Recall that the planner seeks to optimize bias parameters over the set of Nash equilibria rather than over all possible strategy distributions in the large population of agents. The objective of the planner, therefore, is in this sense restricted. But if we accept this as an admissible objective, as is the case in the contest literature, then we can also provide a deeper theoretical reason why the large population setting may be more appropriate to examine the question of optimal bias parameters in Tullock contests. We show that irrespective of the bias parameters, the aggregate welfare of the contestants in the associated Nash equilibrium remains unchanged. This result arises only in the large population context and not in finite player Tullock contests.[6] With this result, we can then interpret the objective to maximize aggregate strategy across all possible Nash

[4]In fact, it is the fully linear case which presents the most complications in a large population model and needs to be analyzed as a special case, which we do.

[5]See Proposition 5 of Fu and Wu [17]. The responsiveness of the CSF to effort is measured by their parameter $r \in (0, 1]$. If $r = 1$, then there is linearity in the CSF.

[6]Given non–linearities and asymmetries, the Nash equilibrium aggregate payoff in finite player Tullock contests cannot be analytically established. But numerical estimates can show that changes in bias parameters can change the equilibrium aggregate payoff.

equilibrium as choosing the Pareto optimal Nash equilibrium. From a theoretical viewpoint, this parsimonious characterization of the planner's objective provides another justification for adopting a large population perspective towards the problem.

Apart from the large population context, the other important feature of our paper is that it also considers an incomplete information environment. In particular, we examine the question whether the planner can implement the optimal bias parameter when he does not have complete information about the agents. It is certainly plausible that the planner will lack such information, which makes the problem a relevant one. We adopt a standard mechanism design approach to this problem, appropriately extended to a large population model, and establish conditions under which the optimal bias parameters can be truthfully implemented in dominant strategies.[7] The crucial determinant of our results is the parameter r. If $r = 1$, then the optimal contest is an unbiased one due to which, no one can obtain a strategic advantage by misreporting type. In this case, we show that the optimal bias parameters are truthfully implementable. But if $r \neq 1$, then truthful implementation is possible only under certain conditions. An interesting conclusion is the equivalence of dominant strategy implementation and Nash implementation in our large population context, a feature that is rare in finite player models of mechanism design (Laffont and Maskin [18]). In our model, this feature arises due to the measure zero characteristic of each agent.

Polischuk and Tonis [30] also apply a mechanism design approach to the optimal contest design problem. They apply Bayesian implementation to find the optimal CSF for a planner. They find that under certain type distributions, the optimal CSF has a logit form that has some similarities with the Tullock CSF whereas for other type distributions, the optimal CSF is entirely different. In one sense, their problem is more general since they do no confine themselves to a specific form of CSFs like in Tullock contests. In contrast, we fix the general form of the CSF and try to find the optimal bias parameters. Perhaps one benefit of our approach is that our results are independent of the type distribution which, in practice, may not be known to the planner.

Ours is not the first paper to study large population contests. Polischuk and Tonis [30] themselves extend the optimal form of their CSFs to a continuum of agents as a limiting result of their finite player model. As with the rest of their paper, the details differ greatly from our model. Cornes and Hartley [8] derive certain limiting upper bounds on rent dissipation in Tullock contests with a large number of contestants. As we discuss in Section 3, our large population approach yields a precise estimate of such rent dissipation. In addition, using techniques developed in Olszewski and Siegel [28], Olszewski and Siegel [29] characterize the prize structure that maximizes performance in large population contests where prizes are awarded according to the rank order of the performances. The main differences with our model are that their contest is not a Tullock contest and their focus is on the optimal prize structure whereas ours is on the optimal bias parameters. Bodoh–Creed and Hickman [2] also use large contests, which again differ from Tullock contests, to study college admissions. Lahkar and Sultana [21], who introduce large population Tullock contests, analyze two specific form of bias parameters; one being the standard unbiased contest and another that

[7] See Lahkar and Mukherjee [22] for an application of such an approach to a large population public goods game.

incorporates a specific form of affirmative action. Our paper, therefore, may be regarded as the first one to generalize large population Tullock contests by allowing for other forms of bias parameters.

The rest of the paper is as follows. Section 2 introduces large population Tullock contests with general bias parameters. In Section 3, we characterize the unique Nash equilibrium in such contests given any vector of type specific bias parameters. Section 4 derives the optimal bias parameters given a set of bias parameters. In Section 5, we consider implementing the optimal bias parameters in an environment of incomplete information. Finally, Section 6 concludes.

2 Large Population Tullock Contests

To describe a large population Tullock contest, we consider a continuum of agents of mass 1, which we call a society. We assume that the society is divided into a finite set of populations $\mathcal{P} = \{1, 2, \cdots, n\}$. We interpret each such population as a *type* of agent. The mass of population or type $p \in \mathcal{P}$ is $m_p \in (0, 1)$ with $\sum_{p=1}^{n} m_p = 1$. We refer to this distribution $m = (m_1, m_2, \cdots, m_n)$ as the type distribution. We endow every agent in the society with the common strategy set $\mathcal{S} = [0, \infty) = \mathbf{R}_+$.

Due to the fact that the strategy set in our model is continuous, we will require certain measure theoretic notions to formalize population and social behavior. Much of this notation will, however, remain in the background during our analysis. Let $\mathcal{M}^+_\nu(\mathcal{S})$ be the space of finite positive measures that impose a total mass of $\nu > 0$ on $\mathcal{S}$. A *population state* $\mu_p \in \mathcal{M}^+_{m_p}(\mathcal{S})$ is then the distribution of strategies in population p. Thus, we interpret $\mu_p(A) \in [0, m_p]$ as the mass of agents in population p who play strategies in $A \subseteq \mathcal{S}$. Thus, $\mu_p(\mathcal{S}) = m_p$. Monomorphic population states are population states in which every agent in a population play the same strategy. If every agent in population p play $x_p \in \mathcal{S}$, then we denote the resulting monomorphic population state as $m_p\delta_{x_p}$.[8]

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

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

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.[9] The Nash equilibrium of such a game is 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 define large population Tullock contests, we assume that the society of agents are contesting over a resource of fixed value $V > 0$, which we refer to as the prize in the contest. If an agent of

[8] Here, δ_{x_p} is the Dirac distribution with probability 1 on x_p.

[9] There is perhaps a slight abuse of notation here. Given the description in (1), it would be technically more correct to write the payoff as $F(x, p, \mu)$. But in our view, $F_{x,p}(\mu)$ appears more elegant and we stick to this convention.

type p plays strategy x in the contest, then that agent incurs a cost of $k_p x^\gamma$, where $k_p > 0$ is a type specific parameter and $\gamma \geq 1$ is common for all types. Thus, agents with a higher k_p incur a higher cost of playing a strategy. We can interpret such agents as weaker or disadvantaged agents. Without loss of generality, we assume $k_1 < k_2 < \cdots < k_n$. The condition $\gamma \geq 1$ implies that the cost function is convex with respect to strategy.

Tullock contests are characterized by a contest success function (CSF) which describes the probability of success of an agent in the contest (if V is indivisible) or the share of the agent in the prize (if V if is divisible). To derive the CSF in our model, suppose an agent plays strategy x. We then introduce the mapping $x \to \theta_p x^r$ and refer to as the *strategy impact function.* As the name suggests, this function measures how impactful the strategy of an agent is in the contest. We refer to θ_p as the *bias parameter* for type p and assume that $\theta_p \geq 0$ for all $p \in \mathcal{P}$, with the inequality being strict for at least one type.[10] Thus, all agents of a particular type have the same bias parameter but that parameter can differ across types. In our subsequent analysis, we will assume that θ_p is under the control of a planner who can manipulate it to achieve certain prior objectives. Different bias parameters will mean that the same strategy can have different levels of impact depending upon the type. The common exponent r measures the responsiveness or elasticity of the impact function to strategy.

For most of the paper, we also impose the condition that $r \in (0, \gamma)$. This condition will ensure that best responses are well defined and unique at every state. Notice that since $\gamma \geq 1$, we allow for the possibility that $r > 1$. Hence, the strategy impact function need not be concave. But if $\gamma = 1$, then $r < 1$ in most of our analysis. This does leave out the one case that has been most widely analyzed in the finite player contest literature, the one where both the impact function and cost function are linear (see, for example, Franke et al. [14]). As an exception, therefore, we also discuss the $r = \gamma = 1$ as a special case. But unless otherwise specified, we will have $r \in (0, \gamma)$ and $\gamma \geq 1$. Using the strategy impact functions, we now define an aggregate measure

$$A(\mu) = \sum_{q \in \mathcal{P}} \int_{\mathcal{S}} \theta_q x^r \mu_q(x) dx. \tag{2}$$

Intuitively, $A(\mu)$ measures the aggregate impact of the strategies played by all agents in the society.[11] Hence, we call it the *aggregate strategy impact.* Since $x \in [0, \infty)$ and $\theta_q \geq 0$ for all q, we also have $A(\mu) \in [0, \infty)$.

We now define the contest success function (CSF) in our model. In finite player Tullock contests, the CSF is a fraction between 0 and 1 that measures the probability of success or the share obtained by an agent in the contest.[12] Thus, suppose there had been N players and player i plays strategy i and has bias parameter θ_i. Then, the CSF for player i would be $\frac{\theta_i x_i^r}{\sum_{j=1}^N \theta_j x_j^r}$ if $\sum_{j=1}^N x_j^r \neq 0$ and $\frac{1}{N}$ if $\sum_{j=1}^N \theta_j x_j^r = 0$. The latter fraction reflects the fact that $\sum_{j=1}^N \theta_j x_j^r = 0$ only if $\theta_i x_i^r = 0$ for every

[10] If $\theta_p = 0$ for all types, then the problem becomes trivial. All agents will find it optimal to play $x = 0$ in equilibrium.

[11] If $r = 1$ and $\theta_q = 1$ for all $q \in \mathcal{P}$, then (2) would simply be aggregate strategy in the society.

[12] See, for example, Corchón [7].

i, in which case every player shares the prize equally. The appropriate extension of such a CSF to our large population Tullock contest is then

$$\begin{cases} \frac{\theta_p x^r}{A(\mu)} & \text{if } A(\mu) \neq 0, \\ 1 & \text{if } A(\mu) = 0, \end{cases} \tag{3}$$

where $A(\mu)$ is as defined in (2).

To understand (3), let us first consider $A(\mu) \neq 0$. In that case, the CSF measures how impactful an agent's strategy is relative to the aggregate strategy impact. But unlike the finite player CSF, the CSF here is not a probability or share with value always between 0 and 1. Instead, it is to be interpreted as a density function that measures the likelihood of success or likely share of an agent. Hence, like any density function, it's value can exceed 1. If we integrate the CSF over all agents of a population p, i.e $\int_{p\in\mathcal{P}} \frac{\theta_p x^r}{A(\mu)} \mu_p(dx) \in [0, 1]$, we obtain the probability of a type p agent winning the contest or the aggregate share of all agents in that population.[13] If $A(\mu) = 0$, then the density 1 implies the uniform distribution, which is the large population analogue of every agent getting an equal share $\frac{1}{N}$ in the finite player contest. It is evident that the CSF is homogeneous of degree zero. Therefore, bias parameters are uniquely defined only up to a multiplicative constant. If $\theta_p = 1$ for all $p \in \mathcal{P}$ (or, equivalently, all bias parameters are equal), then (3) would be an unbiased CSF where equal effort would lead to equal share or probability of success. This is the case in the standard Tullock contest which is usually analyzed.

The CSF (3) and the cost functions described above allow us to define a large population Tullock contest as a population game in which the payoff of an agent of type p who plays strategy x is

$$F_{x,p}(\mu) = \begin{cases} \frac{\theta_p x^r}{A(\mu)} V - k_p x^\gamma & \text{if } A(\mu) \neq 0, \\ V - k_p x^\gamma & \text{if } A(\mu) = 0. \end{cases} \tag{4}$$

Thus, in each case, we multiply the "share" of an agent as described by the CSF in (3) with V to measure the benefit of the agent. Subtracting the cost $k_p x^\gamma$ gives us the payoff (4). Depending upon context, x can be interpreted as effort or payment or some other variable. Here, we will continue using the neutral term "strategy".

The large population contest F defined in (4) is an example of an aggregative game (Corchón [6, 7]). This is in the sense that payoffs in F depend upon the own strategy of an agent and an aggregate measure based on individual strategies which, in this case, is $A(\mu)$. This notion of aggregative games has been extended to large populations by Lahkar [19] and Cheung and Lahkar [4] for games like Cournot competition and tragedy of the commons. For these games, the relevant aggregate variable is aggregate strategy (see footnote 11). Lahkar and Sultana [21] also consider large population Tullock contests but for two specific forms of θ_p. One is where $\theta_p = 1$ for all types

[13] If every agent of every population q plays the same strategy α_q, then the aggregate share of population p would be $\frac{m_p \theta_p \alpha_p^r}{\sum_{q\in\mathcal{P}} m_q \theta_q \alpha_q^r}$. The "share" of each player in population p would then be $\frac{1}{m_p} \frac{m_p \theta_p \alpha_p^r}{\sum_{q\in\mathcal{P}} m_q \theta_q \alpha_q^r} = \frac{\theta_p \alpha_p^r}{\sum_{q\in\mathcal{P}} m_q \theta_q \alpha_q^r}$, which is the CSF in this case.

and which, therefore, represents a standard Tullock contest. The other is where $\theta_p = k_p^{\frac{r}{\gamma}}$ and is used to model a contest with affirmative action. The present paper is a generalization since it does not impose any restrictions on the values that the bias parameters can take except that they be positive. Another difference is that we allow the strategy 0 whereas in Lahkar and Sultana [21], the lowest strategy is $\underline{x} > 0$. The presence of $x = 0$ is important in our model as certain types of agents may choose not to participate in the contest at the equilibrium under the optimal bias parameters.

The contest F is also a direct extension of finite player Tullock contests with biased CSFs (see, for example, Cornes and Hartley [8], Franke et al. [14], Fu and Wu [17]). By allowing for bias parameters that differ according to player or type, such contests incorporate asymmetries in the CSF. Indeed, (4) has asymmetries not only in the CSF but also in the cost functions. In addition, (4) also allows for non–linearities in the strategy impact and cost functions. In fact, the large population framework is particularly useful in dealing with such asymmetries and non–linearities. As noted in, for example, Franke [13], finite player Tullock contests with asymmetries cannot be solved in closed form unless the strategy impact and cost functions are linear. But as we will see, such asymmetries and non–linearities can be conveniently handled in the large population case. We also note that Tullock contests can be analyzed equivalently by allowing cost functions to be homogeneous and incorporating heterogeneity in players' valuation of the prize (Fu and Wu [17]). Thus, in (4), such a strategically equivalent formulation will mean all agents have the same cost function x^γ but type specific valuations $\frac{V}{k_p}$.

3 Nash Equilibrium

We now characterize Nash equilibria of our model. Recall the definition of a Nash equilibrium in large population games from Definition 2.1. First, we show that the state where every agent plays the strategy 0 is a Nash equilibrium. For reasons quite evident, we will argue that this is not an interesting equilibrium. Hence, in the rest of this section, we will focus on an alternative Nash equilibrium. We denote the state where all players play 0 as $\mu^0 = (m_1\delta_0, m_2\delta_0, \cdots, m_n\delta_0)$. The proof of the result, stated below, is in the Appendix. It relies on the fact that a deviation by a single agent will not change the social state and, hence, the aggregate measure $A(\mu)$.

Proposition 3.1 *The state μ^0 defined above is a Nash equilibrium of the contest F defined by (4).*

In finite player Tullock contests, the outcome where all agents play 0 is not a Nash equilibrium. At this outcome, all agents would have to divide the prize equally. But a deviation by a single agent to $\epsilon > 0$ would give that agent the entire prize at a marginally higher cost. In the large population case, μ^0 is a Nash equilibrium only due to the technical reason that a deviation by a single agent does not change the social state and, hence, payoffs. But as we will see below, μ^0 is not stable with respect to a deviation by any positive measure, howsoever small, of agents. This rules out this equilibrium as a very compelling solution to our model. Moreover, with all agents playing zero, μ^0 is also not very relevant to us from the point of view of maximizing aggregate strategy. Hence, we now seek an alternative equilibrium in our model.

Our first step is to argue that any other Nash equilibrium of F, it must be that $A(\mu) > 0$, where $A(\mu)$ is as defined in (2). Obviously, if $\mu = \mu^0$, then $A(\mu) = 0$. Is it possible to have a Nash equilibrium $\mu \neq \mu^0$ such that $A(\mu) = 0$? No. At any Nash equilibrium, if $\theta_q = 0$, then every agent in such a population q must be playing $x = 0$. In fact, by (4), $x = 0$ is the strictly dominant strategy for any such agent with $\theta_q = 0$ irrespective of the value of $A(\mu)$. Therefore, to have a Nash equilibrium with $A(\mu) = 0$, it must be that every agent of every type p such that $\theta_p > 0$ must be playing $x = 0$. But this brings us back to μ^0. Therefore, the only Nash equilibrium with $A(\mu) = 0$ is $\mu = \mu^0$. Any other Nash equilibrium must have $A(\mu) > 0$.

To characterize such a Nash equilibrium, we denote $A(\mu)$ as $\alpha \in (0, \infty)$ and write the payoff function (4) as $\frac{\theta_p x^r V}{\alpha} - k_p x^\gamma$. Recall our general assumptions that $\gamma \geq 1$ and $r \in (0, \gamma)$. These assumptions imply that this payoff function is strictly quasiconcave with respect to x and, therefore, has a unique maximizer in $\mathcal{S}$.[14] This maximizer is the unique best response of a type p agent when the aggregate measure $A(\mu) = \alpha$. We denote this best response as $b_p(\alpha) \in (0, \infty)$. Maximizing (4) with respect to x when $\alpha = A(\mu) > 0$, we obtain its form

$$b_p(\alpha) = \left(\frac{\theta_p V r}{k_p \alpha \gamma}\right)^{\frac{1}{\gamma - r}}. \tag{5}$$

An important feature of (5) is the ease with which we are able to derive it. This is because agents are of measure zero. Therefore, the aggregate variable α is not affected by the actions of a single agent and, hence, effectively becomes a constant. Differentiating the payoff function and obtaining (5) then becomes simple.

Under our general assumption that $\gamma > r$, $b_p(\alpha)$ is strictly declining in α. Further, while $b_p(\alpha) = 0$ if $\theta_p = 0$, $b_p(\alpha) > 0$ if $\theta_p > 0$. Thus, for all types with a positive bias parameter, the best response is always strictly positive at any $\alpha > 0$. We note that even when $\alpha > 0$, $b_p(\alpha)$ is well–defined only if $\gamma \neq r$. Thus, in the special case of $r = \gamma = 1$, we cannot describe best responses as in (5). Intuitively, this is because in that case, an agent's payoff becomes $\left(\frac{\theta_p V}{\alpha} - k_p\right) x$. With this linear payoff, there cannot be a unique best response in $(0, \infty)$.[15] As we explain later in Section 3.1, this has implications for the structure of Nash equilibria in the model.

We can also use (5) to explain why the Nash equilibrium μ^0 characterized in Proposition 3.1 is unstable. Recall that at that equilibrium, every agent plays 0 and, hence, $\alpha = A(\mu) = 0$. Suppose now that there is a slight displacement of α from 0 to some positive value arbitrarily close to zero caused by a very small but positive measure of agents playing strategies different from 0. The best response of all players of types p with $\theta_p > 0$ then becomes arbitrarily high, as can be seen from (5). Thus, the Nash equilibrium μ^0 gets dislodged by any positive measure ε of agents playing $x > 0$.

[14] If $r \in (0, 1]$, then the payoff function is strictly concave with respect to x.

[15] If $\frac{\theta_p V}{\alpha} < k_p$, the best response is 0. If $\frac{\theta_p V}{\alpha} > k_p$, the best response is not well defined. If $\frac{\theta_p V}{\alpha} = k_p$, then any strategy in $[0, \infty)$ is a best response.

This renders μ^0 unstable. Moreover, μ^0 is also not robust against small changes in the strategy set. Suppose the strategy set is $[\underline{x}, \infty)$, where $\underline{x}$ is positive but arbitrarily close to zero. Then, the state where all agents play $\underline{x}$ will not be a Nash equilibrium. To see this, note from (5) that $b_p(\underline{x}) > \underline{x}$ for all agents of types p with $\theta_p > 0$ when $\underline{x}$ is positive but arbitrarily small.

Continuing with $\alpha > 0$, we note from (5) that different social states μ which generate the same aggregate measure $\alpha = A(\mu)$ will also generate the same best responses. Further, $b_p(\alpha)$ is identical for all agents of a particular type p but differs from agents of other types. Hence, if all agents in all populations play their respective best responses, then it will generate a social state in which every population state is monomorphic. We denote such a social state as $B(\mu)$ and describe it formally as

$$B(\mu) = \left(m_1 \delta_{b_1(\alpha)}, m_2 \delta_{b_2(\alpha)}, \cdots, m_n \delta_{b_n(\alpha)}\right), \tag{6}$$

where $\alpha = A(\mu)$. Thus, every agent in population p in (6) is playing the unique best response $b_p(\alpha)$ for that population at the social state μ.

We now apply (2) to (6) and use (5) to calculate the aggregate strategy impact at the social state $B(\mu)$. This is

$$\begin{aligned} A(B(\mu)) &= \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} \theta_p x^r m_p \delta_{b_p(\alpha)}(dx) \\ &= \sum_{p \in \mathcal{P}} m_p \theta_p b_p^r(\alpha) \\ &= \left(\frac{Vr}{\alpha\gamma}\right)^{\frac{r}{\gamma-r}} \sum_{p \in \mathcal{P}} m_p k_p^{\frac{-r}{\gamma-r}} \theta_p^{\frac{\gamma}{\gamma-r}}. \end{aligned} \tag{7}$$

Notice that $A(B(\mu))$ is strictly declining in α which, in turn, follows from the fact that $b_p(\alpha)$ is itself strictly declining in α. Using (7), we can now characterize such Nash equilibria of the Tullock contest F where $\alpha > 0$. First, we establish the following lemma. The proof, which is in the Appendix, follows from a straightforward calculation and the fact that $A(B(\mu))$ is strictly declining in α.

Lemma 3.2 *Let $\gamma \geq 1$ and $r \in (0, \gamma)$. Consider the equation*

$$A(B(\mu)) = \alpha, \tag{8}$$

where $A(B(\mu))$ is as defined in (7). This equation has a unique solution

$$\alpha^* = \left(\frac{Vr}{\gamma}\right)^{\frac{r}{\gamma}} \left(\sum_{q \in \mathcal{P}} m_q k_q^{\frac{-r}{\gamma-r}} \theta_q^{\frac{\gamma}{\gamma-r}}\right)^{\frac{\gamma-r}{\gamma}}. \tag{9}$$

Our assumption that $\theta_p > 0$ for at least one type $p \in \mathcal{P}$ implies that $\alpha^* > 0$. The following proposition now identifies the set of Nash equilibria of the contest F with $A(\mu) > 0$. In fact, we show that F has only one Nash equilibrium that satisfies this condition, the characterization of which follows from Lemma 3.2. The proof of the result is in the Appendix.

Proposition 3.3 *Consider $\mu \in \Delta$ such that $\alpha = A(\mu) > 0$ so that $b_p(\alpha)$ as characterized in (5) is well defined. In this set of social states, the Tullock contest F defined by (4) with $\gamma \geq 1$ and $r \in (0, \gamma)$ has a unique Nash equilibrium*

$$\mu^* = \left(m_1\delta_{b_1(\alpha^*)}, m_2\delta_{b_2(\alpha^*)}, \cdots, m_n\delta_{b_n(\alpha^*)}\right), \tag{10}$$

where α^* is as characterized in (9) and

$$b_p(\alpha^*) = \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}} \left(\frac{1}{\sum_{q\in\mathcal{P}} m_q k_q^{\frac{-r}{\gamma-r}} \theta_q^{\frac{\gamma}{\gamma-r}}}\right)^{\frac{1}{\gamma}} \left(\frac{\theta_p}{k_p}\right)^{\frac{1}{\gamma-r}}. \tag{11}$$

Therefore, at this his Nash equilibrium, every agent in population p plays the strategy $\alpha_p^* = b_p(\alpha^*)$ and $A(\mu^*) = \alpha^*$, where $A(\mu)$ is as defined in (2).

The proof of this result relies on the fact that (8) has a unique solution in F. Intuitively, at that solution α^*, (8) implies that agents best responding to the existing social state leaves the aggregate value α^* unchanged. This conclusion is reminiscent of a fixed point result but, as the proof of Lemma 3.2 shows, follows from much simpler arguments due to the aggregative nature of the game. Thus, when the aggregate value is α^*, (5) implies every agent in population p plays $b_p(\alpha^*)$ whereupon, as direct calculations show, the aggregate value remains at α^*. But if every type p agent plays $b_p(\alpha^*)$, $\mu_p^* = m_p\delta_{b_p(\alpha^*)}$ or $\mu^* = B(\mu^*)$ so that μ^* is a Nash equilibrium. Since (8) has a unique solution, we obtain the unique Nash equilibrium μ^* with the characteristic that $A(\mu^*) > 0$.

Propositions 3.1 and 3.3 complete the characterization of Nash equilibria of the Tullock contest F defined in (4). We obtain two Nash equilibria. Of these, it is μ^* characterized in Proposition 3.3 that is of interest to us. At this equilibrium, as can be seen from (11), the planner can ensure that agents play positive strategies by making the bias parameters positive. Hence, it is this equilibrium that is relevant from the point of view of maximizing aggregate strategy. Moreover, μ^* is also stable in the sense that deviations by a positive measure of agents will not dislodge it. Suppose $\alpha > \alpha^*$. The fact that $b_p(\alpha)$ is strictly declining in α implies that all agents will play a lower strategy than at α^*, thereby reducing the aggregate value (2) towards α^*. The reverse happens if $\alpha < \alpha^*$.

The characterization of μ^* in Proposition 3.3 has relied entirely on the fact that F is an aggregative game and is a generalization of the technique used in Lahkar [20], where $A(\mu)$ is simply the aggregate strategy (see footnote 11), to broader notions of aggregation. We should note that this particular method is not the only one to characterize Nash equilibria of large population Tullock contests. For example, Lahkar and Sultana [21] apply the technique of large population potential games (Sandholm [32]) for this purpose. But the present method is more general as it can also handle the zero strategy.[16]

We can also relate Proposition 3.3 to results in Franke [13] and Franke et al. [14]. Those papers consider finite–player Tullock contest models with biased CSFs and characterize the Nash equilibrium of such contests. Due to the finite player context, these papers require strategy impact and cost functions to be linear in order to derive closed form solutions. Even then, the presence of asymmetries in these functions render equilibrium characterization a challenging task. Our large population framework, on the other hand, can accommodate asymmetries and non–linearities very parsimoniously and without encountering any additional complications. The reason is that agents are of measure zero which, as mentioned earlier, greatly simplifies the task of calculating best responses and deriving the Nash equilibrium. Of course, in any real world interaction, the number of agents are always finite. Therefore, as with any result involving a continuum of agents, the caveat remains that these conclusions are valid only in a limiting sense when the number of agents involved are sufficiently large.

3.1 Nash Equilibria: $r = \gamma = 1$

Proposition 3.3 excluded our special case $r = \gamma = 1$. The reason, as we now discuss, is that the structure of Nash equilibria in this case is completely different. Recall that if $r = \gamma = 1$ and $\alpha = A(\mu) > 0$, then the payoff becomes $\left(\frac{\theta_p V}{\alpha} - k_p\right) x$. Further, the best response (5) is not well–defined in this case. Hence, we cannot apply the methodology of Proposition 3.3 to this case. Notice though that in any Nash equilibrium under this case, payoff can only be zero.[17] Therefore, in any equilibrium, it must be that either $x = 0$ for an agent or if $x > 0$, then $\frac{V}{\alpha} = \frac{k_p}{\theta_p}$. Hence, any state μ is a Nash equilibrium if $\frac{V}{\alpha} = \frac{k_p}{\theta_p}$ for all p with $\theta_p > 0$ and $\alpha = A(\mu) = \sum_{p:\theta_p>0} \int_{\mathcal{S}} \theta_p x \mu_p(dx) > 0$. If $\theta_p = 0$, then the equilibrium strategy will be zero for all agents in that population.

We then no longer have a unique equilibrium with $A(\mu) > 0$. Instead, given the linearity of $A(\mu)$, we will have a convex set of such equilibria. For example, one possibility is that $\theta_p = k_p$ for all $p \in \mathcal{P}$. In that case, any μ such that $A(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p x \mu_p(dx) = V$ will be a Nash equilibrium. Alternatively, we can have $\theta_1 = 1$ and $\theta_p \leq 1$ for all $p > 1$, with all agents from $p > 1$ playing 0. In that case, any μ with μ_1 such that $\int_{\mathcal{S}} x\mu_1(dx) = \frac{V}{k_1}$ and $\mu_p = m_p\delta_0$ for all $p > 1$ will be a Nash equilibrium. To see this, note that at such a state, $\alpha = A(\mu) = \int_{\mathcal{S}} \theta_1 x \mu_1(dx) = \frac{V}{k_1}$. Then, the payoff of type 1 players is zero from any strategy. For players of types $p > 1$, $\frac{V}{\alpha} = k_1 < k_p \leq \frac{k_p}{\theta_p}$ if $\theta_p \in (0, 1]$, which makes 0 their best response.

An important feature of the Nash equilibrium of finite player Tullock contests is the possibility that agents with high cost parameters do not participate in the contest (Franke [13], Franke et al. [14]). Such players exert zero effort in equilibrium. In our model, if $r < \gamma$, then this can happen only if $\theta_p = 0$ as can be seen from (11) in Proposition 3.3. But if $r = \gamma = 1$, then such

[16] As we discuss in Lahkar and Sultana [21], the potential function cannot be defined if $A(\mu) = 0$. To prevent that, the lowest strategy in that paper is $\underline{x} > 0$. In such a model, we will have a unique equilibrium corresponding to the one in Proposition 3.3 here. As noted earlier, all agents playing $\underline{x}$ will not be a Nash equilibrium.

[17] Note that θ_p, V, α and k_p are constants. Hence, if equilibrium payoff is positive, the concerned agent can deviate to a higher x and increase payoff.

non–participation is possible even with $\theta_p > 0$. Thus, as we saw in the last paragraph, if $\theta_1 = 1$ and $\theta_p \leq 1$ for all $p > 1$, then we have a convex set of Nash equilibria where all agents of type $p > 1$ play 0. Thus, if an agent does not have the lowest cost parameter k_1, then that agent does not participate in the contest. It is even possible that not all agents of type 1 participate as all we require is $\int_{\mathcal{S}} x\mu_1(dx) = \frac{V}{k_1}$ for μ to be a Nash equilibrium.

4 Optimal Bias Parameters

The Nash equilibrium strategies characterized in Proposition 3.3 depends upon the vector of bias parameters $\theta = (\theta_1, \theta_2, \cdots, \theta_n)$. Where do these bias parameters come from? Following the approach of the mechanism design literature, we now assume that there exists a planner who chooses these parameters in order to achieve certain prior objective. Before discussing that objective in detail, we introduce some notational changes to indicate the dependence of the Nash equilibrium on θ. Thus, we denote the scalar α^* in (9) as $\alpha^*(\theta)$, the Nash equilibrium μ^* in (10) as $\mu^*(\theta)$ and the equilibrium strategy α^*_p in (11) as $\alpha^*_p(\theta)$. Using these notations, we can now discuss the planner's objective that will determine the choice of the bias parameters. The following corollary calculates the key variable that we require for this purpose. Before stating the corollary, we note that we are once again focusing on the case $r < \gamma$ and $\gamma \geq 1$, which generates the Nash equilibrium in Proposition 3.3. We will consider the special case $r = \gamma = 1$ separately in Section 4.1.

Corollary 4.1 *Consider the Nash equilibrium $\mu^*(\theta)$ characterized in Proposition 3.3. The aggregate strategy at this Nash equilibrium is*

$$AS(\mu^*(\theta)) = \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}} \left(\frac{1}{\sum_{q\in\mathcal{P}} m_q k_q^{\frac{-r}{\gamma-r}} \theta_q^{\frac{\gamma}{\gamma-r}}}\right)^{\frac{1}{\gamma}} \sum_{l\in\mathcal{P}} m_l \left(\frac{\theta_l}{k_l}\right)^{\frac{1}{\gamma-r}}. \tag{12}$$

Proof. The aggregate strategy at a social state μ is simply (2) with $r = 1$. Hence, (10) implies that

$$AS(\mu^*(\theta)) = \sum_{l\in\mathcal{P}} \int_{\mathcal{S}} x\mu^*_l(\theta)(dx) = \sum_{l\in\mathcal{P}} m_l \alpha^*_l(\theta).$$

The result then follows from the form of $\alpha^*_l(\theta)$ in (11). ■

Recall our postulate that the bias parameters are chosen by a planner to meet some prior objective. We now assume that the *objective* of the planner is to maximize the aggregate equilibrium strategy $AS(\mu^*(\theta))$ characterized in (12). He, therefore, chooses the bias parameters $\theta = (\theta_1, \cdots, \theta_n)$ accordingly.

In finite player contests, such an objective has been justified as the planner seeking to maximize total payment or effort in contexts like lobbying, labor market tournaments or a sports tournament (Nti [27]). Or, having benevolent concerns, the planner may wish to level the field in favor of weaker contestants and determine conditions when such leveling will maximize total effort (Fu and Wu [17]). Such motives remain valid in our large population contest particularly if the number of contestants are significantly large. But in addition, the large population context also allows us to justify the planner's objective on the basis of the more abstract notion of maximizing total welfare. To present this argument, we introduce the notion of the *aggregate payoff.* In a population game F as defined in (1), the aggregate payoff at a social state μ is $\bar{F}(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} F_{x,p}(\mu)\mu_p(dx)$. We now establish the following result on the Nash equilibrium level of aggregate payoff in our model. The proof, which follows from direct calculation, is in the Appendix.

Proposition 4.2 *Consider the Tullock contest F defined by (4) for $\gamma \geq 1$ and $r \in (0,\gamma)$. Suppose $\theta_p \geq 0$ for all $p \in \mathcal{P}$ and $\theta_p > 0$ for at least one $p \in \mathcal{P}$. Then, the aggregate payoff in F at the Nash equilibrium $\mu^*(\theta)$ characterized in Proposition 3.3 is*

$$\bar{F}(\mu^*(\theta)) = V\left(1 - \frac{r}{\gamma}\right). \tag{13}$$

Before elaborating the planner's objective, we briefly describe the implications of Proposition 4.2, which is an important result. It shows that aggregate payoff of the contestants at the Nash equilibrium $\mu^*(\theta)$ is independent of the vector θ of the bias parameters or, for that matter, even the cost parameters k_p and the type distribution m. The equilibrium level of aggregate payoff depends only upon V and the ratio $\frac{r}{\gamma} < 1$. Of course, the individual payoff of agents of different types will depend upon the choice of the bias parameters.[18] But that is a distributional issue. The aggregate welfare of agents, as measured by the aggregate payoff (13), itself is unchanged by the bias parameters, cost parameters or the type distribution.

To understand Proposition 4.2, note that aggregate payoff $\bar{F}(\mu) = \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} F_{x,p}(\mu)\mu_p(dx)$ in our contest (4) takes the form $V - \sum_p \int_{\mathcal{S}} k_p x^\gamma \mu_p(dx)$, which is simply the fixed value minus the aggregate strategy cost. At the Nash equilibrium $\mu^*(\theta)$, $\mu_p = m_p\delta_{\alpha_p^*(\theta)}$ where $\alpha_p^*(\theta)$ is as defined in (11). Hence, the aggregate cost is $\sum_p m_p k_p(\alpha_p^*(\theta))^\gamma = \frac{Vr}{\gamma}$. Thus, the aggregate cost at equilibrium is the same irrespective of cost or bias parameters and the type distribution. Hence, at any such Nash equilibrium $\mu^*(\theta)$, the aggregate payoff is $V - \frac{Vr}{\gamma} = V\left(1 - \frac{r}{\gamma}\right)$. In some ways, this result is reminiscent of the *revenue equivalence theorem* from auction theory. All agents have the same valuation V. Therefore, every contest, irrespective of the bias parameters, is equally efficient. A revenue equivalence type argument suggests that the expenditure of the contestants, which is the aggregate cost in equilibrium, will also be the same. Hence, the aggregate payoff is also the same.

In rent–seeking models of Tullock contest, there is the well–known concept of *rent dissipation* (Nitzan [26]). This is that part of the original value V that gets wasted in equilibrium due to unproductive effort exertion measured as a proportion of V. In our context, we can use (13) to calculate rent dissipation as $\frac{V - V\left(1-\frac{r}{\gamma}\right)}{V} = \frac{r}{\gamma}$. Once again, we notice the importance of the ratio $\frac{r}{\gamma}$ in our large population Tullock contests. This calculation generalizes the findings in Lahkar and

[18] See (30) in the Appendix for the value of individual equilibrium payoffs.

Sultana to an arbitrary choice of bias parameters. In finite player Tullock contests, rent dissipation (or aggregate equilibrium payoff) cannot be explicitly calculated unless the contest is homogeneous. Cornes and Hartley [8] does derive an upper bound on rent dissipation in heterogeneous contests. They consider $N-$player unbiased Tullock contests in which the payoff of player i is $\frac{y_i}{\sum_{j=1}^{N} y_j} V - k_i y_i^{\gamma}$, where $y_j \in [0, \infty)$ is the strategy of player j and $\gamma > 1$. Note that $r = 1$ in this payoff function. Theorem 5 in Cornes and Hartley [8] then derive the upper bound $\frac{1}{\gamma}\frac{N-1}{N} \to \frac{1}{\gamma}$ as $N \to \infty$. This limiting value is, of course, the rent dissipation in our model had $r = 1$. But ours is an exact value and not just an upper bound. Moreover, it holds for biased contests as well and not just unbiased ones. The reason our large population approach works is because it is tractable enough to provide closed form expressions of Nash equilibrium and equilibrium payoffs.[19]

We now return to planner's objective. The planner's welfare is the aggregate strategy $AS(\mu^*(\theta))$ expended by the agents (in the form of, for example, total effort or total payment). The combined welfare of the planner and the contestants at the Nash equilibrium $\mu^*(\theta)$ is, therefore, $AS(\mu^*(\theta)) + V\left(1 - \frac{r}{\gamma}\right)$. By seeking to maximize $AS(\mu^*(\theta))$, the planner also maximizes this combined welfare. Therefore, the objective of planner can be equivalently interpreted as implementing the Pareto optimal Nash equilibrium by choosing the appropriate vector of bias parameters.

Of course, the planner could have had other objectives as well. For example, he could have tried to achieve a social state that maximizes welfare across all social states rather than all possible Nash equilibria. One reason why it may be justifiable to focus on implementing the optimal Nash equilibrium is that the planner does not wish to or does not have the capability to interfere too much with the natural outcome in a society, which is, of course, a Nash equilibrium. Another possibility is that instead of leaving agents with an assured aggregate welfare of $V\left(1 - \frac{r}{\gamma}\right)$, the planner could have sought to drive that to zero and extract the entire surplus for himself. The reason why the planner doesn't is that he may be benevolent enough to allow agents to retain the aggregate welfare that would have resulted through their natural interaction, which is the Nash equilibrium level of welfare. Without disturbing that welfare, the planner wishes to maximize his own benefit. By itself, this is an interesting problem to study and the reason why it is even feasible to consider it is our particular context of large population contests which ensures a constant Nash equilibrium level of aggregate welfare.

We now characterize the bias parameters that does achieve the planner's objective. Since bias parameters are unique only upto a multiplicative constant, we normalize $\theta_1 = 1$ in all our subsequent results. Thus, given our assumption that $k_1 < k_2 < \cdots < k_n$, agents with the lowest cost parameter have the bias parameter 1. The following proposition states the relevant result. The proof is in the Appendix. Notice that this result is valid only for $\gamma > 1$. We consider $\gamma = 1$ in a later corollary.

Proposition 4.3 *Consider the aggregate equilibrium strategy $AS(\mu^*(\theta))$ calculated in (12). Normalize $\theta_1 = 1$ and suppose that $\gamma > 1$, $r \in (0, \gamma)$. The vector of bias parameters that maximize*

[19] This calculation of rent dissipation is only valid for the Nash equilibrium $\mu^*(\theta)$. For the other Nash equilibrium μ^0 characterized in Proposition 3.1, aggregate payoff is V as all agents receive V. Hence, rent dissipation is 0. For reasons already explained, we are discounting that equilibrium. Since rent dissipation is zero, μ^0 is also the efficient state of the model in the sense of maximizing aggregate payoff.

$AS(\mu^*(\theta))$ is $\theta^* = (\theta_1^*, \theta_2^*, \cdots, \theta_n^*)$, where

$$\theta_1^* = 1; \theta_p^* = \left(\frac{k_1}{k_p}\right)^{\frac{1-r}{\gamma-1}}, \text{ for } p \in \{2, \cdots, n\}. \tag{14}$$

The aggregate strategy (12) at the resulting Nash equilibrium is then

$$AS(\mu^*(\theta^*)) = \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}} \frac{\sum_{p\in\mathcal{P}} m_p k_p^{\frac{-1}{\gamma-1}}}{\left(\sum_{p\in\mathcal{P}} m_p k_p^{\frac{-1}{\gamma-1}}\right)^{\frac{1}{\gamma}}}. \tag{15}$$

We refer to the Tullock contest (4) generated by applying the optimal bias parameters (14) as the *optimal Tullock contest.* Notice that θ_p^* is not well–defined for $p > 1$ if $\gamma = 1$. This is the reason we excluded this case here. Before taking up that case, we briefly discuss the implications of this result.

We can distinguish three cases in Proposition 4.3. First, if $r = 1$, then irrespective of γ, $\theta_p^* = 1$ for all $p \in \mathcal{P}$. In that case, the optimal Tullock contest is the standard Tullock contest where there is no bias and all agents' strategies are weighed equally. Second, if $r < 1$, then the optimal bias parameters are strictly decreasing in k_p. Hence, $\theta_1^* > \theta_2^* > \cdots > \theta_n^*$. In this case, the contest gets biased in favor of the stronger types, i.e. the types with the lower cost parameters. Finally, if $r \in (1, \gamma)$, then $\theta_1^* < \theta_2^* < \cdots < \theta_n^*$. The optimal bias parameters increase as k_p increases. Thus, the contest gets biased in favor of weaker contestants. A broad conclusion that, therefore, arises is that if $r \neq 1$, then the optimal Tullock contest is not the standard one. The optimal contest would incorporate bias in favor of some types of agents. We also note that in all three cases, the optimal biases are positive for all agents and, therefore, by (11), all agents participate in the contest by playing positive strategies.

The intuition behind why the order of the optimal bias parameters depends upon r is as follows. Recall that r measures the responsiveness of the impact of a strategy to the strategy. When $r < 1$, this responsiveness is low. In that case, inducing high cost agents to play a high strategy would require biasing the contest too much in their favor, which would then cause low cost agents to reduce their strategy inordinately. Instead, it is optimal to create a bias in favor of low cost agents and encourage them to increase their strategy even though it ends up reducing the equilibrium strategy of high cost agents as compared to an unbiased contest. The opposite happens when $r > 1$ so that the impact of a strategy is high. Low cost agents then anyway have a natural inclination to play a high strategy. Therefore, it becomes optimal to create incentives for high cost agents by biasing the contest in their favor. In the borderline case of $r = 1$, no such incentives are required for any agent leading to an unbiased contest.

Proposition 4.3 has implications for when leveling the field can increase aggregate effort. Leveling the field means biasing the contest in favor of weaker contestants which, in our context, means agents with higher cost parameters k_p. It is natural that any such leveling will increase effort by high cost agents but lower that of low cost agents. Whether the net effect is positive or negative is a concern in models of affirmative action in Tullock contests (Franke [13], Lahkar and Sultana [21]). Our result shows that total effort will increase if if two conditions are satisfied; the cost function is strictly convex ($\gamma > 1$) and effort is sufficiently impactful ($r \in (1, \gamma)$). Only then will the increase in effort by high cost agents be sufficient to compensate for the decline in effort by low cost agents. We now extend Proposition 4.3 to $\gamma = 1$ in the following corollary. The proof is in the Appendix.

Corollary 4.4 *Suppose $\gamma = 1$ and $r < 1$. Normalize $\theta_1 = 1$. Then, the aggregate equilibrium strategy $AS(\mu^*(\theta))$ calculated in (12) is maximized at*

$$\theta_1^* = 1; \theta_p^* = 0, \text{ for } p \in \{2, \cdots, n\}. \tag{16}$$

Intuitively, Corollary 4.4 follows from taking the limit of θ_p^* in (14) as $\gamma \to 1$. More formal arguments rely on the fact that the objective function (12) is continuous in $\gamma \geq 1$ when $r < 1$. Hence, by Berge's maximum theorem, the bias parameters that maximize (12) are continuous in γ. In Proposition 4.3, all bias parameters are strictly positive and, therefore, all agents play strictly positive strategies at the corresponding Nash equilibrium. In contrast, in Corollary 4.4, the bias parameters for all types except the lowest cost ones zero. At such bias parameters, their equilibrium strategy is zero and they will not participate in the contest. Only type 1 agents will play a positive strategy and, hence, participate. We try to provide some intuition of this finding in Section 4.1 where we discuss optimal contests when $r = \gamma = 1$. There too, we will find that aggregate strategy maximization requires participation to be restricted to type 1 agents.

We can compare Corollary 4.4 to results in Fu and Wu [17] who characterize the aggregate strategy (total effort) maximizing bias parameters in a finite player Tullock contest with $r \leq 1$ and $\gamma = 1$. Proposition 4 of Fu and Wu [17] identifies the optimal bias parameters and shows that even in the finite player context, players with a low valuation or, equivalently, with a high cost remain inactive once the optimal bias parameters are applied. The bias parameters of such inactive agents are fixed at zero. Our result that $\theta_p^* = 0$ for all $p > 1$ is similar. Proposition 5 of that paper shows that if r is close to 1, then the optimal bias parameters favour the disadvantaged players among the ones who are active. But if r is sufficiently below 1, then the optimal parameters are biased in favor of the stronger active players. In contrast, in the large population case, all active agents belong to the lowest cost type and are, hence, identical. So they all have the same bias parameter for any $r < 1$. Intuitively, this is because in the large population case, no matter how small the mass m_1 of type 1 is, the number of players of that type are still sufficiently large that even after eliminating all other players, the contest designer can generate sufficient competition within type 1 contestants so as to maximize aggregate strategy. This also suggests that even in finite player models like in Fu and Wu [17], as the number of contestants increase, the optimal bias parameters of all but the strongest players will either be zero or close to zero effectively rendering them non–participants in the contest.

4.1 Optimal Bias: $r = \gamma = 1$

Recall our discussion of Nash equilibria of our model in Section 3.1 when $r = \gamma = 1$. We concluded that there exists a convex set of equilibria such that $A(\mu) > 0$. We now argue that in any such equilibria, aggregate strategy is maximized when only type 1 agents exert positive effort.

Let μ be a Nash equilibrium with $A(\mu) > 0$ such that only agents in population 1 play a positive strategy at that equilibrium.[20] Recall from Section 3.1 that $\theta_1 = 1$ and $\theta_p \leq 1$ for $p > 1$ will induce such an equilibrium. Also recall that at such an equilibrium, payoffs are zero for all agents. Hence, from the payoff function $\left(\frac{\theta_p V}{\alpha} - k_p\right) x = 0$, we conclude that if $x > 0$ and $\theta_1 = 1$, $\frac{V}{\alpha} = k_1$. Moreover, with $x = 0$ for all $p > 1$ and $\theta_1 = 1$, $\alpha = A(\mu) = \int_{\mathcal{S}} x\mu_1(dx)$ is the aggregate strategy. Therefore, the aggregate strategy in a Nash equilibrium where only type 1 agents play a positive strategy is $\frac{V}{k_1}$.

Now consider another equilibrium $\hat{\mu}$ with $\alpha = A(\hat{\mu}) > 0$ where agents from multiple populations, possibly including but not limited to population 1, are playing a positive strategy. The zero equilibrium payoff condition then implies that for all types p where a positive measure of agents play $x > 0$ in equilibrium, $\frac{V}{\alpha} = \frac{k_p}{\theta_p}$. Normalizing $\theta_1 = 1$, we then obtain $\theta_p = \frac{k_p}{k_1}$ for all such $p > 1$. This also implies $\frac{V}{\alpha} = k_1$ or

$$\begin{aligned} \frac{V}{k_1} &= \alpha \\ \Rightarrow \frac{V}{k_1} &= \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} \theta_p x \hat{\mu}_p(dx) \\ \Rightarrow V &= \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} k_p x \hat{\mu}_p(dx). \end{aligned} \tag{17}$$

The last equality in (17) follows from $\theta_p = \frac{k_p}{k_1}$. But because $k_1 < k_2 < \cdots < k_n$, $k_1 \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} x\hat{\mu}_p(dx) < \sum_{p\in\mathcal{P}} \int_{\mathcal{S}} k_p x\hat{\mu}_p(dx)$. Hence, (17) implies

$$\begin{aligned} V &> k_1 \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} x \hat{\mu}_p(dx) \\ \Rightarrow \frac{V}{k_1} &> \sum_{p \in \mathcal{P}} \int_{\mathcal{S}} x \hat{\mu}_p(dx). \end{aligned} \tag{18}$$

But as we argued in previous paragraph, the LHS of (18) is the aggregate strategy in an equilibrium where only type 1 agents play a positive strategy. The RHS of (18) is the aggregate strategy at $\hat{\mu}$ where agents from other populations also play a positive strategy. Hence, to maximize aggregate strategy, only type 1 agents should play a positive strategy. The bias parameters that ensure this outcome are $\theta_1 = 1$ and $\theta_p \leq 1$ for $p > 1$. Notice that once we normalize $\theta_1 = 1$, the other optimal parameters cannot be uniquely determined. We summarize this discussion in the following

[20] It suffices that any positive measure of agents, and not necessarily all agents, in population 1 play a positive strategy.

proposition.

Proposition 4.5 *Consider the Tullock contest F defined by (4) and suppose $r = \gamma = 1$. For any assignment of bias parameters, this contest has a convex set of Nash equilibria such that $A(\mu) > 0$. The optimal bias parameters that maximize aggregate strategy at any such Nash equilibrium is $\theta_1 = 1$ and $\theta_p \leq 1$ for $p > 1$. Only agents of type 1 play a positive strategy at a Nash equilibrium under these bias parameters. The resulting aggregate strategy level is $\frac{V}{k_1}$.*

We can also compare our main conclusion here with Corollary 4.1 where $r < 1$ and $\gamma = 1$. Thus, when the cost functions are linear, aggregate strategy maximization requires that only type 1 agents should participate for all $r \leq 1$. For contestants of type $p > 1$ to participate in the optimal contest, we require $\gamma > 1$ (Proposition 4.3). When the cost function is strictly convex, the planner cannot maximize total strategy by restricting participation to the lowest cost agents. Intuitively, the cost of strategy rises so fast that type 1 agents do not find it optimal to play a sufficiently high strategy so as to maximize aggregate strategy. Instead, the planner should also rely on low cost agents to also play positive, albeit of lower magnitude, levels of strategy so as to enable low cost agents to reduce their effort and still maximize aggregate strategy. When the cost function is linear, the rate of increase in the strategy cost is sufficiently low that type 1 agents alone can maximize the aggregate strategy.

Franke et al. [14] characterize the aggregate strategy maximizing bias parameters in finite player Tullock contests with linear impact and cost functions, i.e. $r = \gamma = 1$. How does our findings compare with theirs? They also find that agents with high cost do not participate at the Nash equilibrium under the optimal bias parameters. Moreover, the optimal bias parameters of such non–participating agents are indeterminate. These conclusions are similar to ours. Further, within the active contestants, Franke et al. [14] find that the optimal bias parameters are increasing in the cost parameter thereby favoring weaker agents. Recall from our discussion of Fu and Wu's [17] that they also obtain the same conclusion for r sufficiently close to 1.[21] As we mentioned in the context of that discussion, the optimal bias parameter for all our active agents are identical as they all belong to the same type.

We conclude this section with a few remarks on the difference between the methodology of the current paper and the two papers with which we have compared our results in this section, namely Franke et al. [14] and Fu and Wu [17]. Due to the linear nature of their contest ($r = \gamma = 1$), Franke et al. [14] are able to explicitly characterize the Nash equilibrium for any given vector of bias parameters, insert the equilibrium strategy levels to the aggregate strategy function and maximize for the optimal bias parameters. In principle, this approach is similar to ours. But the characterization of Nash equilibrium in their finite player model is a computationally complex task. Moreover, finding the optimal bias parameters requires Franke et al. [14] to implement a complex bi–level mathematical program. On the other hand, our large population approach greatly

[21]In both Franke et al. [14] and Fu and Wu [17], $\gamma = 1$. The main difference is while in Franke et al. [14], r also equals 1, Fu and Wu [17] allows $r \leq 1$.

simplifies both the computation of Nash equilibria as well as the characterization of the optimal parameters. This is true in Proposition 4.3 (and the accompanying Corollary 4.4) as well as in our discussion of the $r = \gamma = 1$ case. Fu and Wu [17] allow $r \leq 1$ and due to this non–linearity, their finite player model does not allow an explicit computation of Nash equilibrium. Hence, even they have to resort to a programming exercise to get around the lack of a closed form expression of the Nash equilibrium and characterize the optimal bias parameters. In contrast, our large population approach allows us to derive Nash equilibria and write down the aggregate strategy in closed form. Optimizing for the bias parameters then becomes fairly straightforward. The fully linear case does present certain complications due to the convex structure of equilibria but even there, the arguments needed to arrive at the optimal bias parameters are not too daunting.

5 Implementing the Optimal Bias Parameters

We now address the question of implementing the optimal bias parameters θ^* as characterized in Section 4. We first consider the case where the optimal bias parameters θ^* as given by (14). Thus, we assume that $\gamma > 1$ and $r \in (0, \gamma)$. We will consider the two other cases, $r < 1, \gamma = 1$ and $r = \gamma = 1$, separately. The planner's concern is, of course, not directly with θ^* but with the associated equilibrium level of aggregate strategy $AS(\mu^*(\theta^*))$ as defined in (15). Hence, implementing θ^* is equivalent to implementing the Nash equilibrium $\mu^*(\theta^*)$ characterized in Proposition 3.3. In the language of mechanism design theory, $m \mapsto \mu^*(\theta^*)$ is the social choice function the planner wishes to implement. This would then generate the aggregate strategy $AS(\mu^*(\theta^*))$.

If the planner has complete information about the types of the agents, the problem is rather trivial–the planner can simply assign the optimal bias parameters (14) to each agent according to the type and let the contest unfold. The resulting outcome would be the Nash equilibrium $\mu^*(\theta^*)$.[22] Applying (14) to (11), we then obtain the strategy that a type p agent plays in that equilibrium, which is

$$\begin{aligned} \alpha_p^*(\theta^*) &= \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}} \left(\frac{1}{\sum_{q\in\mathcal{P}} m_q k_q^{\frac{-r}{\gamma-r}} \theta_q^{*\frac{\gamma}{\gamma-r}}}\right)^{\frac{1}{\gamma}} \left(\frac{\theta_p^*}{k_p}\right)^{\frac{1}{\gamma-r}} \\ &= \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}} \frac{k_p^{\frac{-1}{\gamma-1}}}{\left(\sum_{q\in\mathcal{P}} m_q k_q^{\frac{-1}{\gamma-1}}\right)^{\frac{1}{\gamma}}}. \end{aligned} \tag{19}$$

The resulting aggregate strategy is (15).

[22]We ignore the possibility of the Nash equilibrium being μ^0 as characterized in Proposition 3.1 (where every agent plays the zero strategy) on grounds of its instability.

5.1 Incomplete Information: Dominant Strategy Implementation

We now consider a scenario where the planner knows V, r, γ and the cost parameters k_p for all $p \in \mathcal{P}$, but not the types of the individual agents. Clearly the planner cannot assign the optimal bias parameters to the respective agents. In fact, we assume that the planner doesn't even know the type distribution. The planner's problem now becomes a "mechanism design problem". Our objective is to characterize conditions under which the planner can implement $\mu^*(\theta^*)$ in this situation. The appropriate solution concept from mechanism design theory depends on the "information structure" amongst the agents. Information structure refers to the assumption we make about mutual knowledge of the types of the players, i.e. whether the agents themselves know each others' types. In the "complete information" case, it is assumed that the every agent's type is a common knowledge. The natural solution concept in this case is Nash implementation: the agents choose to play equilibrium strategies at any type distribution m. However, in our model we do not make any assumption about the information structure. Hence, in this paper, we focus on dominant strategy implementation. Later though, we will make certain comments about the equivalence of Nash implementation and dominant strategy implementation in our model, an equivalence which implies double implementation of our social choice rule.

In dominant strategy implementation, the planner designs a mechanism where every player has a weakly dominant strategy at every possible type. Clearly, this solution concept is very robust and least demanding in terms of the assumption about an agent's knowledge about others' types. In particular, it requires no knowledge among agents of each other's types. Thus, we seek to design a mechanism that implements $\mu^*(\theta^*)$ without any assumption on the agents' (or planner's) information about types. A more demanding solution concept in the incomplete information case is Bayes-Nash implementation which assumes the existence of a commonly-known prior belief about the types of the agents. Further in Bayes-Nash implementation it is also required that the planner knows the prior belief. In our model we have already assumed that the planner does not know the type distribution, i.e. the prior belief is unknown to her. Recall that at present, we only consider the case where $\gamma > 1$ and $r \in (0, \gamma)$.

By the revelation principle, it suffices to consider only direct mechanisms. Thus, the planner asks each agent to report his type (which can take the form of either an announcement of the type p or the type specific cost parameter k_p). Let $\tilde{m} = (\tilde{m}_1, \cdots, \tilde{m}_n)$ be the distribution generated by the reported types, which we will call the reported type distribution. Thus, $\tilde{m}_p$ is the proportion of agents who report their type to be p. Since agents are free to misreport type, it is possible that $\tilde{m}_p \neq m_p$, the true proportion of type p agents. The mechanism the planner constructs now assigns to an agent who reports type to be q (i) the optimal bias parameter $\theta_q^* = \left(\frac{k_1}{k_q}\right)^{\frac{1-r}{\gamma-1}}$ as defined in (14), and (ii) the strategy level $\tilde{\alpha}_q(\theta^*)$ obtained from the Nash equilibrium strategy (11) by using the reported type distribution $\tilde{m}$ and the optimal bias parameter $\theta_l = \theta_l^*$ for every type $l \in \mathcal{P}$. As in (19), this strategy level simplifies to

$$\tilde{\alpha}_q(\theta^*) = \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}} \left(\frac{1}{\sum_{l\in\mathcal{P}} \tilde{m}_l k_l^{\frac{-r}{\gamma-r}} {\theta_l^*}^{\frac{\gamma}{\gamma-r}}}\right)^{\frac{1}{\gamma}} \left(\frac{\theta_q^*}{k_q}\right)^{\frac{1}{\gamma-r}} \tag{20}$$

$$= \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}} \frac{k_q^{\frac{-1}{\gamma-1}}}{\left(\sum_{l\in\mathcal{P}} \tilde{m}_l k_l^{\frac{-1}{\gamma-1}}\right)^{\frac{1}{\gamma}}}. \tag{21}$$

In terms of notation, the planner constructs a direct mechanism $\phi : (k_q, \tilde{m}) \mapsto (\theta_q^*, \tilde{\alpha}_q(\theta^*))$ that takes the reported cost parameter k_q (or, equivalently, reported type q) of an agent and the reported type distribution $\tilde{m}$ and assigns the optimal bias parameter θ_q^* and strategy $\tilde{\alpha}_q(\theta^*)$ to that agent. Applying these bias parameters (14) and the strategy levels (21) to the original payoff function (4), we then compute the payoff of a type p agent who reports his cost parameter to be k_q in the mechanism ϕ as

$$\phi_p(k_q, \tilde{m}) = \frac{\theta_q^* \tilde{\alpha}_q^r(\theta^*)}{\sum_{l\in\mathcal{P}} \tilde{m}_l \theta_l^* \tilde{\alpha}_l^r(\theta^*)} V - k_p \tilde{\alpha}_q^\gamma(\theta^*) \tag{22}$$

$$= \frac{{\theta_q^*}^{\frac{\gamma}{\gamma-r}} k_q^{\frac{-r}{\gamma-r}}}{\sum_{l\in\mathcal{P}} \tilde{m}_l {\theta_l^*}^{\frac{\gamma}{\gamma-r}} k_l^{\frac{-r}{\gamma-r}}} V \left(1 - \frac{k_p}{k_q}\frac{r}{\gamma}\right) \tag{23}$$

$$= \frac{k_q^{\frac{-1}{\gamma-1}}}{\sum_{l\in\mathcal{P}} \tilde{m}_l k_l^{\frac{-1}{\gamma-1}}} V \left(1 - \frac{k_p}{k_q}\frac{r}{\gamma}\right). \tag{24}$$

We now wish to establish conditions such that every agent finds it optimal to report his true type in the mechanism ϕ. Since we wish to check for dominant strategy implementation, truthful revelation should hold for any reported type distribution $\tilde{m}$. Applying (24), we can formally write this incentive for truthful revelation as

$$\begin{aligned} &\phi_p(k_p, \tilde{m}) \geq \phi_p(k_q, \tilde{m}) \\ \Rightarrow &\frac{k_p^{\frac{-1}{\gamma-1}}}{\sum_{l\in\mathcal{P}} \tilde{m}_l k_l^{\frac{-1}{\gamma-1}}} V \left(1 - \frac{r}{\gamma}\right) \geq \frac{k_q^{\frac{-1}{\gamma-1}}}{\sum_{l\in\mathcal{P}} \tilde{m}_l k_l^{\frac{-1}{\gamma-1}}} V \left(1 - \frac{k_p}{k_q}\frac{r}{\gamma}\right) \\ \Rightarrow &k_p^{\frac{-1}{\gamma-1}} \left(1 - \frac{r}{\gamma}\right) \geq k_q^{\frac{-1}{\gamma-1}} \left(1 - \frac{k_p}{k_q}\frac{r}{\gamma}\right), \end{aligned} \tag{25}$$

for all $p \in \mathcal{P}$ and all $q \neq p$. Condition (25) is the incentive compatibility (IC) constraint that ensures that if $\theta_l = \theta_l^*$ for all $l \in \mathcal{P}$, then no agent of type $p \in \mathcal{P}$, will pretend to be of type $q \neq p$. An important characteristic of this condition is that it depends entirely on the cost parameters k_p, k_q and the two other parameters r and γ. In particular, the reported type distribution $\tilde{m}$ plays no role. This distribution enters the inequality preceding (25) through the aggregate value $\sum_{l\in\mathcal{P}} \tilde{m}_l k_l^{\frac{-1}{\gamma-1}}$.

But this value is the same on both sides of this inequality and, therefore, gets cancelled. This, in turn, happens because each agent is of measure zero and an announcement by any single agent cannot affect this aggregate value. We will discuss the implication of this observation further when we consider the *double implementation* of our social choice function. We now state the following lemma that establishes conditions under which (25) is satisfied. The proof is in the Appendix.

Lemma 5.1 *Consider the incentive compatibility condition (25). Suppose $\gamma > 1$.*

1. *If $r = 1$, then (25) is satisfied for every $p \in \mathcal{P}$ and every $q \neq p$. Therefore, no agent of any type p has the incentive to misreport their type to be $q \neq p$.*

2. *If $r \in (0, 1)$, then (25) is satisfied with strict inequality for all $p \in \{1, 2, \cdots, n-1\}$ and all $q > p$. Therefore, no agent of any type $p \in \{1, 2, \cdots, n-1\}$ has the incentive to misreport type to be any $q > p$.*

3. *If $r \in (1, \gamma)$, then (25) is satisfied with strict inequality is satisfied for all $p \in \{2, 3, \cdots, n\}$ and all $q < p$. Therefore, no agent of any type $p \in \{2, 3, \cdots, n\}$ has the incentive to misreport type to be any $q < p$.*

Lemma 5.1, therefore, shows that the extent to which (25) is satisfied depends upon the value of r. If $r = 1$, then (25) is always satisfied. No agent of any type p has the incentive to report type $q \neq p$. The intuition is that if $r = 1$, the optimal bias parameters (14) are all equal. Hence, misreporting cannot affect the bias parameter that is assigned, as can be seen from (22). In that case, it must be optimal for a type p agent to be assigned his true Nash equilibrium strategy level $\alpha_p^*(\theta^*)$ which can only happen through truthful revelation. If $r \in (0, 1)$, then (25) is satisfied if $q > p$. Again, the intuition arises from the optimal bias parameters (14) which, in this case, are strictly declining in p. Agents, therefore, do not find it beneficial to report such types that would cause them to be assigned a lower bias parameter than the one they would obtain through truthful revelation. This also implies that type 1 agents will never misreport type in this case. The converse is true if $r \in (1, \gamma)$. Optimal bias parameters are strictly increasing so that agents do not have the incentive to report $q < p$. Therefore, in this case, type n agents will never misreport. Figure 1 explains the reasoning behind this result diagrammatically.

But Lemma 5.1 still leaves open the possibility that if $r \neq 1$, then agents may misreport type to obtain a higher bias parameter than the one from truthful revelation. Thus, if $r < 1$, then agents of type $p \in \{2, 3, \cdots, n\}$ may report $q < p$ while if $r > 1$, agents of type $p \in \{1, 2, \cdots, n-1\}$ may report $q > p$. The following lemma establishes conditions that precludes such misreporting. The proof is in the Appendix.

Lemma 5.2 *Consider the incentive compatibility condition (25). Let $\gamma > 1$.*

1. *Suppose $r \in (0, 1)$. For every $p \in \{2, 3, \cdots, n\}$, an agent of type p has no incentive to report*

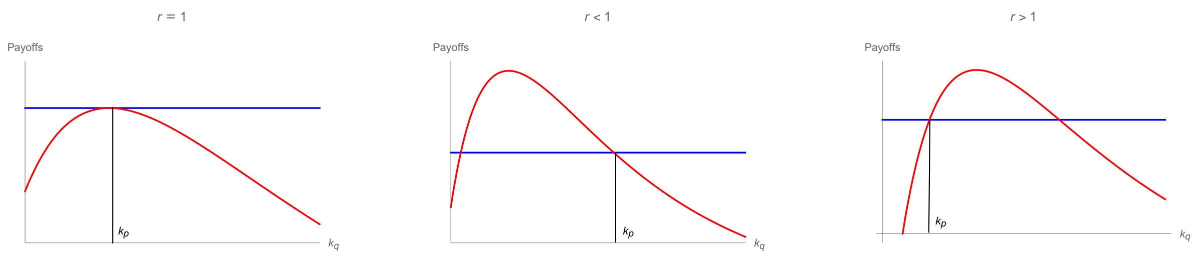

Figure 1: We fix k_p r and γ. Then, in each of these figures, the horizontal (blue) line is the LHS of (25) and the (red) curve depicts the RHS of (25) as a function of k_q. Notice that when $k_q = k_p$, the two curves intersect. (a) In the left panel, $r = 1$. As shown in Lemma 5.1(1), the LHS of (25) is strictly greater than the RHS at all $k_q \neq k_p$. (b) In the middle panel, $r < 1$. If $k_q > k_p$, then the LHS of (25) is strictly greater than the RHS. This is as shown in Lemma 5.1(2). But if $k_q < k_p$, then the RHS can be higher than the LHS unless k_q is much smaller than k_p. This is as implied in Lemma 5.2(1). (c) In the right panel, $r > 1$. Then, as shown in Lemma 5.1(3), the LHS of (25) is strictly greater than the RHS if $k_q < k_p$. If $k_q > k_p$, then the RHS can be larger than the LHS unless, as implied by Lemma 5.2(2), k_q is much larger than k_p.

type $q < p$ if given k_{p-1}, k_p is sufficiently large such that

$$k_p^{\frac{-1}{\gamma-1}}\left(1-\frac{r}{\gamma}\right) \geq k_{p-1}^{\frac{-1}{\gamma-1}}\left(1-\frac{k_p}{k_{p-1}}\frac{r}{\gamma}\right). \tag{26}$$

2. *Suppose $r \in (1, \gamma)$. For every $p \in \{1, 2, \cdots, n-1\}$, an agent of type p has no incentive to report type $q > p$ if given k_{p+1}, k_p is sufficiently small such that*

$$k_p^{\frac{-1}{\gamma-1}}\left(1-\frac{r}{\gamma}\right) \geq k_{p+1}^{\frac{-1}{\gamma-1}}\left(1-\frac{k_p}{k_{p+1}}\frac{r}{\gamma}\right). \tag{27}$$

The two key conditions (26) and (27) in this lemma are simply the original no deviation condition (25) but only for two successive types. If $r < 1$, then part 1 of Lemma 5.2 requires that k_p should be sufficiently larger than k_{p-1} so that no agent in $p \in \{2, 3, \cdots, n\}$ reports $p - 1$. This automatically ensures that there is no incentive to report $p-2$, $p-3$ etc.[23] The argument for part 2, where $r > 1$, is analogous. It requires that k_p be sufficiently smaller than k_{p+1} so that agents in $p \in \{1, 2, \cdots, n-1\}$ do not report $p + 1$. Figure 1 once again explains this result diagrammatically.

Intuitively, in each case in Lemma 5.2, we need to rule out the possibility that agents acquire a higher bias parameter through false reporting than truthful reporting. If $r < 1$, then a type p agent can obtain a higher parameter by reporting $q < p$, whereas if $r > 1$, then the same is possible by reporting $q > p$. In each case though, a higher bias parameter inflates the strategies assigned to them from the one that is optimal for them given their true cost parameter.[24] If this distortion in strategy is sufficiently high, then agents would not have the incentive for false reporting. This would

[23]If k_p is sufficiently larger than k_{p-1}, it must also be sufficiently larger than $k_{p-2} < k_{p-1}$.

[24]Notice from (21) that because $r < \gamma$, higher is θ_q^*, higher is the assigned strategy $\tilde{\alpha}_q(\theta^*)$.

be the case if the cost parameters are sufficiently far apart from each other, which is what (26) and (27) require. We can now combine Lemmas 5.1 and 5.2 to arrive at the following proposition on the implementability of the optimal bias parameters (14) as a Nash equilibrium.

Proposition 5.3 *Consider the cost parameters $\{k_1, k_2, \cdots, k_n\}$ in the direct mechanism ϕ defined by (24) and recall $k_1 < k_2 < \cdots < k_n$. Let $\gamma > 1$. Also recall the optimal bias parameters θ^* from (14) and the associated Nash equilibrium $\mu^*(\theta^*)$ characterized by effort levels $\alpha_p^*(\theta^*)$ as defined in (19) for type $p \in \mathcal{P}$ agents.*

1. *Suppose $r = 1$. Then, for all cost parameters $\{k_1, k_2, \cdots, k_n\}$, the planner is able to implement $\mu^*(\theta^*)$ in dominant strategies in ϕ.*

2. *Suppose $r \in (0, 1)$. Then, the planner is able to implement $\mu^*(\theta^*)$ in dominant strategies in ϕ if and only if for all $p \in \{2, 3, \cdots, n\}$, k_p satisfies (26).*

3. *Suppose $r \in (1, \gamma)$. Then, the planner is able to implement $\mu^*(\theta^*)$ in dominant strategies in ϕ if and only if for all $p \in \{1, 2, \cdots, n-1\}$, k_p satisfies (27).*

Proof. Part 1 follows immediately from Lemma 5.1(1). For part 2, Lemma 5.1(2) establishes that no agent of any type p reports $q > p$. Condition (26) in Lemma 5.2(1) provides the condition such that such agents will not report $q < p$. For part 3, Lemma 5.1(3) establishes that no agent of any type p reports $q < p$. Condition (27) in Lemma 5.2(2) provides the condition such that those agents will not report $q > p$. ■

Proposition 5.3 is the main result of this section. It establishes the conditions under which the planner is able to implement the Nash equilibrium $\mu^*(\theta^*)$ or, equivalently, the optimal bias parameters θ^* defined by (14). Such implementation is always possible if $r = 1$. But if $r \neq 1$, then truthful implementation in dominant strategies happens only if conditions (26) or (27), whichever is relevant, holds. Otherwise, at least some agents will have the incentive to misrepresent type. Whenever truthful revelation is incentive compatible, the planner is able to extract the aggregate strategy (15). We also note that if dominant strategy implementation holds, then, generically, it will be in strictly dominant strategies. Weak dominance will hold only if (26) or (27) is satisfied with equality for some types of agents.

We end this section with a discussion on the link between dominant strategy implementation and Nash implementation in our setting. Notice that in our mechanism, the incentive compatibility condition (25) is independent of the type distribution, whether reported or true. As noted earlier, this arises from the very nature of the large population framework. It follows that the condition for dominant strategy implementation remains the same as the condition for Nash implementation in our mechanism. For Nash implementation, we need to check that truthful revelation is a mutual best response. Each agent is of measure zero. Therefore, when every other agent reports type truthfully, the reported distribution would be m itself. The condition would be the same as (25) except that the distribution in the second of those inequalities would be the real one m. But the effect of that distribution would have cancelled out in exactly the same way as in (25) leading to the same exact conditions as in Lemmas 5.1 and 5.2 for Nash implementation. Thus, the fact that each agent is of measure zero ultimately ensures that our mechanism *double implements* our social choice function. In particular, this shows the equivalence of the two notions of implementation in our mechanism: should dominant strategy implementation not work, as may happen in parts 2 and 3 of Proposition 5.3, then the weaker notion of Nash implementation will also not work.

5.2 Two Special Cases

Proposition 5.3 only considered the case where $\gamma > 1$ and $r \in (0,1)$. We now extend our discussion of implementing the optimal bias parameters to the two special cases we have been discussing in this paper; (i) $r < 1$, $\gamma = 1$ and (ii) $r = \gamma = 1$. As will be clear below, we will need to make a distinction on the basis of whether in the reported type distribution, $\tilde{m}_1 > 0$ or $\tilde{m}_1 = 0$. First, we consider $\tilde{m}_1 > 0$.

We now examine the case $r < 1, \gamma = 1$. Recall from Corollary 4.4 that in this case, the optimal bias parameters are $\theta_1^* = 1$ and $\theta_p^* = 0$, for all $p > 1$. Hence, by (20), the strategy assigned by the planner to any agent who claims to be of type $q > 1$ is 0. By (23), the payoff obtained by such an agent will also be zero. Clearly, no agent of type $p > 1$ has any strong incentive to claim to be of any other type $q > 1$. For agents of type 1, the above argument implies that false reporting yields a payoff of zero. On the other hand, if such an agent reports truthfully, then, by (20), the strategy assigned to him is $\frac{Vr}{\tilde{m}_1 k_1}$ and by (22), his payoff is $\frac{V}{\tilde{m}_1}(1-r) > 0$. Therefore, a type 1 agent doesn't have any incentive to report type to be $q > 1$. The only other case we need to check is whether a type $p > 1$ agent has any incentive to claim to be type 1. Once again, by applying (22), we obtain the payoff of any such type p agent who claims to be of type 1 to be $\frac{V}{\tilde{m}_1}\left(1 - \frac{k_p}{k_1} r\right)$.

Hence, the condition for truthful revelation to be weakly dominant by a type $p > 1$ agent is $0 \geq \frac{V}{\tilde{m}_1}\left(1 - \frac{k_p}{k_1} r\right) \Rightarrow \left(1 - \frac{k_p}{k_1} r\right) \leq 0$, for all $p > 1$. Due to the fact that $k_1 < k_2 < \cdots < k_n$, it is sufficient that for this condition to be satisfied, we have

$$\left(1 - \frac{k_2}{k_1} r\right) \leq 0. \tag{28}$$

Notice that we could have also obtained (28) by applying part 2 of Proposition 4.3 and taking the limit as $\eta \to 1$. Since $r \in (0,1)$, the relevant incentive compatibility condition is (26). We can rewrite that condition for $p = 2$ as $\left(\frac{k_1}{k_2}\right)^{\frac{1}{\gamma-1}}\left(1 - \frac{r}{\gamma}\right) \geq \left(1 - \frac{k_2}{k_1}\frac{r}{\gamma}\right)$. As $k_1 < k_2$, $\lim_{\gamma \to 1}\left(\frac{k_1}{k_2}\right)^{\frac{1}{\gamma-1}} = 0$. Hence, taking the limit on both sides of the inequality gives us (28).

Again assuming $\tilde{m}_1 > 0$, we now consider the second of our special cases; $r = \gamma = 1$. Since $r = 1$, part 1 of Proposition 4.3 suggests that the planner will be able to implement the optimal bias parameters so as to maximize aggregate effort irrespective of the cost parameters. Recall that the reason why this holds in Proposition 4.3 (1) is that when $r = 1$ and $\gamma > 1$, $\theta_p^* = 1$ for all p. With all bias parameters same, agents do not obtain any strategic advantage by misreporting their type. In Section 4.1, we had concluded that the optimal bias parameters in the $r = \gamma = 1$

case would be such that only type 1 agents have the incentive to play a positive strategy in the resulting Nash equilibrium (Proposition 4.5). Further once we fix $\theta_1 = 1$, then any $\theta_p \leq 1$ for $p > 1$ would generate such a Nash equilibrium. Any resulting Nash equilibrium μ with $A(\mu) > 0$ will have aggregate strategy $\frac{V}{k_1}$, which is the highest possible equilibrium level of aggregate strategy in this case.

In particular, the bias vector $(\theta_1, \theta_2, \cdots, \theta_n) = (1, 1, \cdots, 1)$ also induces the highest possible equilibrium level of aggregate strategy. The direct mechanism that the planner can use, therefore, should assign $\theta_q = 1$ to all agents irrespective of reported type q. This eliminates the possibility of any strategy advantage that such agents may obtain from the bias parameter by misreporting type. Since the planner wishes aggregate strategy to be $\frac{V}{k_1}$, he assigns strategy $\frac{V}{\tilde{m}_1 k_1}$ to any agent who reports type 1, where $\tilde{m}_1$ is the mass of agents who make such a report. For agents reporting any other type $q > 1$, the assigned strategy is 0. As the bias vector is $\theta_q = 1$ for all q for all announced types q, the generalized measure (2) resulting from this strategy assignment is the aggregate strategy itself, which is $\tilde{\alpha} = \tilde{m}_1 \frac{V}{\tilde{m}_1 k_1} = \frac{V}{k_1}$.

Applying these bias parameters and this strategy assignment to the original payoff function (4), we obtain the payoff of type p agent who reports type to be 1 to be $\left(\frac{V}{\tilde{\alpha}} - k_p\right) \frac{V}{\tilde{m}_1 k_1} < 0$ if $p > 1$ due to $k_p > k_1$. On the other hand, had a type $p > 1$ agent reported his type truthfully, the assigned strategy and, therefore, payoff would have been 0. Hence, no agent of type $p > 1$ has the incentive to report type to be 1. They also do not have any strong incentive to report any $q \neq p$, $q > 1$. The strategy assigned and, therefore, the payoff will be both 0. We only need to check that agents of type 1 also do not have the incentive to report $q > 1$. Again, there is no strong incentive as both truthful and false reporting will result in 0 payoff. Therefore, truthful revelation is weakly dominant.

The above discussion for both cases required $\tilde{m}_1 > 0$ to make the assigned strategy for type 1 meaningful. But what happens if $\tilde{m}_1 = 0$? The planner can announce that whether (i) $r < 1, \gamma = 1$ or (ii) $r = \gamma = 1$, if $\tilde{m}_1 = 0$, all agents will be assigned the strategy 0. The assigned vector of type specific bias parameter $(\theta_1, \theta_2, \cdots, \theta_n)$ can be arbitrary. If all agents play strategy 0, the relevant aggregate measure (2) is also 0 irrespective of the bias parameters. By (4), every agent then receives the payoff V. But in that case, it is weakly dominant for every agent to announce type truthfully as no single announcement can affect the reported type distribution. Notice that with the true $m_1 > 0$, truthful revelation will never actually result in $\tilde{m}_1 = 0$. Hence, it will not happen that the planner will actually have to assign the 0 strategy to all agents. Nevertheless, specifying such an assignment is required to make our mechanism complete. With truthful revelation, strategy and bias parameter assignment will be as discussed earlier: type 1 agents are assigned strategy $\frac{Vr}{m_1 k_1}$ while all other agents are assigned strategy 0; type 1 agents are assigned $\theta_1 = 1$ while other agents are assigned $\theta_p = 0$ if $r < 1$ and $\theta_p = 1$ if $r = 1$. We summarize our entire discussion in this subsection in the following corollary.

Corollary 5.4 *Suppose $\gamma = 1$. Consider the direct mechanism $\phi : (k_q, \tilde{m}) \mapsto (\theta_q^*, \tilde{\alpha}_q(\theta^*))$, where $\tilde{\alpha}_q(\theta^*)$ is the strategy assigned to an agent who announces type to be q.*

1. *Let $r \in (0,1)$. Define ϕ as follows.*

   (a) *If $\tilde{m}_1 > 0$, then ϕ assigns bias parameter $\theta_1^* = 1$ and strategy $\frac{Vr}{\tilde{m}_1 k_1}$ to an agent who reports his cost parameter to be k_1. To all other agents, it assigns bias parameter 0 and strategy 0.*

   (b) *If $\tilde{m}_1 = 0$, then ϕ assigns strategy 0 to all agents and any arbitrary vector of type specific bias parameters $(\theta_1, \theta_2, \cdots, \theta_n)$.*

   Then ϕ implements the optimal bias parameters in weakly dominant strategies and, therefore, maximizes the Nash equilibrium level of aggregate strategy if (28) is satisfied.

2. *Let $r = 1$. Define ϕ as follows.*

   (a) *If $\tilde{m}_1 > 0$, then ϕ assigns bias parameter $\theta_q^* = 1$ to all agents irrespective of reported type q. The strategy assigned is $\frac{V}{\tilde{m}_1 k_1}$ if reported type is 1 and 0 if reported type is anything else.*

   (b) *If $\tilde{m}_1 = 0$, then ϕ assigns strategy 0 to all agents and any arbitrary vector of type specific bias parameters $(\theta_1, \theta_2, \cdots, \theta_n)$.*

   Then, ϕ implements the optimal parameters $\theta_p^* = 1$ in weakly dominant strategies and, therefore, maximizes the Nash equilibrium level of aggregate strategy.

5.3 Failure of Truthful Implementation: Two Types

Proposition 5.3 is our main result on truthful implementation. But if $r \neq 1$, it holds only if (26) or (27), whichever is relevant, is satisfied. The question that then arises is what can the planner implement if these conditions are violated. For the general case, this question is difficult and we don't have an answer. We are, however, able to characterize the solution completely if there are only two types of agents.

We present our analysis only for the case $r < 1$. Suppose there are two types, 1 and 2 and (26) is violated from type 2. Then, Lemma 5.1(2) and Lemma 5.2(2) imply that while truthful revelation remains dominant for type 1, type 2 will have the incentive to claim to be of type 1. In that case, the optimal bias parameters θ_1^* and θ_2^* as defined in (14) cannot be implemented.

In this situation, denote the Nash equilibrium level of aggregate strategy maximizing bias parameters that can be truthfully implemented as θ_1^{**} and θ_2^{**} and normalize $\theta_1^{**} = 1$. We need to characterize θ_2^{**}. Standard results from mechanism design theory implies that in this case, the IC constraint has to bind tightly from type 2. In terms of the notation introduced in (24) and using the normalized value θ_1^{**}=1, this would mean

$$\phi_2(k_1, m) = \phi_2(k_2, m)$$

$$\Rightarrow \frac{k_1^{\frac{-r}{\gamma-r}}}{\sum_{l \in \mathcal{P}} m_l (\theta_l^{**})^{\frac{\gamma}{\gamma-r}} k_l^{\frac{-r}{\gamma-r}}} V \left(1 - \frac{k_2}{k_1}\frac{r}{\gamma}\right) = \frac{(\theta_2^{**})^{\frac{\gamma}{\gamma-r}} k_2^{\frac{-r}{\gamma-r}}}{\sum_{l \in \mathcal{P}} m_l (\theta_l^{**})^{\frac{\gamma}{\gamma-r}} k_l^{\frac{-r}{\gamma-r}}} V \left(1 - \frac{r}{\gamma}\right)$$

$$\Rightarrow \theta_2^{**} = \left(\frac{k_2}{k_1}\right)^{\frac{r}{\gamma}} \left(\frac{1 - \frac{k_2}{k_1}\frac{r}{\gamma}}{1 - \frac{r}{\gamma}}\right)^{1-\frac{r}{\gamma}}. \tag{29}$$

Thus, in the event of the first best solution θ_2^* failing incentive compatibility for type 2, (29) is the optimal bias parameter that the planner needs to announce for type 2 agents to ensure truthful revelation. Using (14), we can verify that $\theta_2^{**} > \theta_2^*$. The planner, therefore, needs to increase the probability of success for type 2 agents, and correspondingly reduce for type 1 agents, to ensure incentive compatibility.

We can also verify without too much difficulty that given these values of θ_1^{**} and θ_2^{**}, the IC constraint for type 1 agents is satisfied. Using the notation from (24), we need to show that $\phi_1(k_1, m) \geq \phi_1(k_2, m)$ under these bias parameters. Applying $\theta_1^{**} = 1$ and θ_2^{**} as calculated in (29) to (24), we simplify this condition to $\left(1 - \frac{r}{\gamma}\right)^2 \geq \left(1 - \frac{k_2}{k_1}\frac{r}{\gamma}\right)\left(1 - \frac{k_1}{k_2}\frac{r}{\gamma}\right)$, which always holds. Indeed, with $k_2 \neq k_1$, this inequality holds strictly.

6 Conclusion

This paper has considered the design of optimal large population Tullock contests. The planner has the capacity to appropriately bias the likelihood of success of different types of agents in such a large population contest. We have characterized the unique Nash equilibrium for any such vector of bias parameters that generates a positive level of aggregate strategy. The objective of the planner is to select that vector of bias parameters that maximizes this equilibrium level of aggregate strategy. We show that the aggregate payoff of agents is unchanged by the choice of bias parameters. Therefore, choosing to maximize aggregate strategy implies the choice of the Pareto optimal Nash equilibrium if we measure the welfare of the planner as the aggregate strategy (in the form of aggregate effort or aggregate payment) being played by the agents.

We then characterize the optimal bias parameters and identify conditions under which those parameters are increasing or decreasing according to the cost parameter of the different types of agents. If we identify agents with high cost parameters as disadvantaged agents, our results provide insight into when it might be beneficial to level the playing field in favor of such agents in order to maximize aggregate strategy. Significantly, due to the convenience accorded by the large population framework, the computational techniques we need to use to characterize Nash equilibrium and optimal bias parameters are much more straightforward than in the corresponding results in finite player contests. We then identify conditions under which the planner can truthfully implement the optimal parameters in conditions of incomplete information. If the optimal bias parameters are identical for all types so that no strategic advantage can be gained by misreporting, then those parameters are always implementable. Otherwise, implementation is possible only if the cost parameters that distinguish different types of agents are sufficiently distinct.

We can think of two further interesting research questions. One is to generalize the planner's objective function. In this paper, the planner has only sought to maximize aggregate strategy.

But it is possible that the planner may have a broader objective function which, for example, also incorporates the variance of the equilibrium strategy levels. This would be relevant, for example, if the planner also cares about how close the competition is between the contestants. It will be interesting to derive optimal bias parameters under such more general objective functions. Second is the property of double implementation that we found in our discussion of implementation of the optimal bias parameters. The equivalence between Nash and dominant strategy implementation is an interesting topic in mechanism design theory. In finite player models, it has been established that *only* in a sufficiently rich environment are social choice functions that are implementable in Nash strategies also implementable in dominant strategies (Laffont and Maskin [18]). For instance, the class of quasi-linear preferences is not rich and hence the equivalence fails. In this particular model of ours, the two solution concepts turn out to be equivalent and this is due to the measure zero characteristic of every agent. It will be interesting to explore whether such equivalence holds more generally in a large population setting beyond Tullock contests.

A Appendix

Proof of Proposition 3.1: At $\mu = \mu^0$, $A(\mu) = 0$. Hence, by (4), the payoff of every agent at that state is $F_{0,p}(\mu^0) = V$. Suppose a single agent of type p deviates to strategy $x > 0$. Since every agent is of measure zero, the social state doesn't change. Hence, we still have $A(\mu) = 0$. But then, by (4), the payoff of the agent deviating to $x > 0$ is $V - k_p x^\gamma < F_{0,p}(\mu^0)$. Hence, μ^0 is a Nash equilibrium. ■

Proof of Lemma 3.2: The fact that α^* is a solution to $A(B(\mu)) = \alpha$ follows immediately from (7). To establish that this is a unique solution, note that the assumptions on r and γ imply $\gamma > r$. Therefore, it follows from (7) that (5) is continuous and strictly declining in α. Moreover, as $\alpha \to 0$, $A(B(\mu)) \to \infty$ and as $\alpha \to \infty$, $A(B(\mu)) \to 0$. Therefore, $A(B(\mu))$ and α has only one intersection, which is α^*. ■

Proof of Proposition 3.3: First, we show that μ^* is a Nash equilibrium. For this, we show that given (10) and (11), $A(\mu^*) = \alpha^*$, where α^* is as characterized in (9). This follows because $\mu^*_p = m_p \delta_{b_p(\alpha^*)}$, with $b_p(\alpha^*)$ being described by (11). Therefore, by (2), $A(\mu^*) = \sum_{p \in \mathcal{P}} m_p \theta_p b_p^r(\alpha^*)$. The conclusion then follows from direct calculation using (11). Comparing (6) and (10), we then obtain $\mu^* = BR(\mu^*)$. Every agent at μ^* plays the unique best response to μ^*. Hence, μ^* is a Nash equilibrium with $A(\mu^*) = \alpha^* > 0$.

We now show that there is no other Nash equilibrium μ with $A(\mu) > 0$. Consider $\mu \neq \mu^*$ such that $\alpha = A(\mu) > 0$. Lemma 3.2, therefore, implies $A(B(\mu)) \neq \alpha$. In our model, every agent of every type has a unique best response to every state if $\alpha > 0$. Hence, for μ to be a candidate Nash equilibrium, it must be of the form $\mu_p = m_p \delta_{b_p(\alpha)}$. But then that would imply $A(B(\mu)) = \sum_{p \in \mathcal{P}} m_p \theta_p b_p^r(\alpha) = A(\mu) = \alpha$ by the definition of $A(\mu)$ in (2), which contradicts our earlier assumption that $A(B(\mu)) \neq \alpha$. Hence, μ is not a Nash equilibrium. ■

Proof of Proposition 4.2: We can use (4) and (11) to write the Nash equilibrium payoff a type p agent as

$$F_{\alpha_p^*(\theta),p}(\mu^*(\theta)) = \frac{\theta_p(\alpha_p^*(\theta))^r}{\sum_{q\in\mathcal{P}} m_q\theta_q(\alpha_q^*(\theta))^r}V - k_p(\alpha_p^*(\theta))^\gamma$$
$$= \frac{\theta_p^{\frac{\gamma}{\gamma-r}} k_p^{\frac{-r}{\gamma-r}}}{\sum_{q\in\mathcal{P}} m_q\theta_q^{\frac{\gamma}{\gamma-r}} k_q^{\frac{-r}{\gamma-r}}}V\left(1-\frac{r}{\gamma}\right). \tag{30}$$

Recall that the aggregate payoff at a social state μ is $\bar{F}(\mu) = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} F_{x,p}(\mu)\mu_p(dx)$. From (30), we can, therefore, also calculate the aggregate payoff at the Nash equilibrium. Since the equilibrium (10) is in monomorphic population states, this turns out to be

$$\bar{F}(\mu^*(\theta)) = \sum_{p\in\mathcal{P}}\int_{\mathcal{S}} F_{x,p}(\mu^*(\theta))\mu_p^*(\theta)(dx)$$
$$= \sum_{p\in\mathcal{P}} m_p F_{\alpha_p^*(\theta),p}(\mu^*(\theta))$$
$$= V\left(1-\frac{r}{\gamma}\right). \blacksquare$$

Proof of Proposition 4.3: Denote $\left(\frac{\theta_p}{k_p}\right)^{\frac{1}{\gamma-r}} = s_p$. We can then write (12) equivalently as (suppressing the dependence on θ)

$$AS(\mu^*) = \left(\frac{Vr}{\gamma}\right)^{\frac{1}{\gamma}}\left(\frac{1}{\sum_{q\in\mathcal{P}} m_q k_q s_q^\gamma}\right)^{\frac{1}{\gamma}}\sum_{l\in\mathcal{P}} m_l s_l \tag{31}$$

We now maximize (31) with respect to $(s_1,\cdots,s_n)$. Denote the maximizer as $(\hat{s}_1,\cdots,\hat{s}_n)$. Differentiating (31) with respect to s_p, setting $\frac{\partial TE(\mu^*)}{\partial s_p} = 0$ and simplifying, we obtain

$$k_p\hat{s}_p^{\gamma-1} = \frac{\sum_{q\in\mathcal{P}} m_q k_q \hat{s}_q^\gamma}{\sum_{l\in\mathcal{P}} m_l \hat{s}_l}, \text{ for all } p\in\mathcal{P}. \tag{32}$$

Hence, $k_1\hat{s}_1^{\gamma-1} = k_2\hat{s}_2^{\gamma-1} = \cdots = k_n\hat{s}_n^{\gamma-1}$. We can, therefore, express the optimum values of all s_p in terms of $\hat{s}_1$. Thus, $\hat{s}_p = \left(\frac{k_1}{k_p}\right)^{\frac{1}{\gamma-1}}\hat{s}_1$, for all $p\neq 1$. Normalize $\hat{s}_1 = 1$ so that $\hat{s}_p = \left(\frac{k_1}{k_p}\right)^{\frac{1}{\gamma-1}}$, for all $p\neq 1$.

Using the definition of s_p, we denote $\hat{s}_p = \left(\frac{\hat{\theta}_p}{k_p}\right)^{\frac{1}{\gamma-r}}$, where $(\hat{\theta}_1,\cdots,\hat{\theta}_n)$ is the vector of effort maximizing weights that we seek to characterize. Using the normalized values of $\hat{s}_p$, we then obtain $\hat{s}_p = \left(\frac{k_1}{k_p}\right)^{\frac{1}{\gamma-1}} = \left(\frac{\hat{\theta}_p}{k_p}\right)^{\frac{1}{\gamma-r}}$. This implies $\hat{\theta}_1 = k_1$ and $\hat{\theta}_p = \left(\frac{k_1^{\gamma-r}}{k_p^{1-r}}\right)^{\frac{1}{\gamma-1}}$, for all $p\neq 1$. Normalizing further (since the CSF is homogeneous of degree zero), we can write the optimal weights equivalently as $\theta_1^* = 1$ and $\theta_p^* = \left(\frac{k_1}{k_p}\right)^{\frac{1-r}{\gamma-1}}$. Using $\theta = \theta^*$ in (12), we then obtain the aggregate strategy (15). $\blacksquare$

Proof of Corollary 4.4: Due to our assumption that $k_1 < k_2 < \cdots < k_n$, $\frac{k_1}{k_p} < 1$ in (14). Therefore, with $r < 1$, $\lim_{\gamma \to 1} \theta_p^* = 0$ for every $p > 1$ in (14). We now apply Berge's maximum theorem to argue that this limit is also the optimal bias parameter at $\gamma = 1$.

The objective function $AS(\mu^*(\theta))$ defined by (12) is continuous in $\gamma \geq 1$ if $r < 1$. We can write the problem of finding the optimal bias parameter as $\max AS(\mu^*(\theta))$ subject to the constraint set

$$\{\theta = (\theta_1, \theta_2, \cdots, \theta_n) : \theta_i \geq 0, \sum_{i=1}^{n} \theta_i = 1\}, \tag{33}$$

which is compact in $\mathbf{R}_+^n$. Therefore, by Berge's maximum theorem, the solution to this constrained optimization problem is continuous in γ. Applying (14), we can write the solution to this problem if $\gamma > 1$ as

$$\hat{\theta}_p = \frac{\left(\frac{k_1}{k_p}\right)^{\frac{1-r}{\gamma-1}}}{\sum_{q=1}^{n} \left(\frac{k_1}{k_q}\right)^{\frac{1-r}{\gamma-1}}}. \tag{34}$$

As CSFs are homogeneous of degree zero, the solution in (14) is equivalent to the one in (34). Due to the continuity of the maximizer, the solution when $\gamma = 1$ must be $\lim_{\gamma \to 1} \hat{\theta}_p$, which is 1 if $p = 1$ and 0 if $p > 1$. ■

Proof of Lemma 5.1: Consider an agent of type p who is seeking to report cost parameter k_q and observe that (25) holds with equality if $k_q = k_p$. Interpret k_q as a variable k. We differentiate the RHS of (25) with respect to k and obtain

$$\frac{d}{dk} k^{\frac{-1}{\gamma-1}} \left(1 - \frac{k_p}{k} \frac{r}{\gamma}\right) = \frac{k_p r - k}{\gamma - 1} k^{\frac{1}{\gamma-1} - 2}. \tag{35}$$

1. If $r = 1$, then using (35), we conclude that the RHS of (25) is maximized at $k = k_p$. Therefore, in this case, (25) holds with strict inequality at all $k_q \neq k_p$. The agent, therefore, has no incentive to report $k_q \neq k_p$.

2. If $r < 1$, then (35) implies that the RHS of (25) is maximized at $k < k_p$. In fact for $k_q \geq k_p$, the derivative in (35) is strictly negative. Therefore, the payoff from misreporting $k_q > k_p$ is strictly lower than the payoff from truthful revelation. Hence, the agent has no incentive to claim $k_q > k_p$ or $q > p$.

3. If $r > 1$, then (35) implies that the RHS of (25) is maximized at $k > k_p$. In fact for $k_q \leq k_p$, the derivative in (35) is strictly positive. Therefore, the payoff from misreporting $k_q < k_p$ is strictly lower than the payoff from truthful revelation. Hence, the agent has no incentive to claim $k_q < k_p$ or $q < p$. ■

Proof of Lemma 5.2:

1. Fix k_{p-1}. First, we argue that the two sides of (26) hold with equality at two points. One is obviously $k_p = k_{p-1}$. Given our assumptions about the cost parameters, we know $k_p \neq k_{p-1}$. Second, there exists another value of k_p, which we denote k_p^h, such that the two sides of (26) holds with equality. To see this, note that both sides of (26) are strictly declining in k_p. It is easy to verify that given $r < 1$, the left hand side of the inequality declines faster than the right hand side at $k_p = k_{p-1}$. Moreover, as $k_p \to \infty$, the left hand side converges to zero but remains strictly positive while the right hand side declines to $-\infty$. Together, these properties must imply that there exists $k_p^h > k_{p-1}$ such that if $k_p \in (k_{p-1}, k_p^h)$, then $k_p^{\frac{-1}{\gamma-1}}\left(1-\frac{r}{\gamma}\right) < k_{p-1}^{\frac{-1}{\gamma-1}}\left(1-\frac{k_p}{k_{p-1}}\frac{r}{\gamma}\right)$, but for all $k_p > k_p^h$, $k_p^{\frac{-1}{\gamma-1}}\left(1-\frac{r}{\gamma}\right) > k_{p-1}^{\frac{-1}{\gamma-1}}\left(1-\frac{k_p}{k_{p-1}}\frac{r}{\gamma}\right)$. Therefore, if $k_p \geq k_p^h$, then (26) holds.

   If (26) is satisfied for all $p \in \{2, 3, \cdots, n\}$, then that suffices to ensure that for all such p, an agent of type p will not claim to be of type $p-2$, $p-3$ and so on, should such types exist. To see this, write (26) for types p and $p-2$.[25] This will be satisfied if $k_p \geq k_{p-1}^h$, where k_{p-1}^h is as defined earlier but for type $p-1$ (i.e., k_{p-1}^h is obtained by writing (26) for types $p-1$ and $p-2$). But by assumption, $k_p > k_{p-1}$ and $k_{p-1} \geq k_{p-1}^h$ if (26) is satisfied for type $p-1$. Thus, an agent of type p wouldn't claim to be of $p-2$. By induction, the argument can be extended to types $p-3$, $p-4$ etc.

2. The argument is analogous. Fix k_{p+1} and treat the two sides of (27). The two sides are equal at $k_p = k_{p+1}$ and $k_p = k_p^l < k_{p+1}$. Our interest is in k_p^l since we know that $k_p < k_{p+1}$. With $r > 1$, the properties of the two functions imply that if $k_p \in (k_p^l, k_{p+1})$, then $k_p^{\frac{-1}{\gamma-1}}\left(1-\frac{r}{\gamma}\right) < k_{p+1}^{\frac{-1}{\gamma-1}}\left(1-\frac{k_p}{k_{p+1}}\frac{r}{\gamma}\right)$, but for all $k_p < k_p^l$, $k_p^{\frac{-1}{\gamma-1}}\left(1-\frac{r}{\gamma}\right) > k_{p+1}^{\frac{-1}{\gamma-1}}\left(1-\frac{k_p}{k_{p+1}}\frac{r}{\gamma}\right)$. Therefore, if $k_p \leq k_p^l$, then (27) holds. An agent of type p will not report $p+1$. By induction, the argument can be extended to not reporting $p+2$, $p+3$ etc., should such types exist. ■

References

[1] Baye M, Kovenock D, de Vries CG (1993) Rigging the lobbying process: an application of the all-pay auction. Am. Econ. Rev. 83: 289–294.

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

[3] Baye MR, Kovenock D, de Vries CG (1993) Rigging the lobbying process: an application of the all-pay auction. Am. Econ. Rev. 83:289–294.

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

[25] The relevant inequality will be $k_p^{\frac{-1}{\gamma-1}}\left(1-\frac{r}{\gamma}\right) \geq k_{p-2}^{\frac{-1}{\gamma-1}}\left(1-\frac{k_p}{k_{p-2}}\frac{r}{\gamma}\right)$.

[5] Congleton RD, Hillman AL, Konrad KA (eds.) (2008) 40 Years of Research on Rent Seeking (Vols. 1, 2). Berlin: Springer.

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

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

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

[9] Clark DJ, Riis R (1998). Contest success functions: an extension. Econ Theory 11:201–204.

[10] Dasgupta A, Nti KO (1998). Designing an optimal contest. Eur. J. Political Econ. 14:769—781.

[11] Epstein GS, Mealem Y, Nitzan S (2011) Political culture and discrimination in contests. J. Public Econ. 1:88–93.

[12] Fang H (2002) Lottery versus all-pay auction models of lobbying. Public Choice 112:351–371.

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

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

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

[16] Fu W, Lu J (2012) The optimal multi-stage contest. Econ. Theory 51:351–382.

[17] Fu Q, Wu Z (2020) On the optimal design of biased contests. Theor. Econ. 15:1435–1470.

[18] Laffont JJ, Maskin E (1982) Nash and Dominant Strategy Implementation in Economic Environments. J. Math. Econ. 10:17–47.

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

[20] Lahkar R (2019) Elimination of non-individualistic preferences in large population aggregative games. J. Math. Econ. 84:150–165.

[21] Lahkar R, Sultana R (2020). Affirmative Action in Large Population Contests. Working paper.

[22] Lahkar R, Mukherjee S (2020) Dominant Strategy Implementation in a Large Population Public Goods Game. Econ. Lett. 197:109616.

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

[24] Moldovanu B, Sela A (2001) The optimal allocation of prizes in contests. Am. Econ. Rev. 91: 542—558.

[25] Mealem Y, Nitzan S (2016) Discrimination in contests: a survey. Rev. Econ. Design 20:145–172.

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

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

[28] Olszewski W, Siegel R (2016) Large Contests. Econometrica 84:835–854.

[29] Olszewski W, Siegel R (2020) Performance–maximizing large contests. Theor. Econ. 15:57–88.

[30] Polischuk L, Tonis A (2013) Endogenous contest success functions: a mechanism design approach. Econ. Theory 52:271–297.

[31] Runkel M (2011) Revenue sharing, competitive balance and the contest success function. German Econ. Rev. 12:256–273.

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

[33] Skaperdas S (1995) Contest success functions. Econ. Theory 7:283–290

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

[35] Tsoulouhas T, Knoeber CR, Agrawal A (2007) Contests to become CEO: incentives, selection and handicaps. Econ. Theory 30: 195–221.

[36] 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.

[37] 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.