Optimal sequential contests*

Toomas Hinnosaar†

February 2018‡

First version: December 2016

Abstract

I study sequential contests where the efforts of earlier players may be disclosed to later players by nature or by design. The model has a range of applications, including rent seeking, R&D, oligopoly, public goods provision, and tragedy of the commons. I show that information about other players' efforts increases the total effort. Thus, the total effort is maximized with full transparency and minimized with no transparency. I also study the advantages of moving earlier and the limits of large contests.

JEL: C72, C73, D72, D82, D74

Keywords: contest design, oligopoly, public goods, rent-seeking, R&D

---

*I am grateful to Simon Board, Federico Boffa, Jeff Ely, Alex Frankel, Andrea Gallice, Dan Garrett, Dino Gerardi, Marit Hinnosaar, Johannes Hörner, Martin Jensen, Kai Konrad, Dan Kovenock, Nenad Kos, Ignacio Monzon, Peter Norman, Marco Ottaviani, Mallesh Pai, Ron Siegel, Andy Skrzypacz, Martin Szydlowski, and Jean Tirole as well as seminar participants at Bocconi University, Collegio Carlo Alberto, Universidad Carlos III de Madrid, University of Bern, Rice University, University of North Carolina at Chapel Hill, Canadian Economic Theory Conference, Conference on Economic Design, Contests: Theory and Evidence Conference, SAET, PET, Stony Brook Game Theory Festival, SITE at Stanford, EARIE, Midwest Economic Theory Meeting, Lancaster Game Theory Conference, and Markets with Informational Asymmetries workshop (Turin) for their comments and suggestions. I would also like to acknowledge the hospitality of the CMS-EMS at the Kellogg School of Management at Northwestern University where some of this work was carried out.

†Collegio Carlo Alberto, `toomas@hinnosaar.net`.

‡The latest version of the paper is available at `toomas.hinnosaar.net/contests.pdf`. A previous version of this paper was circulated under the title "Dynamic common-value contests".

1 Introduction

Many economic interactions have contest-like structures, where payoffs are increasing in players' own efforts and decreasing in the total effort. Examples include oligopolies, public goods provision, tragedy of the commons, rent seeking, R&D, advertising, and sports. Most of the previous literature assumes that effort choices are simultaneous. Simultaneous contests often have convenient properties: the equilibrium is unique, in pure strategies, and is relatively easy to characterize.

In this paper, I study contests where the effort choices are not necessarily simultaneous. In most real-life situations some players can observe their competitors' efforts. Later movers can respond appropriately to the choices earlier movers make. However, earlier movers can anticipate these responses and affect the behavior of later movers. Each additional period in a sequential contest adds a level of complexity, which might explain why most of the previous work studies only simultaneous and two-period models. I characterize equilibria for an arbitrary sequential contest and analyze how the information about other players' efforts affects the equilibrium behavior.

Contests may be sequential by nature or by design. Contest designers often can choose how much information about players' efforts to disclose. For example, in rent-seeking contests, firms lobby to achieve market power. One tool that regulators can use to minimize rent-seeking is to design a disclosure policy.[1] A non-transparent policy would lead to simultaneous effort choices, full transparency to a fully sequential contest, and there may be potentially intermediate solutions such as revealing information only occasionally. In research and development, the probability of a scientific breakthrough may be proportional to research efforts, which are typically considered socially desirable. The question is again how to organize the disclosure rules. A transparent policy could be implemented as a public leaderboard or early working papers, whereas a non-transparent policy would encourage teams to work in isolation.

I provide two main results. First, in the characterization theorem (Theorem 1) I characterize all equilibria for any given disclosure rule. The standard backward-induction approach requires finding best-response functions every period and substituting them recursively. I show that this approach is not generally tractable, or even feasible. Instead, I introduce an alternative approach, in which I characterize best-response functions by their inverses. This allows me to pool all the optimality conditions into one necessary

[1] In the last decades, many countries have introduced new legislation regulating transparency in lobbying, including the United States (Lobbying Disclosure Act, 1995; Honest Leadership and Open Government Act, 2007), the European Union (European Transparency Initiative, 2005), and Canada (Lobbying Act, 2008). However, there are significant cross-country differences in regulations—for example, in the US, lobbying efforts must be reported quarterly, whereas in the EU, reporting is arranged on a more voluntary basis and on yearly frequency.

condition and solve the resulting equation just once. I prove that the equilibrium exists and is unique, and I show how to compute it.

In my second main result (Theorem 2) I show that the information about other players' efforts strictly increases the total effort. This means the optimal contest is always one of the extremes. When efforts are desirable (as in R&D competitions) the optimal contest is one with full transparency, whereas when the efforts are undesirable (as in rent-seeking), the optimal contest is one with hidden efforts. The intuition behind this result is simple. Near the equilibrium, the efforts are strategic substitutes. Therefore, players whose actions are disclosed have an additional incentive to exert more effort to discourage later players. This discouragement effect is less than one-to-one near the equilibrium because if players can increase their effort in a way that diminishes total effort, this provides them with a profitable deviation.

The information about other players' effort is important both qualitatively and quantitatively. Under typical assumptions (Tullock contest payoffs), the sequential contest with 5 players ensures higher total effort than the simultaneous contest with 24. The differences become even larger with larger contests. For example, a contest with 14 sequential players achieves higher total effort than a contest with 16,000 simultaneous players. Therefore, the information about other players' efforts is at least as important as other characteristics of the model, such as the number of players.

I also generalize the first-mover advantage result from Dixit (1987). Dixit showed that a player who can pre-commit chooses greater effort and obtains a higher payoff than the followers. The leader has two advantages: he moves earlier and has no direct competitors. With the new characterization, I can explore this idea further and compare players' payoffs and efforts in an arbitrary sequential contest. I show that there is a strict *earlier-mover advantage*—earlier players choose greater efforts and obtain higher payoffs than later players.

I provide additional results for large contests. Although the characterization result holds for any number of players, calculation of equilibrium becomes cumbersome with large contests, especially when the number of periods becomes large. I show that there is a convenient approximation method in which the equilibrium efforts can be directly computed using a simple formula. This allows me to study the rates of convergence under different disclosure policies. I show that the convergence is much faster in the case of sequential versus simultaneous contests.

My results have applications in various branches of economic literature, including oligopoly theory, contestability, rent-seeking, research and development, public goods provision, and parimutuel betting. For example, they provide a natural foundation for the contestability theory: if the moves are sequential, then a market could be highly concentrated but still be very close to competitive equilibrium. The early players produce most of the output and get most of the profits, but later players behave almost as an endogenous competitive fringe by being ready to produce more as soon as earlier players try to exploit their market power. Similarly, my results provide an explanation to a paradox in the rent-seeking literature: to explain rent dissipation with strategic agents, we need an unrealistically large number of players. I show that in a sequential rent-seeking contest it is sufficient to have only a small number of players to achieve almost full rent dissipation.

Literature: The simultaneous version of the model has been studied extensively in many branches of the literature, starting from Cournot (1838). I use the Tullock contest as my leading example. The literature on this type of contest was initiated by Tullock (1967, 1974) and motivated by rent-seeking (Krueger, 1974; Posner, 1975).[2] Bell et al. (1976) provided an axiomatization under which market shares are proportional to advertising spending so that this model could be directly applied to advertising. The most general treatment of simultaneous contests is provided by the literature on aggregative games (Jensen, 2010; Acemoglu and Jensen, 2013; Martimort and Stole, 2012).[3] My model is an aggregative game only in the simultaneous case.

Two-period contests have also been studied extensively, starting with von Stackelberg (1934), who studied quantity leadership in an oligopoly. In Tullock contests, the outcomes in a two-player sequential contest coincide with that of a simultaneous contest (Linster, 1993). With more than two players, a first mover has a strict advantage (Dixit, 1987).[4] Moreover, when the order of moves is endogenous, sequential contests arise in equilibrium in both two-player Tullock contests (Leininger, 1993; Morgan, 2003) and in oligopoly (Spence, 1977; Fudenberg and Tirole, 1985; Anderson and Engers, 1992; Amir and Grilo, 1999).

Relatively little is known about contests with more than two periods. The only paper I am aware of that has studied sequential Tullock contests with more than two periods is Glazer and Hassin (2000), who characterized the equilibrium in the sequential three-player Tullock contest. The only class of contests where equilibria are fully characterized for sequential contests are oligopolies with linear demand (Daughety, 1990; Ino and Matsumura, 2012; Julien et al., 2012). Linear demand implies quadratic payoffs; this is the only case where the first-order conditions are linear in all variables and therefore the

[2]See Tullock (2001); Nitzan (1994); Konrad (2009) for literature reviews on contests.

[3]See Jensen (2017) for a literature review on aggregative games.

[4]In asymmetric contests even the two-player sequential contest differs from the simultaneous contest (Morgan, 2003; Serena, 2017).

equilibrium is simple to characterize.[5]

More is known about large contests. Perfect competition (Marshall equilibrium) is a standard assumption in economics and it is important to understand its foundations. Novshek (1980) showed that although under some assumptions the Cournot equilibrium may not exist with a finite number of players, in large markets the Cournot equilibrium exists and converges to the Marshall equilibrium. Robson (1990) provided further foundations for Marshall equilibrium by proving an analogous result for large sequential oligopolies. In this paper, I take an alternative approach and under stronger assumptions about payoffs I provide a full characterization of equilibria with any number of players and any disclosure structure, which includes simultaneous and sequential contests as extremes. This allows me not only to show that the large contest limit is Marshall equilibrium but study the rates of convergence under any contest structures.

There are three somewhat distinct branches of literature on contests: Tullock contests, all-pay contests, and tournaments.[6] These three branches differ in terms of the contest success function, i.e., the criteria for allocating prizes. First, in Tullock contests the probabilities of receiving the prizes are proportional to efforts (Tullock, 1967, 1974, 2001). In the static framework, they typically give unique equilibria, which is in pure strategies. This model is often used to study rent-seeking, R&D races, advertising, and elections. My model extends this class of models to sequential settings.

The second class of contests includes all-pay auctions and wars of attrition, where the player with the highest effort always wins. These models are often used to study lobbying, military activities, and auctions. The equilibria in this class are typically in mixed strategies. Baye et al. (1996) characterize the equilibria in the static common-value (first-price) all-pay auction and Hendricks et al. (1988) in the second-price all-pay auction (also called the war of attrition). Siegel (2009) provides a general payoff characterization for the static all-pay contests. The third class of contests is rank-order tournaments in which prizes are allocated according to the highest output rather than the highest effort. Output is a noisy measure of effort. Tournaments were introduced by Lazear and Rosen (1981) and Rosen (1986) and are most often used to model principal-agent relationships and contract design in personnel economics and labor economics. My assumptions exclude all-pay contests and tournaments.

The paper also contributes to the contest design literature. Previous papers on contest design include Glazer and Hassin (1988), Taylor (1995), Che and Gale (2003), Moldovanu and Sela (2001, 2006), Olszewski and Siegel (2016), and Bimpikis et al. (2017), which

[5]However, linearity is a special assumption with perhaps undesirable consequences. For example, the leader's quantity in Stackelberg model with linear demand is independent of the number of followers (Anderson and Engers, 1992). See appendix H for more detailed discussion of contests with linear payoffs.

[6]For detailed literature reviews, see Konrad (2009) and Vojnovic (2015).

have focused on contests with private information. In this paper I study contest design on a different dimension: how to disclose other players' efforts optimally to minimize or maximize total effort.

The paper is organized as follows. Section 2 introduces the model and discusses the main assumptions. Section 3 discusses the difficulties with the standard solution method, describes the approach introduced in this paper, and provides the characterization result. Section 4 gives the second main result, the information theorem, which shows that the total effort is increasing in information in a general sense. It also discusses the corollaries of the result, including its implications for optimal contests. Section 5 studies earlier-mover advantage and section 6 discusses large contests. Section 7 analyses applications and policy implications of the results. Finally, section 8 concludes. All proofs are presented in the appendixes.

2 Model

Players and timing: There are n identical players $\mathcal{N} = \{1, \dots, n\}$ who arrive to the contest sequentially and make effort choices on arrival. At $T-1$ points in time, the sum of efforts by previous players is publicly disclosed.[7] These disclosures partition players into T groups, denoted by $\boldsymbol{\mathcal{I}} = (\mathcal{I}_1, \dots, \mathcal{I}_T)$. In particular, all players in $\mathcal{I}_1$ arrive before the first disclosure and therefore have no information about other players' efforts. All players in $\mathcal{I}_t$ arrive between disclosures $t-1$ and t and therefore have exactly the same information: they observe the total effort of players arriving prior to disclosure $t-1$. I refer to the time interval in which players in group $\mathcal{I}_t$ arrived as period t. As all players are identical, the disclosure rule of the contest is fully described by the vector $\mathbf{n} = (n_1, \dots, n_T)$, where $n_t = \#\mathcal{I}_t$ is the number of players arriving in period t.[8]

Efforts: I assume that each player i chooses an individual effort $x_i \geq 0$ at the time of arrival. I denote the profile of effort choices by $\mathbf{x} = (x_1, \dots, x_n)$, the total effort in the contest by $X = \sum_{i=1}^{n} x_i$, and the cumulative effort after period t by $X_t = \sum_{s=1}^{t} \sum_{i \in \mathcal{I}_s} x_i$. As the payoffs only depend on the sum of other players' efforts, it is sufficient to keep track of the cumulative effort, even if the players are able to observe individual efforts. By construction, the cumulative effort before the contest is $X_0 = 0$, and the cumulative

[7] The points of disclosure could be exogenous or chosen by a contest designer. In the following sections I first characterize the equilibria of arbitrary contests with fixed disclosures and then study the relationship between the equilibrium efforts and disclosures.

[8] Equivalently the model can be stated as follows: there are T time periods and n players are distributed among these periods and either exogenously or by the contest designer.

effort after period T is the total effort exerted during the contest, i.e. $X_T = X$. Figure 1 summarizes the notation with an example of a four-period contest.

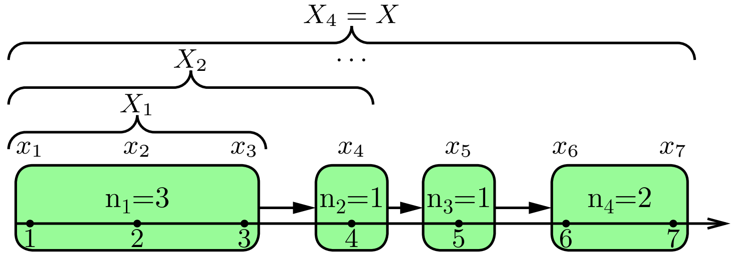

Figure 1: A contest with 7 players and 3 disclosures. Players 1 to 3 choose efforts x_1, x_2, and x_3 independently; player 4 observes $X_1 = x_1+x_2+x_3$, player 5 observes $X_2 = X_1+x_4$, and players 6 and 7 observe $X_3 = X_2 + x_5$.

Payoffs: As the leading example of this paper I use *normalized Tullock payoffs*, with

$$u_i(\mathbf{x}) = \frac{x_i}{X} - x_i. \tag{1}$$

This can be interpreted as a contest where players compete for prizes of total size one, the probability of winning is proportional to efforts, and the marginal cost of effort is one. Alternatively, this could be a model of an oligopoly with unit-elastic inverse demand function $P(X) = \frac{1}{X}$ and marginal cost $c = 1$, or a public goods provision (or tragedy of the commons) with the marginal benefit of private consumption $MB(X) = \frac{1-X}{X}$.

Note that all of the results apply to a more general class of contests than Tullock contests. As the proofs of the general results are analogous, it is useful to focus on the example with Tullock payoffs first. I discuss general payoff functions in section 7.

Equilibrium concept: I study pure-strategy subgame-perfect equilibria (SPE), which is a natural equilibrium concept in this setting: there is no private information and earlier arrivals can be interpreted as having greater commitment power. I show that there always exists a unique pure-strategy SPE.

Restrictive assumptions: Throughout this paper, I maintain a few assumptions that simplify the analysis. First, there is no private information. Second, the arrival times and the disclosure rules are fixed and common knowledge. Third, each player makes an effort choice just once—on arrival. Fourth, disclosures make cumulative efforts public. In section 8, I discuss the extent to which the results rely on each of these assumptions.

3 Characterization

3.1 The problem with standard backward induction

Before describing the solution, let me use a simple example to illustrate the difficulty of using the standard backward-induction approach. I will then describe the alternative approach I introduce in this paper.

The standard Tullock contest with identical players is a simple game. If n players make their choices in isolation, i.e., no information is revealed to them, then each player i chooses effort x_i simultaneously to maximize the payoff (1). The optimal efforts have to satisfy the first-order condition

$$\frac{1}{X} - \frac{x_i}{X^2} - 1 = 0. \tag{2}$$

Combining the optimality conditions leads to a total equilibrium effort $X^* = \frac{n-1}{n}$ and individual efforts $x_i^* = \frac{n-1}{n^2}$. The equilibrium is unique, in pure strategies, easy to compute, and easy to generalize in various directions, which may explain the widespread use of this model in various branches of economics.

Consider next a three-player version of the same contest, but the players arrive sequentially and their efforts are instantly publicly disclosed. That is, players 1, 2, and 3 make their choices x_1, x_2, and x_3 after observing the efforts of previous players. I will first try to find equilibria using the standard backward-induction approach.

Player 3 observes the total effort of the previous two players, $X_2 = x_1 + x_2 \in [0, 1)$[9] and maximizes the payoff (1), which in this case is

$$\max_{x_3 \geq 0} \frac{x_3}{X_2 + x_3} - x_3.$$

The optimality condition for player 3 is

$$\frac{1}{X_2 + x_3} - \frac{x_3}{(X_2 + x_3)^2} - 1 = 0 \quad \Rightarrow \quad x_3^*(X_2) = \sqrt{X_2} - X_2. \tag{3}$$

Now, player 2 observes $x_1 \in [0, 1)$ and knows $x_3^*(X_2)$, and therefore maximizes

$$\max_{x_2 \geq 0} \frac{x_2}{x_1 + x_2 + x_3^*(x_1 + x_2)} - x_2 = \max_{x_2 \geq 0} \frac{x_2}{\sqrt{x_1 + x_2}} - x_2.$$

[9] Players get non-positive payoffs in histories with $X_2 \geq 1$ or $x_1 \geq 1$, which cannot be an equilibrium.

The optimality condition is

$$\frac{1}{\sqrt{x_1+x_2}} - \frac{x_2}{2(x_1+x_2)^{\frac{3}{2}}} - 1 = 0.$$

For each $x_1 \in [0,1)$, this equation defines a unique best-response,

$$x_2^*(x_1) = \frac{1}{12} - x_1 + \frac{\left(8\sqrt{27x_1^3(27x_1+1)} + 216x_1^2 + 36x_1 + 1\right)^{\frac{2}{3}} + 24x_1 + 1}{12\left(8\sqrt{27x_1^3(27x_1+1)} + 216x_1^2 + 36x_1 + 1\right)^{\frac{1}{3}}}. \tag{4}$$

Finally, player 1's problem is

$$\max_{x_1 \geq 0} \frac{x_1}{x_1 + x_2^*(x_1) + x_3^*(x_1 + x_2^*(x_1))} - x_1,$$

where $x_2^*(x_1)$ and $x_3^*(X_2)$ are defined by equations (3) and (4). Although the problem is not complex, it is not tractable. Moreover, it is clear that the direct approach is not generalizable for an arbitrary number of players.[10] In fact, the best response function typically does not have an explicit representation for contests with a larger number of periods.[11]

3.2 Inverted best-response approach

In this paper I introduce a different approach. Instead of characterizing individual best-responses $x_i^*(X_{t-1})$ or the total efforts induced by X_{t-1}, denoted by $X^{*t-1}(X_{t-1})$, I characterize the inverse of this function. For any level of total effort X, the inverted best-response function $f_{t-1}(X)$ is the cumulative effort X_{t-1} prior to period t, that is consistent with total effort being X, given that the players in periods $t, \dots, T$ behave optimally.

To see how this characterization works, consider the three-player sequential contest again. In the last period, player 3 observes X_2 and chooses x_3. Equivalently we can think of his problem as choosing the total effort $X \geq X_2$ by setting $x_3 = X - X_2$. His maximization problem is

$$\max_{X \geq X_2} \frac{X - X_2}{X} - (X - X_2) \quad \Rightarrow \quad \frac{1}{X} - \frac{X - X_2}{X^2} - 1 = \frac{X_2}{X^2} - 1 = 0,$$

[10] Glazer and Hassin (2000) characterized the equilibrium in the contest with three sequential players. To my knowledge, no existing papers have characterized equilibria for sequential contests with more than three periods.

[11] The representation from theorem 1 implies that best-response functions include roots of higher-order polynomials that are not solvable (see appendix C for details). Therefore, the best-response functions $x_i^*(X_{t-1})$ cannot be expressed explicitly in terms of standard mathematical operations.

which implies $X_2 = X^2$. That is, if the total effort in the contest is X, then before the move of player 3, the cumulative effort had to be $f_2(X) = X^2$; otherwise, player 3 is not behaving optimally.

We can now think of player 2's problem as choosing $X \geq X_1$, which he can induce by making sure that the cumulative effort after his move is $X_2 = f_2(X)$, which means setting $x_2 = f_2(X) - X_1$. Therefore, his maximization problem can be written as

$$\max_{X \geq X_1} \frac{f_2(X) - X_1}{X} - (f_2(X) - X_1) \quad \Rightarrow \quad \frac{f_2'(X)}{X} - \frac{f_2(X) - X_1}{X^2} - f_2'(X) = 0. \tag{5}$$

This is the key equation to examine in order to see the advantages of the inverted best-response approach. Equation (5) is non-linear in X and therefore in x_2, which causes the difficulty for the standard backward-induction approach. Solving this equation every period for the best-response function leads to complex expressions, and the complexity accumulates with each step of the recursion. But equation (5) is linear in X_1, and it is therefore easy to derive the inverted best-response function

$$f_1(X) = X_1 = f_2(X) - f_2'(X)X(1 - X) = X^2(2X - 1).$$

Again, we know that if the total effort of the contest is X, then after player 1 the cumulative effort has to be $X_1 = f_1(X)$; otherwise, at least one of the later players is not behaving optimally.

Moreover, note that $X < \frac{1}{2}$ cannot be induced by any X_1, as even if $X_1 = 0$, the total effort chosen by players 2 and 3 would be $\frac{1}{2}$. Inducing total effort below $\frac{1}{2}$ would require player 1 to exert negative effort, which is not possible. Therefore, $f_1(X)$ is defined over the domain $\left[\frac{1}{2}, 1\right]$, and it is strictly increasing in this interval.

Finally, player 1 maximizes a similar problem of choosing $X \geq \frac{1}{2}$, which he can induce by setting $x_1 = f_1(X)$, to maximize

$$\max_{X \geq \frac{1}{2}} \frac{f_1(X)}{X} - f_1(X) \quad \Rightarrow \quad \frac{f_1'(X)}{X} - \frac{f_1(X)}{X^2} - f_1'(X) = 0,$$

which implies

$$0 = f_0(X) = f_1(X) - f_1'(X)X(1 - X) = X^2(6X^2 - 6X - 1). \tag{6}$$

Solving equation (6) gives three candidates for the total equilibrium effort X^*. It is either 0, $\frac{1}{2} - \frac{1}{2\sqrt{3}}$, or $\frac{1}{2} + \frac{1}{2\sqrt{3}}$. As we already found that in equilibrium $X^* \geq \frac{1}{2}$, only the highest root $\frac{1}{2} + \frac{1}{2\sqrt{3}} \approx 0.7887$ constitutes an equilibrium.[12]

The main advantage of the inverted best-response approach is that instead of finding the solution of a non-linear and increasingly complex equation, the approach allows for combining all of the first-order necessary conditions into one, which is then solved only once.

Note that in each period, there is a simple recursive dependence that determines how the inverted best-response function evolves. At the end of the contest, i.e., after period 3, the total effort is $f_3(X) = X$. In each of the previous periods it is equal to $f_{t-1}(X) = f_t(X) - f_t'(X)X(1-X)$. Extending the analysis from three sequential players to four or more sequential players is straightforward. It simply requires applying the same rule more times and solving the somewhat more complex equation at the end.

3.3 Characterization theorem

Theorem 1 formalizes this approach, generalizes the result for more general payoff functions, and allows for multiple players who make simultaneous decisions. It shows that each contest $\mathbf{n}$ has a unique equilibrium. The total equilibrium effort X^* is the highest root of $f_0(X)$, and the individual equilibrium effort of player i from period t can be computed directly as $x_i^* = \frac{1}{n_t}[f_t(X^*) - f_{t-1}(X^*)]$.

The characterization builds on the recursively defined inverted best-response functions $f_0, \ldots, f_T$, where $X_t = f_t(X)$ is the cumulative effort after period t, consistent with the total effort in the contest being X. The total effort after the last period T is $f_T(X) = X$, and the cumulative efforts in all previous periods are recursively defined by

$$f_{t-1}(X) = f_t(X) - n_t f_t'(X)g(X), \quad \forall X \in [\underline{X}_t, 1], \forall t = 1, \ldots, T, \tag{7}$$

where $\underline{X}_t$ is the highest root of $f_t(X) = 0$ in $[0,1]$ and $g(X) = X(1-X)$.

The recursive rule is similar to the example, but with two natural differences. First, it allows simultaneous decisions. When $n_t > 1$ players make simultaneous decisions in period t, then each of them has only a fractional impact on the best-responses of the following players, which means that the impact on the inverted best-response function is magnified by n_t. Second, in the proof of the theorem, I allow more general payoff functions, which are captured by function $g(X)$.[13]

The first part of the proof shows that the inverted best-response functions are well-behaved. As figure 2 illustrates, all of these functions take positive values 1 at $X = 1$,

[12] I did not check that the individual necessary conditions are also sufficient in this example, but in the characterization result in the next section, I verify this for arbitrary contests.

[13] I discuss sufficient conditions on general payoffs in section 7.

their highest roots $\underline{X}_t$ are always in $[0, 1)$ and decreasing in t (i.e., moving closer to 1 as t decreases). Moreover, each f_t is strictly increasing between its highest roots and 1, which means that in this relevant range it is invertible and therefore defines the best-response function uniquely, and f_t is strictly negative in the between the two consecutive highest roots, i.e. $[\underline{X}_{t+1}, \underline{X}_t)$, which excludes all the other candidates for equilibria. These properties are formalized as condition 1.

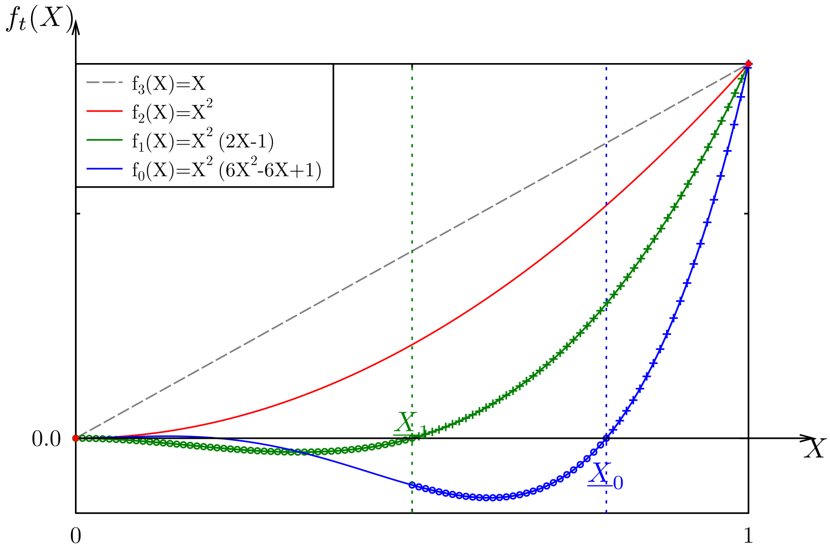

Figure 2: Illustration of condition 1 in three-player sequential Tullock contest. Example: for $t = 0$, we have (1) $\underline{X}_0 \approx 0.7887 > \underline{X}_1 = 0.5$, (2) $f_0(X) < 0$ for all $X \in [\underline{X}_1, \underline{X}_0)$ (line segment marked with circles), and (3) $f_0(X)$ is strictly increasing in $[\underline{X}_0, 1]$ (line segment marked with plusses). For calculations, see section 3.2.

Condition 1 (Inverted best responses are well-behaved)**.** *For all $t = 0, \ldots, T-1$, the function f_t has the following properties:*

1. *$f_t(X) = 0$ has a root in $[\underline{X}_{t+1}, 1]$.*[14] *Let $\underline{X}_t$ be the highest such root.*
2. *$f_t(X) < 0$ for all $X \in [\underline{X}_{t+1}, \underline{X}_t)$.*
3. *$f_t'(X) > 0$ for all $X \in [\underline{X}_t, 1]$.*

Moreover, $\underline{X}_0 \in (0, 1)$.

The following proposition 1 shows that condition 1 is satisfied in case of Tullock payoffs.

Proposition 1. *If $g(X) = X(1-X)$, then condition 1 is satisfied.*

[14]Function $f_T(X) = X$ has only one root $\underline{X}_T = 0$.

The second part of the proof shows that the properties highlighted in condition 1 are sufficient for the existence and uniqueness of the equilibrium, and the equilibrium is characterized as described above. The idea of the proof is analogous to the example in section 3.2.

Theorem 1 (Characterization theorem)**.** *Suppose condition 1 holds. Each contest* $\mathbf{n}$ *has a unique equilibrium. The equilibrium strategy of player i in period t is*[15]

$$x_i^*(X_{t-1}) = \begin{cases} \frac{1}{n_t}\left[f_t(f_{t-1}^{-1}(X_{t-1})) - X_{t-1}\right] & \forall X_{t-1} < 1, \\ 0 & \forall X_{t-1} \geq 1. \end{cases} \tag{8}$$

In particular, the total equilibrium effort is $X^* = \underline{X}_0$*, i.e., the highest root of* $f_0(X) = 0$*, and the equilibrium effort of player* $i \in \mathcal{I}_t$ *is* $x_i^* = \frac{1}{n_t}\left[f_t(X^*) - f_{t-1}(X^*)\right]$*.*

The characterization theorem provides a straightforward method for computing the equilibria in any contest $\mathbf{n}$. Appendix K provides several examples demonstrating how it can be applied in the case of Tullock payoffs and other payoff functions.

4 Information and effort

In this section, I present the second of the two main result of this paper. I show that the information increases total effort in contests. The result has strong implications for both the comparative statics and for optimal contests.

Before giving the formal result, I need to introduce the notation to keep track of the relevant information in contest $\mathbf{n}$. Let me use a contest $\mathbf{n} = (1, 2, 1)$ to illustrate the construction. This contest has four players: a first-mover, two simultaneous followers, and a last-mover. At the most basic level, all players observe their own efforts (regardless of the disclosure rule). There are $n = 4$ observations of this kind and they clearly affect the outcomes of the contest. I call this the first level of information and denote as $S_1 = 4$. More importantly, some players directly observe the efforts of some other players. In the example, players 2 and 3 observe the effort of player 1 and player 4 observes the efforts of all three previous players. Therefore, there are five direct observations of other players' efforts. I call this the second level of information and denote as $S_2 = 5$. Finally, player 4 observes players 2 and 3 observing player 1. There are two indirect observations of this kind, which I call the third level of information and denote as $S_3 = 2$. In contests with

[15]Function $f_t^{-1}(X_t)$ denotes the inverse function of $f_t(X)$ in the interval $[\underline{X}_t, 1]$, where $\underline{X}_t$ is the highest root of f_t in $[0, 1]$. By condition 1, $f_t(X)$ is strictly increasing in this interval.

more periods, there would be more levels of information—observations of observations of observations and so on.

I call a vector $\mathbf{S}(\mathbf{n}) = (S_1(\mathbf{n}), \ldots, S_T(\mathbf{n}))$ the measure of information in a contest $\mathbf{n}$. In the example described above, $\mathbf{S}((1,2,1)) = (4,5,2)$. Formally, $S_k(\mathbf{n})$ is the sum of all products of k-combinations of set $\{n_1, \ldots, n_T\}$.[16] Note that we can also think of $\mathbf{S}(\mathbf{n})$ as an infinite sequence with $S_k = 0$ for all $k > T$.

Theorem 2 shows that the total effort is strictly increasing in vector $\mathbf{S}(\mathbf{n})$. The key step in proving this is showing that the inverted best-response function $f_0(X)$ can be expressed as

$$f_0(X) = X - \sum_{k=1}^{T} S_k(\mathbf{n}) g_k(X), \tag{9}$$

where $g_1, \ldots, g_T$ are recursively defined as $g_1(X) = g(X) = X(1-X)$ and $g_{k+1}(X) = -g_k'(X)g(X)$. Each $g_k(X)$ is independent of $\mathbf{n}$ and clearly $S_k(\mathbf{n})$ is independent of X. As the total equilibrium effort, X^* is the highest root of f_0; it therefore depends on $\mathbf{n}$ only through information measures $\mathbf{S}(\mathbf{n})$.

Each g_k describes the substitutability between efforts. In particular, if $g_2(X^*) > 0$, then the efforts are strategic substitutes near equilibrium in the standard sense, i.e., the effort of an earlier player i discourages the effort of a later j player who observes player i. If it is also the case that $g_3(X^*) > 0$, then the indirect impact through observations of observations also leads to discouragement, i.e., not only does player i discourage a later player l directly, but if this player l moves after player j, then the indirect effect of player i's effort through player j on player l is discouraging. Similarly $g_k(X^*) > 0$ describes the discouragement effect through k-th level of information.[17]

In particular, a sufficient condition that guarantees X^* is strictly increasing in $\mathbf{S}(\mathbf{n})$ is that each $g_k(X^*) > 0$, i.e., the efforts are *higher-order strategic substitutes* near equilibrium, which is formalized by condition 2.

Condition 2 (Higher-order strategic substitutes)**.** $g_k(X^*) > 0$ *for all* $k = 2, \ldots, T$.

In the case of Tullock payoffs, $g(X) = X(1-X)$, which means that with low total effort $X < \frac{1}{2}$, the efforts are strategic complements; with high total effort $X > \frac{1}{2}$, the efforts are strategic complements; and in the knife-edge case $X = \frac{1}{2}$, they are independent. This is why information about the first mover's effort in the two-player Tullock contest does not change the equilibrium efforts. Indeed, in both $\mathbf{n} = (2)$ and $\hat{\mathbf{n}} = (1,1)$ the total effort is $\frac{1}{2}$ and the individual efforts are $\frac{1}{4}$. I show that for any other contests the total equilibrium

[16] For example, in a sequential n-player contest $\mathbf{n} = (1, 1, \ldots, 1)$, $S_k(\mathbf{n})$ is simply the number of all k-combinations, i.e., $S_k(\mathbf{n}) = \frac{n!}{k!(n-k)!}$.

[17] My assumptions also guarantee that $g_1(X^*) = g(X^*) > 0$, i.e., individual effort discourages one's own effort, which is a standard concavity assumption that guarantees the interior optimum.

effort is strictly higher than $\frac{1}{2}$ and therefore, the efforts are strategic substitutes in the standard sense.

As the following proposition shows, this observation generalizes to n-player contests. The efforts are higher-order strategic substitutes for any contest except in the fully sequential case, where the highest $g_n(X^*) = 0$. However, if there are at least three players, this exception does not affect the conclusions because the fully sequential contest also has strictly higher lower-level information measures than any other contest. For example, when $n = 3$, $\mathbf{S}((1,1,1)) = (3,3,1)$, $\mathbf{S}((2,1)) = \mathbf{S}((1,2)) = (3,2)$ and $\mathbf{S}((3)) = (3)$, i.e., contest $(1,1,1)$ has a strictly higher S_2 than any other three-player contest.

Proposition 2. *Suppose $g(X) = X(1-X)$. Then*

1. *if $\mathbf{n} = (1,1,\dots,1)$ and $k = T = n$, then $g_k(X^*) = 0$; otherwise,*

2. *if $\mathbf{n} \neq (1,\dots,1)$ or $k < T$, then $g_k(X^*) > 0$.*

Now, if condition 2 holds, then the increase[18] in $\mathbf{S}(\mathbf{n})$ would lower $f_0(X^*)$ near the original equilibrium and therefore, due to condition 1, the total equilibrium effort must increase. These arguments jointly imply the second main result of the paper, theorem 2. It shows that the total equilibrium effort of any contest $\mathbf{n}$ only depends on the information $\mathbf{S}(\mathbf{n})$ in the contest and is strictly increasing in each level of information.

Theorem 2 (Information theorem)**.** *Suppose conditions 1 and 2 hold. Total effort is a strictly increasing function $X^*(\mathbf{S}(\mathbf{n}))$.*

Therefore, information in contests defines a partial order over contests—if $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$ then $X^*(\mathbf{S}(\widehat{\mathbf{n}})) > X^*(\mathbf{S}(\mathbf{n}))$. Theorem 2 has many direct implications. I highlight some of the most important ones in the following corollaries. Note that these corollaries hold whenever conditions 1 and 2 are satisfied, which includes all contests with Tullock payoffs and at least three players.

Corollary 1. *Under conditions 1 and 2*

1. *Comparative statics of $\mathbf{n}$: if $\widehat{\mathbf{n}} > \mathbf{n}$[19], then $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$ and therefore $X^*(\mathbf{S}(\widehat{\mathbf{n}}))) > X^*(\mathbf{S}(\mathbf{n}))$.*

2. *Independence of permutations: if $\widehat{\mathbf{n}}$ is a permutation of $\mathbf{n}$, then $\mathbf{S}(\widehat{\mathbf{n}}) = \mathbf{S}(\mathbf{n})$ and therefore $X^*(\mathbf{S}(\widehat{\mathbf{n}})) = X^*(\mathbf{S}(\mathbf{n}))$.*

3. *Informativeness increases total effort: if $\widehat{\boldsymbol{\mathcal{I}}}$ is a finer partition than $\boldsymbol{\mathcal{I}}$, then $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$ and therefore $X^*(\mathbf{S}(\widehat{\mathbf{n}})) > X^*(\mathbf{S}(\mathbf{n}))$.*

[18] $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$ means that $S_k(\widehat{\mathbf{n}}) \geq S_k(\mathbf{n})$ for all k and $S_k(\widehat{\mathbf{n}}) > S_k(\mathbf{n})$ for at least one k.

[19] Including when $n_t = 0 < \widehat{n}_t$, i.e., $\widehat{\mathbf{n}}$ has more periods with strictly positive number of players.

4. *Homogeneity increases total effort: if* $\widehat{n} = n$ *and there exist* t, t' *such that* $\widehat{n}_t \widehat{n}_{t'} > n_t n_{t'}$ *and* $\widehat{n}_s = n_s$ *for all* $s \neq t, t'$*, then* $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$ *and therefore* $X^*(\mathbf{S}(\widehat{\mathbf{n}})) > X^*(\mathbf{S}(\mathbf{n}))$.

5. *Full dissipation in large contests:* $\lim_{n\to\infty} X^*(\mathbf{S}(\mathbf{n})) = 1$.

The first implication is natural—if the number of players increases in any particular period or a period with a positive number of players is added, then the total effort increases. Note that this does not mean that the total effort is a strictly increasing function of the number of players. For example, as we saw above, the three-player sequential Tullock contest gives total effort 0.7887, whereas the four-player simultaneous Tullock contest gives total effort $\frac{4-1}{4} = 0.75$. The conclusion only holds when we add players while keeping the positions of other players unchanged, because this increases information.

The second implication is perhaps much more surprising—reallocating disclosures in a way that creates a permutation of $\mathbf{n}$ (or equivalently, reordering periods together with the corresponding players) does not affect the total effort. For example, a contest with a first mover and $n-1$ followers leads to the same total effort as a contest with $n-1$ first movers and one last mover. The result comes from the property that all observations of the same level have the same impact on the total effort; i.e., it does not matter whether k players observe one player or one player observes k players.

The third implication is perhaps the most important in terms of its consequences. It shows that disclosures strictly increase total effort. In particular, adding disclosures makes the contest strictly more informative: all players observe everything that they observed before, but some observe the efforts of more players. Formally, the new contest is a finer partition of players than the old contest.

To see the intuition of this result, consider contests $\mathbf{n} = (1, 2, 1)$ and $\widehat{\mathbf{n}} = (1, 1, 1, 1)$. Now, player 2's effort is made visible for player 3. Therefore, in addition to all marginal costs and benefits of effort, player 2 has an additional benefit of effort: as efforts are strategic substitutes, exerting more effort discourages player 3. This discouragement effect leads to added effort by player 2 and reduced effort by player 3, and the remaining question is how these effects compare. Player 2's payoff is $u_2(\mathbf{x}) = x_2 h(X)$, where $h(X) = \frac{1}{X} - 1$ is a strictly decreasing function of total effort, which means that if he could increase his own effort without increasing the total effort, he would certainly do so and it would not be an equilibrium. Therefore, near equilibrium we would expect the discouragement effect to be less than one-to-one, i.e., player 2 exerts more effort and player 3 exerts less but the total increases. Of course, this argument works only for small changes, keeping the efforts of indirectly affected players 1 and 4 unchanged. However, since the efforts are strategic substitutes of higher order, the indirect effects have the same signs and therefore the result still holds.

The fourth implication gives even clearer implications for the optimal contest. Namely, more homogeneous contests give higher total effort. Intuitively, a contest is more homogeneous if its disclosures are spread out more evenly (or equivalently, players are divided more evenly across periods). For example, a contest $\widehat{\mathbf{n}} = (2,2)$ is more homogeneous than $\mathbf{n} = (1,3)$. It has also more direct observations as $2 \times 2 = 4 > 3 = 1 \times 3$. I define the more homogeneous contest as one which can be achieved by pairwise increases of products of group sizes while keeping everything else fixed. Therefore, by construction, it increases S_2 strictly and all other S_k measures weakly.

The final implication is that in large contests rents are fully dissipated. This result follows from the fact that simultaneous contests are the least informative and give total effort $\frac{n-1}{n}$, which converges to 1 as $n \to \infty$. Therefore, the total effort from any contest converges to 1.[20]

These results give strong implications for the contest design, which I summarize in the following corollary.

Corollary 2. *Assume that conditions 1 and 2 hold and fix n. A simultaneous contest $\mathbf{n} = (n)$ minimizes the total effort, and a fully sequential contest $\mathbf{n} = (1,1,\ldots,1)$ maximizes the total effort. Moreover, if the contest designer can only make a fixed number of disclosures[21], then contests that allocate players into groups that are as equal as possible maximize the total effort.*

Therefore, if the goal is to minimize the total effort (such as in rent-seeking contests), then the optimal policy is to minimize the available information, which is achieved by a simultaneous contest. Transparency gives earlier players incentives to increase efforts to discourage later players, but this discouragement effect is less than one-to-one and therefore increases total effort.

On the other hand, if the goal is to maximize the total effort (such as in research and development), then the optimal contest is fully sequential as it maximizes the incentives to increase efforts through this discouragement effect. If the number of possible disclosures is limited (for example, collecting or announcing information is costly), then it is better to spread the disclosures as evenly as possible.[22]

Vector comparison $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$ defines a partial order over contests. To complete the order, we would have to know how to weigh different measures of information. Equa-

[20] This limit result is known in simultaneous contests since Novshek (1980) and (fully) sequential contests since Robson (1990).

[21] Or equivalently, there is a fixed number of periods.

[22] It is easy to see that parts 2 and 4 of corollary 1 imply that the contest that maximizes total effort with $T-1$ disclosures (T periods) and n players is such that there are $\widehat{T} = Tk-n$ periods with $n_t = \lceil n/T \rceil$ players and $T - \widehat{T}$ periods with $n_t = \lfloor n/T \rfloor$ players.

tion (9) provides a clue: the correct weights are $g_k(X^*)$; i.e., by magnitudes of discouragement effects near equilibrium. Unfortunately, these weights are endogenous and generally depend on a specific contest, but there are a few cases where we can say more.

I argue in section 6 that if the number of players is large, then $g_k(X^*)$ is approximately $\alpha^k(1-X^*)$, with $\alpha = -g'(1) = 1$ in Tullock payoffs case. Therefore, in large sequential Tullock contests all measures of information are produced approximately equally, which gives a complete order on contests: $X^*(\mathbf{S}(\widehat{\mathbf{n}})) > X^*(\mathbf{S}(\mathbf{n}))$ if and only if $\sum_k S_k(\widehat{\mathbf{n}}) > \sum_k S_k(\mathbf{n})$.

However, with a smaller number of players, the following lemma shows that lower information measures have a higher weight.

Lemma 1. *For $g(X) = X(1-X)$, for each $k \geq 2$, $g_{k-1}(X^*) > g_k(X^*)$.*

For example, let us compare two 10-player contests $\mathbf{n} = (5,5)$ and $\widehat{\mathbf{n}} = (8,1,1)$. Then $\mathbf{S}(\mathbf{n}) = (10,25)$ and $\mathbf{S}(\widehat{\mathbf{n}}) = (10,17,8)$. The two contests cannot be ranked according to the information measures, because the first contest $\mathbf{n}$ has more second-order information, whereas the second contest $\widehat{\mathbf{n}}$ has one more disclosure and thus more third-order information. However, the sum of all information measures is $10 + 25 = 10 + 17 + 8 = 35$. Since the weights are higher in lower-order information, this implies that the total effort is higher in the first contest. Indeed, direct application of the characterization theorem confirms this, as $X^* = \frac{13+\sqrt{41}}{20} \approx 0.9702 > \widehat{X}^* = \frac{31+\sqrt{241}}{48} \approx 0.9693$.

5 Earlier-mover advantage

In this section, I revisit Dixit's first-mover advantage result. Dixit showed that in a contest with at least three players, if one player can pre-commit, this first mover chooses a strictly higher effort and achieves a strictly higher payoff than the followers. Using the tools developed in this paper, I can explore this idea further. Namely, in Dixit's model, the first mover has two advantages compared to the followers. First, he moves earlier, and his action may impact the followers. Second, he does not have any direct competitors in the same period.

I can now distinguish these two aspects. For example, what would happen if $n-1$ players chose simultaneously first and the remaining player chose after observing their efforts? Or more generally, in an arbitrary sequence of players, which players choose the highest efforts and which ones get the highest payoffs? The answer to all such questions turns out to be unambiguous. As proposition 3 shows, it is always better to move earlier.

Proposition 3 (Earlier-mover advantage)**.** *Suppose conditions 1 and 2 hold. The efforts and payoffs of earlier*[23] *players are strictly higher than for later players.*

Let us consider payoff comparisons first. Note that the equilibrium payoff of a player $i \in \mathcal{I}_t$ is in the form $u_i(\mathbf{x}^*) = x_i^* h(X^*)$, and since X^* is the same for all the players, payoffs are proportional to efforts. Therefore, it suffices to show that the efforts of earlier players are strictly higher. In the proof, I show that we can express the difference between the efforts of players i and j from consecutive periods t and $t+1$ as

$$x_i^* - x_j^* = \sum_{k=1}^{T-t} [S_k(\mathbf{n}^t) - S_k(\mathbf{n}^{t+1})] g_{k+1}(X^*), \tag{10}$$

where $\mathbf{n}^{t+1} = (n_{t+2}, \ldots, n_T)$ is the sub-contest starting after period $t+1$ and $\mathbf{n}^t = (n_{t+1}, \mathbf{n}^{t+1})$ is the sub-contest starting after period t. Clearly, $S_k(\mathbf{n}^t) > S_k(\mathbf{n}^{t+1})$ for each $k = 1, \ldots, T-t$; i.e., there is more information on all levels in a strictly longer contest. By condition 2, $g_{k+1}(X^*) > 0$ for each k as well, and therefore the whole sum is strictly positive.

The intuition of the result is straightforward: players in earlier periods are observed by strictly more followers than the players from the later periods. Therefore, in addition to the incentives that later players have, the earlier players have additional incentives to exert more effort to discourage later players.

6 Large contests

In this section, I show that as the number of players becomes large, the total effort X^* converges to 1[24] much faster in sequential contests than in simultaneous contests. Figure 3 illustrates that although all contest types X^* converge to 1, the rate of convergence depends significantly on the type of contest.

To compare the rates of convergence formally I need to discuss how to compute equilibria in large contests. Although the characterization theorem holds for an arbitrary contest, since $f_0(X)$ is a polynomial of degree $T+1$, its highest root may sometimes be difficult to compute. In appendix H I show that the total equilibrium effort and individual

[23] By earlier I mean players who belong to the strictly earlier group, i.e., player $i \in \mathcal{I}_t$ is earlier than $j \in \mathcal{I}_s$ if and only if $s < t$.

[24] Corollary 1 showed that the total effort converges to 1 in any contest.

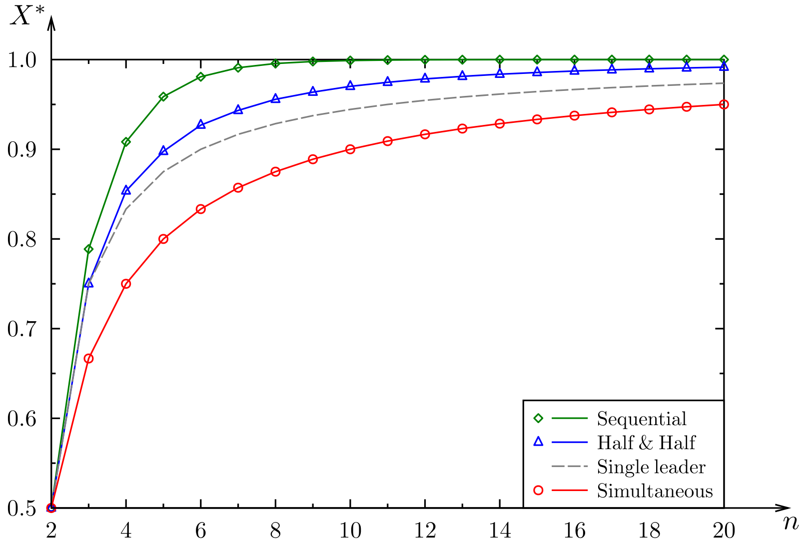

Figure 3: The total equilibrium effort in different contests with Tullock payoffs: fully sequential $\mathbf{n}^n = (1, \ldots, 1)$, half & half $\mathbf{n}^n = (\lceil n/2 \rceil, \lfloor n/2 \rfloor)$, single leader $\mathbf{n}^n = (1, n-1)$, and simultaneous $\mathbf{n}^n = (n)$.

equilibrium efforts with a large number of players can be approximated by[25]

$$1 - X^* \approx \frac{1}{S(\mathbf{n})} \approx \frac{1}{\prod_{t=1}^{T}(1+n_t)}, \quad x_i^* \approx \frac{1}{\prod_{s=1}^{t}(1+n_s)}, \forall i \in \mathcal{I}_t, \forall t = 1, \ldots, T. \tag{11}$$

That is, the total equilibrium effort X^* converges to a strictly increasing function of $S(\mathbf{n}) = \prod_{t=1}^{T}(1+n_t) - 1 = \sum_{k=1}^{T} S_k(\mathbf{n})$. Therefore, in large contests, information of all degrees carries equal weight.

The reason for this result is simple. As the payoff function is smooth, it can be closely approximated by a linear function. More precisely, payoffs are represented by function $g(X) = X(1-X)$. In large contests, the total effort $X^* \approx 1$, and therefore near X^* we have that $g_1(X) = g(X) \approx 1-X$ and $-g_1'(X) \approx 1$. This gives us $g_2(X) = -g_1'(X)g(X) \approx 1-X$ and so on; i.e., each $g_k(X) \approx 1 - X$ near $X^* \approx 1$. Equation (9) can be approximated by

$$f_0(X^*) \approx 1 - (1-X^*)S(\mathbf{n}) = 0 \quad \Rightarrow 1 - X^* \approx \frac{1}{S(\mathbf{n})}.$$

From (11) we can make a few observations. As already argued, all measures of information carry approximately equal weight in determining the total equilibrium effort. Moreover, from (11) we see that the convergence $1-X^* \to 0$ is with rate $S(\mathbf{n})$. With simultaneous contests the sum of all levels of information is simply $S(\mathbf{n}) = S_1(\mathbf{n}) = n$, so that the convergence is linear. On the other extreme, with sequential contests $\mathbf{n} = (1, \ldots, 1)$,

[25]In case of general payoff functions, if $\alpha = -g'(1) \neq 1$, then these formulas are adjusted by a constant α. See appendix H for details.

we get that $S(\mathbf{n}) = \prod_{t=1}^{n}(1+1) - 1 = 2^n - 1$, so that the convergence is exponential. When there are at most T periods, the rate of convergence is bounded by n^T by the same argument.

These results highlight the importance of information in contests. The information provided is at least as important as the number of players in determining the total effort in contests. For example, the total effort in the simultaneous contest with ten players is 0.9, whereas the total effort with four sequential players is 0.9082. Adding a fifth sequential player increases the total effort to 0.9587, and to achieve this in a simultaneous contest with the same payoffs, we need 24 players. This comparison becomes even more favorable for the sequential contests with a large number of players, as figure 4 illustrates.

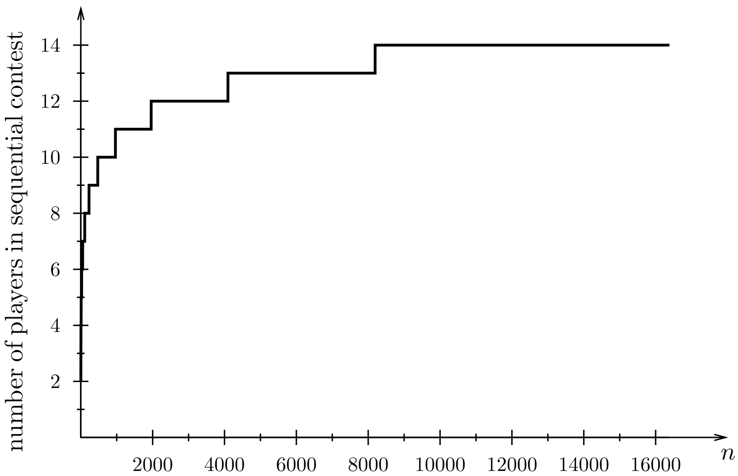

Figure 4: Number of players in a sequential Tullock contest that leads to the same total effort as a simultaneous Tullock contest with n players

For large contests, we can use the approximation from equation (11) to get a relationship between the number of players in simultaneous contests n^{sim} and the number of players in sequential contests n^{seq} with approximately the same total revenue:

$$\frac{1}{1+n^{sim}-1} = \frac{1}{n^{sim}} \approx 1 - X^* \approx \frac{1}{(1+1)^{n^{seq}}} \quad \Rightarrow \quad n^{sim} \approx 2^{n^{seq}}.$$

7 Applications

In this section, I discuss some implications of the results in the various branches of economic literature. To apply the results to a wider range of applications, it is useful to note that all the results presented in the earlier sections apply more generally than Tullock payoffs. The approach and the results generalize to payoff functions in the form

$$u_i(\mathbf{x}) = x_i h(X), \tag{12}$$

where $h(X)$ is a strictly decreasing function that satisfies sufficient conditions 1 and 2. In particular, $g(X) = -\frac{h(X)}{h'(X)}$ and $\alpha = -g'(1) > 0$. Appendix I discusses the sufficient conditions in detail and describes a class of payoff functions, where the conditions are satisfied.[26] Appendix K provides some examples of specific functional forms.

7.1 Oligopolies

In case of an oligopoly with a homogeneous product, the profit function of firm i is

$$u_i(\mathbf{x}) = x_i P(X) - cx_i = x_i h(X),$$

where x_i is firm's own quantity, X the total quantity, $P(X)$ the inverse demand function, and c the marginal cost of production, which is assumed to be constant and equal for all firms. The simultaneous contest $\mathbf{n} = (n)$ is the standard Cournot oligopoly [27] and $\mathbf{n} = (1, n-1)$ is the Stackelberg quantity-leader model. In general, the model is a hierarchical oligopoly model, where firms in earlier periods have more market power.

By theorem 1, the equilibrium exists and is unique, and the total equilibrium quantity is always below $\overline{X}$.[28] Proposition 3 shows that earlier firms produce strictly more and earn strictly higher profits than the later firms. Corollary 2 shows that the market structure that maximizes total quantity is fully sequential and the market structure that minimizes is simultaneous (Cournot).

The social planner maximizes the total surplus

$$S(X) = \int_0^X (P(t) - c)dt = \int_0^X h(t)dt. \tag{13}$$

The surplus is maximized at $\overline{X}$ and is strictly increasing in X below that. By theorem 2 the total equilibrium quantity is strictly increasing in information in this range; therefore, the socially optimal oligopoly would be fully sequential.[29]

For the firms, this outcome is not as appealing. The total profit of all firms is

$$U(X) = \sum_{i=1}^{n} u_i(\mathbf{x}) = \sum_{i=1}^{n} x_i(P(X) - c) = Xh(X). \tag{14}$$

[26] A program that verifies condition 1 and, if it is satisfied, computes the equilibrium for any contest is available at `http://toomas.hinnosaar.net/contests/`.

[27] In the case of Tullock payoffs, it is also known as the Cournot-Puu oligopoly after Puu (1991).

[28] The threshold $\overline{X}$ is determined by as $h(\overline{X}) = 0$ or equivalently $P(\overline{X}) = c$, so it can be interpreted as the Marshall equilibrium quantity. It can be normalized to $\overline{X} = 1$ without loss of generality.

[29] This implication has been studied in an oligopoly with linear demand and constant marginal costs by Daughety (1990).

The joint profit maximization would lead to optimality condition

$$\frac{d\pi(X)}{dX} = h(X) + Xh'(X) = h'(X)\left[X - g(X)\right] = 0, \tag{15}$$

where $g(X) = -\frac{h(X)}{h'(X)}$; so the monopoly's optimal total quantity satisfies $X^m = g(X^m)$. The joint profit is strictly decreasing in $X > X^m$. Equation (9) and condition 2 imply that the total equilibrium quantity $X^* > X^m$ regardless of market structure, and therefore, a collusive agreement would choose a market structure that minimizes X^*, which is the simultaneous contest (Cournot).

In practice, we may think of quantity competition as capacity competition.[30] The results in this paper imply that credible commitments to large capacities may lead to larger total capacity and thus to lower prices and higher welfare. Pre-commitments to large capacities allow discouraging capacity investments by the later movers. Since this discouragement effect is less than one-to-one, i.e., increased earlier-mover capacity is not fully canceled out through capacity reduction by later movers, the pre-commitments increase total capacity. Kalyanaram et al. (1995) survey the widespread empirical evidence that a negative relationship exists between brands' entry to the market and market share. The negative relationship holds in many mature markets, including pharmaceutical products, investment banks, semiconductors, and drilling rigs. For example, Bronnenberg et al. (2009) studied brands of typical consumer packaged goods and found a significant early entry advantage. The advantage is strong enough to drive the rank order of market shares in most cities.

7.2 Contestability

The theory of contestability (Baumol, 1982; Baumol et al., 1988) is widely used in practice; it postulates that with frictionless reversible entry, a market may be concentrated and contestable at the same time. In particular, if firms have no entry and exit barriers, no sunk costs, and access to the same technology, then the market is contestable. In a contestable market, firms operate at a zero-profit level, regardless of the number of incumbent firms in the market. The reason is that when they would try to use their dominance to extract rents, the competitive fringe would enter for a short period, undercut them, and capture the extra rents. Therefore, the theory postulates that the threat of entry works as a disciplining device, and that firms operate at zero profits.

This theory has been widely used, but also criticized (Brock, 1983; Shepherd, 1984; Dasgupta and Stiglitz, 1988; Gilbert, 1989), partly because it is sensitive to even small

[30] This idea was formalized by Kreps and Scheinkman (1983).

sunk costs and partly because it requires the existence of a competitive fringe that is not observed in equilibrium.

The results in this paper provide a natural foundation for contestability. As a benchmark, a competitive equilibrium would mean a large number of identical players who each choose low quantities. In contrast, in a fully sequential contest, we need a much lower number of firms to achieve the same total quantity and the individual behavior is very different. The first firm chooses a quantity which is about $\frac{\alpha}{1+\alpha}$ of the total quantity.[31] The second firm chooses $\frac{\alpha}{(1+\alpha)^2}$ and so on. Therefore, the market is highly concentrated[32] and the first few firms have significant market power which they use to achieve higher profits than the followers. However, the market is quite competitive and this competitiveness comes from the later players. They produce very little in equilibrium, but as soon as earlier players deviate to lower quantities, the later players respond with higher quantities. Therefore, they act as a competitive fringe but in a standard oligopoly model. In this paper, I do not include fixed costs and entry decisions, but it would be straightforward to extend the analysis in this direction and endogenize the number of players.[33]

7.3 Rent dissipation in rent-seeking

Early papers by Tullock (1967), Krueger (1974), and Posner (1975) proposed that competitive rent-seeking leads to full dissipation of rents. However, strategic modeling of rent-seeking (Tullock, 1974, 2001) showed that strategic behavior leads to rent underdissipation.

As the results in this paper show, the significant underdissipation result may be an implication of mainly focusing on simultaneous rent-seeking contests. As discussed in section 6, it suffices to have 5–10 players in a sequential rent-seeking contest to achieve outcomes that are very close to full dissipation, whereas in a simultaneous contest the same outcomes are achieved with a very high number of players.[34]

7.4 Research and development

Closely related models have been used to study patent races. The classic models, such as Loury (1979) and Dasgupta and Stiglitz (1980), assume that n firms compete to innovate.

[31] In the Tullock payoffs case $\alpha = -g'(1) = 1$, so the first player's quantity is about $\frac{1}{2}$ of the total.

[32] Herfindahl-Hirschman Index is $HHI \approx \frac{\alpha^2}{(1+\alpha)^2-1}$. With Tullock payoffs, $HHI \approx \frac{1}{3}$. For any $\alpha \geq \frac{1}{2}$, $HHI \geq \frac{1}{5}$, which is a highly concentrated market. In contrast, in a simultaneous contest, $HHI \approx 0$.

[33] The observation that high concentration does not necessarily coincide with low competitiveness has been made by Demsetz (1968), Daughety (1990), Ino and Matsumura (2012) among others.

[34] According to Murphy et al. (1993), rent-seeking activities exhibit increasing returns, which means that the differences between simultaneous and sequential behavior may be magnified even further.

Each firm chooses lump-sum investment x_i, and its probability of making a discovery on or before time t is $1 - e^{-\rho(x_i)t}$ with some hazard $\rho(x)$ that is constant over time and increasing and concave function of investment. The probability that a firm makes a discovery at time t is therefore $\sum_{j=1}^{n} \rho(x_j) e^{-\sum_{j=1}^{n} \rho(x_j)t}$, and the conditional probability that it is the firm i who makes the discovery is $\frac{\rho(x_i)}{\sum_{j=1}^{n} \rho(x_j)}$. The first firm that innovates gets a patent with value v, discounted at rate r, and the payoff of the other firms is normalized to 0. Therefore, the expected payoff of firm i is

$$\pi_i(\mathbf{x}) = \int_0^\infty v\rho(x_i)e^{-\sum_{j=1}^{n} \rho(x_j)t} e^{-rt} dt - x_i. \tag{16}$$

If the hazard rate is a linear function of x, i.e., $\rho(x) = ax$ and firms are very patient, i.e., $r \to \infty$, we get Tullock payoffs

$$\pi_i(\mathbf{x}) = v \frac{\rho(x_i)}{\sum_{j=1}^{n} \rho(x_j)} - x_i = v \frac{x_i}{\sum_{j=1}^{n} x_j} - x_i. \tag{17}$$

Therefore, the patent race game is strategically equivalent to Tullock contest with patient firms.[35]

It is natural to ask what changes if instead of making the investment decisions simultaneously, some firms observe the investments made by some of their opponents.[36] Suppose that the firms enter the market over time (in deterministic order) and at some (again deterministic) points of time the current total investments are made public. Then, if the firms are very patient, these initial timing changes do not affect the expected payoffs, and the expected payoff of firm i is still given by (17). However, the earlier firms have a strategic impact on the investments of later firms. We can apply the results from this paper to characterize equilibrium investments. Moreover, we know that the total investments to R&D are strictly increasing in information provided to the firms and that earlier firms choose higher investments and expect higher profits.

There is some empirical evidence of behavior along these lines. For example, in case of new product introductions, Dranove and Gandal (2003) provided evidence that product pre-announcements had a significant impact on the outcomes of DVD standards.

[35] This equivalence was first proven by Baye and Hoppe (2003).

[36] Dasgupta and Stiglitz (1980) discuss a similar extension, but with full commitment and ability to revise decisions. In this case, the first firm to enter is a monopolist and all other firms stay inactive. In the event they choose a positive investment, the monopolist is committed to increasing investment to a very high level, which deters any additional entry. They do not discuss subgame-perfect Nash equilibria.

7.5 Public goods provision

Suppose each player i has resource w_i that she can divide between private consumption x_i and contribution to public good g_i. The payoff of player i is as in Romano and Yildirim (2005)[37]:

$$u_i(\mathbf{x}, \mathbf{g}) = x_i B(G), \quad G = \sum_{i=1}^{n} g_i, \tag{18}$$

where $B(G)$ is the marginal benefit of private consumption, which depends on public good provision (rule of law, public amenities, etc.). Then $G = \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} x_i = W - X$. We can thus denote $h(X) = B(W - X)$ and apply the results in this paper.

The social planner would maximize total payoff $\sum_i u_i = Xh(X)$. By the same arguments as before, he would then choose disclosures that minimize total private consumption X^*, which implies the simultaneous choices. Allowing players to publicly commit to decisions would lead to earlier players choosing high quantities of private consumption and therefore to free-riding on later players.

7.6 Parimutuel betting

Parimutuel betting (or pool betting)[38] has been studied both as a model of a marketplace that aggregates information[39] and as a field experiment that highlights cognitive biases[40]. There is also an extensive literature on equilibrium behavior in parimutuel betting.[41]

To make a connection to my model, let me assume that there are a number of potential bets (a classical example is betting on horses). In particular, n bettors consider making a particular bet. As in the rest of the paper, I assume that there is no private information, i.e., those n players all believe that the bet wins with probability p. The total pool is the sum of their bets $X = \sum_{i=1}^{n} x_i$ plus the expected sum of all other bets, denoted by Y. If their bet wins, then the pool $X + Y$ is divided proportionally to their bets; otherwise, they lose their bets. Then, the expected payoff of player i is

$$u_i(\mathbf{x}, Y) = p\frac{x_i}{X}(X + Y) - x_i = \frac{x_i}{X}pY - (1 - p)x_i, \tag{19}$$

which is a Tullock payoff function with prize $v = pY$ and marginal cost $1 - p$.

[37] They studied a two-player version of this application, but under more general assumptions.

[38] I am grateful to Marco Ottaviani for suggesting this application.

[39] For example, recently Gillen et al. (2017) documented a field experiment Intel Corporation that used a parimutuel-like mechanism to aggregate information about the beliefs within the company and to forecast the outcomes. They found that the forecasts from the mechanism were more accurate than the benchmark forecasts.

[40] See Thaler and Ziemba (1988) for an overview.

[41] For example, Ottaviani and Sørensen (2009) show that the often-documented favorite-longshot bias may arise as an equilibrium behavior with rational, privately informed bettors.

My results imply that the total bet on X is increasing in the information provided to bettors about the other bets. It is minimized if players make their bets independently and maximized if the current pool is public.

Empirical work by Lemus and Marshall (2017) found that in prediction contests the disclosure rule has a significant impact on the outcomes—providing more information about other players' choices reduces overall participation but increases the upper tail of the distribution, and thus improves overall outcomes. These results provide some confirmation of my model's predictions, where disclosures incentivize the effort of earlier movers, reduce the payoffs of later movers, and increase the total equilibrium effort.

8 Discussion

In this paper, I characterize equilibria for all sequential contests and study their properties. Each contest has a unique equilibrium; it is in pure strategies and simple to compute. The calculation only requires finding the highest root of a recursively defined function.

The total equilibrium effort is strictly increasing in information in a general sense. This implies that the simultaneous contest minimizes the total effort and the fully sequential contest maximizes it. There is an earlier-mover advantage: players in earlier periods exert strictly greater effort and obtain strictly higher payoffs than players in later periods. As the number of players increases, total effort converges to the perfectly competitive value in all contests. The convergence is linear in the case of simultaneous contests and exponential in the case of fully sequential contests.

Some of the assumptions made for tractability here can be relaxed. First, I assume that the efforts are made public after each period. It is natural to ask whether players would want to hide their efforts. The answer in the setting studied here is that they would not. In a more general model, where players can make both public efforts (those studied above) and hidden efforts (those known only to themselves), all efforts would be public and therefore the equilibrium would remain unchanged.

Second, I assume that each player exerts effort only during the arrival period. Another natural extension is to study a setting where players can exert efforts multiple times following their arrival. Although making this argument formally requires more notation and a more careful analysis, there is no reason to expect the results to change. To see this, suppose that at the end of the contest all players have a chance to add to their previous efforts. Then again, each player can choose whether to make an effort publicly at arrival or secretly at the end. And since they can discourage the opponents by making public efforts, they don't hide their efforts. Yildirim (2005) and Romano and Yildirim (2005) proved a similar result in a more general model with two players.[42] Leininger and Yang (1994) showed that when players can react infinitely many times, it allows collusion to emerge in equilibrium.

Third, although the payoff structure in this model is quite general, it assumes linearity in players' own efforts and makes specific assumptions about the marginal benefit function. Relaxing this structure is not straightforward. I use both assumptions extensively in the characterization, and without these assumptions, it would be more difficult to determine an explicit characterization of equilibria. However, all results (with the exception of independence of permutations in corollary 1) are with strict inequalities, so I would expect the results to be robust to small perturbations in the assumptions, at least within some small range of parameter values.

Fourth, I assumed symmetry and no private information. It is well-known in the literature[43] that with private information there are no general closed-form characterizations of equilibria in these types of contests.[44] Note that it is not crucial for the contest designer to know the payoffs exactly since most of the results hold for any payoffs satisfying the sufficient conditions. The approach is generalizable to asymmetric (and publicly known) payoffs and entry costs but is beyond the scope of this paper.

Fifth, I assume that disclosures are public, which means that later players observe the sum of the efforts of more competitors than earlier players. This, combined with the assumption that there is no private information, means that no signaling or learning exists in this model. In particular, if the designer had an option for targeted disclosures, it would extend the possibilities significantly.

References

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

Amir, R. and I. Grilo (1999, January). Stackelberg versus Cournot equilibrium. *Games and Economic Behavior 26*(1), 1–21.

[42]They also showed that in a model with non-identical players, the possibility of revising efforts changes the equilibrium.

[43]See for example Wasser (2013) and Gallice (2017) for details.

[44]Recent work has shown significant progress in embedding private information to these types of models. For example, Lambert et al. (2017) show how to add private information to simultaneous quadratic games, which corresponds to linear h and $T = 1$ in my model. Bonatti et al. (2017) study a model dynamic of Cournot oligopoly, where firms learn their competitors' information over time, and the play converges to full-information static Cournot game.

Anderson, S. P. and M. Engers (1992, March). Stackelberg versus Cournot oligopoly equilibrium. *International Journal of Industrial Organization 10*(1), 127–135.

Baumol, W. J. (1982). Contestable markets: An uprising in the theory of industry structure. *The American Economic Review 72*(1), 1–15.

Baumol, W. J., J. C. Panzar, and R. D. Willig (1988, January). *Contestable Markets and the Theory of Industry Structure*. Harcourt.

Baye, M. R. and H. C. Hoppe (2003, August). The strategic equivalence of rent-seeking, innovation, and patent-race games. *Games and Economic Behavior 44*(2), 217–226.

Baye, M. R., D. Kovenock, and C. G. de Vries (1996, June). The all-pay auction with complete information. *Economic Theory 8*(2), 291–305.

Bell, D., R. Keeney, and J. Little (1976). A market share theorem. In *Mathematical Models in Marketing*, Volume 132 of *Lecture Notes in Economics and Mathematical Systems*, pp. 482–484. Springer Berlin Heidelberg.

Bimpikis, K., S. Ehsani, and M. Mostagir (2017). Designing dynamic contests. Stanford University, mimeo.

Bonatti, A., G. Cisternas, and J. Toikka (2017, April). Dynamic oligopoly with incomplete information. *The Review of Economic Studies 84*(2), 503–546.

Brock, W. A. (1983). Contestable markets and the theory of industry structure: A review article. *Journal of Political Economy 91*(6), 1055–1066.

Bronnenberg, B. J., S. K. Dhar, and J. H. Dubé (2009, February). Brand history, geography, and the persistence of brand shares. *Journal of Political Economy 117*(1), 87–115.

Che, Y.-K. and I. Gale (2003). Optimal design of research contests. *The American Economic Review 93*(3), 646–671.

Cournot, A.-A. (1838). *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot.* Chez L. Hachette.

Dasgupta, P. and J. E. Stiglitz (1980). Uncertainty, industrial structure, and the speed of R&D. *The Bell Journal of Economics 11*(1), 1–28.

Dasgupta, P. and J. E. Stiglitz (1988, March). Potential competition, actual competition, and economic welfare. *European Economic Review 32*(2-3), 569–577.

Daughety, A. F. (1990). Beneficial concentration. *The American Economic Review 80*(5), 1231–1237.

Demsetz, H. (1968). Why regulate utilities? *The Journal of Law & Economics 11*(1), 55–65.

Dixit, A. (1987). Strategic behavior in contests. *The American Economic Review 77*(5), 891–898.

Dranove, D. and N. Gandal (2003, September). The DVD-vs.-Divx standard war: Empirical evidence of network effects and preannouncement effects. *Journal of Economics & Management Strategy 12*(3), 363–386.

Fudenberg, D. and J. Tirole (1985). Preemption and rent equalization in the adoption of new technology. *The Review of Economic Studies 52*(3), 383–401.

Gallice, A. (2017, January). An approximate solution to rent-seeking contests with private information. *European Journal of Operational Research 256*(2), 673–684.

Gilbert, R. J. (1989, August). The role of potential competition in industrial organization. *Journal of Economic Perspectives 3*(3), 107–127.

Gillen, B. J., C. R. Plott, and M. Shum (2017, June). A pari-mutuel-like mechanism for information aggregation: A field test inside Intel. *Journal of Political Economy 125*(4), 1075–1099.

Glazer, A. and R. Hassin (1988, January). Optimal contests. *Economic Inquiry 26*(1), 133–143.

Glazer, A. and R. Hassin (2000). Sequential rent seeking. *Public Choice 102*(3-4), 219–228.

Hendricks, K., A. Weiss, and C. Wilson (1988). The war of attrition in continuous time with complete information. *International Economic Review 29*(4), 663–680.

Ino, H. and T. Matsumura (2012, November). How many firms should be leaders? beneficial concentration revisited. *International Economic Review 53*(4), 1323–1340.

Jensen, M. K. (2010). Aggregative games and best-reply potentials. *Economic Theory 43*(1), 45–66.

Jensen, M. K. (2017). Aggregative games. In L. C. Corchón and M. A. Marini (Eds.), *Handbook of Game Theory and Industrial Organization, Volume I*. Edward Elgar Publishing.

Julien, L. A., O. Musy, and A. W. Saïdi (2012). On hierarchical competition in oligopoly. *Journal of Economics 107*(3), 217–237.

Kalyanaram, G., W. T. Robinson, and G. L. Urban (1995). Order of market entry: Established empirical generalizations, emerging empirical generalizations, and future research. *Marketing Science 14*(3), G212–G221.

Konrad, K. A. (2009, April). *Strategy and Dynamics in Contests* (1 ed.). London School of Economics Perspectives in Economic Analysis. Oxford University Press.

Kreps, D. M. and J. A. Scheinkman (1983). Quantity precommitment and Bertrand competition yield Cournot outcomes. *The Bell Journal of Economics 14*(2), 326–337.

Krueger, A. O. (1974). The political economy of the Rent-Seeking society. *The American Economic Review 64*(3), 291–303.

Lambert, N. S., G. Martini, and M. Ostrovsky (2017). Quadratic games. Stanford Graduate School of Business, mimeo.

Lazear, E. P. and S. Rosen (1981). Rank-Order tournaments as optimum labor contracts. *Journal of Political Economy 89*(5), 841–64.

Leininger, W. (1993). More efficient rent-seeking — a Münchhausen solution. *Public Choice 75*(1), 43–62.

Leininger, W. and C.-L. Yang (1994, November). Dynamic Rent-Seeking games. *Games and Economic Behavior 7*(3), 406–427.

Lemus, J. and G. Marshall (2017). Dynamic tournament design: An application to prediction contests. University of Illinois at Urbana-Champaign, mimeo.

Linster, B. G. (1993). Stackelberg rent-seeking. *Public Choice 77*(2), 307–321.

Loury, G. C. (1979). Market structure and innovation. *The Quarterly Journal of Economics 93*(3), 395–410.

Martimort, D. and L. Stole (2012, November). Representing equilibrium aggregates in aggregate games with applications to common agency. *Games and Economic Behavior 76*(2), 753–772.

Moldovanu, B. and A. Sela (2001). The optimal allocation of prizes in contests. *The American Economic Review 91*(3), 542–558.

Moldovanu, B. and A. Sela (2006, January). Contest architecture. *Journal of Economic Theory 126*(1), 70–96.

Morgan, J. (2003). Sequential contests. *Public Choice 116*(1-2), 1–18.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1993). Why is Rent-Seeking so costly to growth? *The American Economic Review 83*(2), 409–414.

Nitzan, S. (1994, May). Modelling rent-seeking contests. *European Journal of Political Economy 10*(1), 41–60.

Novshek, W. (1980). Cournot equilibrium with free entry. *Review of Economic Studies 47*(3), 473–486.

Olszewski, W. and R. Siegel (2016, March). Large contests. *Econometrica 84*(2), 835–854.

Ottaviani, M. and P. N. Sørensen (2009, December). Surprised by the parimutuel odds? *American Economic Review 99*(5), 2129–2134.

Posner, R. A. (1975, August). The social costs of monopoly and regulation. *Journal of Political Economy 83*(4), 807–827.

Puu, T. (1991, January). Chaos in duopoly pricing. *Chaos, Solitons & Fractals 1*(6), 573–581.

Robson, A. J. (1990). Stackelberg and Marshall. *The American Economic Review 80*(1), 69–82.

Romano, R. and H. Yildirim (2005, January). On the endogeneity of Cournot-Nash and Stackelberg equilibria: games of accumulation. *Journal of Economic Theory 120*(1), 73–107.

Rosen, S. (1986). Prizes and incentives in elimination tournaments. *The American Economic Review 76*(4), 701–715.

Serena, M. (2017). Sequential contests revisited. *Public Choice 173*(1-2), 131–144.

Shepherd, W. G. (1984). "Contestability" vs. competition. *The American Economic Review 74*(4), 572–587.

Siegel, R. (2009). All-Pay contests. *Econometrica 77*(1), 71–92.

Spence, A. M. (1977). Entry, capacity, investment and oligopolistic pricing. *The Bell Journal of Economics 8*(2), 534–544.

Taylor, C. R. (1995). Digging for golden carrots: An analysis of research tournaments. *The American Economic Review 85*(4), 872–890.

Thaler, R. H. and W. T. Ziemba (1988). Anomalies: Parimutuel betting markets: Racetracks and lotteries. *The Journal of Economic Perspectives 2*(2), 161–174.

Tullock, G. (1967, June). The welfare costs of tariffs, monopolies, and theft. *Economic Inquiry 5*(3), 224–232.

Tullock, G. (1974). *The social dilemma: The economics of war and revolution*. University publications.

Tullock, G. (2001). Efficient rent seeking. In A. Lockard and G. Tullock (Eds.), *Efficient Rent-Seeking*, pp. 3–16. Springer US.

Vojnovic, M. (2015). *Contest Theory: Incentive Mechanisms and Ranking Methods*. Cambridge: Cambridge University Press.

von Stackelberg, H. (1934). *Marktform und gleichgewicht*. J. springer.

Wasser, C. (2013, January). Incomplete information in rent-seeking contests. *Economic Theory 53*(1), 239–268.

Yildirim, H. (2005, April). Contests with multiple rounds. *Games and Economic Behavior 51*(1), 213–227.

Appendices for online publication

A Tullock payoffs satisfy condition 1 (proposition 1)

Before proving proposition 1, let me briefly describe its key idea. The function f_{t+1} is a polynomial, of degree $r = T - t$, so it can have at most r roots. By keeping track of all the roots, I show by induction that all r roots are real and in $[0, 1)$, with the highest being $\underline{X}_{t+1}$. Therefore, all $r - 1$ roots of the derivative f'_t are also real and in $[0, \underline{X}_{t+1})$. Evaluating f_t at $\underline{X}_{t+1}$ and 1, we get

$$f_t(\underline{X}_{t+1}) = \underbrace{f_{t+1}(\underline{X}_{t+1})}_{=0} - n_{t+1} \underbrace{f'_{t+1}(\underline{X}_{t+1})}_{>0} \underbrace{\underline{X}_{t+1}(1 - \underline{X}_{t+1})}_{>0} < 0$$

$$f_t(1) = f_{t+1}(1) - n_{t+1} f'_{t+1}(1) \underbrace{1(1-1)}_{=0} = f_{t+1}(1) = \cdots = f_T(1) = 1 > 0.$$

This implies that f_t must have a root $\underline{X}_t \in (\underline{X}_{t+1}, 1)$. Moreover, since the highest root of its derivative is again below $\underline{X}_t$, it is strictly increasing in $[\underline{X}_t, 1]$. Finally, I show that the second highest root of f_t is strictly below $\underline{X}_{t+1}$, so that $f_t(X) < 0$ for all $[\underline{X}_{t+1}, \underline{X}_t)$. Proving this requires keeping track of all the roots.

Proof of proposition 1. First note that $f_T(X) = X$ is a polynomial of degree 1 and each step of the recursion adds one degree, so $f_t(X)$ is a polynomial of degree $T + 1 - t$, which I denote by r for brevity. The following two technical lemmas describe the values of the polynomials f_t at 1 and the number of roots at 0.

Lemma 2. $f_t(1) = 1$ *for all* $t = 0, \ldots, T$.

Proof. $f_{t-1}(1) = f_t(1) - n_t f'_t(1)1(1-1) = f_t(1) = f_T(1) = 1$. □

Lemma 3. $f_t(0) = 0$ *for all* $t = 0, \ldots, T$*. Depending on* $\mathbf{n}$*, there could be either one or two roots at zero:*

1. *If* $n_s = 1$ *for some* $s > t$*, then* $f_t(X)$ *has exactly two roots at zero.*

2. *Otherwise, i.e., if* $n_s \neq 1$ *for all* $s > t$*, then* $f_t(X)$ *has exactly one root at zero.*

Proof. As $f_t(X)$ is a polynomial of degree $r = T + 1 - t$, it can be expressed as

$$f_t(X) = \sum_{s=0}^{r} c_s^t X^s \quad \Rightarrow \quad f'_t(X) = \sum_{s=1}^{r} c_s^t s X^{s-1},$$

where $c_0^t, \ldots, c_r^t$ are the coefficients. Therefore,

$$f_{t-1}(X) = c_0^t + c_1^t(1-n_t) + \sum_{s=2}^{r} \left[c_s^t(1 - sn_t) + n_t c_{s-1}^t (s-1) \right] X^s + n_t c_{T+1-t}^t (T+1-t) X^{T+2-t}.$$

As $f_T(X) = X$, we have that $c_0^T = 0$ and so $c_0^t = 0$ for all t. Therefore, each f_t has at least one root at 0. Next, $f_{t-1}(X)$ has two roots at zero if and only if $c_1^{t-1} = c_1^t(1 - n_t) = 0$. This can happen only if either $c_1^t = 0$ (i.e., $f_t(X)$ has two roots at zero) or $n_t = 1$. As $f_T(X) = X$, we have that $c_1^T = 1$ and therefore, $f_t(X)$ does indeed have two roots at zero if and only if $n_s = 1$ for some $s > t$.

Finally, $f_{t-1}(X)$ would have three roots at zero only if $c_2^{t-1} = c_1^{t-1} = 0 = c_0^{t-1}$. This would require that $c_2^{t-1} = c_2^t(1 - 2n_t) + n_t c_1^t = c_2^t(1 - 2n_t) = 0$. Since $2n_t \neq 1$, this can happen only when $c_2^t = 0$. But note that $f_{T-1}(X) = n_T X^2 - (1 - n_T)X$, so that $c_2^{T-1} = n_T \neq 0$. Therefore, $f_t(X)$ cannot have more than two roots at zero. □

Lemma 4. *The leading coefficient of f_t is $(T-t)! \prod_{s=t+1}^{T} n_s > 0$.*

Proof. Using the same notation as in lemma 3, the leading coefficient of $f_{t-1}(X)$ is $c_{r+1}^{t-1} = r n_t c_r^t = r! \prod_{s=t}^{T} n_s$. □

Now I can proceed with the proof of proposition 1 itself. The proof uses that fact that the f_t is a polynomial of degree $r = T + 1 - t$ and keeps track of all of its roots. In particular, it can be expressed as

$$f_t(X) = c_t \prod_{s=1}^{r} (X - X_{s,t}), \tag{20}$$

where $c_t > 0$ is the leading coefficient and $X_{1,t}, \ldots, X_{r,t}$ are the r roots. By lemma 3, either one or two of these roots are equal to zero. I show by induction that all other roots are distinct and in $(0, 1)$.

Let us consider the case of a single zero root first, i.e., assume that $0 = X_{1,t} < X_{2,t} < \ldots < X_{r,t} < 1$. We can express the derivative of f_t as

$$f_t'(X) = c_t \sum_{i=1}^{r} \prod_{s \neq i} (X - X_{s,t}).$$

Therefore, at root $X_{j,t}$, the polynomial $f_t'(X)$ takes value

$$f_t'(X_{j,t}) = c_t \prod_{s \neq j} (X_{j,t} - X_{s,t}). \tag{21}$$

In particular, at the highest root, $f_t'(X_{r,t}) > 0$, and at the second highest $f_t'(X_{r-1,t}) < 0$;

therefore, f_t' must have a root $Y_{r-1,t} \in (X_{r-1,t}, X_{r,t})$. By the same argument, there must be a root $Y_{s,t}$ of f_t' between each of the two adjacent distinct roots of f_t. As f_t' is a polynomial of degree $r-1$, this argument implies that all the roots of f_t' are distinct and such that

$$X_{1,t} = 0 < Y_{1,t} < X_{2,t} < Y_{2,t} < \cdots < X_{r-1,t} < Y_{r-1,t} < X_{r,t} < 1.$$

In particular, $\operatorname{sgn} f_t'(X_{s,t}) = \operatorname{sgn} f_t(Y_{s,t})$ for all $s \in \{1, \ldots, r-1\}$. Next, note that $f_t(1) = 1 > 0$ and, as the highest root of f_t' is $Y_{r-1,t} < X_{r,t}$, this implies $f_t'(X_{r,t}) > 0$, and so

$$f_{t-1}(X_{r,t}) = f_t(X_{r,t}) - n_t f_t'(X_{r,t}) X_{r,t}(1 - X_{r,t}) < 0.$$

Therefore, f_{t-1} must have a root $X_{r+1,t-1} \in (X_{r,t}, 1)$. Now, for each $s \in \{2, r-1\}$

$$f_{t-1}(Y_{s,t}) = f_t(Y_{s,t}) \text{ and } f_{t-1}(X_{s,t}) = -n_t f_t'(X_{s,t}) X_{s,t}(1 - X_{s,t}).$$

Hence, $\operatorname{sgn} f_{t-1}(Y_{s,t}) = \operatorname{sgn} f_t(Y_{s,t}) = \operatorname{sgn} f_t'(X_{s,t}) = -f_{t-1}(X_{s,t})$. This means that f_{t-1} must have a root $X_{s+1,t-1} \in (X_{s,t}, Y_{s,t})$. This argument determines $r-2$ distinct roots in $(X_{2,t}, Y_{r-1,t})$. By lemma 3, f_{t-1} also has at least one root $X_{1,t-1} = 0$.

We have therefore found $1 + r - 2 - 1 = r$ distinct real roots of f_{t-1} that is a polynomial of degree $r+1$. Thus, the final root $X_{2,t}$ must also be real. By lemma 3, if $n_t = 1$,then the f_{t-1} must have two roots at zero; so, $X_{2,t} = 0$. Let us consider the remaining case where $n_t > 1$. By lemma 3, $X_{2,t} \neq 0$. To determine its location, consider the function $f_{t-1}^X(X) = f_{t-1}(X)/X$. Note that

$$f_t^X(X) = \frac{f_t(X)}{X} = c_t \prod_{s>0} (X - X_{s,t}) \Rightarrow f_t^X(0) = c_t \prod_{s>0} (-X_{s,t})$$

and

$$f_t'(0) = c_t \prod_{s>0} (-X_{s,t}).$$

Therefore,

$$f_{t-1}^X(0) = f_t^X(0) - n_t f_t'(0)(1 - 0) = c_t \prod_{s>0} (-X_{s,t}) \left[1 - n_t\right] = f_t'(0) \left[1 - n_t\right].$$

We assumed that $n_t > 1$; so, $\operatorname{sgn} f_{t-1}^X(0) = -\operatorname{sgn} f_t'(0)$. Evaluating the function $\operatorname{sgn} f_{t-1}^X$ at $Y_{1,t}$ gives

$$\operatorname{sgn} f_{t-1}^X(Y_{1,t}) = \operatorname{sgn} f_t(Y_{1,t}) = \operatorname{sgn} f_t'(X_{1,t}) = -\operatorname{sgn} f_{t-1}^X(0).$$

Hence, f_{t-1}^X must have a root $X_{2,t-1} \in (0, Y_{1,t})$. As $f_{t-1}(X) = X f_{t-1}^X(X)$, it must be a root of f_{t-1} as well. We have therefore located all $r+1$ roots of f_{t-1}, which are all distinct in this case.

Let us now get back to the case where f_t had two roots at zero. By the same argument as above, there must be a root of f_t' between each positive root of f_t. As there are $r-2$ positive roots, this determines $r-3$ distinct positive roots of f_t'. It is also clear that f_t' must have exactly one root at zero. Polynomial f_t' has $r-1$ roots, and we have determined that $r-2$ of them are real and distinct. Thus, the remaining root must be real. To determine its location, using the above approach, let $f_t^{X'}(X) = \frac{f_t'(X)}{X}$. Then as $f_t'(X_{r,t}) > 0$, we have $f_t^{X'}(X_{r,t}) > 0$. Similarly, $f_t^{X'}(X_{r-1,t}) < 0$, and so on. In particular, $f_t^{X'}(X_{3,t}) < 0$ if r is even, and $f_t^{X'}(X_{3,t}) > 0$ if r is odd. Now,

$$f_t^{X'}(0) = 2c_t \prod_{s>2} (-X_{s,t}),$$

which is strictly positive if r is odd and strictly negative if r is even, so that $\operatorname{sgn} f_t^{X'}(0) = -\operatorname{sgn} f_t^{X'}(X_{3,t})$. Hence, $f_t^{X'}$ must have a root $Y_{2,t} \in (0, X_{3,t})$. Clearly this $Y_{2,t}$ is also a root of $f_t'(X) = X f_t^{X'}(X)$. Now we have found all $r-1$ roots of polynomial f_t' and

$$X_{1,t} = Y_{1,t} = X_{2,t} = 0 < Y_{2,t} < X_{3,t} < \ldots < X_{r-1,t} < Y_{r,t} < X_{r,t}.$$

Again, $\operatorname{sgn} f_t'(X_{s,t}) = \operatorname{sgn} f_t(Y_{s,t})$ for all $s \in \{2, \ldots, r-1\}$.

By the same arguments as above, f_{t-1} has a root $X_{r+1,t-1} \in (X_{r,t}, 1)$ and $r-3$ roots $X_{s+1,t-1} \in (X_{s,t}, Y_{s,t})$ for each $s \in \{3, r-1\}$. Also, by lemma 3, f_{t-1} must have two roots at zero. Therefore, we have determined $1 + r - 3 + 2 = r$ roots of f_{t-1}, and so the final root must also be real. The argument for determining this root is similar to the previous case. Let $f_{t-1}^{X^2}(X) = \frac{f_{t-1}(X)}{X^2}$. Then

$$f_{t-1}^{X^2}(X) = f_t^{X^2}(X) - n_t f_t^{X'}(X)(1-X).$$

Therefore,

$$f_{t-1}^{X^2}(0) = c_t \prod_{s>2} (-X_{s,t})(1 - 2n_t),$$

so that $\operatorname{sgn} f_{t-1}^{X^2}(0) = -\operatorname{sgn} f_t^{X'}(0)$. Also,

$$f_{t-1}^{X^2}(Y_{2,t}) = f_t^{X^2}(Y_{2,t}).$$

Since $Y_{2,t} > 0$ and $X_{3,t} > 0 = X_{2,t}$, we have that

$$\begin{aligned}\operatorname{sgn} f_{t-1}^{X2}(Y_{2,t}) &= \operatorname{sgn} f_t^{X^2}(Y_{2,t}) = \operatorname{sgn} f_t(Y_{2,t}) = -\operatorname{sgn} f_t(Y_{3,t}) \\ &= -\operatorname{sgn} f_t'(X_{3,t}) = -\operatorname{sgn} f_t^{X'}(X_{3,t}) = \operatorname{sgn} f_t^{X'}(0) = -\operatorname{sgn} f_{t-1}^{X2}(0).\end{aligned}$$

Therefore, f_{t-1}^{X2} must have a root in $(0, Y_{2,t})$ which must also be a root of f_{t-1}. Again, we have found all $r+1$ roots of f_{t-1}.

In all cases, we found that

1. $X_{r+1,t-1} \in (X_{r,t}, 1)$; i.e., indeed the highest root of f_{t-1} is between the highest root of f_t and 1.

2. $[X_{r,t}, X_{r+1,t-1}) \subset (X_{r,t-1}, X_{r+1,t-1})$, so that $f_{t-1}(X) < 0$ for all $X \in [X_{r,t}, X_{r+1,t-1})$.

3. By the same argument as above (or by the Gauss-Lucas theorem), $X_{r+1,t-1} > Y_{r,t-1}$, so that $f_{t-1}'(X) > 0$ for all $X \in [X_{r+1,t-1}, 1]$.

$\square$

B Characterization theorem (theorem 1)

Proof of theorem 1. Let $X^{*t}(X_t)$ denote the total effort in the contest, if the cumulative effort after period t is X_t and players in periods $t+1, \dots, T$ behave according to their equilibrium strategies. I show by induction that $X^{*t}(X_t) = f_t^{-1}(X_t)$, which is defined as a mapping from $[0,1]$ to $[\underline{X}_t, 1]$, the region where f_t is strictly increasing. These statements are clearly true for $X^{*T}(X_T) = X_T$.

Let us take any $t < T$. The cumulative effort prior to period t is X_{t-1} and $X_t = X_{t-1} + \sum_{i \in \mathcal{I}_t} x_i$ after period t. So, the total effort is $X^{*t}(X_t) = f_t^{-1}(X_t)$, which is by the induction assumption such that $X^{*t}(0) = \underline{X}_t$, $X^{*t}(1) = 1$, and $\frac{dX^{*t}(X_t)}{dX_t} > 0$ for all $X_t \in [0,1]$.

The rest of the proof is divided into five lemmas as follows:

1. Lemma 5 shows that in all histories where $X_{t-1} < 1$, each player in period t chooses strictly positive effort, but these added efforts in period t are small enough so that the cumulative effort after period t remains strictly below one, $X_t < 1$. On the other hand, in histories where $X_{t-1} \geq 1$, the players in period t exert no effort. Therefore, on the equilibrium path $X_t < 1$ for all t.

2. Lemma 6 shows that $X_{t-1} = f_{t-1}(X)$ is a necessary condition for equilibrium.

3. Lemma 7 shows that under condition 1, $X^{*t-1}(X_{t-1})$ is well-defined and strictly increasing, $X^{*t-1}(0) = \underline{X}_{t-1}$ and $X^{*t-1}(1) = 1$.

4. Lemma 8 shows that the best-response function of player $i \in \mathcal{I}_t$ after cumulative effort X_{t-1} is $x_i^*(X_{t-1}) = \frac{1}{n_t}\left[f_t(f_{t-1}^{-1}(X_{t-1})) - X_{t-1}\right]$ for all $X_{t-1} < 1$ and $x_i^*(X_{t-1}) = 0$ for all $X_{t-1} \geq 0$. On the equilibrium path the individual efforts are $x_i^* = \frac{1}{n_t}\left[f_t(X^*) - f_{t-1}(X^*)\right]$.

5. Finally, lemma 9 verifies that the unique candidate for equilibrium, i.e., $\mathbf{x}^*$ specified in the theorem, is indeed an equilibrium.

Lemma 5. *Depending on X_{t-1}, we have two cases:*

1. If $X_{t-1} < 1$, then $x_i > 0$ for all $i \in \mathcal{I}_t$ and $X_{t-1} < X_t < 1$.

2. If $X_{t-1} \geq 1$, then $x_i = 0$ for all $i \in \mathcal{I}_t$ and $X_t = X_{t-1} \geq 1$.

In other words, if period t starts with cumulative effort $X_{t-1} < 1$, the players exert strictly positive efforts, but the cumulative effort stays below 1. On the other hand, if the cumulative effort is already $X_{t-1} \geq 1$, then all players choose zero effort and therefore $X_t = X_{t-1} \geq 1$. A straightforward implication of this lemma is that the total effort never reaches 1 or above in equilibrium, and the individual efforts on the equilibrium path are always interior (i.e., strictly positive).

Proof. If $X_{t-1} \geq 1$, then if any player i in period t chooses $x_i > 0$, then $X_t > 1$ and therefore $X^{*t}(X_t) \geq X_t > 1$, which means that $u_i(\mathbf{x}) = x_i h(X^{*t}(X_t)) < 0$. Since player i can ensure zero payoff by choosing $x_i = 0$, this is a contradiction. So, $x_i^*(X_{t-1}) = 0$ for all $X_{t-1} \geq 1$ and thus $X_t = X_{t-1} \geq 1$.

Now, take $X_{t-1} < 1$. Suppose by contradiction that it leads to $X^{*t-1}(X_{t-1}) \geq 1$. This implies that in some period $s \geq t$ players chose efforts such that $X_{s-1} < 1$, but $X_s \geq 1$. This means that one or more of these players $i \in \mathcal{I}_s$ chose $x_i > 0$ and gets a payoff of $u_i(\mathbf{x}) = x_i h(X^{*s}(X_{s-1} + \sum_{j \in \mathcal{I}_s} x_j)) \leq 0$. Now, there are two cases. First, if $X^{*s}(X_{s-1} + \sum_{j \in \mathcal{I}_s} x_j) > 1$, then player i's payoff is strictly negative, and the player could deviate and choose $x_i = 0$ to ensure zero payoff. On the other hand, if $X^{*s}(X_{s-1} + \sum_{j \in \mathcal{I}_s} x_j) = 1$, which means that $X_s = 1$, then player i could choose effort $\frac{x_i}{2}$, thus making $X_s < 1$ and therefore $X^{*s}(X_s) < 1$, ensuring a strictly positive payoff. In both cases we arrive at a contradiction. Thus $X_{t-1} < 1$ implies $X_t < 1$ and $X^{*t}(X_t) < 1$.

The last step is to show that $X_{t-1} < 1$ implies $x_i > 0$ for all $i \in \mathcal{I}_t$. Suppose that this is not true, so that $x_i = 0$ for some i. Then player i gets a payoff of 0. But by choosing $\widehat{x}_i \in (0, 1 - X_t)$, he can ensure that the cumulative effort $\widehat{X_t} = X_t + \widehat{x}_i < 1$ and thus $X^{*t}(\widehat{X_t}) < 1$, and the new payoff of player i is strictly positive. This is a contradiction. $\square$

Lemma 6. $X_{t-1} = f_{t-1}(X)$ *is a necessary condition for equilibrium.*

This lemma implies that if X^* is the total effort in equilibrium, then before the last period, the cumulative effort had to be $X^*_{T-1} = f_{T-1}(X^*)$, one period before that $X^*_{T-2} = f_{T-2}(X^*)$, and so on. Generally, after period t, the total effort had to be $X^*_t = f_t(X^*)$. Therefore, X^* must satisfy $f_0(X^*) = X^*_0 = 0$ and $f_t(X^*) - f_{t-1}(X^*) = \sum_{i\in\mathcal{I}_t} x^*_i \geq 0$.

Proof. By lemma 5, we only need to consider the histories with $X_{t-1} < 1$. Moreover, we know that each player $i \in \mathcal{I}_t$ chooses $x_i > 0$, i.e., an interior solution. Player i's maximization problem is

$$\max_{x_i\geq 0} x_i h(X^{*t}(X_t)),$$

where $X^{*t}(X_t)$ is the total effort induced by cumulative effort $X_t = X_{t-1} + \sum_{j\in\mathcal{I}_t} x_j$. Therefore, a necessary condition for optimum is

$$h(X^{*t}(X_t)) + x_i h'(X^{*t}(X_t))\frac{dX^{*t}(X_t)}{dX_t} = 0.$$

It is convenient to rewrite this condition in terms of the total effort X, taking into account that $X = X^{*t}(X_t) = f_t^{-1}(X_t)$, and therefore $\frac{dX^{*t}(X_t)}{dX_t} = \frac{1}{f'_t(X)}$ to get

$$x_i = -f'_t(X)\frac{h(X)}{h'(X)} = f'_t(X)g(X). \tag{22}$$

Now, we can add up these necessary conditions for all players $i \in \mathcal{I}_t$ and take into account that $f_t(X) = X_t = X_{t-1} + \sum_{i\in\mathcal{I}_t} x_i$ to get a necessary condition for the equilibrium

$$X_{t-1} = f_t(X) - n_t f'_t(X)g(X) = f_{t-1}(X).$$

□

Lemma 7. *Under condition 1,* $X^{*t-1}(X_{t-1})$ *is well-defined, strictly increasing,* $X^{*t-1}(0) = \underline{X}_{t-1}$ *and* $X^{*t-1}(1) = 1$.

Proof. Note that $X_t \geq X_{t-1}$, so that the total effort induced by X_{t-1} can never be lower than $X^{*t}(X_{t-1}) \geq X^{*t}(0) \geq \underline{X}_t$. Therefore, $X < \underline{X}_t$ cannot be the total effort following any X_{t-1}.

Moreover, by condition 1, $\underline{X}_{t-1} \geq \underline{X}_t$ and $f_{t-1}(X) < 0$ for all $X \in [\underline{X}_t, \underline{X}_{t-1})$; therefore, total efforts in this range are not consistent with any X_{t-1} either. We get that the only feasible range of the total effort X induced by cumulative effort X_{t-1} is $[\underline{X}_{t-1}, 1]$. By condition 1, the function f_{t-1} is continuously differentiable and strictly increasing in this range; therefore, the inverse is well-defined, continuously differentiable, and strictly increasing. Moreover, since $f_{t-1}(1) = 1$, we have $X^{*t-1}(1) = 1$, and since $\underline{X}_{t-1}$ is a root of f_{t-1}, we have $X^{*t-1}(0) = \underline{X}_{t-1}$. $\square$

Lemma 8. *The best-response function of player $i \in \mathcal{I}_t$ after cumulative effort X_{t-1} is*

$$x_i^*(X_{t-1}) = \begin{cases} \frac{1}{n_t}\left[f_t(f_{t-1}^{-1}(X_{t-1})) - X_{t-1}\right] & \forall X_{t-1} < 1, \\ 0 & \forall X_{t-1} \geq 1. \end{cases} \tag{23}$$

On the equilibrium path the individual efforts are $x_i^* = \frac{1}{n_t}\left[f_t(X^*) - f_{t-1}(X^*)\right]$.

Proof. Lemma 5 proved the claim for any $X_{t-1} \geq 1$, so let us consider the case $X_{t-1} < 1$. Then by lemma 5, the individual efforts are interior, so they have to satisfy the individual first-order conditions (22). We showed that the total effort induced by X_{t-1} is $X^{*t-1}(X_{t-1}) = f_{t-1}^{-1}(X_{t-1})$. Inserting these results into the individual optimality condition for player $i \in \mathcal{I}_t$ we get

$$x_i^*(X_{t-1}) = \frac{1}{n_t}\left[f_t(X^{*t-1}(X_{t-1})) - f_{t-1}(X^{*t-1}(X_{t-1}))\right] = \frac{1}{n_t}\left[f_t(f_{t-1}^{-1}(X_{t-1})) - X_{t-1}\right].$$

In particular, on the equilibrium path, $X = X^*$, and therefore $x_i^* = \frac{1}{n_t}\left[f_t(X^*) - f_{t-1}(X^*)\right]$. $\square$

The arguments so far show that necessary conditions for equilibria lead to a unique candidate for equilibrium—the strategies specified in the theorem. Finally, we have to check that this is indeed an equilibrium. That is, we need to show that all players are indeed maximizing their payoffs.

Lemma 9. $\mathbf{x}^*$ *is an equilibrium.*

Proof. By construction $x_i^*(X_{t-1})$ is a local extremum for player $i \in \mathcal{I}_t$, given that the cumulative effort prior to period t is X_{t-1} and all other players behave according to their equilibrium strategies. Since the local extremum is unique and ensures strictly positive payoff (which is strictly more than zero from corner solution $x_i = 0$), $x_i^*(X_{t-1})$ is also the global maximum. Thus, no player has an incentive to deviate. $\square$

$\square$

C Impossibility of standard backward-induction

In this appendix, I show by a simple example[45], that a standard backward-induction, where we compute best-response functions $x_i^*(X_{t-1})$ backward by plugging them into the optimization problems of earlier players, is not feasible even in the case of Tullock payoffs. That is, I show that we cannot explicitly define $x_i^*(X_{t-1})$ as a function of X_{t-1} using standard mathematical operations.

One counterexample is a five-player sequential contest $\mathbf{n} = (1, 1, 1, 1, 1)$ with Tullock payoffs. By theorem 1, the best-response of player 2 to $x_1 \in [0, 1)$ is

$$\begin{aligned} x_2^*(x_1) &= f_2(f_1^{-1}(x_1)) - x_1, \text{where} \\ f_2(X) &= 6X^4 - 6X^3 + X^2, \\ f_1(X) &= 24X^5 - 36X^4 + 14X^3 - X^2. \end{aligned}$$

Note that to express $x_2^*(x_1)$ explicitly, we would have to be able to express $f_1^{-1}(x_1)$ explicitly. I claim that no such expression exists for all $x_1 \in [0, 1)$. It would have to be a root of polynomial $f_1(X) - x_1$ (in the interval $[\underline{X}_2, 1]$[46]), where the coefficients of polynomial $f_1(X)$ are integers from the expression above, but x_1 can be any number in $[0, 1]$.

Take for example $x_1 = \frac{1}{23}$. Multiplying $f_1(X) - x_1$ by 23 gives us a polynomial with integer coefficients:

$$P(X) = 23f_1(X) - 1 = 552X^5 - 828X^4 + 322X^3 - 23X^2 - 1.$$

This polynomial has three real roots in $(0, 1)$ and two complex roots. Therefore, Galois group $\mathrm{Gal}(P)$ contains a transposition. Reducing it, modulo 5 gives

$$P(X) \bmod 5 = 2X^5 + 2X^4 + 2X^3 + 2X^2 - 1.$$

This polynomial is irreducible in $\mathbb{Z}/p\mathbb{Z}[X]$; therefore, $P(X)$ is irreducible in $\mathbb{Q}[X]$. Hence, $\mathrm{Gal}(P)$ is S_5 and thus not solvable. This means that roots of $P(X)$ are not expressible in radicals. Therefore, $x_2^*(x_1)$ cannot have an explicit formula.

Note that finding the equilibrium with the inverted best-response approach poses no difficulties. The total equilibrium effort is the highest root of

$$f_0(X) = 120X^6 - 240X^5 + 150X^4 - 30X^3 + X^2,$$

[45] I am grateful to Jyotiraditya Singh for help with this example.

[46] The highest root of $f_2(X)$ is $\underline{X}_2 = X = \frac{\sqrt{3}+3}{6} \approx 0.7887$.

which implies

$$X^* = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{\sqrt{7}}{2\sqrt{15}}} \approx 0.9587, \quad \mathbf{x}^* \approx (0.4424, 0.2583, 0.1425, 0.0759, 0.0396).$$

In this example, X^* and hence each x_i^* can be expressed in terms of radicals, but even if they could not, it would not pose problems, since we have to find the roots just once and do not need to substitute the best-response functions to new optimization problems as with the standard best-response approach.

D Tullock payoffs satisfy condition 2 (proposition 2)

Proof of proposition 2. The proof relies on three lemmas that I prove below.

1. Lemma 10 shows that the highest root of g_k is $Z_{k:k} = 1$, the second highest $Z_{k-1:k} \in (Z_{k-2:k-1}, 1)$, and $g_k(X) > 0$ for all X between the highest two roots. Therefore, to prove that $g_k(X^*) > 0$, it suffices to show that $X^* > Z_{k-1:k}$.

2. Lemma 11 establishes a connection between X^* and $Z_{k-1:k}$. It shows that if we take the sequential n-player contest $\mathbf{n} = (1, \dots, 1)$, then $f_{n-k}(X) = g_k(X)\frac{X}{1-X}$ for all $k = 1, \dots, n$. Therefore, if we take the k-player sequential contest, we get $f_0(X) = g_k(X)\frac{X}{1-X}$, and so the total equilibrium effort X^* of this contest is exactly equal to the second highest root of g_k, i.e., $Z_{k-1:k}$.

   This proves the first part of the claim; i.e., if $\mathbf{n}$ is fully sequential and $k = T = n$, then $X^* = Z_{n-1:n}$, which is a root of g_n, and therefore $g_k(X^*) = 0$.

3. Lemma 12 shows directly[47] that X^* is strictly increasing in each n_t. Therefore, if the contest is not sequential (i.e., $n_t > 1$ for at least one t), then the total effort in this contest is strictly higher than in the sequential case, i.e., $X^* > Z_{T-1:T}$, and therefore $g_T(X^*) > 0$.

4. Finally, lemma 10 also shows that the adjacent g_k's are interlaced; i.e., the second highest roots are increasing in k, so that for all $k < T$, $Z_{k-1:k} < Z_{T-1:T} \leq X^*$, and therefore $g_k(X^*) > 0$ for all $k < T$.

□

Lemma 10. *If $g(X) = X(1-X)$, then g_k have the following properties for all $k \geq 1$:*

[47] Note the first part of corollary 1 proves the same claim, but since proposition 2 establishes a sufficient condition for theorem 2 and hence its corollary 1, to avoid a circular argument I prove it here directly.

1. $g_k(1) = 1$.

2. $g_k'(1) = -1$.

3. g_k *can be expressed as*
$$g_k(X) = -\prod_{j=0}^{k}(X - Z_{j:k}), \tag{24}$$
where $0 = Z_{0:k} < Z_{1:k} < \cdots < Z_{k:k} = 1$.

4. $Z_{s:k+1} \in (Z_{s-1:k}, Z_{s:k})$ *for all* $s = 1, \ldots, k$.

Proof. First note that $g_1(X) = g(X) = X(1-X)$ is a polynomial of degree 2. Each step of the recursion gives a polynomial of one degree higher; i.e., $g_k(X)$ is a polynomial of degree $k+1$, so $g_k'(X)$ is a polynomial of degree k, and therefore $g_{k+1}(X) = -g_k'(X)X(1-X)$ is a polynomial of degree $k+2$.

1. $g_{k+1}(1) = -g_k'(1)g(1) = 0$, because $g(1) = 1(1-1) = 0$.

2. $g_k'(1) = -g_{k-1}''(1)g(1) - g'(1)g_{k-1}'(1) = g_{k-1}'(1) = \cdots = g_1'(1) = g'(1) = 1-2\cdot1 = -1$.

3. The claim clearly holds for $g_1(X) = X(1-X)$ with $Z_{0:1} = 0 < Z_{1:1} = 1$. Suppose it holds for k. Since all $k+1$ roots of g_k are real and in $[0,1]$, by the Gauss-Lucas theorem all k roots of g_k' are in $(0,1)$. Then $g_{k+1}(X) = -g_k'(X)X(1-X)$ clearly has roots at 0 and 1 and k roots in $(0,1)$. To see that the roots are all distinct, note that
$$g_k'(X) = -\sum_{s=0}^{k}\prod_{j\neq s}(X - Z_{j:k}).$$
Therefore, $g_k'(Z_{s:k}) = -\prod_{j\neq s}(Z_{s:k} - Z_{j:k})$, which is strictly negative for $s = k$, strictly positive for $s = k-1$, and so on. Therefore, for each $s = 1, \ldots, k$, function g_k'; hence, g_{k+1} also has a root $Z_{s:k+1} = (Z_{s-1:k}, Z_{s:k})$. This determines the k interior roots.

4. The previous argument also proves the last claim.

□

Lemma 11. *If* $\mathbf{n} = (1, \ldots, 1)$, *then* $f_{n-k}(X) = g_k(X)\frac{X}{1-X}$ *for all* $k = 1, \ldots, T$.

Proof. Suppose that $\mathbf{n} = (1, \ldots, 1)$. First, $f_{n-1}(X) = X - X(1-X) = X^2 = g_1(X)\frac{X}{1-X}$. Now, suppose that $f_{n-k}(X) = g_k(X)\frac{X}{1-X}$. Then since
$$\frac{d\frac{X}{1-X}}{dX}X(1-X) = \left[\frac{1}{1-X} - \frac{-X}{(1-X)^2}\right]X(1-X) = \frac{X(1-X)}{(1-X)^2} = \frac{X}{1-X},$$

we get that

$$\begin{aligned}
f_{n-(k+1)}(X) &= f_{n-k}(X) - f'_{n-k}(X)X(1-X) \\
&= g_k(X)\frac{X}{1-X} - g_k(X)\frac{d\frac{X}{1-X}}{dX}X(1-X) - g'_k(X)\frac{X}{1-X}X(1-X) \\
&= g_{k+1}(X)\frac{X}{1-X}.
\end{aligned}$$

□

Lemma 12. *X^* is independent of permutations of $\mathbf{n}$ and strictly increasing in each n_t.*

Proof. Fix a contest $\mathbf{n}$ and a period $t > 1$. To shorten the notation, let $\phi(X) = f'_t(X)g(X)$.

$$\begin{aligned}
f_{t-1}(X) &= f_t(X) - n_t f'_t(X)g(X) = f_t(X) - n_t\phi(X), \\
f'_{t-1}(X) &= f'_t(X) - n_t\phi'(X) = \frac{\phi(X)}{g(X)} - n_t\phi'(X), \\
f_{t-2}(X) &= f_{t-1}(X) - n_{t-1}f'_{t-1}(X)g(X) = f_t(X) - [n_{t-1} + n_t]\phi(X) + n_{t-1}n_t\phi'(X)g(X).
\end{aligned}$$

Switching n_{t-1} and n_t in $\mathbf{n}$ does not affect f_{t-2}, and therefore it also doesn't affect f_0. This means that any such switch leaves X^* unaffected, which means that X^* is independent of permutations of $\mathbf{n}$.

To prove that X^* is strictly increasing in each n_t, it therefore suffices to prove that it is strictly increasing on n_1. Now, suppose $\widehat{\mathbf{n}} = (n_1 + 1, n_2, \ldots, n_T)$. Then f_1 is unchanged and the corresponding $\widehat{f}_0$ at the original equilibrium X^* is

$$\widehat{f}_0(X^*) = f_1(X^*) - (n_1 + 1)f'_1(X^*)g(X^*) = f_0(X^*) - f'_1(X^*)g(X^*) < 0,$$

because $f_0(X^*) = 0$, $f_1(X^*) > 0$ by condition 1 and $g(X^*) > 0$ since $X^* \in (0, 1)$. By condition 1, $\widehat{f}_0$ is strictly increasing between its highest root $\widehat{X}^*$ and 1, thus $\widehat{X}^* > X^*$. □

E Information theorem (theorem 2)

Let $\mathbf{n}^t = (n_{t+1}, \ldots, n_T)$ denote the sub-contest starting after period t. Note that $f_t(X)$ depends only on $\mathbf{n}^t$.

Remember that $g_1, \ldots, g_T$ are recursively defined as $g_1(X) = g(X)$ and $g_{k+1}(X) = -g'_k(X)g(X)$, so they are independent of $\mathbf{n}$. Also, $\mathbf{S}(\mathbf{n}) = (S_1(\mathbf{n}), \ldots, S_T(\mathbf{n}))$ are defined so that $S_k(\mathbf{n})$ is the sum of all products of k-combinations of vector $\mathbf{n}$ and is therefore independent of X.

The key step of the proof of the information theorem involves expressing $f_0(X)$ in terms of information measures $\mathbf{S}$ and functions g_k. This representation is more general, holds for any $f_t(X)$ and is useful for other results as well, so let me prove this first.

Lemma 13. *The function $f_t(X)$ can be expressed as*

$$f_t(X) = X - \sum_{k=1}^{T-t} S_k(\mathbf{n}^t) g_k(X). \tag{25}$$

Proof. Clearly, $f_T(X) = X$ satisfies the condition. Now, suppose that the characterization holds for $f_t(X)$. Then, since $g_{k+1}(X) = -g_k'(X)g(X)$, we get that

$$f_t'(X)g(X) = g(X) - g(X)\sum_{k=1}^{T-t} S_k(\mathbf{n}^t) g_k'(X) = g(X) + \sum_{k=2}^{T-t+1} S_{k-1}(\mathbf{n}^t) g_k(X).$$

Therefore, $f_{t-1}(X) = f_t(X) - n_t f_t'(X) g(X)$ implies that

$$\begin{aligned} f_{t-1}(X) &= X - \sum_{k=1}^{T-t} S_k(\mathbf{n}^t) g_k(X) - n_t g(X) - n_t \sum_{k=2}^{T-t+1} S_{k-1}(\mathbf{n}^t) g_k(X) \\ &= X - [S_1(\mathbf{n}^t) + n_t] g_1(X) - \sum_{k=2}^{T-t} [S_k(\mathbf{n}^t) + n_t S_{k-1}(\mathbf{n}^t)] g_k(X) - n_t S_{T-t}(\mathbf{n}^t) g_{T+1}(X). \end{aligned}$$

Note that $S_1(\mathbf{n}^t) = \sum_{s>t} n_s$ and $g_1(X) = g(X)$, so that $S_1(\mathbf{n}^t) + n_t = S_1(\mathbf{n}^{t-1})$. Similarly, $\mathbf{n}^{t-1} = (n_t, \mathbf{n}^t)$, so $S_k(\mathbf{n}^t)$ includes all k-combinations of $\mathbf{n}^{t-1}$ except the ones involving n_t. Adding $n_t S_{k-1}(\mathbf{n}^t)$ therefore completes the sum, so that $S_k(\mathbf{n}^{t-1}) = S_k(\mathbf{n}^t) + n_t S_{k-1}(\mathbf{n}^t)$. Since $S_{T-t}(\mathbf{n}^t) = n_{t+1} \dots n_T$, we have that $n_t S_{T-t}(\mathbf{n}^t) = n_t \times \dots \times n_T = S_{T-(t-1)}(\mathbf{n}^{t-1})$. Therefore, we can express $f_{t-1}(X)$ as

$$f_{t-1}(X) = X - \sum_{k=1}^{T-(t-1)} S_k(\mathbf{n}^{t-1}) g_k(X).$$

□

With this, the proof of the information theorem is now straightforward.

Proof of theorem 2. By theorem 1, the total equilibrium X^* is the highest root of $f_0(X)$ in $[0,1]$. By lemma 13, we can express $f_0(X^*)$ as

$$f_0(X^*) = X^* - \sum_{k=1}^{T} S_k(\mathbf{n}) g_k(X^*).$$

Now, take another contest $\widehat{\mathbf{n}}$ such that $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$, i.e., $S_k(\widehat{\mathbf{n}}) \geq S_k(\mathbf{n})$ for all k and the inequality is strict for at least one $k \in \{1, \ldots, T\}$. Then, by condition 2, the corresponding $\widehat{f}_0$ is such that at the original X^*,

$$\widehat{f}_0(X^*) = X^* - \sum_{k=1}^{T} S_k(\widehat{\mathbf{n}}) g_k(X^*) < X^* - \sum_{k=1}^{T} S_k(\mathbf{n}) g_k(X^*) = f_0(X^*).$$

Note that the total equilibrium effort of the new contest, $\widehat{X}^*$, is the highest root of $\widehat{f}_0$ and by condition 1, the function $\widehat{f}_0$ is strictly increasing in $[\widehat{X}^*, 1]$. Therefore, $\widehat{X}^* > X^*$. □

F Tullock payoffs imply declining weights (lemma 1)

Proof of lemma 1. By lemma 11, $g_k(X) = \widehat{f}_{\widehat{n}-k}(X)\frac{1-X}{X}$, where $\widehat{f}_{\widehat{n}-k}$ is defined for a sequential $\widehat{n} \geq k$-player contest. Similarly, $g_{k-1}(X) = \widehat{f}_{n+1-k}(X)\frac{1-X}{X}$. Therefore,

$$g_{k-1}(X^*) - g_k(X^*) = [\widehat{f}_{\widehat{n}+1-k}(X^*) - \widehat{f}_{\widehat{n}-k}(X^*)]\frac{1-X^*}{X^*} = \widehat{f}'_{\widehat{n}+1-k}(X^*)(1-X^*)^2.$$

Now, take $\widehat{n} = T$. Then by lemma 12, X^* is weakly higher than the highest root of $\widehat{f}_0$. By condition 1, the highest root of $\widehat{f}_{T+1-k}$ is even (weakly) lower and $\widehat{f}_{T+1-k}$ is strictly increasing above its highest root, so that $\widehat{f}'_{\widehat{n}+1-k}(X^*) > 0$. This proves that $g_{k-1}(X^*) > g_k(X^*)$. □

G Earlier-mover advantage (proposition 3)

Proof of proposition 3. The equilibrium payoff of player i is $u_i(\mathbf{x}^*) = x_i^* h(X^*)$, so the payoffs are ranked in the same order as the individual efforts (in fact they are proportional to individual efforts). Therefore, it suffices to prove that if $i \in \mathcal{I}_t$ and $j \in \mathcal{I}_{t+1}$, then $x_i^* > x_j^*$. We can express the equilibrium efforts as

$$x_i^* = \frac{1}{n_t}[f_t(X^*) - f_{t-1}(X^*)] = f'_t(X^*)g(X^*),$$
$$x_j^* = \frac{1}{n_{t+1}}[f_{t+1}(X^*) - f_t(X^*)] = f'_{t+1}(X^*)g(X^*).$$

Using lemma 13, we get that

$$f'_t(X) = 1 - \sum_{k=1}^{T-t} S_k(\mathbf{n}^t) g'_k(X) \quad \text{and} \quad f'_{t+1}(X) = 1 - \sum_{k=1}^{T-t-1} S_k(\mathbf{n}^{t+1}) g'_k(X).$$

Therefore, we can express the difference as

$$x_i^* - x_j^* = f_t'(X^*)g(X^*) - f_{t+1}'(X^*)g(X^*) = \sum_{k=1}^{T-t} \left[S_k(\mathbf{n}^t) - S_k(\mathbf{n}^{t+1})\right] g_{k+1}(X^*) > 0,$$

because $S_k(\mathbf{n}^t) > S_k(\mathbf{n}^{t+1})$ (there is less information remaining in the game that starts one period later) and as $t \geq 1$, then $k+1 \leq T$, so condition 2 implies that $g_{k+1}(X^*) > 0$.[48] □

H Large contests

In this appendix, I formalize the claim from section 6 and generalize it to arbitrary payoff functions that satisfy sufficient conditions of the characterization theorem, i.e., condition 1. The reason for the result was discussed in section 6. As the number of players grows, the total equilibrium effort $X^* \to 1$. By assumptions, payoffs and therefore $g(X)$ are smooth near X^*; hence, the function $g(X)$ can be well approximated by linear function near X^*.

Generally $-g'(1) = \alpha > 0$ and $g(1) = 0$, so the linear approximation near 1 is $\alpha(1-X)$. Therefore, $g_1(X) \approx \alpha(1-X)$, so $g_1'(X) \approx -\alpha$. Then $g_2(X) = -g_1'(X)g(X) \approx \alpha^2(1-X)$, so $g_2'(X) \approx -\alpha^2$, and so on. This shows that $g_k(X) \approx \alpha^k(1-X)$. Inserting this into (9) gives

$$f_0(X^*) \approx 1 - (1-X^*)\sum_{k=1}^{T} S_k(\mathbf{n})\alpha^k = 0 \quad \Rightarrow \quad 1 - X^* \approx \frac{1}{S(\mathbf{n})},$$

where $S(\mathbf{n}) = \sum_{k=1}^{T} S_k(\mathbf{n})\alpha^k = \prod_{t=1}^{T}(1+\alpha n_t) - 1$ is the weighted sum of information measures.

Using lemma 13 and the fact that $g_k'(X^*) \approx \alpha^k$, the individual equilibrium effort of player i from period t can be expressed as

$$\begin{aligned} x_i^* &= \frac{1}{n_t}\left[f_t(X^*) - f_{t-1}(X^*)\right] = g(X^*)f_t'(X^*) \approx \alpha(1-X^*)\left[1 - \sum_{k=1}^{T-t} S_k(\mathbf{n}^t)\alpha^k\right] \\ &= \alpha(1-X^*)\prod_{s=t+1}^{T}(1+\alpha n_s) = \alpha\frac{\prod_{s=t+1}^{T}(1+\alpha n_s)}{\prod_{s=1}^{T}(1+\alpha n_s)} = \frac{\alpha}{\prod_{s=1}^{t}(1+\alpha n_s)}. \end{aligned}$$

Therefore, the individual equilibrium effort is $x_i^* \approx \frac{\alpha}{S(n_1,\ldots,n_t)}$, where $S(\cdot)$ is again a weighted sum of the information measures, but only in the subcontest that includes players up to period t. This means that in contests with linear g function, players' efforts are

[48] Note that in case of Tullock payoffs and fully sequential contest, $g_{k+1}(X^*) = g_T(X^*) = g_n(X^*) = 0$, but this does not affect the conclusion, because as long as $n > 2$, the sum includes at least one more element and $g_{n-1}(X^*) > 0$. But in the case of $n = 2$, we get that the sum equals $[S_1(\mathbf{n}^1) - S_1(\mathbf{n}^2)]g_2(X^*) = 1 \times 0 = 0$, which we knew, as $x_1^* = x_2^* = \frac{1}{4}$.

independent of players arriving after them[49].

To make this argument formally, let $(\mathbf{n}^n)_{n\in\mathbb{N}}$ be a sequence of contests, such that contest $\mathbf{n}^n = (n_1^n, \ldots, n_{T^n}^n)$ has $T_n \leq n$ periods and $n = \sum_{t=1}^{T^n} n_t^n$ players. Moreover, for each $T \in \mathbb{N}$, let $\mathbf{n}^n|T$, denote a *censored contest*, which has a total of n players like contest $\mathbf{n}^n$ and has the same first $T-1$ disclosures, but all the other disclosures have been removed.[50] That is, $\mathbf{n}^n|T = (n_1^n, \ldots, n_{T-1}^n, n_{T+}^n)$, where $n_{T+}^n = \sum_{t=T}^{T^n} n_t$. Clearly, if $T > T^n$, then T does not affect the contest, i.e., $\mathbf{n}^n|T = \mathbf{n}^n$. But when $T^n > T$, then higher T means a more informative contest, i.e., $\mathbf{S}(\mathbf{n}^n) > \mathbf{S}(\mathbf{n}^n|T) > \mathbf{S}(\mathbf{n}^n|T-1)$.

Proposition 4 (Large contests limit)**.** *Suppose condition 1 holds. For each $T \in \mathbb{R}$*

$$\lim_{n\to\infty} S(\mathbf{n}^n|T)(1 - X^*(\mathbf{n}^n|T)) = 1 \tag{26}$$

and for each $t \leq T$, for each player i in period t,

$$\lim_{n\to\infty} x_i^*(\mathbf{n}^n|T) = \frac{\alpha}{\prod_{s=1}^{t}(1+\alpha n_s)}. \tag{27}$$

Proof. Fix any T and any n. The total equilibrium effort of a censored contest $\mathbf{n}^n|T$ is the highest root of $f_0(X)$, which can be expressed as

$$f_0(X) = X - \sum_{k=1}^{T} S_k(\mathbf{n}^n|T) g_k(X),$$

where $S_k(\mathbf{n}^n|T)$ is the sum of all products of k-combinations of $\mathbf{n}^n|T$. Functions g_k are defined as $g_1(X) = g(X)$ and $g_{k+1}(X) = -g_k'(X)g(X)$. Therefore, for each k, function $g(X)$ is a twice continuously differentiable function, $g_k(1) = 0$, and

$$g_k'(1) = -g_{k-1}''(1)g(1) - g_{k-1}'(1)g'(1) = \alpha g_{k-1}'(1) = \cdots = \alpha^{k-1} g_1'(1) = -\alpha^k, \tag{28}$$

because $g(1) = 0$ and $\alpha = -g'(1)$. Therefore, for all $k > 1$,

$$\lim_{X\to 1} \frac{g_k(X)}{\alpha^{k-1}g(X)} = \lim_{X\to 1} \frac{-g_{k-1}'(X)g(X)}{\alpha^{k-1}g(X)} = \frac{-\lim_{X\to 1} g_{k-1}'(X)}{\alpha^{k-1}} = 1.$$

We can use Taylor's theorem to express $g(X) = [\alpha - r_g(X)](1-X)$, where the remainder

[49]For example, in the Stackelberg oligopoly with linear demand, the leader's quantity is independent of the number of followers. This observation has been made before, for example by Anderson and Engers (1992); Julien et al. (2012). In fact, Julien et al. (2012) derived this equilibrium characterization for arbitrary sequential oligopolies with linear demand (which implies $\alpha = 1$ here).

[50]Equivalently, it is identical to $\mathbf{n}^n$ in the first $T-1$ periods, but all remaining players are collected into period T.

term satisfies $\lim_{X\to 1} r_g(X) = 0$. Therefore, equation (28) implies that

$$\begin{aligned}
\lim_{n\to\infty} X^*(\mathbf{n}^n|T) = 1 = \lim_{n\to\infty} \sum_{k=1}^{T} S_k(\mathbf{n}^n|T) g_k(X^*(\mathbf{n}^n|T)) \\
&= \lim_{n\to\infty} \left[1 + \frac{r_g(X^*(\mathbf{n}^n|T))}{\alpha}\right] S(\mathbf{n}^n|T)(1 - X^*(\mathbf{n}^n)) \\
&\quad \times \sum_{k=1}^{T} \frac{\alpha^k S_k(\mathbf{n}^n|T)}{S(\mathbf{n}^n|T)} \frac{g_k(X^*(\mathbf{n}^n|T))}{\alpha^{k-1} g(X^*(\mathbf{n}^n|T))} \\
&= \lim_{n\to\infty} S(\mathbf{n}^n|T)(1 - X^*(\mathbf{n}^n)),
\end{aligned}$$

where

$$S(\mathbf{n}^n|T) = \sum_{k=1}^{T} \alpha^k S_k(\mathbf{n}^n|T) = \prod_{s=1}^{T-1} (1 + \alpha n_s^n)(1 + \alpha n_{T+}^n) - 1$$

is the weighted sum of all the measures of information. For individual efforts, let us take individual i from period $t \le T$; then

$$x_i^*(\mathbf{n}^n|T) = \frac{1}{n_t^n} \left[f_t(X^*(\mathbf{n}^n|T)) - f_{t-1}(X^*(\mathbf{n}^n|T))\right] = g(X^*(\mathbf{n}^n|T)) f_t'(X^*(\mathbf{n}^n|T)).$$

Note that we can express $f_t(X)$ as

$$f_t(X) = X - \sum_{k=1}^{T-t} S_k((\mathbf{n}^n|T)^t) g_k(X),$$

where $(\mathbf{n}^n|T)^t = (n_{t+1}^n, \dots, n_{T-1}^n, n_{T+}^n)$. By the results above, $\lim_{n\to\infty} S(\mathbf{n}^n|T) \frac{g(X^*(\mathbf{n}^n|T))}{\alpha} = \lim_{n\to\infty} S(\mathbf{n}^n|T)(1 - X^*(\mathbf{n}^n)) = 1$. Therefore, we get that

$$\begin{aligned}
\lim_{n\to\infty} x_i^*(\mathbf{n}^n|T) &= \lim_{n\to\infty} g(X^*(\mathbf{n}^n|T)) - \lim_{n\to\infty} g(X^*(\mathbf{n}^n|T)) \sum_{k=1}^{T-t} S_k((\mathbf{n}^n|T)^t) g_k'(X^*(\mathbf{n}^n|T)) \\
&= \lim_{n\to\infty} S(\mathbf{n}^n|T) g(X^*(\mathbf{n}^n|T)) \sum_{k=1}^{T-t} \frac{\alpha^k S_k((\mathbf{n}^n|T)^t)}{S(\mathbf{n}^n|T)} \frac{g_k'(X^*(\mathbf{n}^n|T)}{-\alpha^k} \\
&= \lim_{n\to\infty} \alpha \sum_{k=1}^{T-t} \frac{\alpha^k S_k((\mathbf{n}^n|T)^t)}{S(\mathbf{n}^n|T)} = \lim_{n\to\infty} \alpha \frac{\prod_{s=t+1}^{T-1}(1 + \alpha n_s^n)(1 + \alpha n_{T+}^n) - 1}{\prod_{s=1}^{T-1}(1 + \alpha n_s^n)(1 + \alpha n_{T+}^n) - 1} \\
&= \frac{\alpha}{\prod_{s=1}^{t}(1 + \alpha n_s^n)}.
\end{aligned}$$

$\square$

I General payoffs

The general payoff function where the results apply is $u_i(\mathbf{x}) = x_i h(X)$, where $h(X)$ is the marginal benefit of individual effort, which is a smooth and decreasing function of the total effort. In particular, I assume that $h : \mathbb{R}_+ \to \mathbb{R}$ is a $(T+1)$-times continuously differentiable strictly decreasing function with $h(\overline{X}) = 0$ for some $\overline{X} < \infty$. Without loss of generality, we can normalize $\overline{X} = 1$, so that $h(1) = 0$.

Instead of the $h(X)$ function, the relevant function for the equilibrium characterization is $g(X) = -\frac{h(X)}{h'(X)}$, which is the function that enters into the optimality conditions. Then $g(X)$ is T-times continuously differentiable, $g(1) = 0$, and $g(X) > 0$ for all $X \in (0, 1)$.

I have already described sufficient conditions for the results. For the characterization theorem, the sufficient condition was condition 1, which guarantees that the inverted best-response functions are well-behaved. Although the condition is not directly defined on primitives, it is a mathematical property of function g that can be easily verified.

Condition 1 is not a necessary condition, but it is constructed in a way that identifies situations where the analysis may fail if the condition is not satisfied. Part 3 (i.e., increasing f_t), guarantees that the inverted best-response functions f_t are invertible and therefore that the best-response functions are well-defined. Parts 1 and 2 guarantee that there is a unique combination of efforts that satisfy the individual first-order conditions which ensures the uniqueness of the equilibrium. The existence is guaranteed by two parts. First, it is crucial that the highest root of $f_0(X)$ is strictly positive because otherwise, all necessary conditions cannot be satisfied at the same time, and thus there are no pure-strategy equilibria. Second, assuming that there exist effort levels that satisfy individual first-order necessary conditions, they are also sufficient, because it is a unique local extremum which guarantees strictly higher payoff than the corner solution (which gives zero payoff) and is thus also global maximizer.

Conditions 1 and 2 jointly are sufficient conditions for theorem 2, all its corollaries, and the earlier-mover advantage. Again, the conditions are not necessary, but they are constructed in a way that makes it clear why the information theorem may not hold without these assumptions. The result relies on the characterization from theorem 1, so if condition 1 is not satisfied, then the result may not hold. Moreover, the proof uses condition 1 directly; it shows that X^* increases because $f_0(X)$ is increasing above its highest root. Condition 2 is used in a transparent way as well: if efforts were strategic complements on some level, then contests that increase information at these levels may reduce the total effort. Appendix K.5 provides an example.

It is natural to ask whether there are standard payoff functions other than Tullock payoffs that satisfy conditions 1 and 2. Appendix J describes a class of standard functional form assumptions for which both conditions are always satisfied: completely monotone functions and T-times monotone functions. Appendix K provides a few examples: appendix K.1 contains Tullock payoffs; in appendices K.2 and K.3 function g is completely monotone, whereas in appendix K.4 it is neither Tullock nor monotone. Finally, in Appendix K.5 condition 1 is violated, and there are no equilibria.

J T-times monotone payoff functions

A class of payoff functions, where conditions 1 and 2 can be relatively easily verified, are payoffs with T-times monotone g functions. Monotonicity assumptions are convenient, since checking the signs everywhere is usually simpler than checking signs in relevant (typically endogenous) ranges, and the same is true here. To make a connection with the standard assumptions, note that in a simultaneous contest where $T = 1$, the assumption only requires that $g(X) = -\frac{h(X)}{h'(X)}$ is decreasing in X. Since $\frac{d \log h(X)}{dX} = \frac{h'(X)}{h(X)} = -\frac{1}{g(X)}$, this assumption is equivalent to assuming that $h(X)$ is log-concave. This is a standard assumption in most applications with simultaneous decisions and guarantees that each player chooses an interior optimum.

In dynamic contests we have to extend the assumption to higher-order impacts, and T-times monotonicity is easy to verify and is satisfied in many standard functional form assumptions, including many polynomial, exponential, and logarithmic functions. Also, sums and products of T-times monotone functions are T-times monotone, so the class of functional forms satisfying the assumption is very large.

A function g is *T-times monotone*[51] (or *multiply monotone*) in $[0, 1]$ if it is T-times continuously differentiable in $[0, 1]$, and satisfies

$$(-1)^k \frac{d^k g(X)}{dX^k} \geq 0, \quad \forall X \in [0, 1], \forall k \leq T. \tag{29}$$

Function g is *completely monotone* (or *totally monotone*) if it is m-times monotone for any $m \in \mathbb{N}$.

There are many functions that are completely monotone or T-times monotone. Lemmas 14 and 15 give a few examples, and lemma 16 shows that all sums and products of T-times monotone functions are also T-times monotone.

Lemma 14. *The following functions are completely monotone in $[0, 1]$ for any $\alpha > 0$*

1. $g(X) = \alpha(1 - X^m)$ *for any* $m \in \mathbb{N}$

[51] By definition, a function could be either increasing or decreasing and T-times monotone, but since I assume that $g(X) > 0$ for $X \in (0, 1)$ and $g(X) < 0$ for $X > 1$, this means that g must be decreasing.

2. $g(X) = \alpha(1-X)^m$ *for any* $m \in \mathbb{N}$,

3. $g(X) = \alpha((X+\gamma)^s - (1+\gamma)^s)$ *for any* $s < 0, \gamma > 0$,

4. $g(X) = \alpha[e^{-rX} - e^{-r}]$ *for any* $r > 0$,

5. $g(X) = -\alpha \log(X)$.

Lemma 15. *Function* $g(X) = \alpha[(1+\gamma-X)^{m-\varepsilon} - \gamma^{m-\varepsilon}]$ *with* $\alpha > 0, \gamma > 0$ *and* $\varepsilon \in (0,1)$ *is* m*-times monotone, but not* $(m+1)$*-times monotone.*

Let me also introduce some few very useful well-known properties of m-times monotone functions.

Lemma 16. *If* $g(X)$ *and* $f(X)$ *are decreasing* m*-times monotone functions, then*

1. $g(X)$ *is a decreasing* $\widehat{m}$*-times monotone function for each* $\widehat{m} < m$.

2. $ag(X)$ *is a decreasing* m*-times monotone function for all* $a \geq 0$ *and an increasing* m*-times monotone function for all* $a \leq 0$.

3. $g(X) + f(X)$ *is a decreasing* m*-times monotone function.*

4. $g(X)f(X)$ *is an decreasing* m*-times monotone function.*

5. $-g'(X)$ *is a decreasing* $(m-1)$*-times monotone function.*

6. $\int_X^a g(t)dt$ *is a decreasing* $(m+1)$*-times monotone function in* $[0,a]$ $g(X) \geq 0$ *for all* $X \in [0,a]$.

Proposition 5 (T-times monotone g satisfies condition 1)**.** *If* g *is* T*-times monotone in* $[0,1]$*, then condition 1 is satisfied.*

Proof. It is first useful to show that each f_t inherits the monotonicity properties of g.

Lemma 17. *If* g *is* T*-times monotone in* $[0,1]$*, then each* $f_t(X)$ *is a strictly increasing* $(t+1)$*-times monotone function in* $[0,1]$.

Proof. First, $f_T(X) = X$ is strictly increasing and completely monotone, so also $(T+1)$-times monotone. Suppose that $f_t(X)$ is a strictly increasing $(t+1)$-times monotone function. Then $f_t'(X)$ is a decreasing t-times monotone function. Therefore, $f_t'(X)g(X)$ is a product of a two decreasing t-times monotone function and thus is also a decreasing t-times monotone function. Then $-n_t f_t'(X)g(X)$ is an increasing t-times monotone function. Finally,

$$f_{t-1}(X) = f_t(X) - n_t f_t'(X)g(X)$$

is the sum of two increasing t-times monotone functions. Finally, since $-n_t f_t'(X)g(X)$ is an increasing function, we have that

$$f_{t-1}'(X) = f_t'(X) + \frac{d(-n_t f_t'(X)g(X))}{dX} \geq f_t'(X) \geq \cdots \geq f_T'(X) = 1 > 0.$$

Therefore, $f_{t-1}(X)$ is a strictly increasing t-times monotone function. $\square$

With this, we can prove the claim by induction. Clearly all assumptions of condition 1 are satisfied for $f_T(X) = X$. Suppose that they are satisfied for f_t and we need to show that they are also then satisfied for $f_{t-1}(X)$. By lemma 17, each $f_t(X)$ is a strictly increasing ($(t+1)$-times monotone) function in $[0,1]$, thus $f_t'(1) < \infty$, we get

$$f_{t-1}(1) = f_t(1) - n_t f_t'(1)g(1) = f_t(1) = \ldots = f_T(1) = 1$$

Moreover, at $\underline{X}_t$,

$$f_{t-1}(\underline{X}_t) = f_t(\underline{X}_t) - n_t f_t'(\underline{X}_t)g(\underline{X}_t) \leq 0,$$

as $f_t(\underline{X}_t) = 0$, $f_t'(\underline{X}_t) > 0$ (since f_t is strictly increasing) and $g(\underline{X}_t) \geq 0$. As f_{t-1} is strictly increasing and continuous in $[0,1]$, it must have exactly one root in $\underline{X}_{t-1} \in [\underline{X}_t, 1)$.

Finally, since $f_{t-1}(X)$ is strictly increasing in $[0,1]$, we must have that $f_{t-1}(X) < f_{t-1}(\underline{X}_{t-1}) = 0$ for all $X \in [\underline{X}_t, \underline{X}_{t-1})$ and $f_{t-1}'(X) > 0$ for all $X \in [\underline{X}_{t-1}, 1]$. $\square$

Proposition 6 (T-times monotone g satisfies condition 2)**.** *If g is T-times monotone in $[0,1]$, then condition 2 is satisfied.*

Proof. I first show by induction that g_k is a strictly decreasing $(T+1-k)$-times monotone function in $[0,1]$. It is clearly true for $g_1(X) = g(X)$. Suppose that this is true for k. Then $g_{k+1}(X) = -g_k'(X)g(X)$ is strictly decreasing, as

$$g_{k+1}'(X) = -g_k''(X)g(X) - g_k'(X)g'(X) < 0,$$

because $g_k''(X) \geq 0$, $g(X) \geq 0$, $g_k'(X) < 0$ and $g'(X) < 0$. It is also $(T-k)$-times monotone as for all $m \leq T-k$

$$(-1)^m \frac{d^m g_{k+1}(X)}{dX^m} = (-1)^{m+1} \frac{d^{m+1} g_k(X)}{dX^{m+1}} g(X) (-1)^m \frac{d^m g_{k+1}(X)}{dX^m} (-g'(X)) \geq 0,$$

because $(-1)^{m+1} \frac{d^{m+1} g_k(X)}{dX^{m+1}} \geq 0$ and $(-1)^m \frac{d^m g_{k+1}(X)}{dX^m} (-g'(X)) \geq 0$, since g_k is at least $(m+1)$-times monotone, $g(X) \geq 0$, and $-g'(X) > 0$.

Now, note that $g_k(1) = 0$ for all k. Since $X^* \in (0,1)$ and g_k is strictly decreasing in $[0,1]$, we get that $g_k(X^*) > g_k(1) = 0$. $\square$

K Examples

K.1 Example: Tullock contest with four players

Let us take a three-period four-player contest $\mathbf{n} = (1, 2, 1)$. Starting from $f_3(X) = X$ and applying the formula $f_{t-1}(X) = f_t(X) - n_t f_t'(X)X(1 - X)$ three times, we get

$$\begin{aligned}
&f_2(X) = X - X(1 - X) = X^2,\\
&f_1(X) = X^2 - 2(2X)X(1 - X) = X^2(4X - 3),\\
&f_0(X) = X^2(4X - 3) - (12X^2 - 6X)X(1 - X) = X^2(12X^2 - 14X + 3).
\end{aligned}$$

The highest root of $f_0(X)$ is the total equilibrium effort $X^* = \frac{1}{12}\left(\sqrt{13} + 7\right) \approx 0.8838$. The individual efforts can be computed as $x_i^* = \frac{1}{n_t}[f_t(X^*) - f_{t-1}(X^*)]$. In particular,

$$\begin{aligned}
x_1^* &= \frac{1}{432}\left(\sqrt{13} - 2\right)\left(\sqrt{13} + 7\right)^2 \approx 0.4180,\\
x_2^* = x_3^* &= \frac{1}{864}\left(5 - \sqrt{13}\right)\left(\sqrt{13} + 7\right)^2 \approx 0.1815,\\
x_4^* &= \frac{1}{144}\left(5 - \sqrt{13}\right)\left(\sqrt{13} + 7\right) \approx 0.1027.
\end{aligned}$$

K.2 Example: linear g

Perhaps the simplest class of payoff functions is $h(X) = d\sqrt[\alpha]{1 - X}$ with $d, \alpha > 0$, because this implies linear $g(X) = \alpha(1 - X)$, which is clearly completely monotone. By the argument in appendix H, the total equilibrium effort in this case can be simply (exactly) calculated as $X^* = 1 - \frac{1}{S(\mathbf{n})}$, where $S(\mathbf{n}) = \sum_{k=1}^{T} S_k(\mathbf{n})\alpha^k$, and the individual equilibrium efforts are $x_i^* = \frac{\alpha}{\prod_{s=1}^{t}(1+\alpha n_s)}$.

K.3 Example: completely monotone g

Suppose that $\mathbf{n} = (1, 2, 1)$ and $h(X) = d\sqrt[\alpha]{c - b^X}$, where $\alpha > 0, b > 1, c > 0$, and $d > 0$. Then $h(\overline{X}) = 0$ defines $\overline{X} = \log_b c$. Without loss of generality, we can rescale X such that $\overline{X} = 1$, which implies $c = b$. Then

$$g(X) = -\frac{h(X)}{h'(X)} = -\frac{c(b - b^X)^{\frac{1}{a}}}{c\frac{1}{a}(b - b^X)^{\frac{1}{a}-1}(-\log b)b^X} = \alpha\frac{b^{1-X} - 1}{\log b} \tag{30}$$

and $(-1)^k \frac{d^k g(X)}{dX^k} = \alpha(\log b)^{k-1}b^{1-X} > 0$ for all X and all k. Therefore, g is completely monotone and condition 1 is satisfied. Note that $\alpha = -g'(1)$. Applying the characterization result gives

$$\begin{aligned}
f_3(X) &= X, \\
f_2(X) &= X - \alpha \frac{b^{1-X} - 1}{\log b}, \\
f_1(X) &= X - \left(3\alpha + 2\alpha^2 b^{1-X}\right) \frac{b^{1-X} - 1}{\log b} \\
f_0(X) &= X - \left(4\alpha + 5\alpha^2 b^{1-X} + 2\alpha^3 (2b^{1-X} - 1) b^{1-X}\right) \frac{b^{1-X} - 1}{\log b}.
\end{aligned}$$

Since g is completely monotone, lemma 17 implies that each f_t is strictly increasing and proposition 5 implies that the roots satisfy $0 = \underline{X}_3 < \underline{X}_2 < \underline{X}_1 < \underline{X}_0 = X^* < 1$.

For particular parameter values, it is straightforward to compute the equilibrium numerically. For example, let $\alpha = \frac{1}{2}$ and $b = 2$, so that $g(X) = \alpha \frac{b^{1-X}-1}{\log b} = \frac{2^{-X}-2^{-1}}{\log 2}$, then $\underline{X}_2 \approx 0.3841$, $\underline{X}_1 \approx 0.7015$, and $\underline{X}_0 \approx 0.8030$. Therefore, the total equilibrium effort is $X^* = \underline{X}_0 \approx 0.8030$, and the individual efforts are $\mathbf{x}^* \approx (0.3653, 0.1661, 0.1661, 0.1056)$. Figure 5 illustrates the functions in this case.

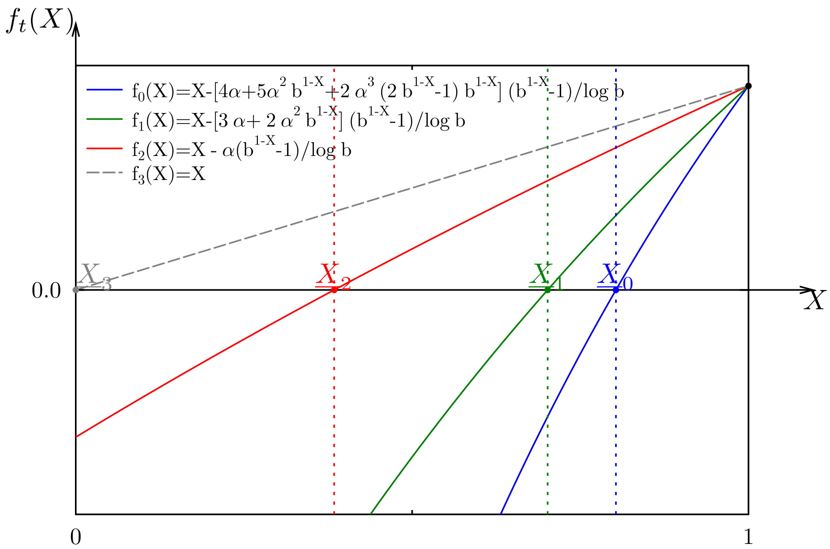

Figure 5: Illustration in the case $\mathbf{n} = (1, 2, 1)$, $g(X) = \alpha \frac{b^{1-X}-1}{\log b}$, $\alpha = \frac{1}{2}$, and $b = 2$.

K.4 Example: an oligopoly with logarithmic demand

Take an oligopoly with n_1 leaders and n_2 followers, inverse demand $P(X) = 1 - \log X$, and marginal cost $c = 1$. Then $h(X) = -\log X$ and $g(X) = -\frac{h(X)}{h'(X)} = -X \log X$, which is neither a Tullock payoff nor monotone. Therefore, we have to verify condition 1 directly.

Applying the recursive rule gives

$$\begin{aligned} f_2(X) &= X, \\ f_1(X) &= X(1 + n_2 \log X), \\ f_0(X) &= X(1 + (n_1 + n_2 + n_1 n_2) \log X + n_1 n_2 (\log X)^2). \end{aligned}$$

Figure 6 depicts the functions. The highest root of $f_1(X)$ is the solution to $1 + n_2 \log X = 0$, which is $\underline{X}_1 = e^{-\frac{1}{n_2}} > \underline{X}_2 = 0$, and $f_1(X)$ is negative in $[0, \underline{X}_1]$ and strictly increasing in $[\underline{X}_1, 1]$.

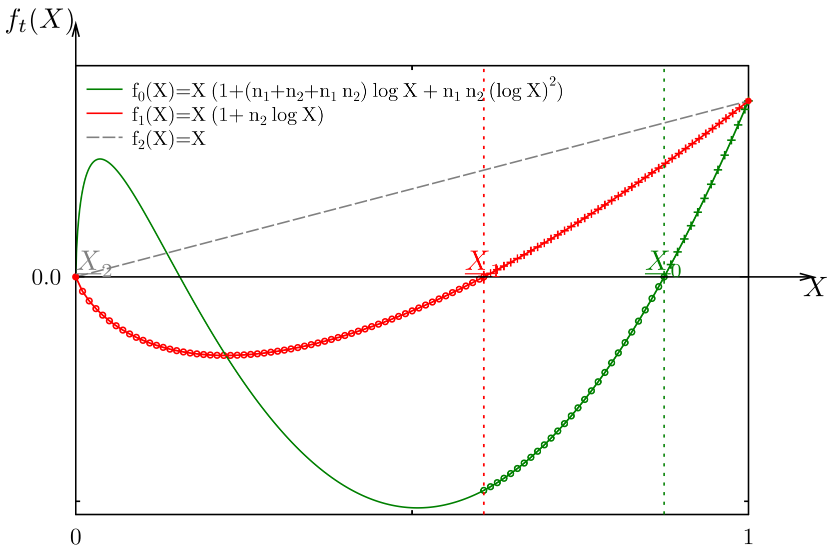

Figure 6: Illustration of condition 1 in the case $\mathbf{n} = (2, 2)$ and $g(X) = -X \log X$. Line segments marked with circles are negative and line segments marked with plusses are strictly increasing.

Now, the expression in the parenthesis of $f_0(X)$ is a quadratic function of $\log X$ that has two roots, both in $(0, 1)$. The function f_0 is strictly increasing above its highest root, so to verify condition 1 it suffices to verify that at $f_0(\underline{X}_1) < 0$. Since $\log \underline{X}_1 = -\frac{1}{n_2}$, this is equivalent to

$$f_0(\underline{X}_1) = \underline{X}_1 \frac{1}{n_2} (n_2 - n_1 - n_2 - n_1 n_2 + n_1) = -n_1 \underline{X}_1 < 0$$

Now, the highest root of $f_0(X)$ is the solution to

$$1 + (n_1 + n_2 + n_1 n_2) \log X + n_1 n_2 (\log X)^2 = 0,$$

which is a quadratic equation of $\log X$ and gives

$$\log \underline{X}_0 = \frac{\sqrt{(n+p)^2 - 4p} - (n+p)}{2p} \quad \iff \quad \underline{X}_0 = e^{\frac{\sqrt{(n+p))^2-4p}-(n+p)}{2p}},$$

where $n = n_1 + n_2$ and $p = n_1 n_2$. Then $\underline{X}_0 > \underline{X}_1$, $f_0(X) < 0$ between and strictly increasing above $\underline{X}_0$. The equilibrium effort is $X^* = \underline{X}_0$

For example, if $n_1 = n_2 = 2$, then $X^* = e^{\frac{\sqrt{3}}{2}-1} \approx 0.8746$. On the other hand, if all 4 players were to make simultaneous decisions, then the equilibrium would be $X^* = e^{-\frac{1}{4}} \approx 0.7788$. If $\mathbf{n} = (1,3)$ or $\mathbf{n} = (3,1)$, then $X^* = e^{\frac{\sqrt{37}-7}{6}} \approx 0.8582$.[52]

K.5 Example: payoffs that do not satisfy necessary conditions

Suppose that the payoff function is $h(X) = \left[\frac{1-X}{X}\right]^2$. Then

$$\begin{aligned}
g_1(X) = g(X) &= \frac{1}{2}X(1-X), \\
g_2(X) = -g_1'(X)g(X) &= \frac{1}{4}X(2X-1)(1-X), \\
g_3(X) = -g_2'(X)g(X) &= \frac{1}{8}X(6X^2-6X+1)(1-X), \\
g_4(X) = -g_3'(X)g(X) &= \frac{1}{16}X(2X-1)(12X^2-12X+1)(1-X).
\end{aligned}$$

With these functions, it is straightforward to illustrate multiple possibilities illustrating how the results may not hold if condition 1 or condition 2 is not satisfied.

Consider first the two-player simultaneous contest $\mathbf{n} = (2)$. Then

$$f_0(X) = X - 2g_1(X) = X^2.$$

Therefore, the highest root is $\underline{X}_0 = 0$, which means that condition 1 is not satisfied, and thus theorem 1 does not apply. Indeed, the individual optimization problem is

$$\max_{x_i \geq 0} x_i \left[\frac{1}{x_1+x_2} - 1\right]^2 \quad \Rightarrow \quad x_i^*(x_{-i}) = \frac{1}{2}\left[\sqrt{8x_{-i}+1} - 2x_{-i} - 1\right], \forall x_{-i} \in (0,1].$$

From this, it is straightforward to check that there are no equilibria in pure strategies with $x_1, x_2 > 0$. However, by the assumption payoff from $x_i = 0$ is 0, so $(0,0)$ cannot be an equilibrium either, because any $x_i > 0$ is then a profitable deviation.

[52] The corresponding individual efforts are $\mathbf{x}^*((2,2)) \approx (0.3201, 0.3201, 0.1172, 0.1172)$, $\mathbf{x}^*((4)) \approx (0.1947, 0.1947, 0.1947, 0.1947)$, $\mathbf{x}^*((1,3)) \approx (0.4646, 0.1312, 0.1312, 0.1312)$, and $\mathbf{x}^*((3,1)) \approx (0.2423, 0.2423, 0.2423, 0.1312)$.

Consider now, a simultaneous four-player contest $\mathbf{n} = (4)$, such that

$$f_0(X) = X - 4g_1(X) = 2X\left(X - \frac{1}{2}\right).$$

The total equilibrium effort is $X^* = \frac{1}{2}$ and the individual efforts are $x_i^* = \frac{1}{8}$. If we take a four-player two-period contest, i.e. $\widehat{\mathbf{n}} = (\widehat{n}_1, \widehat{n}_2)$ such that $\widehat{n}_1 + \widehat{n}_2 = 4$, then

$$\widehat{f}_0(X) = X - 4g_1(X) - n_1n_2g_2(X),$$

which has the same highest root (i.e., $\widehat{X}^* = \frac{1}{2}$, and individual efforts are $x_i^* = \frac{1}{8}$. It is straightforward to verify that in all these cases, condition 1 is satisfied, but since $g_2(\frac{1}{2}) = 0$, condition 2 is not satisfied. Indeed, we see that the additional disclosure in $\widehat{\mathbf{n}}$ compared to $\mathbf{n}$ does not increase total effort and there is no earlier-mover advantage.

The result is even starker if we add another disclosure, for example when $\widehat{\mathbf{n}} = (1, 2, 1)$, then $\mathbf{S}(\widehat{\mathbf{n}}) = (4, 5, 2)$ and so

$$\widehat{f}_0(X) = X - 4g_1(X) - 5g_2(X) - 3g_3(X) = \frac{3}{2}X^3\left(X - \frac{1}{3}\right).$$

Therefore, the total equilibrium effort is $\widehat{X}^* = \frac{1}{3} < \frac{1}{2}$ and the individual efforts are $x_1^* = \frac{1}{27} < x_2^* = x_3^* = \frac{5}{54} < x_4^* = \frac{1}{9}$. That is, the additional disclosure decreases total effort and there is a strict later-mover advantage. Again, it is straightforward to check that condition 1 is satisfied, but condition 2 is not, as $g_2\left(\frac{1}{3}\right) = -\frac{1}{54} < 0$ and $g_3\left(\frac{1}{3}\right) = -\frac{1}{108} < 0$.

Finally, if we add the final possible disclosure to get the sequential four-player contest $\widehat{\mathbf{n}} = (1, 1, 1, 1)$, then $\mathbf{S}(\widehat{\mathbf{n}}) = (4, 6, 4, 1)$, so that

$$\widehat{f}_0(X) = X - 4g_1(X) - 6g_2(X) - 4g_3(X) - g_4(X) = \frac{X}{16}(24X^4 - 12X^3 + 2X^2 - 31X + 33).$$

The polynomial $24X^4 - 12X^3 + 2X^2 - 31X + 33 > 0$ for all $X \in \mathbb{R}$, so it cannot have any real roots in $[0, 1]$. Therefore, the highest root of $f_0(X) = 0$ and thus there are no equilibria in pure strategies (and condition 1 is not satisfied).