Theoretical Economics 19 (2024), 207–244 1555-7561/20240207

Optimal sequential contests

Toomas Hinnosaar
School of Economics, University of Nottingham and CEPR

I study sequential contests where the efforts of earlier players may be disclosed to later players by nature or by design. The model has many 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 show that in addition to the first-mover advantage, there is an earlier-mover advantage. Finally, I derive the limits for large contests and discuss the limit to perfectly competitive outcomes under different disclosure rules.

Keywords. Contest design, oligopoly, public goods, rent-seeking, R&D.

JEL classification. C72, C73, D72, D74, D82.

1. Introduction

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

In this paper, I study contests where the effort choices are not necessarily simultaneous. In many real-life situations, some players can observe their competitors' efforts and respond appropriately to those choices. However, earlier movers can also anticipate these subsequent responses and, therefore, influence the behavior of later movers. Each

Toomas Hinnosaar: toomas@hinnosaar.net

I am grateful to Stefano Barbieri, 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 Monzón, Peter Norman, Marco Ottaviani, Mallesh Pai, Alessandro Pavan, Alex Possajennikov, Debraj Ray, Christian Seel, Marco Serena, Ron Siegel, Andy Skrzypacz, Martin Szydlowski, Jean Tirole, and Asher Wolinsky 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, University of Bristol, University of Bonn, Humboldt University of Berlin, 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, Markets with Informational Asymmetries workshop (Turin), IIOC, SEA, Lingnan Workshop (Guangzhou), QuaGaTCo virtual conference, and China Meeting of the Econometric Society for their comments and suggestions. I would also like to thank the Kellogg School of Management at Northwestern University and the Collegio Carlo Alberto where some of this work was carried out.

Available at https://econtheory.org. https://doi.org/10.3982/TE5536

additional period in a sequential contest adds complexity to the analysis, which might explain why previous studies have focused mainly on simultaneous models. I characterize equilibria for a general class of sequential contests and analyze how the information about other players' efforts influences the equilibrium behavior.

Contests may be sequential by nature or by design. For example, in rent-seeking contests, firms lobby the government to achieve market power. One tool that regulators can use to minimize such rent-seeking is a disclosure policy. A nontransparent disclosure policy would lead to simultaneous effort choices, but a full transparency policy would lead to a fully sequential contest. There may be potentially intermediate solutions as well, where the information is revealed only occasionally. Over the last few decades, many countries have introduced new legislation regulating transparency in lobbying activity. This list includes 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, lobbying efforts in the US must be reported quarterly, whereas in the EU, reporting occurs annually and on a more voluntary basis.

Another classic example of a contest is research and development (R&D), where the probability of a scientific breakthrough is proportional to agents' research efforts. The question is how to best organize the disclosure rules to maximize aggregate research efforts. In some academic fields, it is common to present early findings in working papers and conferences. In other fields, these efforts are kept confidential until the work has been vetted and published in a journal. Similarly, when announcing an R&D contest, the organizer can choose a transparency level: whether to use a public leaderboard or perhaps keep the entries secret until the deadline.

To address such questions, I study a model of sequential contests. First, I characterize all equilibria for any given sequential contest, i.e., for any fixed disclosure rule. The standard backward-induction approach requires finding best-response functions every period and substituting them recursively. This solution method is not generally tractable or even feasible. Instead, I use an alternative approach, in which I characterize best-response functions by inverse functions. This method pools all the optimality conditions into one necessary condition and solves the resulting equation just once. I prove that for any contest the equilibrium exists and is unique. Importantly, the characterization theorem shows how to compute the equilibrium.

The main result of the paper shows that the information about other players' efforts strictly increases the total effort. Consequently, 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. When the efforts are undesirable (as in rent-seeking), the optimal contest is one with hidden efforts. The intuition behind this result is simple. While players' efforts could be strategic substitutes or complements, I show that efforts are strategic substitutes sufficiently close to the equilibrium. Therefore, earlier-moving players have an additional incentive to exert effort to discourage later players' efforts. If the discouragement effect were strong enough to reduce the total effort, this would offer profitable deviations for some players. Therefore, the discouragement effect is less than one-to-one. It increases earlier-movers efforts more than it reduces later-movers efforts, therefore increasing total effort. While there could be indirect effects that change the conclusions, I show that (again, near the equilibrium) efforts are higher-order strategic substitutes and, therefore, the result still holds.

The information about other players' efforts is important both qualitatively and quantitatively. For example, the sequential contest with 5 players ensures a higher total effort than the simultaneous contest with 24 players. The differences become even larger with larger contests. For example, a contest with 14 sequential players achieves a 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 by Dixit (1987), who showed that a player who pre-commits chooses a greater level of effort and obtains a higher payoff than his followers. This leader exploits two advantages: he moves earlier and has no direct competitors. With the characterization result, I can further explore this question and compare players' payoffs and effort levels in sequential contests. I show that there is a strict *earlier-mover advantage*—earlier players choose greater efforts and obtain higher payoffs than later players.

Finally, I provide insights for large contests. I derive an approximation result for contests with an infinitely large number of players. This result allows me to show that as the number of players becomes large, the total effort converges to the prize's value (or perfectly competitive outcome more generally) regardless of the contest structure. However, the speed of convergence to this level is different under different disclosure policies. In simultaneous contests, the rate of convergence is linear, whereas in sequential contests it is exponential.

These results paint a different picture of highly competitive strategic interactions. In simultaneous contests, a high degree of competitiveness requires a large number of players, all choosing a minuscule effort level. Contrastingly, in a sequential contest, the same total effort requires a much smaller number of players, each exerting different effort levels. The first player chooses a much higher effort than anyone else, the second one much higher than the first, and so on. By any definition, this is a highly concentrated market. However, the early movers cannot capitalize on their position, as later movers would react by increasing their efforts. Therefore, despite the different effort levels, their payoffs are still close to zero. These results thus provide an alternative foundation for the contestability theory (Baumol (1982)). Instead of introducing a separate class of inactive players—the competitive fringe—in this model, the competitive fringe arises endogenously through the order of moves.

Literature: The simultaneous version of the model has been studied extensively, starting from Cournot (1838). The literature on Tullock contests was initiated by Tullock (1967, 1974) and motivated by rent-seeking (Krueger (1974)).[1] The most general treatment of simultaneous contests is provided by the literature on aggregative games (Selten

[1] See Nitzan (1994), Konrad (2009), and Vojnović (2015) for literature reviews on contests.

(1970), Acemoglu and Jensen (2013), Jensen (2018)). My model is an aggregative game only in the simultaneous case.

The only sequential contest that has been studied extensively is the first-mover contest. It was introduced by von Stackelberg (1934), who studied quantity leadership in an oligopoly. Dixit (1987) showed that there is a first-mover advantage in contests. Relatively little is known about (Tullock) contests with more than two periods. The only paper prior to this that studied sequential Tullock contests with more than two periods is Glazer and Hassin (2000), which characterized the equilibrium in the sequential three-player Tullock contest. Kahana and Klunover (2018) is an independent and concurrent work that uses a similar approach to characterize the equilibrium in an important special case of my model: an n-player fully sequential Tullock contest. In contrast to my paper, they do not study any of the questions that are the main focus of my paper, such as the optimal contests, earlier-mover advantage, and large contests. Moreover, as I argue in Section 8, the characterization alone is not sufficient to answer these questions. The only class of contests where equilibria are fully characterized for sequential contests are oligopolies with linear demand.[2]

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

The paper also contributes to the contest design literature. Previous papers on contest design include Taylor (1995), Che and Gale (2003), Moldovanu and Sela (2001, 2006), and Olszewski and Siegel (2016), which have focused on contests with private information. Halac, Kartik, and Liu (2017) studied contest design in the presence of informational externalities when players learn about the feasibility of the project. In this paper, I study contest design on a different dimension: how to optimally disclose other players' efforts, when players move sequentially, to minimize or maximize total effort.[3]

Similar connections between disclosures and subsequent actions have been found in other settings. For example, Fershtman and Nitzan (1991), Varian (1994), and Wirl (1996) used a model of dynamic voluntary public goods provision to show that if contributions are adjusted after observing earlier contributions, this may increase the free-riding problem. Admati and Perry (1991) and Bonatti and Hörner (2011) showed similar

[2]Daughety (1990) used such a model to show that an oligopoly where players are divided between two periods is more concentrated but also closer to competitive equilibrium than an oligopoly where all players move at once. Hinnosaar (2021) provides a literature review and shows that the linear oligopoly model has unique properties that fail when the demand is not linear.

[3]Recently, Ely, Georgiadis, Khorasani, and Rayo (2022) studied feedback design in a continuous-time model where the designer wants to prolong participation for as long as possible and contestants do not observe their successes.

effects in dynamic team production problems. While the driving forces in these papers are similar to the discouragement effect studied herein, none of these works addressed higher-order effects and their implications for resulting equilibria.[4]

The paper also helps to explain empirical findings. For example, there is widespread empirical evidence of earlier-mover advantage in consumer goods markets. According to a survey by Kalyanaram, Robinson, and Urban (1995), there is a negative relationship between a brand's entry time and the brand's market share in many mature markets, including pharmaceutical products, investment banks, semiconductors, and drilling rigs. For example, Bronnenberg, Dhar, and Dubé (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. Lemus and Marshall (2021) used observational data and a lab experiment to study the impact of public leaderboards in prediction contests. They found that public leaderboards encouraged some players and discouraged others, but the overall effect was positive, improving the prediction contest's quality.

The rest of the paper unfolds as follows. Section 2 introduces the model. Section 3 uses a three-player example to illustrate why the standard backward induction is not tractable and shows how the inverted best-response approach solves the tractability problem. Section 4 provides the characterization result. Section 5 discusses the second main result, connecting information and total effort, and discusses its implications. Section 6 studies earlier-mover advantage and Section 7 analyzes large contests. Section 8 shows how the analysis applies to a broader class of models. Finally, Section 9 concludes. All proofs are in Appendix A.

2. Model

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. 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 $t-1$.[5] I refer to the time interval in which players in the 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.[6]

[4]More broadly, there is a connection with the sequential information design literature. For example, Doval and Ely (2020) and Makris and Renou (2023) provide characterization results in sequential models where information design may involve signals about players' actions in addition to unknown types or states. Li and Norman (2021) study a sequential persuasion model and find that players generally want to move only once.

[5]As the payoffs depend on the total effort of other players and not their individual efforts, observing the sum of previous players' efforts is equivalent to observing their individual efforts.

[6]Equivalently, the model can be stated as follows: n players are divided across T time periods, either exogenously or by the contest designer.

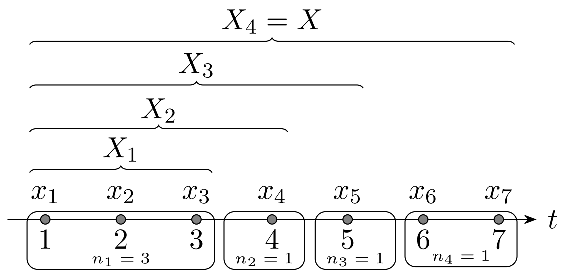

Figure 1. A contest with 7 players and 3 disclosures. Players 1 to 3 choose efforts x_1, x_2, and x_3 independently; player 4 observes $X_1 = x_1 + x_2 + x_3$, player 5 observes $X_2 = X_1 + x_4$, and players 6 and 7 observe $X_3 = X_2 + x_5$.

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 up (and including) to period t by $X_t = \sum_{s=1}^{t} \sum_{i \in \mathcal{I}_s} x_i$. By construction, the cumulative effort before the contest is $X_0 = 0$, and the cumulative effort after period T is the total effort exerted during the contest, i.e., $X_T = X$. Figure 1 illustrates the notation with an example of the four-period contest $\mathbf{n} = (3, 1, 1, 2)$:

Players compete for a prize of size one, the probability of winning is proportional to the level of effort, and the marginal cost of effort is one. I therefore assume the *normalized Tullock payoffs*, with

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

I study pure-strategy subgame-perfect equilibria, 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 equilibrium. Throughout the 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 upon arrival. Fourth, disclosures make cumulative efforts public.[7] In Sections 8 and 9, I discuss the extent to which the results rely on each of these assumptions and explain how the results extend to more general sequential games.

3. Example

The standard Tullock contest has n identical players who make their choices in isolation. Each player i chooses effort x_i to maximize payoff (1). The optimal efforts have to satisfy the first-order condition

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

[7] Specifically, each player observes the sum of earlier-movers' efforts with certainty and unconditionally. More complex disclosure rules would change the conclusions. For example, probabilistic disclosures may limit the earlier-movers commitment power (Bagwell (1995)) and conditional disclosures may substantially expand the set of possible outcomes (Bizzotto, Hinnosaar, and Vigier (2023)).

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

3.1 *The problem with standard backward induction*

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, $X_2 = x_1 + x_2 < 1$ and maximizes the payoff. The optimality condition for player 3 is

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

Solving it for x_3 gives the best-response function $x_3^*(X_2) = \sqrt{X_2} - X_2$.[8] Now, player 2 observes $x_1 < 1$ and knows $x_3^*(X_2)$ and, therefore, solves the maximization problem

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

The optimality condition for player 2 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

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

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, the direct approach is not generalizable for an arbitrary number of players. In fact, the best response function does not have an explicit representation for contests with a larger number of periods.

[8] In this example, I focus only on interior solutions. It is straightforward to verify that corner solutions cannot occur in equilibrium, as they require that at least one player chooses an effort level giving inducing a nonpositive payoff, and there is always a deviation with a strictly positive payoff.

3.2 *Inverted best-response approach*

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

To see how the 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 $x_3 = X - X_2$, i.e.,

$$\max_{X \geq X_2} \frac{X - X_2}{X} - (X - X_2).$$

Differentiating the objective with respect to X gives us the optimality condition

$$\frac{1}{X} - \frac{X - X_2}{X^2} - 1 = \frac{X_2}{X^2} - 1 = 0,$$

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

We can now think of player 2's problem as choosing $X \geq X_1 = x_1$, which he can induce by making sure that the cumulative effort up to his move is $X_2 = f_2(X)$, setting $x_2 = f_2(X) - X_1$. Therefore, his maximization problem can be written as

$$\max_{X \geq X_1} \frac{f_2(X) - X_1}{X} - \big(f_2(X) - X_1\big).$$

Again, differentiating with respect to X, we get the optimality condition

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

This is the key equation that shows the advantage of the inverted best-response approach. Equation (5) is nonlinear 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 increases with each step of the recursion. However, (5) is linear in X_1, making it easy to derive the inverted best-response function

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

The condition $X_1 = f_1(X)$ aggregates the two necessary conditions of equilibrium into one, by capturing the best responses of players 2 and 3. It simply states that if the total effort at the end of the contest is X, then the cumulative effort X_1 had to be $f_1(X)$ after player 1. Otherwise, either player 2 or player 3 is not behaving optimally.

Note that $X < 1/2$ cannot be induced by any x_1, as even if $x_1 = 0$, the total effort chosen by players 2 and 3 would be $1/2$. Inducing total effort below $1/2$ would require player 1 to exert negative effort, which is not possible. Therefore, $f_1(X)$ is defined over the domain $[1/2, 1]$, and it is strictly increasing in this interval.

Player 1 knows that he can induce total effort $X \geq 1/2$ by choosing effort $x_1 = f_1(X)$. Therefore, we can write his maximization problem as

$$\max_{X \geq \frac{1}{2}} \frac{f_1(X)}{X} - f_1(X),$$

which implies optimality condition

$$\begin{aligned} 0 &= \frac{f_1'(X)}{X} - \frac{f_1(X)}{X^2} - f_1'(X) = 0 \quad \iff \\ 0 &= f_0(X) = f_1(X) - f_1'(X)X(1-X) = X^2\bigl(6X^2 - 6X + 1\bigr). \end{aligned} \tag{6}$$

Equation (6) has a simple interpretation again—the total equilibrium quantity X^* must be consistent with the optimal behavior of all three players and the fact that the cumulative effort before the move of the first player is $X_0 = 0$. The expression on the right-hand side of the equation, denoted by $f_0(X)$, captures the optimal behavior of all three players and, therefore, a necessary condition for X^* to be an equilibrium quantity is that $f_0(X^*) = 0$. Equation (6) gives three candidates for the total equilibrium effort X^*. It is either 0, $1/2 - 1/(2\sqrt{3}) < 1/2$, or $1/2 + 1/(2\sqrt{3}) > 1/2$. Only the highest root $X^* = 1/2 + 1/(2\sqrt{3}) \approx 0.7887$ constitutes an equilibrium.[9]

The advantage of the inverted best-response approach is that instead of finding solutions to nonlinear equations that become increasingly complex with each recursion, this method combines all of the first-order necessary conditions into a single equation that is then solved only once.

There is a simple recursive dependence in each period 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 requires applying the same rule more times and solving a somewhat more complex equation at the end.

4. Characterization

The characterization theorem (Theorem 1) in this section shows that each contest $\mathbf{n} = (n_1, \dots, n_T)$ has a unique equilibrium and characterizes it using the inverted best-response functions $f_0, \dots, f_T$. These functions are recursively defined according to the same rule as in the previous example. The function $f_T(X)$ specifies the cumulative effort up to and including the last period T that is consistent with total effort X, which is clearly $f_T(X) = X$.

[9]The other positive root implies negative effort by player 2. From $X = 0$, players would have profitable deviations: for example, player 2 would deviate to strictly positive effort defined by equation (4).

Suppose that the efforts after period t are characterized by a differentiable function $f_t(X)$. That is, if the sum of efforts after period t is $X_t = X_{t-1} + \sum_{i\in\mathcal{I}_t} x_i$, then the total effort at the end of the contest will be $X = f_t^{-1}(X_t)$.[10] Player i in period t solves

$$\max_{x_i\geq 0} \frac{x_i}{f_t^{-1}(X_t)} - x_i.$$

The first-order condition determining effort x_i is

$$\frac{1}{f_t^{-1}(X_t)} - \frac{x_i}{\left[f_t^{-1}(X_t)\right]^2}\frac{df_t^{-1}(X_t)}{dX_t}\frac{dX_t}{dx_i} - 1 = 0.$$

Inserting $X = f_t^{-1}(X_t)$, $df_t^{-1}(X_t)/dX_t = 1/f_t'(X)$, and $dX_t/dx_i = 1$ to this expression and rearranging the terms gives a necessary condition for optimality $x_i = f_t'(X_t)X(1-X)$. Summing these constraints for all n_t players in period t, we obtain

$$f_t(X) - f_{t-1}(X) = \sum_{i\in\mathcal{I}_t} x_i = n_t f_t'(X_t)X(1-X).$$

Therefore, $f_{t-1}(X)$ is characterized by a recursive rule that is analogous to the expression we saw above

$$f_{t-1}(X) = f_t(X) - n_t f_t'(X)X(1-X), \quad \forall t\in\{1,\dots,T\}, \text{where } f_T(X) = X. \tag{7}$$

The only difference with the example is the term n_t. If there are multiple players in period t, then each of them has only a fractional impact on the followers' optimal responses. This means that the effect on inverted best-responses is multiplied by n_t.

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

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

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

The proof in Appendix A starts by showing that the polynomials f_t have some helpful properties. Let $\underline{X}_t$ be the highest root of $f_t(X) = 0$. These highest roots are ordered according to t as $0 = \underline{X}_T < \underline{X}_{T-1} < \cdots < \underline{X}_1 < \underline{X}_0$. Moreover, for all $X \in [\underline{X}_t, \underline{X}_{t-1})$ the function $f_{t-1}(X)$ is strictly less than zero and for all $X \in [\underline{X}_t, 1]$ the derivative $f_t'(X)$ is strictly positive. Then the arguments above imply that there are two types of necessary conditions for equilibria. First, total equilibrium effort X^* must be consistent with cumulative effort before the contest being zero, i.e., X^* must be a root $f_0(X) = 0$. Second,

[10]Function f_t^{-1} is the inverse of $f_t(X)$ in the interval $[\underline{X}_t, 1]$, where $\underline{X}_t$ is the highest root of f_t. The proof of the theorem shows that $\underline{X}_t < 1$ and $f_t(X)$ is strictly increasing this interval.

cumulative effort cannot decrease, i.e., $f_t(X^*) \geq f_{t-1}(X^*)$ for all t. These conditions together with the properties of f_t functions imply that the highest root of $f_0(X^*)$ is the only candidate for equilibrium. Finally, assuming the followers behave according to the strategies characterized by f_t functions, each player has a unique interior local optimum for each $X_t < 1$. Alternatively, choosing a corner outcome (either no effort or an effort lever such that all followers choose no effort) clearly gives the player a nonpositive payoff. Therefore, the interior optimum is also a global optimum, and the candidate for equilibrium determined above is indeed an equilibrium.

5. Information and effort

This section includes the main results of the paper. I show that information increases total effort in sequential contests. Before the formal result, let me give an example. Consider contests $\mathbf{n} = (1, 2, 1)$ and $\widehat{\mathbf{n}} = (1, 1, 1, 1)$. The second contest $\widehat{\mathbf{n}}$ is more informative as the added disclosure after player 2 creates a finer partition of players. Let X^* and $\widehat{X}^*$ denote the corresponding total equilibrium efforts in the two contests. Direct application of Theorem 1 gives equilibrium efforts $X^* = (7+\sqrt{13})/12 < \widehat{X}^* = (6+\sqrt{24})/12$, i.e., total equilibrium effort in the more informative contest is strictly higher.

The intuition for this ranking is the following. While efforts could be strategic complements or strategic substitutes, in equilibrium the efforts are high enough to make the individual efforts strategic substitutes. Compared to contest $\mathbf{n}$, the additional disclosure of player 2's effort in contest $\widehat{\mathbf{n}}$ gives player 2 a new reason to increase his effort: it discourages player 3. Therefore, we would expect the effort of player 2 to be higher and the effort of player 3 to be lower than in contest $\mathbf{n}$.

The remaining question is: which of these two effects is larger? The payoff function of player 2 is $u_2(\mathbf{x}) = x_2(1/X - 1)$, which is strictly increasing in player's own effort x_2 and strictly decreasing in total effort X. If player 2 could increase his effort x_2 in a way that the discouragement effect is so large that total effort decreases, player 2 would happily exploit this opportunity, and the outcome would not be an equilibrium. Therefore, the discouragement effect is less than one-to-one, implying that the total effort is increased.

The full comparison of the two contests must also consider how players 1 and 4 respond to the change of game conditions. Their incentives are driven by indirect effects. For example, player 1 may want to influence player 2 to exert more or less effort in the more informative contest, depending on how this second-order impact affects other players.

Capturing the indirect effects requires some new notation. Let me use a contest $\mathbf{n} = (1, 2, 1)$ again to illustrate the construction of relevant variables. All four players in this contest observe their own efforts (regardless of the disclosure rule), and the number of players clearly affects the outcomes of the contest. I call this the first level of information and denote it as $S_1 = n = 4$. More importantly, some players directly observe the efforts of some other players. 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' effort levels. I call this the second level of information and denote it 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 more periods, there would be more levels of information—observations of observations of observations and so on.

I call a vector $\mathbf{S}(\mathbf{n}) = (S_1(\mathbf{n}), \dots, S_T(\mathbf{n}))$ 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 $\{n_1, \dots, n_T\}$.[11] For example, in a sequential n-player contest $\mathbf{n} = (1, 1, \dots, 1)$, $S_k(\mathbf{n})$ is simply the number of all k-combinations, i.e., $S_k(\mathbf{n}) = n!/(k!(n-k)!)$.

With this notation, I can now state the second main result. Theorem 2 defines a partial order on all contests—if contests can be ranked about the measures of information $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$, i.e., $S_k(\widehat{\mathbf{n}}) \geq S_k(\mathbf{n})$ for all k and the inequality is strict at least for one k, then $X^*(\mathbf{S}(\widehat{\mathbf{n}})) > X^*(\mathbf{S}(\mathbf{n}))$. In the example above, increasing informativeness in contests by adding public disclosures increases vector $\mathbf{S}$ and, therefore, total effort, or more concretely $\mathbf{S}(1, 2, 1) = (4, 5, 2) < \mathbf{S}(1, 1, 1, 1) = (4, 6, 4, 1)$.[12]

Theorem 2 (Information Theorem). *Total effort in contest* $\mathbf{n}$ *is a strictly increasing function* $X^*(\mathbf{S}(\mathbf{n}))$.

There are two key steps in the proof. The first step shows by induction that the inverted best-response functions can be expressed using the measures of information as

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

where $\mathbf{n}^t = (n_{t+1}, \dots, T)$ is a vector of integers describing the subcontest that starts after period t and $\mathbf{S}(\mathbf{n}^t)$ denotes its measures of information, i.e., $S_k(\mathbf{n}^t)$ is the sum of all products of k-combinations of $\mathbf{n}^t$. The functions $g_k(X)$ are defined recursively and independently of the contest $\mathbf{n}$ as $g_1(X) = X(1-X)$, and for all $k > 0$, $g_{k+1}(X) = -g_k'(X)X(1-X)$. In particular, $f_0(X)$ takes the following form:

$$f_0(X) = X - \sum_{k=1}^{T} S_k(\mathbf{n}) g_k(X). \tag{10}$$

Remember that total equilibrium effort X^* is the highest root of $f_0(X)$ and this function is strictly increasing above its highest root. The second key step of the proof shows that functions $g_k(X^*) > 0$ at the equilibrium value X^*. Now, increasing $\mathbf{S}(\mathbf{n})$ decreases the value of $f_0(X^*)$ at the original equilibrium value. Therefore, the highest root of the new function $f_0(X)$ must be higher than the original one.

[11] Let $\binom{\mathbf{n}}{k}$ denote the set of all k-combinations of set $\{n_1, \dots, n_T\}$. For example, $\boldsymbol{s}^3 \equiv \{n_1, n_3, n_4\} \in \binom{\mathbf{n}}{3}$. Then $S_k(\mathbf{n}) = \sum_{\boldsymbol{s}^k \in \binom{\mathbf{n}}{k}} \prod_{s \in \boldsymbol{s}^k} s$ for all $k \leq T$ and $S_k(\mathbf{n}) = 0$, otherwise.

[12] $S_k(\mathbf{n}) = 0$ for all $k > T$, so that $\mathbf{S}(\mathbf{n})$ is an infinite-dimensional vector with only zeroes after element T. I have left zeroes out for brevity. In this example, more detailed comparison would be $\mathbf{S}(1, 2, 1) = (4, 5, 2, 0, 0, 0, \dots) < \mathbf{S}(1, 1, 1, 1) = (4, 6, 4, 1, 0, 0, \dots)$.

Equation (10) also sheds some light on the reason for this result. The positive weights on the measures of information, $g_k(X^*)$, can be interpreted as higher-order strategic substitutability terms. In particular, $g_1(X)$ captures the concavity of payoff functions—as agent i increases his effort, the incentive to increase the effort further decreases and is positive for all X. With slight abuse of terminology, we can think concavity of payoffs as a player's effort being "a substitute" for his own effort. The term $g_2(X)$ captures the standard strategic substitutability. In the example above, if player 2 increases effort, player 3 who observes this deviation has an incentive to decrease effort. The next term $g_3(X)$ captures an indirect incentive: player 4, who observes the response of player 3 has an incentive to decrease effort as well (beyond the direct effect of responding to player 2). We can think of this as second-order strategic substitutability. Similarly, $g_4(X)$ would capture the incentives to respond to second-order effects and, therefore, describe the third-order strategic substitutability, and so on. The fact that near equilibrium $g_k(X^*) > 0$ for all k, means that all these effects move the equilibrium outcomes to the same direction—in more informative contests (in the sense of $\mathbf{S}(\mathbf{n})$) earlier movers exert more effort, later movers less effort, but the total equilibrium effort is higher.

Equation (10) therefore also allows to complete order defined by Theorem 2. To complete the order, we would have to know how to weigh different measures of information. Equation (10) shows that correct weights are $g_k(X^*)$; i.e., by magnitudes of discouragement effects near equilibrium. The following lemma shows that lower information measures have a higher weight.

Lemma 1 (Decreasing Weights). *$g_{k-1}(X^*) > g_k(X^*)$ for each $k \geq 2$.*

While direct effects have a larger impact than indirect ones, the indirect effects are not qualitatively unimportant. Compare, for example, two seven-player contests $\mathbf{n} = (3, 4)$ and $\widehat{\mathbf{n}} = (1, 1, 5)$. The first contest $\mathbf{n}$ has 12 direct observations whereas $\widehat{\mathbf{n}}$ has only 11. Nevertheless, the total effort in $\mathbf{n}$ is lower than in contest $\widehat{\mathbf{n}}$. This is because of indirect effects, $S_3(\widehat{\mathbf{n}}) = 5 > 0$. Intuitively, in the contest $\widehat{\mathbf{n}}$ player 1 knows that in addition to influencing all followers directly, there is also an effect to the five last movers through the behavior change of player 2. This indirect effect is missing in $\mathbf{n}$.

Theorem 2 has several direct implications summarized by the following corollary.

Corollary 1 (Implications of the Information Theorem). *Take two contests $\mathbf{n}$ and $\widehat{\mathbf{n}}$, with corresponding partitions $\mathcal{I}$ and $\widehat{\mathcal{I}}$, and let $X = X^*(S(\mathbf{n}))$ and $\widehat{X} = X^*(S(\widehat{\mathbf{n}}))$ be the corresponding total equilibrium efforts.*

(a) Comparative statics of $\mathbf{n}$: if $\mathbf{n} < \widehat{\mathbf{n}}$, then $X < \widehat{X}$. This includes the case when $n_t = 0 < \widehat{n}_t$ for some t, i.e., $\widehat{\mathbf{n}}$ has more periods than $\mathbf{n}$.

(b) Independence of permutations: if $\mathbf{n}$ is a permutation of $\widehat{\mathbf{n}}$, then $X = \widehat{X}$.

(c) Disclosures increase effort: if $\mathcal{I}$ is a coarser partition than $\widehat{\mathcal{I}}$, then $X < \widehat{X}$.

(d) Homogeneity increases effort: if $\sum_t n_t = \sum_t \widehat{n}_t$ and there exist t, t' such that $n_t n_{t'} < \widehat{n}_t \widehat{n}_{t'}$ and $n_s = \widehat{n}_s$ for all $s \neq t, t'$, then $X < \widehat{X}$.

(e) *For any n, the simultaneous contest $\mathbf{n} = (n)$ minimizes total effort, and the fully sequential contest $\mathbf{n} = (1, 1, \dots, 1)$ maximizes the total effort. For fixed number of periods T, contests that allocate players into groups that are as equal as possible maximize the total effort.*

The first implication (a) is intuitive, adding players to any period or adding periods to any contest increases the total effort. However, this does not imply that the total effort increases with the total number of players. Total effort in the three-player sequential contest is 0.7887, whereas in the four-player simultaneous contest, it is 0.75. The second implication (b) is more surprising—reallocating disclosures in a way that creates a permutation of $\mathbf{n}$ does not affect the total effort. For example, a first-mover contest $(1, n-1)$ gives the same total effort as the last-mover contest $(n-1, 1)$. The third implication (c) that disclosures increase effort was already discussed above. The fourth implication (d) gives even clearer implications for the optimal contest. Namely, more homogeneous contests give higher total effort. Intuitively, a contest is more homogeneous if players are divided more evenly across periods. For example, a contest $\widehat{\mathbf{n}} = (2, 2)$ is more homogeneous than $\mathbf{n} = (1, 3)$ and also has more direct observations of efforts as $2 \times 2 = 4 > 3 = 1 \times 3$.

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-movers incentives to increase efforts to discourage later players, but this discouragement effect is less than one-to-one and therefore increases total effort. Conversely, 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 (e.g., collecting or announcing information is costly), then it is better to spread the disclosures as evenly as possible.

6. Earlier-mover advantage

Dixit (1987) showed that there is a first-mover advantage. If one player can pre-commit, the first-mover chooses a strictly higher effort and achieves a strictly higher payoff than the followers. Using the tools developed here, I can explore this result further. Namely, 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? 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—there is a strict earlier-mover advantage.

Proposition 1 (Earlier-Mover Advantage). *The efforts and payoffs of earlier players are strictly higher than for later players.*

The equilibrium payoff of a player i is in $u_i(\mathbf{x}^*) = x_i^*(1/X^* - 1)$, 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. Using Theorem 1 and equation (9), I can express the difference between the equilibrium efforts of players i and j from consecutive periods t and $t+1$ as

$$x_i^* - x_j^* = \sum_{k=1}^{T-t} \big[S_k(\mathbf{n}^t) - S_k(\mathbf{n}^{t+1})\big] g_{k+1}(X^*) \tag{11}$$

where $\mathbf{n}^{t+1} = (n_{t+2}, \ldots, n_T)$ is the subcontest starting after period $t+1$ and $\mathbf{n}^t = (n_{t+1}, \mathbf{n}^{t+1})$ is the subcontest starting after period t. Clearly, $S_k(\mathbf{n}^t) > S_k(\mathbf{n}^{t+1})$ for all k; i.e., there is more information on all levels in a strictly longer contest. As $g_{k+1}(X^*) > 0$, for each k 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 an additional incentive to exert more effort to discourage later players.

7. Large contests

Numeric comparison of simultaneous and sequential contests highlights that the information about other players' efforts is at least as important in determining the total effort as other parameters, such as the number of players. For example, the total effort in the simultaneous contest with 10 players is 0.9, whereas the total effort with four sequential players is 0.9082. A fifth sequential player increases the total effort to 0.9587. A simultaneous contest with the same total effort requires 24 players. Figure 2 shows that the comparison becomes even more favorable for sequential contests with large n.

The following proposition gives the reason for this connection. As the number of players becomes large, the total effort converges to 1 no matter the contest structure, but the convergence is different depending on the structure. For large simultaneous contests, the convergence is linear, with $1 - X^* \approx 1/n$, while for large sequential contests, the convergence is exponential, with $1 - X^* \approx 1/2^n$.[13] It is also worth noting that, although the individual payoffs converge to zero, the individual efforts may not.

Proposition 2 (Large Contests). *Fix $T \in \mathbb{N}$ and a sequence of contests $(\mathbf{n}^n)_{n=3}^{\infty}$, such that contest $\mathbf{n}^n$ is n-player contest with at most T periods. Let $X^n = X^*(\mathbf{S}(\mathbf{n}^n))$ and for each player i, let x_i^n the equilibrium effort in contest $\mathbf{n}^n$. For all $t \le T$ and all $i \in \mathcal{I}_t^n$,*

$$\lim_{n\to\infty}\left[1 - X^n - \frac{1}{\prod_{t=1}^{T}(1+n_t^n)}\right] = 0 \quad \textit{and} \quad \lim_{n\to\infty}\left[x_i^n - \frac{1}{\prod_{s=1}^{t}(1+n_s^n)}\right] = 0. \tag{12}$$

[13] Proposition 2 is stated for arbitrary fixed T. Therefore, it is straightforward to apply it to the limit of contests where T itself becomes infinitely large (e.g., large fully sequential contests) by taking another limit with respect to T.

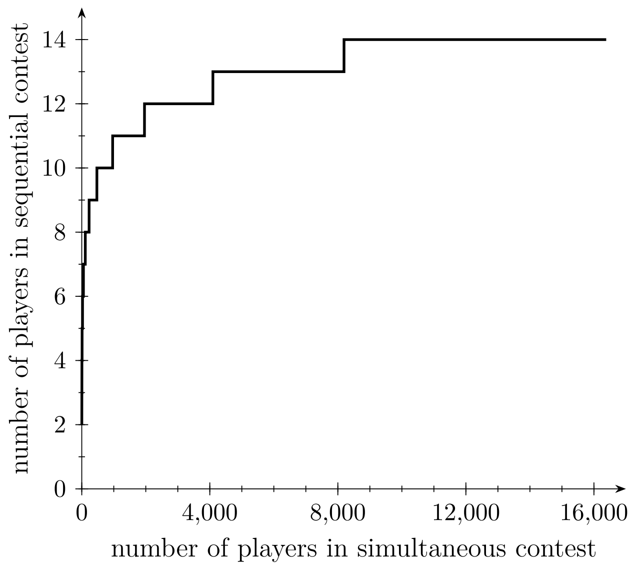

FIGURE 2. Number of players in a sequential contest that leads to the same total effort as a simultaneous contest with n players.

These results shed new light on the meaning of a highly competitive contest or market. In a large simultaneous contest, each contestant chooses a minuscule effort level. Such a market is clearly not concentrated. For example, with $n = 16{,}000$ the standard measure of concentration, the Herfindahl–Hirschman Index is $HHI^{\text{sim}} \approx 0$.

In contrast, a sequential contest requires only a limited number of players to achieve the same aggregate results, and players behave differently. In a large sequential contest, the individual equilibrium efforts are $\mathbf{x}^* \approx (1/2, 1/4, \dots, 1/2^n)$. The earlier movers choose much larger efforts and achieve larger payoffs than the followers. For example, with $n = 14$ sequential players, the corresponding concentration index $HHI^{\text{seq}} \approx 1/3$, which is typically interpreted as a highly concentrated market. However, in terms of outcomes, this market is highly competitive: as total effort is close to one, we have full dissipation of rents, and thus all players earn equilibrium profits that are close to zero.

This effect is similar to contestability theory (Baumol, Panzar, and Willig (1988)), where a small number of firms cannot capitalize on their market power due to the presence of a competitive fringe—a large number of potential competitors, who could frictionlessly enter when a profit opportunity arises. In my model, the later movers are endogenously taking the role of the competitive fringe. In equilibrium, they decide to put in very little effort. However, if the earlier movers were to try and exploit their position by reducing their efforts, the later movers would be there to respond.

8. GENERALIZATION

In this section, I discuss how to implement the methodology for a general class models. I also provide sufficient conditions under which the results above remain unchanged.

Specifically, I define a class of linearly multiplicative payoff functions and show that if it satisfies Property 1, Theorem 1 remains valid without any modifications. By adding another sufficient condition, Property 2, nearly all other results in the paper hold as well. The differences between Property 1 and Property 2 also suggest that Theorem 2 and most other results in the paper are not direct implications of Theorem 1.

Suppose that each player chooses an action x_i from a set $\mathcal{X}_i$ and if the profile of actions is $\mathbf{x} = (x_1, \dots, x_n)$, then player i gets a payoff

$$U_i(\mathbf{x}) = u_i(x_i, X). \tag{13}$$

Take a player i from the last period T. Player i observes cumulative effort X_{T-1} before period T and knows that other players in period T are choosing efforts simultaneously to him. Therefore, he solves the maximization problem

$$\max_{x_i \in \mathcal{X}_i} u_i\left(x_i, x_i + X_{T-1} + \sum_{j \in \mathcal{I}_T \setminus \{i\}} x_j\right).$$

The standard best-response function would be $x_i^*(X_{T-1})$.[14] But suppose we can express the optimal effort x_i choice as a function of total effort, $\phi_i(X)$. Then adding up individual efforts in period T consistent with total effort X gives us a necessary condition for equilibrium,

$$X_{T-1} = X - \sum_{i \in \mathcal{I}_T} \phi_i(X).$$

I denote the function on the right-hand side by $f_{T-1}(X)$. Its inverse function (assuming it exists), $f_{T-1}^{-1}(X_{T-1})$ is the total effort induced by cumulative effort X_{T-1}, if all players in period T behave optimally.[15]

Suppose by induction that the same argument holds starting from period t, i.e., if cumulative effort after t is X_t then the total effort induced is $f_t^{-1}(X_t)$. Then player i in period t solves the following problem:

$$\max_{x_i \in \mathcal{X}_i} u_i\big(x_i, f_t^{-1}(X_t)\big).$$

If again, we can express the optimal x_i only as a function $\phi_i(X)$, then adding up the conditions would give us a necessary condition for equilibrium

$$X_{t-1} = X_t - \sum_{i \in \mathcal{I}_t} x_i = f_t(X) - \sum_{i \in \mathcal{I}_t} \phi_i(X),$$

which I denote by $f_{t-1}(X)$. Finally, in the beginning of the game cumulative total action is $X_0 = 0$, which gives us an equilibrium condition for the whole game.

---

[14]This function is also called the reduced best-response function as it only depends on the sum.

[15]When $T = 1$, the game becomes a linearly aggregative game, as introduced in Selten (1970), with a known equilibrium condition $X = \sum_{i=1}^n \phi_i(X)$, where $\sum_{i=1}^n \phi_i(X)$ is the aggregate backward correspondence. See Jensen (2018) for a literature review. If $T > 1$, the game is not aggregative, so the analysis presented here is a dynamic generalization of linearly aggregative games.

There are some gaps in this analysis that need to be filled. I have already shown that with the Tullock contest payoffs, $u_i(x_i, X) = x_i/X - x_i$ and $\mathcal{X}_i = \mathbb{R}_+$, this approach characterizes the unique equilibrium. It is equally clear that the approach is not valid for all payoff functions, as interior optimums may not exist or be unique. Next, I introduce a more restricted class of payoff functions and sufficient conditions where all results hold and the analysis remains tractable.

Linearly multiplicative payoffs: Assume that the payoff functions are identical and the utility is linearly multiplicative with respect to players' own actions,

$$u_i(x_i, X) = x_i h(X), \quad x_i \in \mathcal{X}_i = \mathbb{R}_+. \tag{14}$$

For Tullock contest payoffs, $h(X) = v/X - c$, where v represents the prize value and c denotes the marginal cost of effort. This class of games also includes oligopolies with linear costs, where $h(X) = P(X) - c$, with x_i as the firm's own quantity, X as the total quantity, $P(X)$ as the inverse demand function, and c as the marginal cost. Additionally, this class includes public goods games, in which x_i denotes private consumption and $h(X)$ represents the marginal benefit of private consumption, which decreases with public good contributions and, therefore, with total private consumption.

It is natural to assume in these applications that $h(X)$ is strictly decreasing up to some upper bound $\overline{X}$, at which it takes value $h(\overline{X}) = 0$ and above which $h(X) \leq 0$. Therefore, effectively the action space is $\mathcal{X}_i = [0, \overline{X}]$. Without loss of generality, we can change the scale of actions so that $\overline{X} = 1$.

The first-order optimality condition for players in period T is then

$$h(X) + x_i h'(X) = 0 \quad \Longleftrightarrow \quad x_i = g_1(X),$$

where $g_1(X) = -h(X)/h'(X)$. Therefore, we can write the inverted best-response function as

$$f_{T-1}(X) = X - n_T g_1(X).$$

Similarly, if the inverted best-response functions at period t is $f_t(X)$, which is invertible in the relevant range, the payoff function of player i in period t is $u_i(x_i, f_t^{-1}(X_t)) = x_i h(f_t^{-1}(X_t))$ and, therefore, the first-order condition for players in period t is

$$h(X) + x_i h'(X) \frac{1}{f_t'(X)} = 0 \quad \Longleftrightarrow \quad x_i = g_1(X) f_t'(X). \tag{15}$$

Therefore, $f_{t-1}(X) = f_t(X) - n_t f_t'(X) g_1(X)$. This shows that we can use the characterization derived in the paper, with two modifications. First, instead of specific expression $X(1-X)$, we have a function $g_1(X) = -h(X)/h'(X)$. And second, we need to impose some conditions on the function $h(X)$ so that the conditions for the existence and uniqueness are satisfied.

In Appendix A, I define Property 1, which is a sufficient condition for all f_t functions to be well behaved so that the characterization theorem (Theorem 1) holds without any modifications. Intuitively, Property 1 puts two restrictions on f_t functions. First, for sufficiently high X, they are strictly increasing and, therefore, invertible in the relevant range. Second, at least one of the f_t functions is taking a negative value for lower values of X, which eliminates such X as a candidate for equilibrium. Proposition 3 in Appendix A proves that Tullock payoffs satisfy Property 1 and below I discuss some other cases when it is satisfied.

Therefore, under Property 1, the equilibrium is still unique and can be computed as the highest root of $f_0(X)$ in [0, 1]. Moreover, the limit for large contests (Proposition 2) holds as well, with a particular adjustment in formulas. Let $\alpha = -g_1'(1) > 0$. Then the formulas in equation (12) would be adjusted as

$$\lim_{n\to\infty}\left[1 - X^n - \frac{1}{\prod_{t=1}^{T}\left(1+\alpha n_t^n\right)}\right] = 0 \quad \text{and} \quad \lim_{n\to\infty}\left[x_i^n - \frac{\alpha}{\prod_{s=1}^{t}\left(1+\alpha n_s^n\right)}\right] = 0. \tag{16}$$

For the information theorem (Theorem 2) its corollaries (Corollary 1), as well as the earlier-mover advantage result (Proposition 1) I need an additional assumption. First, let us adjust g_k functions by defining these as $g_1(X) = -h(X)/h'(X)$ and $g_{k+1}(X) = -g_k'(X)g_1(X)$ for all k. The additional assumption, Property 2 in Appendix A, states essentially that each $g_k(X^*) > 0$ near equilibrium. This assumption can be interpreted as actions being higher-order strategic substitutes. Proposition 4 proves that Tullock payoffs satisfy Property 2 and below I discuss some functional forms that satisfy this assumption.

The only result that does not generalize is Lemma 1 that showed that with Tullock payoffs, the weights $g_k(X^*)$ are decreasing in k. It is easy to see that this result depends on the function $h(X)$. For example, consider the case when $h(X) = \sqrt[\alpha]{1-X}$ for all $X \in [0, 1]$ and 0 otherwise, where $\alpha > 0$ is a constant. Then $g_1(X) = \alpha(1-X)$, $g_2(X) = \alpha^2(1-X)$, and so on, $g_k(X) = \alpha^k(1-X)$. Whenever $\alpha > 1$, the weights are increasing in this case.

The remaining question is when are properties 1 and 2 satisfied? For example, one special class of functions where these assumption are satisfied, is the class of functions, where $g_1(X) = -h(X)/h'(X)$ is *completely monotone*, i.e., $(-1)^k d^k g_1(X)/dX^k \geq 0$ for all $k \in \mathbb{N}$.[16] This includes many functions, including linear $h(X)$, power function $h(X) = \sqrt[\alpha]{1-X}$, but also many other natural functions. For example, the following functions are all completely monotone: $g(X) = \alpha(1-X^m)$, $g(X) = \alpha(1-X)^m$, for all $m \in \mathbb{N}$, $g(X) = \alpha((X+\gamma)^s - (1+\gamma)^s)$ for all $s < 0$, $\gamma > 0$, $g(X) = \alpha[e^{-rX} - e^{-r}]$ for all $r > 0$, and $g(X) = -\alpha\log(X)$, all with any $\alpha > 0$. Also, all sums and products of completely monotone functions are completely monotone.[17]

[16] It suffices that $g_1(X)$ is only T-times monotone, which is less restrictive, but perhaps harder to verify.

[17] In the working paper version (https://arxiv.org/pdf/1802.04669.pdf), I give more examples: (1) An oligopoly with logarithmic demand, where the analysis can be directly extended, even with non-monotonic $g_1(X)$; (2) An example where properties 1 and 2 are violated, and equilibrium may not exist or be unique;

Note that linearly multiplicative payoffs combined with properties 1 and 2 are sufficient and convenient assumptions, but they are not necessary. These assumptions ensure that the first-order conditions are linear in the cumulative action of preceding players, and as a result, the inverted best-response functions can be easily characterized. This raises the question: is the functional form assumption relevant only for tractability or does it have economic implications? A simple continuity argument demonstrates that the results will not change with small perturbations near the original model.[18]

9. Discussion

I showed that each contest has a unique equilibrium. It is in pure strategies and simple to compute. The main result of the paper shows that the total equilibrium effort is strictly increasing in information. This implies that the optimal contest for maximizing total effort is fully sequential, e.g., R&D contests benefit from full transparency. On the other hand, if the goal is to minimize the total effort, such as rent-seeking contests, the optimal contest is nontransparent, i.e., the simultaneous contest. Further, there is a strict earlier-mover advantage: players in earlier periods exert strictly greater efforts and obtain strictly higher payoffs. Total effort converges to full dissipation linearly with the number of players in large simultaneous contests but exponentially in large sequential contests.

The results in this paper hold much more generally than the model discussed herein. In addition to the generalization discussed in Section 8, some assumptions about the timing of arrivals can be relaxed. I assumed that players exert efforts only at their arrival and that their efforts are publicly observable for players in the following periods. Given that players benefit from the discouragement effect, they would not hide or delay their actions. Thus, the outcomes would be unchanged if players could take hidden actions or take actions over multiple periods. This was shown by Yildirim (2005) in the two-player case.

The analysis can also be extended to heterogeneous players. However, there is a new complication: players may find it optimal to stay inactive at different thresholds. This means that earlier movers may sometimes find it optimal to deter entry by followers and the order of players becomes an important determinant of outcomes. Xu, Zhang, and Zhang (2020) use the approach introduced here to study the three-player asymmetric sequential contests.[19] Hinnosaar (2023) extends the methodology to another type of player heterogeneity, where the game is played on a network. Players only observe the choices of players they are linked to. This analysis shows that there is a connection between weighted measures of information and standard centrality measures from network theory.

(3) An example where efforts may be direct strategic substitutes in the standard sense but strategic complements due to indirect effects. In this instance, the equilibrium with two periods behaves as one would expect with strategic substitutes, while introducing a third period alters the conclusions as a result of indirect effects.

[18] In the working paper version, I show that in Tullock contests with quadratic costs, the analysis still applies when the parameter multiplying the quadratic term is sufficiently close to zero.

[19] The two-player case has been studied by Morgan (2003) and Serena (2017), who also considered endogenous order of moves.

Appendix A: Proofs

A.1 *Proof of the characterization theorem (Theorem 1)*

Before proving Theorem 1, it is useful to define the following property.

Property 1 (Inverted Best Responses are Well Behaved). Clearly, $f_T(X) = X$ has unique root $\underline{X}_T = 0$. For all $t = 0, \dots, T-1$, the function f_t has the following properties:

(a) $f_t(X) = 0$ has a root in $[\underline{X}_{t+1}, 1]$. Let $\underline{X}_t$ be the highest such root.

(b) $f_t(X) < 0$ for all $X \in [\underline{X}_{t+1}, \underline{X}_t)$.

(c) $f_t'(X) > 0$ for all $X \in [\underline{X}_t, 1]$.

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

The proof of Theorem 1 has two parts. The first part is Proposition 3 in Appendix A.2 that shows that f_t functions satisfy Property 1. The proof relies on keeping track of the roots of f_t functions. The second part in Appendix A.3 establishes the theorem's claims. Briefly, it shows that behavior where each player i in each period t behaves according to equation (8) and expects that total effort induced by cumulative effort X_t to be $f_t^{-1}(X)$, is an equilibrium and in fact it is the only equilibrium. The proof is divided into five lemmas:

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 $X_t = f_t(X)$ is a necessary condition for equilibrium. In particular, $f_0(X) = 0$ is a necessary condition for equilibrium and, therefore, X^* must be a root of $f_0(X)$.

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

4. Lemma 8 shows that the best-response function of player $i \in \mathcal{I}_t$ after cumulative effort X_{t-1} is $x_i^*(X_{t-1}) = n_t^{-1}[f_t(f_{t-1}^{-1}(X_{t-1})) - X_{t-1}]$ 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 $x_i^* = n_t^{-1}[f_t(X^*) - f_{t-1}(X^*)]$. Note that this step in the proof implicitly also shows that strategies of all players in period t are identical.

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

The combination of these results proves Theorem 1.

A.2 *Proof that Property 1 is satisfied (Proposition 3)*

Proposition 3. *Inverted best responses $f_0, \dots, f_T$ defined by equation (7) are well behaved.*

Before giving the proof of Proposition 3, 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 3. 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* $t = 0, \dots, T$.

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* $t = 0, \dots, T$. *Depending on* $\mathbf{n}$, *there could be either one or two roots at zero:*

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

(b) 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 $c_0^t, \dots, c_r^t$ are the coefficients. Therefore,

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

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 3 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

$$f_t(X) = c_t \prod_{s=1}^{r} (X - X_{s,t}), \tag{17}$$

where $c_t > 0$ is the leading coefficient and $X_{1,t}, \dots, X_{r,t}$ 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 $0 = X_{1,t} < X_{2,t} < \cdots < X_{r,t} < 1$. 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

$$f_t'(X_{j,t}) = c_t \prod_{s \neq j} (X_{j,t} - X_{s,t}). \tag{18}$$

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 $s \in \{1, \dots, r-1\}$. 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,

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

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 $f_t^{X'}(X) = f_t'(X)/X$. 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 $f_t'(X) = Xf_t^{X'}(X)$. Now we have found all $r-1$ roots of polynomial f_t' and

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

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

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 $f_{t-1}^{X^2}(X) = f_{t-1}(X)/X^2$. 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

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

Therefore, $f_{t-1}^{X^2}$ 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

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

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

(c) 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 $X \in [X_{r+1,t-1}, 1]$. □

A.3 *Proof of Theorem 1 using Property 1*

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

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

(b) 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. Alternatively, 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, $X \geq X_t > 1$, which means that $u_i(\mathbf{x}) < 0$. 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 $X \geq 1$. 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 at least one player i in period s chose $x_i > 0$ and gets a payoff of $u_i(\mathbf{x}) \leq 0$. Now, there are two cases. First, if the induced total effort $X > 1$, 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 $X = 1$, which means that $X_s = 1$, then player i could choose effort $x_i/2$, thus making $X_s < 1$ and, therefore, $X < 1$, ensuring a strictly positive payoff. In both cases, we arrive at a contradiction. Thus, $X_{t-1} < 1$ implies $X_t < 1$ and $X < 1$.

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 the induced total effort $\widehat{X} < 1$, and the new payoff of player i is strictly positive. This is a contradiction. □

Lemma 6. *$X_t = f_t(X)$ is a necessary condition for equilibrium.*

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

$$\max_{x_i \geq 0} \frac{x_i}{f_t^{-1}(X_t)} - x_t$$

where $X_t = X_{t-1} + \sum_{j \in \mathcal{I}_t} x_j$. Therefore, a necessary condition for optimum is

$$\frac{1}{f_t^{-1}(X_t)} - 1 + \frac{-x_i}{\left[f_t^{-1}(X_t)\right]^2} \frac{d f_t^{-1}(X_t)}{d X_t} = 0.$$

It is convenient to rewrite this condition in terms of the total effort X, taking into account that $X = f_t^{-1}(X_t)$ and $df_t^{-1}(X_t)/dX_t = 1/f_t'(X)$ to get

$$x_i = f_t'(X) X (1 - X). \tag{19}$$

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, $X_{t-1} = f_{t-1}(X)$, defined by equation (7). □

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

PROOF. First, note that even if X_t would be 0, the total effort induced by it would not be zero. In fact, by recursion it is straightforward to show that it would be $\underline{X}_t$. For any X_{t-1}, therefore, $f_{t-1}^{-1}(X_{t-1}) \geq \underline{X}_t$. Consequently, $X < \underline{X}_t$ cannot be the total effort following any X_{t-1}.

Moreover, by Property 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 $[\underline{X}_t, \underline{X}_{t-1})$ 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 Property 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 $f_{t-1}^{-1}(1) = 1$, and since $\underline{X}_{t-1}$ is a root of f_{t-1}, we have $f_{t-1}^{-1}(0) = \underline{X}_{t-1}$. □

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

$$x_i^*(X_{t-1}) = \begin{cases} \dfrac{1}{n_t}\big[f_t\big(f_{t-1}^{-1}(X_{t-1})\big) - X_{t-1}\big] & \forall X_{t-1} < 1, \\ 0 & \forall X_{t-1} \geq 1. \end{cases} \tag{20}$$

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

PROOF. Lemma 5 proved the claim for any $X_{t-1} \geq 1$. Take $X_{t-1} < 1$. Then by Lemma 5, the individual efforts are interior, so they have to satisfy the individual first-order conditions (19). I showed that the total effort induced by X_{t-1} must be $f_{t-1}^{-1}(X_{t-1})$. Inserting these results into the individual optimality condition for player $i \in \mathcal{I}_t$, I get

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

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

So far, the arguments 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. □

A.4 *Proof of the information theorem (Theorem 2)*

For any contest $\mathbf{n} = (n_1, \dots, n_T)$ and any period $t \le T$, let $\mathbf{n}^t = (n_{t+1}, \dots, n_T)$ denote the subcontest starting after period t. In particular, $\mathbf{n}^T = \emptyset$ and $\mathbf{n}^0 = \mathbf{n}$. Note that $f_t(X)$ depends only on $\mathbf{n}^t$.

Remember that $g_1, \dots, g_T$ are recursively defined as $g_1(X) = X(1-X)$ and $g_{k+1}(X) = -g_k'(X)X(1-X)$, so they are independent of $\mathbf{n}$. Also, $\mathbf{S}(\mathbf{n}) = (S_1(\mathbf{n}), \dots, S_T(\mathbf{n}))$ 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. Similarly, $\mathbf{S}(\mathbf{n}^t)$ is defined in the same way for each subcontest.

The proof has two key steps. The first step (Lemma 10) shows that we can express the inverted best-response functions through a weighted sums of measures of information as in equation (9).

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. The following Property 2 formalizes this idea with a small caveat. Namely, as we will see below, in the case of fully the sequential contest $g_n(X^*) = 0$ and, therefore, the strict version of this definition is not satisfied. However, as I will prove, for any $n > 2$, it is sufficient that efforts are weak strategic substitutes of level n as defined below.

Property 2 (Kth-Order Strategic Substitutes Near Equilibrium). Efforts are (weak) strategic substitutes of level K near equilibrium, if $g_K(X^*) \ge 0$ and $g_k(X^*) > 0$ for all $k \in \{2, \dots, K-1\}$ at the equilibrium level of total effort X^*. Efforts are strict strategic substitutes of level K near equilibrium, if they are strategic substitutes of level K and $g_K(X^*) > 0$.

Lemma 10. *The function $f_t(X)$ can be expressed as*

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

Proof. By construction, the subcontest starting after period T has no players, so $S_k(\mathbf{n}^T) = 0$ for any k. Therefore, $f_T(X) = X$ satisfies equation (9). Now, suppose that the characterization holds for $f_t(X)$. Then, since $g_{k+1}(X) = -g_k'(X)X(1-X)$, we get that

$$\begin{aligned} f_t'(X)X(1-X) &= X(1-X) - X(1-X)\sum_{k=1}^{T-t} S_k(\mathbf{n}^t) g_k'(X) \\ &= g_1(X) + \sum_{k=2}^{T-t+1} S_{k-1}(\mathbf{n}^t) g_k(X). \end{aligned}$$

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

$$f_{t-1}(X) = X - \sum_{k=1}^{T-t} S_k(\mathbf{n}^t) g_k(X) - n_t g_1(X) - n_t \sum_{k=2}^{T-t+1} S_{k-1}(\mathbf{n}^t) g_k(X)$$

$$= X - [S_1(\mathbf{n}^t) + n_t] g_1(X) - \sum_{k=2}^{T-t} [S_k(\mathbf{n}^t) + n_t S_{k-1}(\mathbf{n}^t)] g_k(X)$$

$$- n_t S_{T-t}(\mathbf{n}^t) g_{T+1}(X).$$

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

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

Proposition 4 (Efforts are Higher-Order Strategic Substitutes). *Take a contest* $\mathbf{n}$ *with* $T \le n$ *periods with a positive number of players. Then:*

(a) If $T < n$*, efforts are strict strategic substitutes of level* T *near equilibrium.*

(b) If $T = n$*, efforts are weak strategic substitutes of level* n *near equilibrium.*

As the proof is long, I include the proof as a separate subsection below. With these results, the proof of the information theorem is now straightforward.

Proof of Theorem 2. Take two contests $\mathbf{n}$ and $\widehat{\mathbf{n}}$ such that $\mathbf{S}(\widehat{\mathbf{n}}) > \mathbf{S}(\mathbf{n})$, i.e., $S_k(\widehat{\mathbf{n}}) \ge S_k(\mathbf{n})$ for all k and the inequality is strict for at least one k. Let T and $\widehat{T}$ be the number of periods with strictly positive number of players in contests $\mathbf{n}$ and $\widehat{\mathbf{n}}$, respectively. Notice that by assumptions, $T \le \widehat{T}$ and $S_k(\mathbf{n}) = 0$ for all $k > T$. By Theorem 1, the total equilibrium X^* is the highest root of $f_0(X)$ in $[0, 1]$. By Lemma 10, we can express $f_0(X^*)$, as

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

Similarly, let $\widehat{X}^*$ be the total equilibrium effort in contest $\widehat{\mathbf{n}}$. It is the highest root of $\widehat{f}_0(X^*)$ in $[0, 1]$, which we can write as

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

Suppose by contradiction that the claim of the theorem does not hold and so $\widehat{X}^* \le X^*$. Since by Property 1 $\widehat{f}_0$ is strictly increasing in $[\widehat{X}^*, 1]$, we get that $0 = \widehat{f}_0(\widehat{X}^*) \le \widehat{f}_0(X^*)$.

Therefore,

$$0 \le \widehat{f}_0(X^*) - f_0(X^*) = -\sum_{k=1}^{\widehat{T}} [S_k(\mathbf{n}) - S_k(\mathbf{n})] g_k(X^*)].$$

As $S_k(\mathbf{n}) \ge S_k(\mathbf{n})$ and $g_k(X^*) \ge 0$ for each $k \in \{1, \dots, K\}$, the sum on the right-hand side is nonpositive. Moreover, of at least one k, we have $S_k(\mathbf{n}) > S_k(\mathbf{n})$. Now, by Proposition 4, if $\widehat{T} < n$, the efforts are strict strategic substitutes, so $g_k(X^*) > 0$. Therefore, we get a contradiction.

Finally, suppose that $\widehat{T} = n$. As the only n-player contest with a positive number of players in n periods is the fully sequential contest, we must have $\widehat{\mathbf{n}} = (1, 1, \dots, 1)$. It is straightforward to verify that then $S_n(\widehat{\mathbf{n}}) = 1$ and $S_{n-1}(\widehat{\mathbf{n}}) = n$. Now, notice that since the contest $\mathbf{n}$ is strictly less informative than $\widehat{\mathbf{n}}$, it must have at least one period with two players. Let us replace it with a new contest $\mathbf{n}'$, where we have split all players into separate periods and left only one period with two players, i.e., the contest $\mathbf{n}'$ is a permutation of $(2, 1, 1, \dots, 1)$. Clearly, the contest $\mathbf{S}(\mathbf{n}') \ge \mathbf{S}(\mathbf{n})$. As in both contests $\mathbf{n}$ and $\mathbf{n}'$, the number of periods is strictly less than n, the part of the theorem we already proved implies that the corresponding equilibrium effort $X^{*'} \ge X^*$.

In contest $\mathbf{n}'$, $S_n(\mathbf{n}') = 0 < S_n(\widehat{\mathbf{n}})$, $S_{n-1}(\mathbf{n}') = 2 < nS_{n-1}$, and $S_k(\mathbf{n}') \le S_k(\widehat{\mathbf{n}})$ for all $k < n-1$. As by Proposition 4, the efforts are weak strategic substitutes of level n near equilibrium, and this proves that $X^* \le X^{*'} < \widehat{X}^* \le X^*$. This is a contradiction. □

Remark. The last paragraph of the proof shows why we need the assumption that $n > 2$. Otherwise, in two-player contest, the sequential contest implies $\mathbf{S}(1, 1) = (2, 1)$ and simultaneous contest $\mathbf{S}(2) = (2, 0)$. These two contests only differ by the measure of information of level 2. As the efforts are only weakly strategic substitutes at X^*, the proof would not be valid. Indeed, it is straightforward to check that with $n = 2$, $X^* = 0.5$ and $g_2(0.5) = 0$, so the two contests would give the same total effort. With $n = 3$, this issue does not arise, as $S_2(1, 1, 1) = 3$ and with any other three-player contest $S_2(\mathbf{n}) \le 2$.

A.5 *Proof that efforts are higher-order strategic substitutes (Proposition 4)*

Remember that $g_1(X) = X(1-X)$ and $g_k(X) = -g'_{k-1}(X)X(1-X)$ for all $k > 1$. Therefore, $g_1(X)$ is a second-degree polynomial, $g_2(X)$ third-degree, and so on. In particular, $g_k(X)$ is a polynomial of degree $k+1$ and, therefore, has up to $k+1$ real roots. In the following, I show that all roots are real and in $[0, 1]$. Let these roots be denoted as $Z_{0:k} \le Z_{1:k} \le \cdots \le Z_{k:k}$. The proof keeps track of the order and locations of these roots and their comparison with X^*.

Proof of Proposition 4. The proof relies on three lemmas that I prove below:

1. Lemma 11 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 12 establishes a connection between the total equilibrium effort X^* and $Z_{k-1:k}$. It shows that if we take the sequential n-player contest $\mathbf{n} = (1, \dots, 1)$, then $f_{n-k}(X) = g_k(X)X/(1-X)$ for all $k = 1, \dots, n$. Therefore, if we take the fully sequential contest with n players, we get $f_0(X) = g_n(X)X/(1-X)$, and so the total equilibrium effort X^* of this contest is exactly equal to the second highest root of g_n, i.e., $Z_{n-1:n}$.

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

3. Lemma 13 shows directly[20] that X^* is strictly increasing in each n_t. Therefore, if the contest is not sequential ($n_t > 1$ for some t), then the total effort in this contest is strictly higher than in the fully sequential T-player contest. Thus, $X^* > Z_{T-1:T}$ and $g_T(X^*) > 0$.

4. Finally, Lemma 11 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 11. *Each g_k has the following properties:*

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

(b) g_k *can be expressed as*

$$g_k(X) = -\prod_{j=0}^{k}(X - Z_{j:k}), \tag{21}$$

where $0 = Z_{0:k} < Z_{1:k} < \cdots < Z_{k:k} = 1$.

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

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 $g(1) = 1(1-1) = 0$. Therefore, $g_k'(1) = -g_{k-1}''(1)g(1) - g'(1)g_{k-1}'(1) = g_{k-1}'(1) = \cdots = g_1'(1) = g'(1) = 1 - 2 \cdot 1 = -1$.

2. 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

$$g_k'(X) = -\sum_{s=0}^{k}\prod_{j \neq s}(X - Z_{j:k}).$$

[20] Note the first part of Corollary 1 proves the same claim, but since Proposition 4 establishes a sufficient condition for Theorem 2, and hence its Corollary 1, to avoid a circular argument I prove it here directly.

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.

3. The previous argument also proves the last claim. □

Lemma 12. *If* $\mathbf{n} = (1, \ldots, 1)$, *then* $f_{n-k}(X) = g_k(X)X/(1-X)$ *for all* $k = 1, \ldots, T$.

Proof. Suppose that $\mathbf{n} = (1, \ldots, 1)$. First, $f_{n-1}(X) = X - X(1-X) = X^2 = g_1(X)X/(1-X)$. Now, suppose that $f_{n-k}(X) = g_k(X)X/(1-X)$. Then since

$$\frac{d\dfrac{X}{1-X}}{dX}X(1-X) = \left[\frac{1}{1-X} - \frac{-X}{(1-X)^2}\right]X(1-X) = \frac{X}{1-X},$$

we get that

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

□

Lemma 13. *X^* is strictly increasing in each n_t.*

Proof. I first show that X^* is independent of permutations of $\mathbf{n}$. Fix a contest $\mathbf{n}$ and a period $t > 1$. To shorten the notation, let $\phi_t(X) = f'_t(X)X(1-X)$:

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

Switching n_{t-1} and n_t in $\mathbf{n}$ does not affect f_{t-2} and, therefore, it also does not 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. Take $\widehat{\mathbf{n}} = (n_1 + 1, n_2, \ldots, n_T)$. Then f_1 is unchanged and the corresponding $\widehat{f_0}$ at the original equilibrium X^* is

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

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

A.6 *Proof of decreasing weights lemma (Lemma 1)*

This lemma allows to order some contests, which cannot be ranked according to their information measures. For example, two 10-player contests $\mathbf{n} = (5, 5)$ and $\widehat{\mathbf{n}} = (8, 1, 1)$ have corresponding information measures $\mathbf{S}(\mathbf{n}) = (10, 25)$ and $\mathbf{S}(\widehat{\mathbf{n}}) = (10, 17, 8)$. Contest $\mathbf{n}$ has more second-order information, but $\widehat{\mathbf{n}}$ 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 Theorem 1 confirms this, as $X^* = (13 + \sqrt{41})/20 \approx 0.9702 > \widehat{X}^* = (31 + \sqrt{241})/48 \approx 0.9693$.

Proof of Lemma 1. By Lemma 12, $g_k(X) = \widehat{f}_{\widehat{n}-k}(X)(1 - X)/X$, where $\widehat{f}_{\widehat{n}-k}$ is defined for a sequential $\widehat{n} \geq k$-player contest. Similarly, $g_{k-1}(X) = \widehat{f}_{n+1-k}(X)(1 - X)/X$. Therefore,

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

Now, take $\widehat{n} = T$. Then by Lemma 13, X^* is weakly higher than the highest root of $\widehat{f}_0$. By Property 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 $g_{k-1}(X^*) > g_k(X^*)$. □

A.7 *Proofs of implications of the information theorem (Corollary 1)*

Proof of Corollary 1. Take two contests $\mathbf{n}$ and $\widehat{\mathbf{n}}$ and let X and $\widehat{X}$ be the corresponding total equilibrium efforts.

1. Suppose that $\mathbf{n} < \widehat{\mathbf{n}}$. Then $\mathbf{S}(\mathbf{n}) < \mathbf{S}(\widehat{\mathbf{n}})$ and, therefore, $X < \widehat{X}$.

2. If $\mathbf{n}$ is a permutation of $\widehat{\mathbf{n}}$, then $\mathbf{S}(\mathbf{n}) = \mathbf{S}(\widehat{\mathbf{n}})$ and, therefore, $X = \widehat{X}$.

3. If $\mathcal{I}$ is a coarser partition than $\widehat{\mathcal{I}}$, then $\mathbf{S}(\mathbf{n}) < \mathbf{S}(\widehat{\mathbf{n}})$ and, therefore, $X < \widehat{X}$.

4. If $\sum_t n_t = \sum_t \widehat{n}_t = n$ and there exist t, t' such that $n_t n_{t'} < \widehat{n}_t \widehat{n}_{t'}$ and $n_s = \widehat{n}_s$ for all $s \neq t, t'$, then by construction $S_1(\mathbf{n}) = S_1(\widehat{\mathbf{n}}) = n$ and $S_k(\mathbf{n}) < S_k(\widehat{\mathbf{n}})$ for all $k > 1$. Therefore, $X < \widehat{X}$.

5. Let $\mathbf{n} = (n)$. Then for any $\widehat{\mathbf{n}} \neq \mathbf{n}$, $\mathbf{S}(\mathbf{n}) < \mathbf{S}(\widehat{\mathbf{n}})$, so indeed X is the unique minimum of X^* over all contests. Similarly, if $\widehat{\mathbf{n}} = (1, 1, \dots, 1)$, any other contest has strictly lower measures of information and, therefore, $\widehat{X}$ is the unique maximum of X^* over all contests.

   To establish the final claim of the optimality of equal division of players, let $\mathbf{n}$ be n-player contests where players are distributed among at most T periods. Suppose by contradiction that the corresponding total equilibrium effort X^* is a maximum over all such contests and $\mathbf{n}$ does not split players as equally as possible. In particular, let $k = \lfloor n/T \rfloor$. Equal split requires that each period has either $n_t \in \{k, k + 1\}$ players. Since this is not the case, there exists a period t where $n_t \leq k - 1$ and a period s where $n_s \geq k+1$ (or t, s such that $n_t \leq k$ and $n_s \geq k+2$, then the proof is analogous).

We can now construct a new contest, $\widehat{\mathbf{n}}$, where we have moved one player from period s to period t. Then as $n_s - 1 \geq k > n_t$,

$$\widehat{n}_t \widehat{n}_s = (n_t + 1)(n_s - 1) = n_t n_s - n_t + n_s - 1 > n_t n_s.$$

Therefore, the contest $\widehat{\mathbf{n}}$ is more homogeneous than $\mathbf{n}$ and so $X < \widehat{X}$ by the previous step. Thus, we found a contradiction with the assumption that X is a maximal total effort among such contests. □

A.8 *Proof of the earlier-mover advantage (Proposition 1)*

Proof of Proposition 1. The equilibrium payoff of player i is $u_i(\mathbf{x}^*) = x_i^*(1/X^* - 1)$, 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 $x_i^* > x_j^*$. Using Theorem 1 and equation (9), the difference in equilibrium efforts can be expressed as

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

Now, note that $\mathbf{S}(\mathbf{n}^t) \geq \mathbf{S}(\mathbf{n}^{t+1})$ as there is less information remaining in the game that starts one period later. Moreover, $S_1(\mathbf{n}^t) > S_1(\mathbf{n}^{t+1})$ as $\mathbf{n}^t$ includes player j, whereas $\mathbf{n}^{t+1}$ does not. Finally, note that by Proposition 4, $g_2(X^*) > 0$ and, therefore, $x_i^* - x_j^* > 0$. □

A.9 *Proof of the large contests limit (Proposition 2)*

Proof of Proposition 2. By Theorem 1, each $X^n < 1$. Meanwhile, by Theorem 2, $X^n \geq (n-1)/n$, which is the total equilibrium effort of the simultaneous n-player contest (see Section 3). Therefore, $\lim_{n\to\infty} X^n = 1$.

The total equilibrium effort of a censored contest $\mathbf{n}^n$ is the highest root of $f_0(X)$, which can be expressed by equation (10) as

$$X^n = \sum_{k=1}^{T} S_k(\mathbf{n}^n) g_k(X^n). \tag{22}$$

For each k, function $g_k(X)$ is a twice continuously differentiable function (a polynomial), $g_k(1) = 0$, and $g'_k(1) = -g''_{k-1}(1)g(1) - g'_{k-1}(1)g'(1) = g'_{k-1}(1) = \cdots = g'_1(1) = -1$, as $g_1(X) = X(1-X)$. Therefore, for all $k > 1$,

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

Taking limits from both sides of equation (22) and using the result that $\lim_{n\to\infty} X^n = 1$,

$$1 = \lim_{n\to\infty} X^n = \lim_{n\to\infty} \sum_{k=1}^{T} S_k(\mathbf{n}^n) \frac{g_k(X^n)}{X^n(1-X^n)} X^n(1-X^n) = \lim_{n\to\infty} (1-X^n) \sum_{k=1}^{T} S_k(\mathbf{n}^n).$$

To shorten the notation, let $S^n = \sum_{k=1}^{T} S_k(\mathbf{n}^n)$. Rearranging the previous equation gives

$$0 = \lim_{n\to\infty}\left[1-(1-X^n)S^n\right] = \lim_{n\to\infty}\left[X^n - \left(1-\frac{1}{S^n}\right)\right]S^n. \tag{23}$$

We can express $S^n = \sum_{k=1}^{T} S_k(\mathbf{n}^n) = \prod_{t=1}^{T}(1+n_t^k) - 1$. As $\lim_{n\to\infty} S^n = \infty$, equation (23) implies that

$$\lim_{n\to\infty}\left[X^n - \left(1-\frac{1}{S^n}\right)\right] = \lim_{n\to\infty}\left[X^n - \left(1-\frac{1}{\prod_{t=1}^{T}(1+n_t^n)}\right)\right] = 0.$$

For individual effort of player $i \in \mathcal{I}_t^n$, we can use Theorem 1 and equation (9) to get

$$x_i^n = g_1(X^n) + \sum_{k=1}^{T-t} S_k(\mathbf{n}^t) g_{k+1}(X^n).$$

Taking the limit, again using the facts that $X^n \to 1$ and $g_{k+1}(X^n)/[X^n(1-X^n)] \to 1$,

$$\lim_{n\to\infty} x_i^n = \lim_{n\to\infty}(1-X^n)\left[1+\sum_{k=1}^{T-t} S_k(\mathbf{n}^t)\right].$$

Now, note that $1+\sum_{k=1}^{T-t} S_k(\mathbf{n}^t) = \prod_{s=t}^{T}(1+n_s^n)$. Therefore, using the result from above, we can express the last equation as

$$0 = \lim_{n\to\infty}\left[x_i^n - (1-X^n)\left(1+\sum_{k=1}^{T-t} S_k(\mathbf{n}^t)\right)\right] = \lim_{n\to\infty}\left[x_i^n - \frac{1}{\prod_{s=t}^{T}(1+n_s^n)}\right].$$

$\square$

References

Acemoglu, Daron and Martin K. Jensen (2013), "Aggregate comparative statics." *Games and Economic Behavior*, 81, 27–49. [210]

Admati, Anat R. and Motty Perry (1991), "Joint projects without commitment." *Review of Economic Studies*, 58, 259–276. [210]

Bagwell, Kyle (1995), "Commitment and observability in games." *Games and Economic Behavior*, 8, 271–280. [212]

Baumol, William J. (1982), "Contestable markets: An uprising in the theory of industry structure." *American Economic Review*, 72, 1–15. [209]

Baumol, William J., John C. Panzar, and Robert D. Willig (1988), *Contestable Markets and the Theory of Industry Structure*. Harcourt. Published: Paperback. [222]

Bizzotto, Jacopo, Toomas Hinnosaar, and Adrien Vigier (2023), *The Limits of Commitment*. Available at SSRN: http://dx.doi.org/10.2139/ssrn.4106613. [212]

Bonatti, Alessandro and Johannes Hörner (2011) *Collaborating*. *American Economic Review*, 101, 632–663. [210]

Bronnenberg, Bart J., Sanjay K. Dhar, and Jean-Pierre H. Dubé (2009), "Brand history, geography, and the persistence of brand shares." *Journal of Political Economy*, 117, 87–115. [211]

Che, Yeon-Koo and Ian Gale (2003), "Optimal design of research contests." *American Economic Review*, 93, 646–671. [210]

Cournot, Antoine-Augustin (1838), *Mathematiques de la Theorie des Richesses*. Chez L. Hachette. [209]

Daughety, Andrew F. (1990), "Beneficial concentration." *American Economic Review*, 80, 1231–1237. [210]

Dixit, Avinash (1987), "Strategic behavior in contests." *American Economic Review*, 77, 891–898. [209, 210, 220]

Doval, Laura and Jeffrey C. Ely (2020), "Sequential information design." *Econometrica*, 88, 2575–2608. [211]

Ely, Jeffrey C., George Georgiadis, Sina Khorasani, and Luis Rayo (2022), "Optimal feedback in contests." *Review of Economic Studies*, 1–25. [210]

Fershtman, Chaim and Shmuel Nitzan (1991), "Dynamic voluntary provision of public goods." *European Economic Review*, 35, 1057–1067. [210]

Glazer, Amihai and Refael Hassin (2000), "Sequential rent seeking." *Public Choice*, 102, 219–228. [210]

Halac, Marina, Navin Kartik, and Qingmin Liu (2017), "Contests for experimentation." *Journal of Political Economy*, 125, 1523–1569. [210]

Hinnosaar, Toomas (2021), "Stackelberg independence." *Journal of Industrial Economics*, 69, 214–238. [210]

Hinnosaar, Toomas (2023), *Price Setting on a Network*. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3172236. [226]

Jensen, Martin K. (2018), "Aggregative games." In *Handbook of Game Theory and Industrial Organization* volume I (Luis C. Corchón and Marco A. Marini, eds.). Edward Elgar Publishing. [210, 223]

Kahana, Nava and Doron Klunover (2018), "Sequential lottery contests with multiple participants." *Economics Letters*, 163, 126–129. [210]

Kalyanaram, Gurumurthy, William T. Robinson, and Glen L. Urban (1995), "Order of market entry: Established empirical generalizations, emerging empirical generalizations, and future research." *Marketing Science*, 14, G212–G221. [211]

Konrad, Kai A. (2009), *Strategy and Dynamics in Contests London School of Economics Perspectives in Economic Analysis*, first edition. Oxford University Press. Published: Paperback. [209]

Krueger, Anne O. (1974), "The political economy of the rent-seeking society." *American Economic Review*, 64, 291–303. [209]

Lemus, Jorge and Guillermo Marshall (2021), "Dynamic tournament design—evidence from prediction contests." *Journal of Political Economy*, 129, 383–420. [211]

Li, Fei and Peter Norman (2021), "Sequential persuasion." *Theoretical Economics*, 16, 639–675. [211]

Makris, Miltiadis and Ludovic Renou (2023), "Information design in multi-stage games." *Theoretical Economics*, 18, 1475–1509. [211]

Moldovanu, Benny and Aner Sela (2001), "The optimal allocation of prizes in contests." *American Economic Review*, 91, 542–558. [210]

Moldovanu, Benny and Aner Sela (2006), "Contest architecture." *Journal of Economic Theory*, 126, 70–96. [210]

Morgan, John (2003), "Sequential contests." *Public Choice*, 116, 1–18. [226]

Nitzan, Shmuel (1994), "Modelling rent-seeking contests." *European Journal of Political Economy*, 10, 41–60. [209]

Novshek, William (1980), "Cournot equilibrium with free entry." *Review of Economic Studies*, 47, 473–486. [210]

Olszewski, Wojciech and Ron Siegel (2016), "Large contests." *Econometrica*, 84, 835–854. [210]

Robson, Arthur J. (1990), "Stackelberg and Marshall." *American Economic Review*, 80, 69–82. [210]

Selten, Reinhard (1970), *Preispolitik der Mehrproduktenunternehmung in der statischen Theorie*. Springer-Verlag. [209, 210, 223]

Serena, Marco (2017), "Sequential contests revisited." *Public Choice*, 173, 131–144. [226]

Taylor, Curtis R. (1995), "Digging for golden carrots: An analysis of research tournaments." *American Economic Review*, 85, 872–890. [210]

Tullock, Gordon (1967), "The welfare costs of tariffs, monopolies, and theft" *Economic Inquiry*, 5, 224–232. [209]

Tullock, Gordon (1974), *The Social Dilemma: The Economics of War and Revolution*. University publications. [209]

Varian, Hal R. (1994), "Sequential contributions to public goods." *Journal of Public Economics*, 53, 165–186. [210]

Vojnović, Milan (2015), *Contest Theory: Incentive Mechanisms and Ranking Methods*. Cambridge University Press, Cambridge. [209]

von Stackelberg, Heinrich (1934), *Marktform und Gleichgewicht*. J. Springer. [210]

Wirl, Franz (1996), "Dynamic voluntary provision of public goods: Extension to nonlinear strategies." *European Journal of Political Economy*, 12, 555–560. [210]

Xu, Jin, Chi Zhang, and Jianghua Zhang (2020), "Three-player sequential contests with asymmetric valuations." *Operations Research Letters*, 48, 635–640. [226]

Yildirim, Huseyin (2005), "Contests with multiple rounds." *Games and Economic Behavior*, 51, 213–227. [226]

Co-editor Simon Board handled this manuscript.

Manuscript received 19 December, 2022; final version accepted 22 March, 2023; available online 24 March, 2023.