Centre for
Economic
and Financial
Research
at
New Economic
School

January 2019

The Shape of Luck and Competition in Tournaments

Mikhail Drugov
Dmitry Ryvkin

Working Paper No 251

CEFIR / NES Working Paper series

The shape of luck and competition in tournaments*

Dmitry Ryvkin† Mikhail Drugov‡

This version: January 7, 2019

Abstract

Tournaments are settings where agents' performance is determined jointly by effort and luck, and top performers are rewarded. We study the impact of the "shape of luck" – the details of the distribution of performance shocks – on incentives in tournaments. The focus is on the effect of competition, defined as the number of rivals an agent faces, which can be deterministic or stochastic. We show that individual and aggregate effort in tournaments are affected by an increase in competition in ways that depend critically on the shape of the density and failure (hazard) rate of shocks. When shocks have heavy tails, aggregate effort can decrease with stronger competition.

Keywords: tournament, competition, heavy tails, stochastic number of players, unimodality, log-supermodularity, failure rate

JEL codes: C72, D72, D82

*We are grateful to Johannes Hörner, Bruno Jullien, Michel Le Breton, Margaret Meyer, Marc Möller, Armin Schmutzler, Marco Serena, participants of various seminars and conferences. Parts of this paper were written when the second author was visiting the Toulouse School of Economics, and he is very thankful for its hospitality.

†Department of Economics, Florida State University, Tallahassee, FL 32306-2180, USA, dryvkin@fsu.edu.

‡New Economic School and CEPR, mdrugov@nes.ru.

1 Introduction

[...] because the contests that mete out society's biggest prizes are so bitterly competitive, talent and effort alone are rarely enough to ensure victory. In almost every case, a substantial measure of luck is also necessary.

Robert H. Frank,
"Success and Luck: Good Fortune and the Myth of Meritocracy"

Luck, or lack thereof, plays a crucial role in people's lives. The success stories we observe in business, academia, sports or the arts can often be traced back to a "lucky moment" or an unlikely sequence of events that defined the future path of success. Notable examples are the stories of Bill Gates and Microsoft, Da Vinci's *Mona Lisa* and actor Bruce Willis. In fact, a lucky break, or sequence of breaks, underlying a success story is not an exception but a rule (Mlodinow, 2009; Frank, 2016).[1]

Luck is especially important in *winner-take-all* (WTA) settings where rewards accrue to the select few. Examples include R&D competition, admission to top universities, job applications for an attractive position or competition for promotion in organizations. Careers in professional sports or the arts are predicated almost entirely on WTA incentives. When many hard-working, equally able people are trying to achieve the same thing, success requires a nontrivial amount of luck. As the number of competitors increases, so does the chance that *someone else* will get a better draw, which should discourage individual effort. Yet, economists typically believe that competition provides incentives in markets, at least on the aggregate, leading to larger output, lower prices and higher efficiency (e.g., Ruffin, 1971). In symmetric auctions, revenue increases in the number of bidders (McAfee and McMillan, 1987*a*). For R&D competition, empirical evidence shows a positive effect of competition on investment in innovation even at the individual firm level (Vives, 2008).

In this paper, we study the effect of competition on incentives in WTA settings with a significant luck component. To do so, we utilize the classic rank-order tournament model

[1]When Bill Gates was growing up, he was one of 50 or so students *in the world* who, by sheer chance, had access to a programming terminal allowing to run code with instant feedback (Frank, 2016). The key contract between IBM and Microsoft, which transformed the latter into a world-dominating software company, was signed due to a series of random events; Microsoft did not even develop the initial version of its famous operating system DOS (Mlodinow, 2009). *Mona Lisa* was not considered an exceptional work of art until it was stolen from Louvre in 1911. The newspaper coverage of the painting's theft and recovery two years later created its global fame (Watts, 2011). Bruce Willis acted for seven years in small roles in New York, his main income coming from bartending. He flew to Los Angeles for personal reasons, went to a few television auditions, and got a role in *Moonlighting* far from being everyone's top choice. The first season flopped, but the second one became a hit, and the rest is history (Mlodinow, 2009).

of Lazear and Rosen (1981). Agents' output is given by effort distorted by additive noise, and the agent whose output is the highest wins the tournament and receives a fixed prize. The idiosyncratic noise is synonymous with luck in this model, and different distributions of noise allow for different "shapes of luck."

The contribution of this paper is to answer a very basic question: How are individual and aggregate effort in tournaments affected by an increase in the *number* of competitors? We consider both deterministic and stochastic settings. We show that the shape of the distribution of noise – specifically, of its density and failure (hazard) rate – is crucial for any prediction about the effect of competition on effort. There is not a single comparative static, either for individual or aggregate effort, that cannot be reversed for at least some distribution of noise. Individual and aggregate equilibrium effort can be increasing, decreasing or nonmonotone in the number of players. We systematize and provide new general results for these effects, both when the number of players is known and when it is random. The results have many testable implications, as well as far-reaching applications for tournament design.

In order to cleanly delineate the effects of the number of players and the distribution of noise, we focus on a setting with symmetric players; that is, we assume away differences in ability. While these differences undoubtedly play a critical role in success across the society at large, the most intense competition takes place, and the impact of luck is especially pronounced, in stratified sub-tournaments among (roughly) equally able contestants.[2] Then, in the symmetric pure-strategy equilibrium, there are no differences in effort, and the winner is the luckiest player, i.e., the one with the highest realization of noise.

A general intuition for our results is the following. Suppose there are n players, and let $X_{(i:n)}$ denote the i-th order statistic among n realizations of noise. A marginal increase in a player's effort in the symmetric equilibrium is pivotal – that is, it makes this player the winner – if his noise realization is equal to the highest realization of noise among the other $n-1$ players. This means, formally, that the equilibrium effort is determined by the expectation of the density of noise with respect to $X_{(n-1:n-1)}$, which is stochastically increasing in n. The comparative statics for monotone densities then follow

[2]For example, each year thousands of top high school graduates compete for admission to elite universities; the presence of unqualified applicants in the mix is largely irrelevant. A similar stratification happens naturally in the job market for academic positions or in competition among papers submitted to top journals. Even if quality varies substantially in the initial pool, the actual competition boils down to a subset where quality is very close and, inevitably, luck comes into play. It is also widely believed that tournaments become inefficient as agents' heterogeneity increases (Lazear and Rosen, 1981). Thus, tournament-based incentives are most likely to emerge in settings with symmetric agents.

immediately. For example, in the case of the uniform noise distribution the number of players does not affect the individual equilibrium effort. More generally, adapting results from Athey (2002), we show that the *unimodality* of the distribution of noise leads to the individual equilibrium effort being unimodal in the number of players. We provide a general characterization of the equilibrium comparative statics for unimodal noise distributions, from which all existing results follow as special cases.

Turning to aggregate effort, we start with the case where costs of effort are quadratic. Using arguments similar to the ones in the previous paragraph, we show that aggregate effort can be written as the expectation of the failure (hazard) rate of noise with respect to $X_{(n-1:n)}$. Then, aggregate effort is increasing in the number of players if the noise distribution has increasing failure rate (IFR), such as the normal, uniform and Gumbel distributions. The comparative statics are reversed for distributions with decreasing failure rates (DFR), such as Pareto. We then generalize these results for cost functions more or less convex than quadratic in the sense of the convex transform order (Shaked and Shanthikumar, 2007).

The most counterintuitive results are obtained for aggregate effort in the presence of a *heavy tail* in the distribution of noise. Such noise distributions, most notably *power laws* (Gabaix, 2016), are typically characterized by a decreasing or (interior) unimodal failure rate. Our results then imply a reduction in aggregate effort with the number of players, at least in sufficiently large tournaments. Thus, under heavy-tailed shocks the standard intuition about the effects of competition on effort breaks down. A principal whose goal is to maximize aggregate effort or investment, e.g., in a promotion tournament or an R&D race, would benefit from restricting the number of participants. Heavy-tailed fluctuations are common in many areas often associated with tournament incentives, such as sales of creative and innovative products or the financial sector. Our results suggest that restricting competition can be beneficial in these settings.

We then extend the analysis to WTA tournaments with a stochastic number of players. In many situations, the number of competitors is unknown to the tournament participants at the time they decide how much to invest in competition. This is the case, for example, in coding contests where an unknown and potentially very large number of coders submit their solutions such as The Netflix Prize;[3] in hiring tournaments where a job seeker does

[3] The Netflix Prize competition where the task was to improve the Netflix recommendation algorithm for movies ran for three years overall and about 40,000 teams registered at some point. The final stage lasted 30 days and the two best teams tied in terms of the score. One of the teams won because it submitted its solution *twenty minutes* before the rival (The New York Times, 2009).

not know how many others she is up against; or in promotion tournaments where an employee may not know how many of her colleagues the management is considering for a senior position.

Following the literature on auctions with a stochastic number of bidders (e.g., McAfee and McMillan, 1987*b*; Harstad, Kagel and Levin, 1990; Levin and Ozdenoren, 2004), we assume an arbitrary distribution of the number of players and explore the effects on equilibrium effort of changes in the parameters of the distribution leading to first-order stochastic dominance (FOSD); that is, we explore the effects of a stochastic increase in the number of players. Similar to the case with a known number of players, the unimodality of the distribution of noise plays a key role in robust comparative statics. We show that the preservation of unimodality under uncertainty requires an additional log-supermodularity condition imposed on the distribution of the number of players. This condition follows from arguments similar to those of Athey (2002) for the preservation of single-crossing under uncertainty. The condition is rather weak; it is satisfied, for example, by the family of power series distributions which includes the distributions usually used in the literature with stochastic number of players such as Poisson, binomial, negative binomial and logarithmic distributions.

Finally, we study the design of optimal tournaments with stochastic participation and look at two issues. The first issue is whether the uncertainty about the number of players increases or decreases equilibrium effort. The second issue is whether the contest designer should disclose the realized number of players. Among the 40,000 teams registered for The Netflix Prize competition many were not active. Should Netflix have disclosed the number of active participants?

Relation to prior literature

Starting with the seminal contributions of Tullock (1980) and Lazear and Rosen (1981), there is by now a large theoretical literature on tournaments using the respective models.[4] An important feature of these models distinguishing them from "perfectly discriminating" contests or all-pay auctions (e.g., Hillman and Riley, 1989; Baye, Kovenock and De Vries, 1996; Siegel, 2009; Moldovanu and Sela, 2001) is the presence of uncertainty, or "noise," in the winner determination process.

Yet, the existing analysis of *general* tournament models is quite scarce. For tractability reasons, most of the literature uses either the Tullock CSF (also known as the lottery contest) and its lottery-form generalizations satisfying the axioms of Skaperdas (1996),

[4]See surveys by, e.g., Konrad (2009), Connelly et al. (2014), Corchón and Serena (2018).

or the Lazear-Rosen tournament with two players.[5] Relatively little is known about the basic comparative statics of the WTA tournament model in general. While the symmetric equilibrium effort decreases in the number of players in the Tullock contest (see, for example, surveys by Nitzan, 1994; Corchón and Serena, 2018), it is independent of the number of players in a Lazear-Rosen tournament when the distribution of noise is uniform. For general tournaments with a fixed number of players, Gerchak and He (2003) provide an important first step showing that the equilibrium effort is decreasing in the number of players when the noise density is decreasing or unimodal and symmetric, and increasing when the density is increasing. Even less is known about the behavior of aggregate effort beyond the Tullock contest and Lazear-Rosen tournament with uniformly distributed noise where it is increasing in the number of players.

There is no study of general rank-order tournaments with a stochastic number of players. The previous literature is restricted to the Tullock contest model (and its lottery-form generalizations), which we generalize. Myerson and Wärneryd (2006) compare aggregate equilibrium effort in the case of an arbitrary distribution of group size with expectation μ with the case when the number of players is equal to μ with certainty. Münster (2006) and Lim and Matros (2009) study the comparative statics of effort when the distribution of contest size is binomial.[6] Fu, Jiao and Lu (2011) study the effect of disclosure of the number of participating players on aggregate effort. Boosey, Brookins and Ryvkin (2018) provide results on the effects of disclosure in contests between groups with stochastic sizes. More generally, our paper is related to the literature on games with population uncertainty, including auctions[7] and Poisson games.[8]

The rest of the paper is organized as follows. Section 2 sets up the WTA tournament model with additive or multiplicative noise. Section 3 focuses on tournaments with a deterministic number of players. Section 4 analyzes the case of a stochastic number of players. Section 5 provides technical results on the existence of equilibrium and the preservation of unimodality under uncertainty that are used throughout the paper. Section 6 concludes. The proofs are contained in Appendix A.

[5] Notable exceptions are the papers analyzing optimal prize structures in tournaments with risk-averse players (Nalebuff and Stiglitz, 1983; Green and Stokey, 1983; Krishna and Morgan, 1998; Akerlof and Holden, 2012) and heterogeneity (Balafoutas et al., 2017). See also a survey of the earlier literature by McLaughlin (1988).

[6] Münster (2006) also explores the effect of risk-aversion in the same setting.

[7] For a theoretical analysis of auctions with a stochastic number of bidders see, e.g., McAfee and McMillan (1987*b*), Harstad, Kagel and Levin (1990) and Levin and Ozdenoren (2004).

[8] See, e.g., Myerson (1998, 2000); Makris (2008, 2009); De Sinopoli and Pimienta (2009); Mohlin, Östling and Wang (2015); Kahana and Klunover (2015, 2016).

2 Model setup

2.1 Preliminaries

There are $k \geq 2$ identical, risk-neutral players indexed by $i \in \mathcal{K} = \{1, \ldots, k\}$. All players simultaneously and independently choose efforts $e_i \in \mathbb{R}_+$. The cost of effort e_i to player i is $c(e_i)$, where function $c : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing,[9] strictly convex, and twice differentiable on $(0, e_{\max}]$, where $e_{\max} \equiv c^{-1}(1) < \infty$. Furthermore, $c(0) = c'(0) = 0$. Efforts e_i are perturbed by random additive shocks X_i to generate the players' output levels $y_i = e_i + X_i$. Shocks X_i are i.i.d. with cumulative distribution function (cdf) $F(\cdot)$ and probability density function (pdf) $f(\cdot)$ defined on interval support $U = [\underline{x}, \overline{x}]$, where the bounds $\underline{x}$ and $\overline{x}$ may be finite or infinite.[10] We assume that $f(\cdot)$ is atomless, continuous, piecewise differentiable, and square-integrable. The winner of the tournament is the player whose output is the highest.[11] The winner receives a prize normalized to one, whereas all other players receive zero.[12]

For a given vector of efforts $(e_1, \ldots, e_k)$, the probability of player $i \in \mathcal{K}$ winning the tournament is given by

$$\Pr(i \text{ wins}) = \Pr(y_i > y_j \, \forall j \in \mathcal{K} \setminus \{i\}) = \Pr(e_i + X_i > e_j + X_j \, \forall j \in \mathcal{K} \setminus \{i\})$$
$$= \int_U \left[\prod_{j \in \mathcal{K} \setminus \{i\}} F(e_i - e_j + x) \right] dF(x). \tag{1}$$

Consider a symmetric pure strategy Nash equilibrium in which all players choose effort $e^* > 0$. Using (1), the expected payoff of player $i \in \mathcal{K}$ from some deviation effort e_i is

$$\pi_i(e_i, e^*) = \int_U F(e_i - e^* + x)^{k-1} dF(x) - c(e_i). \tag{2}$$

[9] Throughout this paper, unless noted otherwise, "increasing" will mean nondecreasing and "decreasing" will mean nonincreasing. When distinctions are important, "strictly increasing" and "strictly decreasing" will be used.

[10] In this type of models, it is typically assumed that the shocks are zero-mean. While this assumption can be made without loss of generality, it is not necessary because the probability of winning is determined by differences in shocks. Moreover, shocks can be i.i.d. conditional on an additive common component.

[11] Ties are broken randomly but, under the assumption of atomless $f(\cdot)$, occur with probability zero.

[12] A more general setting could involve up to n distinct prizes; however, in this paper we are not concerned with optimal contract design, and use the simplest "winner-take-all" prize structure. In a follow-up paper (Drugov and Ryvkin, 2018), we extend the techniques developed here to study the optimal allocation of prizes and related issues.

The symmetric first-order condition for payoff maximization, $\left.\frac{\partial \pi_i(e_i,e^*)}{\partial e_i}\right|_{e_i=e^*} = 0$, gives

$$c'(e^*) = b_k \equiv (k-1)\int_U F(x)^{k-2} f(x) dF(x). \tag{3}$$

Note that $c'(\cdot)$ is a strictly increasing function; therefore, if Eq. (3) has a solution then it is positive and unique for $k \geq 2$. Such a solution, denoted e_k^*, is a natural candidate for the symmetric pure strategy equilibrium effort. In Section 5.1, we discuss the existence of equilibrium in detail and formulate sufficient conditions on the primitives of the model for e_k^* to be the equilibrium.

Assuming it exists, the symmetric equilibrium effort e_k^* is determined entirely by coefficients b_k defined in (3), and most of the analysis that follows revolves around the properties of these coefficients. Let $F^{-1}(z) = \inf\{x \in U : F(x) \geq z\}$ denote the quantile function of the distribution of noise. Introduce also an unnormalized density function $m(z) = f(F^{-1}(z))$, known as the inverse quantile density function (Parzen, 1979). Inheriting its properties from $f(\cdot)$, function $m(\cdot)$ is continuous, piecewise differentiable, and integrable. Using the probability integral transformation $z = F(x)$, it will sometimes be convenient to rewrite b_k in Eq. (3) as

$$b_k = (k-1)\int_0^1 z^{k-2} m(z) dz = \int_0^1 m(z) dz^{k-1} = \mathrm{E}(m(Z_{(k-1:k-1)})). \tag{4}$$

Here, $Z_{(k-1:k-1)}$ is the $(k-1)$-th order statistic of $k-1$ i.i.d. uniform random variables on $[0,1]$. Representation (4) separates the effects of the number of players, k, from the effects of the distribution of noise. The latter are contained entirely in the inverse quantile density $m(\cdot)$, while the former are determined by a family of FOSD-ordered highest order statistics of the uniform distribution with cdfs z^{k-1}.

2.2 Multiplicative noise

Via simple transformations of the distribution of noise and the cost of effort, the additive noise model above accommodates tournaments with multiplicative noise where player i's output is given by $y_i = e_i X_i$ and X_i are i.i.d. with a nonnegative support. The probability of player i winning the tournament of k players can then be written as

$$\Pr(i \text{ wins}) = \Pr(e_i X_i > e_j X_j \ \forall j \in \mathcal{K} \setminus \{i\}) = \Pr(\hat{e}_i + \hat{X}_i > \hat{e}_j + \hat{X}_j \ \forall j \in \mathcal{K} \setminus \{i\}),$$

where $\hat{e}_i = \ln e_i$ and $\hat{X}_i = \ln X_i$. Note that $\hat{e}_i$ is no longer restricted to nonnegative values. Defining $\hat{F}(x) = F(\exp(x))$ as the cdf of the transformed shocks $\hat{X}_i$, and $\hat{c}(\hat{e}) = c(\exp(\hat{e}))$ as the cost function for the transformed effort $\hat{e}$, this model reduces to a tournament model with additive noise, and all results go through.

Specifically, the first-order condition (3) for the transformed equilibrium effort, $\hat{e}^*_k = \ln e^*_k$, is $\hat{c}'(\hat{e}^*_k) = \hat{b}_k$, where $\hat{b}_k$ is based on distribution $\hat{F}$. Interestingly,

$$\hat{c}'(\hat{e}) = c'(\exp(\hat{e}))\exp(\hat{e}) = c'(e)e;$$

therefore, the first-order condition for the original equilibrium effort is $c'(e^*_k)e^*_k = \hat{b}_k$. This leads to the following proposition.

Proposition 1 *The symmetric equilibrium effort in a tournament with multiplicative noise is the same as in the tournament with additive noise distributed with cdf $\hat{F}(x) = F(\exp(x))$ and the cost of effort $c_{\mathrm{m}}(e) = \int_0^e c'(t)tdt$.*

Tullock contests

As an illustration, consider contests with the CSF of Tullock (1980). The probability of player i winning the contest of size k is given by $\frac{e_i^r}{\sum_{j=1}^k e_j^r}$, where $r > 0$ is a parameter measuring the level of noise (the "discriminatory power" of the contest) such that a lower r corresponds to higher noise. The cost of effort is linear, $c(e) = c_0 e$. Following Jia (2008), this probability of winning can be written as $\Pr(e_i X_i > e_j X_j \ \forall j \in \mathcal{K} \setminus \{i\})$ where $X_j > 0$ are i.i.d. with the generalized inverse exponential distribution with cdf $F(x) = \exp(-x^{-r})$.

That is, the Tullock contest can be represented as a tournament with multiplicative noise. We can now use Proposition 1 to find the corresponding tournament with additive noise. The transformed shocks $\hat{X}_i = \ln X_i$ have the generalized extreme value type-I (or Gumbel) distribution with cdf $\hat{F}(x) = F(\exp(x)) = \exp[-\exp(-rx)]$ and pdf $\hat{f}(x) = r\exp[-rx - \exp(-rx)]$ (see Jia, Skaperdas and Vaidya, 2013). This pdf is unimodal, with a maximum at zero, and skewed to the right. The transformed cost of effort is $c_{\mathrm{m}}(e) = \int_0^e c_0 t dt = \frac{c_0 e^2}{2}$. The first-order condition then takes the form $c_0 e^*_k = \hat{b}_k$, where $\hat{b}_k$ is given by Eq. (4) with $m(z) = \hat{f}(\hat{F}^{-1}(z)) = -rz\ln z$:

$$\hat{b}_k = -r(k-1)\int_0^1 z^{k-1}\ln z dz = \frac{r(k-1)}{k^2}, \tag{5}$$

which gives the well-known equilibrium effort in the Tullock contest, $e^*_k = \frac{r(k-1)}{c_0 k^2}$.

This approach can be further generalized to cover contests with a CSF of the form $\frac{g(e_i)}{\sum_{j=1}^{k} g(e_j)}$, where $g(\cdot)$ is a strictly increasing "impact function," and a possibly nonlinear cost of effort $c(e_i)$. By introducing effective efforts $w_i = g(e_i)$ and costs $C(x_i) = c(g^{-1}(w_i))$, such models are reduced to the Tullock contest with $r = 1$, and the results apply. Specifically, Proposition 1 implies that the symmetric equilibrium level of effective effort, w^*, satisfies the equation $\frac{k-1}{k^2} = C'(w^*)w^*$, where $C'(w) = \frac{c'(g^{-1}(w))}{g'(g^{-1}(w))}$. Substituting back $w^* = g(e_k^*)$, we obtain for the equilibrium effort $\frac{k-1}{k^2} = \frac{c'(e_k^*)g(e_k^*)}{g'(e_k^*)}$.

3 Tournaments with deterministic group size

3.1 Individual equilibrium effort

Before formulating our main results, we summarize the existing results and develop some intuition. As discussed in Section 2, the properties of the symmetric equilibrium effort are determined by coefficients b_k, see Eq. (3). These coefficients represent the marginal benefit of effort in equilibrium, and can be written as

$$b_k = \int_U f(x) dF(x)^{k-1} = \int_U f(x) f_{(k-1:k-1)}(x) dx, \tag{6}$$

where $F(x)^{k-1}$ is the cdf of the $(k-1)$-th order statistic among $k-1$ i.i.d. draws from distribution F, and $f_{(k-1:k-1)}(x) = \frac{d}{dx} F(x)^{k-1}$ is the corresponding pdf. Indeed, in the symmetric equilibrium player i wins the tournament if her realization of noise, X_i, exceeds $X_{(k-1:k-1)} = \max_{j \neq i} X_j$ – the largest shock among the other $k-1$ players. A marginal increase in the player's effort is then pivotal when there is a tie between the two shocks, i.e., it is determined by the probability density of $X_i - X_{(k-1:k-1)}$ at zero, cf. Eq. (6). This representation immediately leads to comparative statics results for monotone pdfs $f(x)$.

Lemma 1 *(i) If $f(x)$ is increasing (decreasing) then e_k^* is increasing (decreasing) for $k \geq 2$.*
(ii) e_k^* is constant for $k \geq 2$ if and only if $f(x)$ is a uniform distribution.

Indeed, the order statistics $X_{(k-1:k-1)}$ are FOSD-increasing in k; therefore, the realizations of noise from the upper tail of $f(x)$ become more important as k increases. For example, if $f(x)$ is increasing then the probability of having relevant noise realizations increases with k, resulting in a higher equilibrium effort. Part (i) and the "if" part of part (ii) of

Lemma 1 have been proved by Gerchak and He (2003). The "only if" part of part (ii) is proved in Appendix A using the representation (4) for coefficients b_k.

The intuition behind representation (6) allows us to also obtain large-k asymptotic results for an arbitrary $f(x)$. As discussed above, as k increases, b_k is determined by increasingly higher order statistics $X_{(k-1:k-1)}$ whose probability density shifts to the right; hence, the asymptotic behavior of b_k is determined by the shape of the upper tail of pdf $f(x)$. Specifically, a decreasing (increasing) upper tail of $f(x)$ will lead to a decreasing (increasing) b_k for large k. The following lemma states the result formally.

Lemma 2 *Define* $\hat{x} = \inf\{x' \in U : f(x) \text{ is monotone for } x > x'\}$. *If* $f(x)$ *is decreasing (increasing) for* $x > \hat{x}$, *then there exists a* $\hat{k}$ *such that* e^*_k *is decreasing (increasing) for all* $k > \hat{k}$.

Point $\hat{x}$ defined in Lemma 2 determines the location of the "last" interior peak or dip of $f(x)$. If pdf $f(x)$ is monotone, $\hat{x} = \underline{x}$ and b_k is either decreasing or increasing for all $k \geq 2$, by Lemma 1. If $f(x)$ is nonmonotone, b_k is asymptotically decreasing or increasing depending on the behavior of the "last" monotone part of $f(x)$. Lemma 2 is proved in Appendix A using the representation (4) for coefficients b_k.

We now turn to the main results of this section. *Unimodal* distributions are an important class for which universal global properties of coefficients b_k can be established. To this end, we turn to representation (4) of coefficients b_k as expectations of the inverse quantile density, and make use of Lemma 7 in Section 5.2 on the properties of expectations of unimodal functions. Note that $m(z)$ has the same monotonicity as $f(x)$, and for a higher k the weights in the expectation $\mathrm{E}(m(Z_{(k-1:k-1)}))$ shift to the right; that is, the same intuition as in representation (6) applies.

Proposition 2 *(i) If* $f(x)$ *is unimodal then* e^*_k *is unimodal for* $k \geq 2$.
(ii) If $f(x)$ *is unimodal and symmetric then* $e^*_2 = e^*_3$, *and* e^*_k *is decreasing for* $k \geq 3$.
(iii) If $f(x)$ *is symmetric (not necessarily unimodal) then* $e^*_2 = e^*_3$.

Part (i) of Proposition 2 is the main result of this section, and it follows directly from Lemma 7. Indeed, considering the representation (4) with cdf $H(z, k) = z^{k-1}$, it is straightforward to show that $-H_k(z, k) = z^{k-1}(1 - z)$ is log-supermodular. Parts (ii) and (iii) are special cases, which have been proved by Gerchak and He (2003) (we provide a direct proof in Appendix A for completeness).

Part (i) of Proposition 2 shows that the unimodality of b_k (and hence, of the equilibrium effort e^*_k) is a common property of unimodal noise distributions. Note that it relies only on the FOSD-ordering of cdfs $H(z,k) = z^{k-1}$ and the log-supermodularity of $-H_k(z,k)$, but not on the specific order-statistic structure of $H(z,k)$. The unimodality result can, therefore, be extended to other settings, such as the case when the number of players is stochastic (see Section 4). In contrast, parts (ii) and (iii) are more specialized and rely on the functional form of $H(z,k)$.

Additionally, Proposition 2 allows us to characterize the behavior of b_k for *U-shaped* distributions such that $-f(x)$ is unimodal. Of interest is the case when $f(x)$ is U-shaped and nonmonotone (when f is monotone, Lemma 1 applies).

Corollary 1 *(i) If $f(x)$ is U-shaped and nonmonotone then e_k^* is U-shaped for $k \geq 2$. (ii) If $f(x)$ is U-shaped, nonmonotone and symmetric then $b_2 = b_3$, and e_k^* is increasing for $k \geq 3$.*

Part (i) is a direct corollary of Proposition 2(i), while part (ii) is a special case that follows from part (ii) of the proposition and has been proved by Gerchak and He (2003).

For an example of an interior unimodal sequence b_k, consider the type I generalized logistic distribution, which has cdf $F(x) = \frac{1}{(1+\exp(-x))^a}$ with parameter $a > 0$ (Johnson, Kotz and Balakrishnan, 1995). The standard logistic distribution is obtained for $a = 1$. Then, $b_k = \frac{a(k-1)}{k(ak+1)}$. Since $b_{k+1} - b_k \propto 1 + a - ak(k-1)$ is decreasing in k, b_k is either monotonically decreasing or interior unimodal. In particular, b_k reaches the maximum at $\hat{k}$ if $a = \frac{1}{\hat{k}^2 - \hat{k} - 1}$; cf. Figure 1. Figure 2 shows an example of an interior U-shaped sequence b_k.

The unimodality of f is not necessary for the unimodality of b_k (and e_k^*), but it is a tight condition. That is, a non-unimodal distribution of noise can produce a non-unimodal sequence b_k. For an example, consider $m(z) = 0.22z^3 - 0.39z^2 + 0.2z$ which gives rise to a non-unimodal b_k, see Figure 3.[13] At the same time, a non-unimodal $f(x)$ does not necessarily lead to a non-unimodal sequence b_k. For example, a bimodal mixture of two normal distributions with $f(x) = \frac{1}{2} f_{N(-12,4)}(x) + \frac{1}{2} f_{N(12,4)}(x)$, where $f_{N(\mu,\sigma^2)}(x)$ is the normal pdf with mean μ and variance σ^2, generates a decreasing sequence b_k.

More generally, one may ask whether there is any "higher-order" universality in the behavior of b_k (and e_k^*) for multimodal densities. The answer is yes, to some extent. Part (i) of Proposition 2 relies on the fact that $m'(z)$ is single-crossing and $z^{k-1}(1-z)$ is log-supermodular in (z,k). As mentioned in Section 5.2, log-supermodularity is also known

[13] This function $m(z)$ corresponds to the quantile function $F^{-1}(z) = -\frac{5}{2} \ln\left(22\,z^2 - 39\,z + 20\right) + \frac{195\sqrt{239}}{239} \arctan\left(\frac{(44\,z-39)\sqrt{239}}{239}\right) + 5 \ln z$; there is no closed-form expression for $F(x)$.

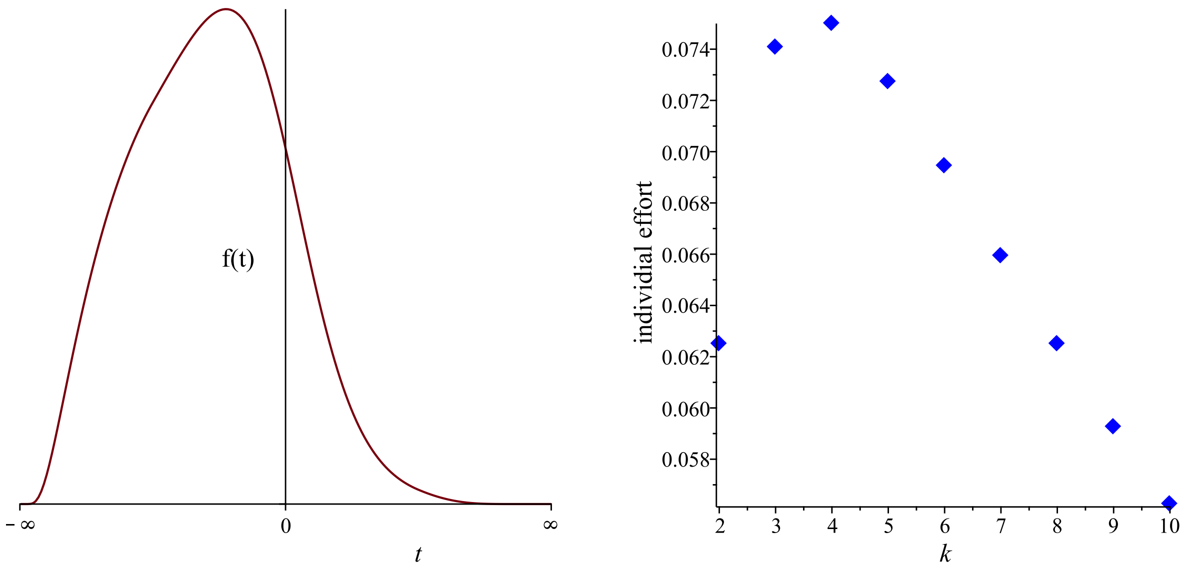

Figure 1: *Left*: The pdf $f(x)$ of the type I generalized logistic distribution with $a = \frac{1}{6}$. *Right*: Individual equilibrium effort e_k^* as a function of k for effort cost function $c(e) = \frac{1}{2}e^2$.

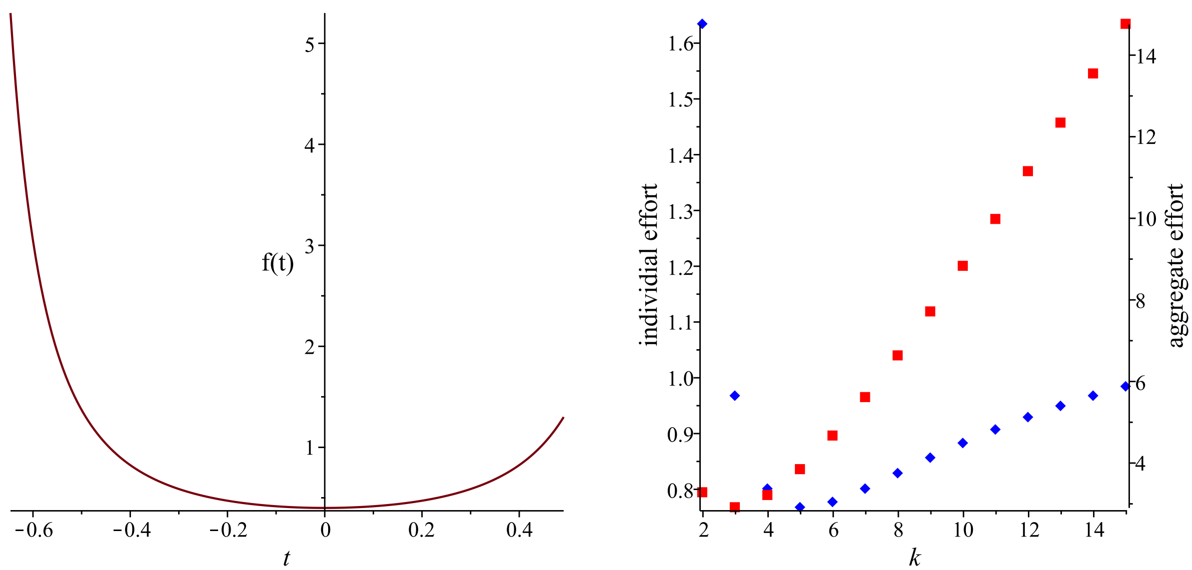

Figure 2: *Left*: The pdf $f(x)$ of a distribution with cdf $F(x) = 0.2\tan(2x) + 0.7$ defined on $[-0.646, 0.491]$. *Right*: Individual equilibrium effort e_k^* (blue diamonds, left scale) and aggregate equilibrium effort E_k^* (red squares, right scale) as a function of k for effort cost function $c(e) = \frac{1}{2}e^2$.

as total positivity of order 2 (TP_2), a special case of total positivity of order r (TP_r). The *variation-diminishing property* of totally positive kernels (Karlin, 1968) states that

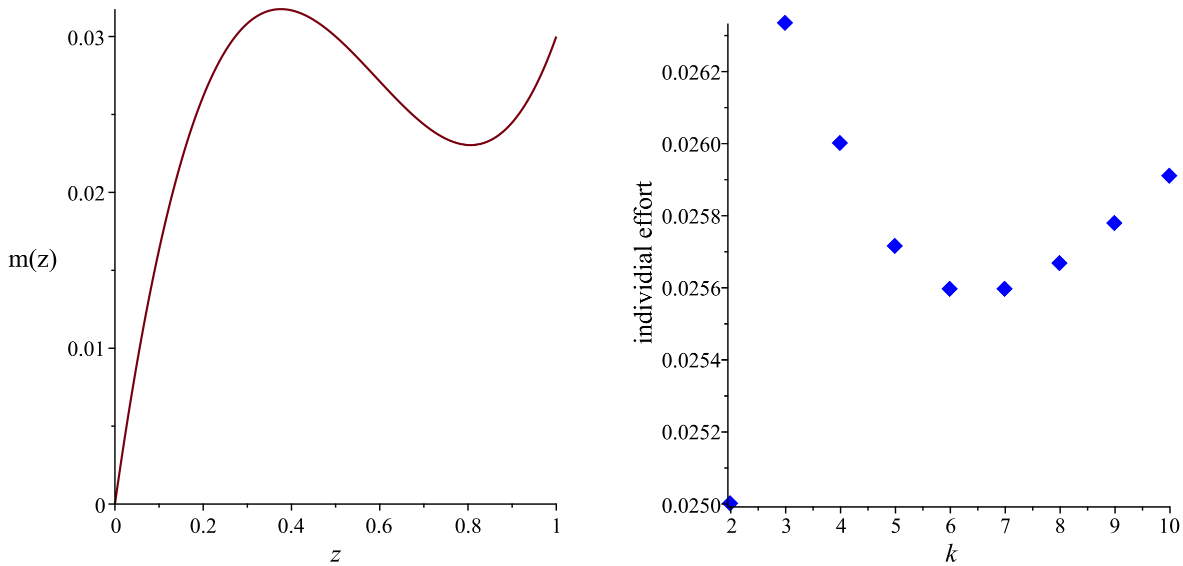

Figure 3: *Left*: $m(z) = 0.22z^3 - 0.39z^2 + 0.2z$. *Right*: Individual equilibrium effort e_k^* as a function of k for effort cost function $c(e) = \frac{1}{2}e^2$.

if function $v : S_1 \times S_2 \to \mathbb{R}$, with $S_1, S_2 \subseteq \mathbb{R}$, is TP_r[14] and function $\phi : S_2 \to \mathbb{R}$ changes sign $j \leq r-1$ times on S_2 then function $\tilde{\phi}(x) = \int_{S_2} v(x,y)\phi(y)dy$ changes sign at most j times on S_1. Moreover, if $\tilde{\phi}$ changes sign exactly j times then it follows the same sequence of sign changes as ϕ. It can be shown that $z^{k-1}(1-z)$ is, in fact, TP_∞ (Marshall, Olkin and Arnold, 2011, p. 759); therefore, if $m'(z)$ has any number j of sign changes then $b_{k+1} - b_k$ will have at most j sign changes. We conclude that if $f(x)$ has j modes, b_k will have at most j modes, and if b_k has exactly j modes then the sequence of local minima and maxima of b_k will follow the shape of $f(x)$. The case of unimodal (or U-shaped) $f(x)$ is special because $m'(z)$ has at most one sign change, and hence b_k is either monotone or interior unimodal (or U-shaped). Figure 3 illustrates a case when b_k and $f(x)$ both have two modes, and b_k follows the shape of $f(x)$.

[14]Function v is TP_r if for all $l = 1, \ldots, r$ and all sequences $x_1 < \ldots < x_l$, $y_1 < \ldots < y_l$ ($x_i \in S_1$, $y_j \in S_2$),

$$\det \begin{pmatrix} v(x_1,y_1) & \ldots & v(x_1,y_l) \\ \vdots & & \vdots \\ v(x_l,y_1) & \ldots & v(x_l,y_l) \end{pmatrix} \geq 0.$$

3.2 Aggregate equilibrium effort

From the tournament designer's perspective, it is of eminent interest how aggregate equilibrium effort $E_k^* = ke_k^*$ changes with the number of players. To gain some intuition, note that the change in aggregate effort when the number of players increases from $k-1$ to k, $\Delta E_k^* = E_k^* - E_{k-1}^*$, can be written as $\Delta E_k^* = e_k^* + (k-1)\Delta e_k^*$, where $\Delta e_k^* = e_k^* - e_{k-1}^*$ is the change in individual effort. An increase in the number of players affects aggregate effort in two ways: The direct positive effect, represented by the term e_k^*, and the indirect equilibrium effect, $(k-1)\Delta e_k^*$, which can be positive or negative. Obviously, aggregate effort will increase in k when $e_k^* \geq e_{k-1}^*$, i.e., whenever individual effort is increasing in k. It is, however, also possible to have aggregate effort increasing in k when e_k^* is decreasing or nonmonotone. For example, in the Tullock contest with linear costs individual effort $e_k^* = \frac{r(k-1)}{k^2}$ is decreasing but aggregate effort $E_k^* = \frac{r(k-1)}{k}$ is increasing in k. Also, unlike the comparative statics of e_k^*, the comparative statics of E_k^* can be sensitive to the shape of the cost function $c(e)$. The reason is that $E_k^* = kc'^{-1}(b_k)$, where $c'^{-1}(\cdot)$ is the inverse marginal cost of effort.

In fact, if $\Delta e_k^* < 0$ for some k (i.e., $b_k < b_{k-1}$), it is always possible to find parameters such that aggregate effort will be decreasing in k as well. To see this, consider a cost function of the form $c(e) = c_0 e^{\xi}$, $\xi > 1$, which leads to the individual effort $e_k^* = \left(\frac{b_k}{c_0\xi}\right)^{\frac{1}{\xi-1}}$ and

$$\Delta E_k^* = (k-1)\left(\frac{b_{k-1}}{c_0\xi}\right)^{\frac{1}{\xi-1}}\left[\frac{k}{k-1}\left(\frac{b_k}{b_{k-1}}\right)^{\frac{1}{\xi-1}} - 1\right], \tag{7}$$

which immediately implies the following result.

Lemma 3 *Suppose $c(e) = c_0 e^{\xi}$, $\xi > 1$. Then $E_k^* \geq E_{k-1}^*$ if and only if $\frac{b_k}{b_{k-1}} \geq \left(\frac{k-1}{k}\right)^{\xi-1}$.*

One consequence of Lemma 3 is that for any $k \geq 3$ it is always possible to find a sufficiently large ξ such that $E_k^* \geq E_{k-1}^*$. The intuition is that a higher ξ makes the cost function more convex and hence, reduces the sensitivity of the equilibrium effort to its marginal benefit, i.e., b_k. Then, for a sufficiently high ξ the direct positive effect of a higher number of players dominates the indirect equilibrium effect. On the other hand, ξ can be arbitrarily close to 1 in which case the equilibrium effort becomes infinitely sensitive to b_k;[15] therefore, if $b_k < b_{k-1}$ for some k, it is always possible to find a $\xi > 1$ such that the condition of Lemma 3 does not hold and hence $E_k^* < E_{k-1}^*$.

[15] As ξ gets closer to 1, it becomes more difficult to satisfy the equilibrium existence conditions, but for any given ξ they can always be satisfied for a sufficiently high c_0 and/or a sufficiently dispersed distribution of noise (see Section 5.1).

For illustration, compare tournaments with group sizes $k = 2$ and 3. It follows from Proposition 2 that $b_3 \geq b_2$, and hence $E_3^* > E_2^*$, when $f(x)$ is symmetric or increasing. However, if $f(x)$ is decreasing (and nonconstant), we have $b_3 < b_2$, in which case $E_3^* < E_2^*$ for $\xi < 1 + \frac{\ln\left(\frac{b_2}{b_3}\right)}{\ln\left(\frac{3}{2}\right)}$. For example, consider the distribution of noise with cdf $F(x) = x^\alpha$ and pdf $f(x) = \alpha x^{\alpha-1}$ on $[0,1]$, with $\alpha > \frac{1}{2}$.[16] This gives $m(z) = \alpha z^{\frac{\alpha-1}{\alpha}}$ and $b_k = \frac{\alpha^2(k-1)}{\alpha k-1}$; therefore, $\frac{b_3}{b_2} = \frac{2(2\alpha-1)}{3\alpha-1} < 1$ if and only if $\alpha < 1$, i.e., $f(x)$ is decreasing. For $\alpha = \frac{3}{4}$, we obtain $E_3^* < E_2^*$ for $\xi < 1 + \frac{\ln\left(\frac{5}{4}\right)}{\ln\left(\frac{3}{2}\right)} \approx 1.55$.

Despite the presence of these two often countervailing effects, there are very simple and powerful sufficient conditions for the monotonicity of aggregate effort with respect to the number of players. Before we proceed, let us remind the reader of some basic concepts from duration analysis. For a random variable X with pdf $f(x)$ and cdf $F(x)$, the failure (or hazard) rate is defined as $h(x) = \frac{f(x)}{1-F(x)}$. A distribution is characterized as having increasing failure rate (IFR) if $h(x)$ is increasing, and decreasing failure rate (DFR) if $h(x)$ is decreasing. IFR is implied by the log-concavity of pdf $f(x)$, while DFR is implied by the log-convexity of $f(x)$ provided $f(\overline{x}) = 0$. The exponential distribution, with $f(x) = \lambda \exp(-\lambda x)$, has a constant failure rate λ and hence is both IFR and DFR. Most standard distributions fall into one of the monotone failure rate classes. As we show below, the behavior of aggregate effort is determined by the failure rate.

Note that b_k in (4) can be rewritten as

$$b_k = \frac{1}{k}\int_0^1 \frac{m(z)}{1-z} dF^B(z; k-1, 2) = \frac{1}{k}\mathrm{E}(h_q(Z_{(k-1:k)})). \tag{8}$$

Here, $F^B(z; k-1, 2)$ is the cdf of the beta distribution with parameters $k-1$ and 2, which is also the cdf of $Z_{(k-1:k)}$, the $(k-1)$-th order statistic from k i.i.d. draws of the uniform distribution on $[0,1]$. Function $h_q(z) = \frac{m(z)}{1-z}$ is the hazard quantile function, which is a quantile representation of failure rate $h(x)$ and has the same monotonicity properties.

We start with the simplest case when the cost of effort is quadratic, $c(e) = c_0 e^2$. In this case aggregate effort is proportional to kb_k, and from Eq. (8) we have $kb_k = \mathrm{E}(h_q(Z_{(k-1:k)}))$. Since order statistics $Z_{(k-1:k)}$ are FOSD-ordered in k, this immediately implies the monotonicity of E_k^* for monotone failure rates. Moreover, it can be shown that the cdf $F^B(z; k-1, 2)$ satisfies the appropriate log-supermodularity condition, and hence Lemma 7 can be applied for unimodal and U-shaped failure rates. The results are summarized as follows.

[16] The restriction $\alpha > \frac{1}{2}$ ensures that $m(z)$ is integrable on $[0,1]$.

Lemma 4 *Suppose the cost of effort is quadratic, $c(e) = c_0 e^2$. Then*
(i) If $f(x)$ is IFR (DRF) then E_k^* is increasing (decreasing) for $k \geq 2$.
(ii) If $f(x)$ is exponential, with $f(x) = \lambda \exp(-\lambda x)$, then $E_k^* = \frac{\lambda}{2c_0}$ is constant for $k \geq 2$.
(iii) If $f(x)$ has a unimodal (U-shaped) failure rate then E_k^* is unimodal (U-shaped) for $k \geq 2$.

To understand the role of the failure rate in Lemma 4, note that representation (8) can be rewritten through the original distribution of noise in the form $b_k = \frac{1}{k}\mathrm{E}(h(X_{(k-1:k)}))$, where $X_{(k-1:k)}$ is the second highest order statistic among k noise realizations. The intuition, therefore, is similar to the one for coefficients $b_k = \mathrm{E}(f(X_{(k-1:k-1)}))$ discussed in Section 3.1. Indeed, winning the tournament can be interpreted as both surpassing $X_{(k-1:k-1)}$, the highest realization among the other $k-1$ players, and surpassing $X_{(k-1:k)}$, the second highest realization among all k players. Note that the failure rate can be written as $h(x) = \frac{f(x)}{1-F(x)} = f(x|X \geq x)$, i.e., the pdf of noise at $X = x$ conditional on $X \geq x$. However, only realizations of noise exceeding $X_{(k-1:k)}$ can lead to winning; therefore, $\mathrm{E}(h(X_{(k-1:k)})) = \mathrm{E}(f(X_{(k-1:k)}|X \geq X_{(k-1:k)}))$ gives exactly the relevant conditional expectation. In order to obtain b_k, it needs to be multiplied by $\Pr(X \geq X_{(k-1:k)}) = \frac{1}{k}$, cf. (8), which makes this representation suitable for aggregate effort as it conveniently subsumes the effect of multiplier k.

Part (i) of Lemma 4 for IFR distributions generalizes the result for the Tullock contest with linear costs. Indeed, as shown in Section 2.2, the properties of equilibrium in such a contest are equivalent to those of a tournament with a quadratic cost and Gumbel distribution of noise, which is IFR. To understand the behavior of E_k^* for DFR distributions, note that such a distribution has a decreasing density which falls faster than its cdf is increasing. Hence, individual effort is decreasing (see Lemma 1(i)), and so fast that aggregate effort decreases too. For a simple example, consider the $F_{2,2}$-distribution whose pdf and cdf are $f(x) = \frac{1}{(1+x)^2}$ and $F(x) = \frac{x}{1+x}$ defined for $x \geq 0$. Then, $b_k = \frac{2}{k(k+1)}$ and aggregate effort $E_k^* = \frac{2}{k+1}$ is strictly decreasing with the number of players.

For part (iii), examples of distributions with interior unimodal failure rates include the F-distribution and beta distribution for some parameters, and the lognormal distribution (for details, see Bagnoli and Bergstrom, 2005). Figure 2 provides an example of a distribution with a U-shaped failure rate, which generates a U-shaped aggregate effort.

For noise distributions with multimodal failure rates, the same generalization as discussed at the end of Section 3.1 applies.

Lemma 3 shows that a "more convex" cost function is more likely to lead to aggregate effort increasing in the number of players. While in Lemma 3 this result is restricted to power cost functions, which are naturally ordered in convexity by parameter ξ, it is, in fact, very general. Specifically, Lemma 4 can be extended to cost functions that are, in a well-defined sense, more or less convex than quadratic. This leads to Proposition 3 below which is the main result of this section.

For two strictly increasing functions c_1 and c_2, we define function c_1 to be more convex than c_2 if $c_1(c_2^{-1}(\cdot))$ is convex. This definition is equivalent to requiring that there exists a strictly increasing, convex function u such that $c_1(e) = u(c_2(e))$; indeed, defining $t = c_2(e)$, obtain $u(t) = c_1(c_2^{-1}(t))$. This partial order is related to the likelihood ratio order of random variables, whereby a random variable X is said to be smaller than random variable Y if the ratio of pdfs $\frac{f_Y(x)}{f_X(x)}$ is increasing in x. An equivalent condition is that $F_Y(F_X^{-1}(z))$ is convex (Shaked and Shanthikumar, 2007). In our case, it implies that the ratio of marginal costs $\frac{c_1'(e)}{c_2'(e)}$ is increasing in e. The definition of a less convex function is analogous.

It follows that a cost function $c(e)$ is more convex than quadratic if $c(\sqrt{t})$ is convex in t or, equivalently, the ratio $\frac{c'(e)}{e}$ is increasing. Thus, a cost function is more convex than quadratic if the marginal cost increases faster than linear. For thrice differentiable functions, this condition implies $c''' \geq 0$, and is equivalent to it provided $c'(0) = 0$. Indeed, the condition that $\frac{c'(e)}{e}$ is increasing is equivalent to $c''(e)e \geq c'(e)$, which implies that $c'(e)$ is convex. Conversely, if $c'(0) = 0$, the convexity of $c'(e)$ implies $c''(e)e \geq c'(e)$. A less convex than quadratic function has $c''' \leq 0$.

We are now in a position to formulate the main result.

Proposition 3 *If $f(x)$ is IFR and $c(e)$ is more convex than quadratic (DFR and $c(e)$ is less convex than quadratic), then E_k^* is increasing (decreasing) for $k \geq 2$.*

Proposition 3 generalizes part (i) of Lemma 4 and provides very general sufficient conditions for monotonicity of aggregate effort in WTA tournaments. When the distribution of noise is IFR and effort costs are sufficiently sensitive, a tournament designer can benefit from additional participants; the opposite is true, i.e., aggregate effort is maximized by $k = 2$, if noise is DFR and effort costs are not very sensitive. If the conditions of Proposition 3 do not hold, the competing direct and indirect effects of the number of players can lead to nonmonotonicities in aggregate effort.

4 Tournaments with stochastic group size

Consider now a setting in which the number of players in the tournament, K, is a random variable taking nonnegative integer values. The maximal possible number of players $n \geq 2$ can be finite or infinite. Let $p = (p_0, p_1, \ldots, p_n)$ denote the probability mass function (pmf) of K, where $p_k = \Pr(K = k)$ is the probability of having k players in the tournament, with $\sum_{k=0}^{n} p_k = 1$. The expected number of players $\bar{k} = \sum_{k=0}^{n} k p_k$ is finite. Operationally, it is convenient to think about a set of potential participants $\mathcal{N} = \{1, \ldots, n\}$ from which a subset $\mathcal{K} \subseteq \mathcal{N}$ is randomly drawn such that $\Pr(|\mathcal{K}| = k) = p_k$, and subsets of the same cardinality $|\mathcal{K}|$ have the same probability of being drawn. Each player is informed if she is selected, but is not informed about the value of K.

Let S_i denote a random variable equal to 1 if player $i \in \mathcal{N}$ is selected for participation and zero otherwise, and let $\tilde{K} = (K|S_i = 1)$ denote the random number of players in the tournament from the perspective of a participating player. The distribution of $\tilde{K}$ is updated as (see, e.g., Harstad, Kagel and Levin, 1990)

$$\tilde{p}_k = \Pr(\tilde{K} = k) = \frac{p_k k}{\bar{k}}, \quad k = 1, \ldots, n. \tag{9}$$

Equation (9) can be understood as follows (cf. Myerson and Wärneryd, 2006). Suppose n is finite (for an infinite n, a similar argument applies in the limit $n \to \infty$). For a given k, the probability for player i to be selected for participation is $\Pr(S_i = 1|K = k) = \frac{k}{n}$; thus,

$$\tilde{p}_k = \Pr(K = k|S_i = 1) = \frac{\Pr(S_i = 1|K = k)p_k}{\sum_{l=0}^{n} \Pr(S_i = 1|K = l)p_l} = \frac{\frac{k}{n} p_k}{\sum_{l=0}^{n} \frac{l}{n} p_l},$$

which gives (9).

Consider a symmetric pure strategy equilibrium in which all participating players choose effort $e^* > 0$. From Eq. (2), the expected payoff of a participating player i from some deviation effort e_i is

$$\pi_i(e_i, e^*) = \sum_{k=1}^{n} \tilde{p}_k \int_U F(e_i - e^* + x)^{k-1} dF(x) - c(e_i). \tag{10}$$

The symmetric first-order condition for payoff maximization, $\left.\frac{\partial \pi_i(e_i, e^*)}{\partial e_i}\right|_{e_i = e^*} = 0$, gives

$$c'(e^*) = B_p \equiv \sum_{k=1}^{n} \tilde{p}_k (k-1) \int_U F(x)^{k-2} f(x) dF(x). \tag{11}$$

Changing the variable of integration to $z = F(x)$, obtain, similar to (4),

$$B_p = \sum_{k=1}^{n} \tilde{p}_k (k-1) \int_0^1 z^{k-2} m(z) dz = \int_0^1 m(z) d\tilde{G}(z). \tag{12}$$

Here, $\tilde{G}(z) = \sum_{k=1}^{n} \tilde{p}_k z^{k-1}$ denotes the probability-generating function (pgf) of distribution $\tilde{p}$.

Let e_p^* denote the unique positive solution of (11), assuming that it exists and it is the symmetric pure strategy equilibrium.[17] When p is degenerate at some k, Eq. (11) reduces to the deterministic group size case, Eq. (3). As before, since $c'(e^*)$ is strictly increasing in e^*, the comparative statics of equilibrium effort e_p^* with respect to parameters of distribution p are determined entirely by coefficients B_p.

Using Eqs. (12) and (9), and the definition of b_k, Eq. (3), coefficients B_p can also be written as

$$B_p = \sum_{k=1}^{n} \tilde{p}_k b_k = \mathrm{E}_{\tilde{p}}(b_K) = \frac{1}{\bar{k}} \sum_{k=2}^{n} p_k k b_k = \frac{1}{\bar{k}} \mathrm{E}_p(K b_K | K \geq 2) \mathrm{Pr}_p(K \geq 2). \tag{13}$$

Here, $\mathrm{E}_p(\cdot)$ and $\mathrm{Pr}_p(\cdot)$ denote expectation and probability with respect to distribution p. Note that the summation in (13) can start with $k = 2$ instead of $k = 1$ because $b_1 = 0$. Representation (13) shows, as expected, that only group sizes $k \geq 2$ contribute to the equilibrium effort.

The uniform distribution of noise

The effects of stochastic participation are straightforward when the distribution of noise is uniform. In this case, $b_k = b_2$ for any $k \geq 2$ (see Lemma 1(ii)). Equation (12) then gives

$$B_p = b_2 \left(\tilde{G}(1) - \tilde{G}(0) \right) = b_2 \left(1 - \frac{p_1}{\bar{k}} \right), \tag{14}$$

leading to the following result.

Lemma 5 *Suppose $f(x)$ is a uniform distribution. Then $e_p^* \leq e_k^*$ for any $k \geq 2$, with equality if and only if $p_1 = 0$.*

Lemma 5 states that for a uniform distribution of noise the individual equilibrium effort of participating players in a tournament with stochastic group size cannot be higher

[17] Conditions similar to those in Proposition 9 can be formulated to ensure existence.

than with deterministic group size, and is strictly lower if the probability for a player to be alone in the tournament is not zero.

4.1 Individual equilibrium effort

We are interested in the effects of changes in distribution p on coefficients B_p, which then monotonically map into equilibrium effort e_p^*. In particular, we explore how B_p responds to a stochastic increase (in an appropriate sense) in the number of players in the tournament. To this end, consider a parameterized family of (updated) group size distributions $\{\tilde{p}(\theta)\}_{\theta\in\Theta}$, where $\Theta \subseteq \mathbb{R}$ is an interval of the real line or an ordered set of discrete numbers. Let $\tilde{P}(\theta)$ and $\tilde{G}(z,\theta)$ denote the corresponding cmf and pgf, respectively.

Suppose an increase in θ leads to a stochastic increase in the number of players in the FOSD sense; that is, assume that $\tilde{P}_k(\theta)$ is decreasing in θ for all $k = 1, \ldots, n$. The following Lemma covers situations where in the deterministic case individual equilibrium effort is monotone in the number of players. It is a straightforward extension of Lemma 1(i) and Proposition 2(ii-iii) to the stochastic case.

Lemma 6 *Suppose an increase in θ leads to a stochastic increase in $\tilde{K}$.*
(i) If $f(x)$ is increasing then e_p^* is increasing in θ.
(ii) If $f(x)$ is decreasing and $p_1(\theta) = 0$ for all $\theta \in \Theta$ then e_p^* is decreasing in θ.
(iii) If $f(x)$ is interior unimodal and symmetric, $p_1(\theta) = 0$ for all $\theta \in \Theta$, and $n \geq 4$, then e_p^* is decreasing in θ.[18]

Part (i) of Lemma 6 is the simplest case that does not require any additional restrictions. Indeed, increasing $f(x)$ implies that the sequence $\{b_k\}_{k=2}^n$ is increasing and hence $\{b_k\}_{k=1}^n$ is increasing as well because $b_1 = 0$. Parts (ii) and (iii) of Lemma 6 provide conditions for when $\{b_k\}_{k=2}^n$ is decreasing. However, since $b_1 = 0$, the sequence $\{b_k\}_{k=1}^n$ is then nonmonotone unless one-player tournaments are ruled out, that is, $p_1 = 0$. It is a rather common provision in many tournaments that competition will be canceled if fewer than a pre-specified number of participants sign up.

Parts (ii) and (iii) of Lemma 6 point to the following observation. When $\{b_k\}_{k=2}^n$ is decreasing, *the only way* e_p^* can be nonmonotone with respect to an upward probabilistic

[18] If $f(x)$ is symmetric (not necessarily unimodal), then $e_2^* = e_3^*$, see Proposition 2(iii). Hence, if $n = 3$, then e_p^* is independent of θ. If $p_1(\theta) > 0$ for some θ, this makes the sequence $\{b_k\}_{k=1}^3$ increasing and then e_p^* is increasing in θ.

shift in $\tilde{p}$ is if $p_1 > 0$. Put differently, the possibility for a player to find herself alone in the tournament is the only mechanism through which the individual equilibrium effort can be nonmonotone in θ. One example is the Tullock contest, for which $b_k = \frac{r(k-1)}{k^2}$ decreases monotonically for $k \geq 2$, and Lim and Matros (2009) found that the individual equilibrium effort is nonmonotone in q for $K \sim \text{Binomial}(n, q)$. This is a consequence of the fact that $p_1(q) = nq(1-q)^{n-1} > 0$. If the distribution of group size is replaced with a truncated binomial distribution such that $p_1(q) = 0$ for all $q \in [0, 1]$, the nonmonotonicity goes away. Of course, the nonmonotonicity can still arise even when $p_1 = 0$ if $\{b_k\}_{k=2}^n$ is nonmonotone; for example, if it is interior unimodal, see Proposition 4 below.

Extending the main result of the deterministic setting, namely, that a unimodal $f(x)$ results in a unimodal sequence $\{e_k\}_{k=2}^n$ (Proposition 2(i)), requires an additional assumption imposed on the distribution of updated group size, $\tilde{p}$. Let $\tilde{G}_\theta(z, \theta) \leq 0$ denote the derivative or the first difference of its pgf with respect to θ. The following proposition is an application of Lemmas 8 and 9 in Section 5.2 on the preservation of unimodality under uncertainty.

Proposition 4 *Suppose an increase in θ leads to a stochastic increase in $\tilde{K}$ and*
(a) $f(x)$ is unimodal;
(b) $-\tilde{G}_\theta(z, \theta)$ is log-supermodular; that is, the ratio $R(z, \theta, \theta') = \frac{\tilde{G}_\theta(z,\theta')}{\tilde{G}_\theta(z,\theta)}$ is increasing in z for all $\theta' > \theta$.
Then e_p^* is unimodal in θ.

The log-supermodularity condition (b) of Proposition 4 is satisfied by the two distributions used most prominently in the literature to model population uncertainty – the Poisson and binomial distributions. These distributions, along with the negative binomial and logarithmic distributions, belong to a family known as *power series distributions* (PSD) that are characterized by pmfs of the form

$$p_k(\theta) = \frac{a_k \theta^k}{A(\theta)}. \tag{15}$$

Here, a_k are nonnegative numbers, $\theta \geq 0$ is a parameter, and $A(\theta) = \sum_{k=0}^{\infty} a_k \theta^k$ (it is assumed that the sum exists) is the normalization function (Johnson, Kemp and Kotz, 2005). The pgf of PSD distributions is $G(z, \theta) = \frac{A(\theta z)}{A(\theta)}$. Proposition 4 is applicable to the whole PSD family due to the following three properties.

Proposition 5 *For any pmf p in the PSD family (15)*

(i) the updated pmf $\tilde{p}$ is also in the PSD family;
(ii) $G_\theta(z, \theta) \le 0$;
(iii) $-G_\theta(z, \theta)$ is log-supermodular.

Property (i) states that the PSD family is closed under the participation updating (9). In some cases, the updated distribution is of the same type as the initial distribution. For example, for $K \sim \text{Binomial}(n, q)$ we have $p_k = \binom{n}{k} q^k (1-q)^{n-k}$ (for $k = 0, \ldots, n$) and $\tilde{p}_k = \binom{n-1}{k-1} q^{k-1}(1-q)^{n-k}$ (for $k = 1, \ldots, n$); that is, $(\tilde{K} - 1) \sim \text{Binomial}(n-1, q)$. Similarly, for $K \sim \text{Poisson}(\lambda)$ we have $p_k = \frac{\exp(-\lambda)\lambda^k}{k!}$ (for $k = 0, 1, \ldots$) and $\tilde{p}_k = \frac{\exp(-\lambda)\lambda^{k-1}}{(k-1)!}$ (for $k = 1, 2, \ldots$); that is, $(\tilde{K} - 1) \sim \text{Poisson}(\lambda)$. It is possible, however, for the updated distribution to be of a different type (albeit still within the PSD family). For example, for $K \sim \text{Logarithmic}(\theta)$, where $\theta \in (0, 1)$, we have $p_k = -\frac{\theta^k}{k \ln(1-\theta)}$, $\bar{k} = -\frac{\theta}{(1-\theta)\ln(1-\theta)}$, and $\tilde{p}_k = (1-\theta)\theta^{k-1}$; that is, $\tilde{K}$ has the geometric distribution with parameter $1 - \theta$.

Property (ii) shows that PSD distributions are FOSD-ordered by parameter θ. Finally, property (iii) ensures that condition (b) of Proposition 4 is satisfied, and hence e_p^* is unimodal in θ for any PSD distribution provided $f(x)$ is unimodal.

4.2 Aggregate effort

In this section, we explore the effects of changes in the distribution of the number of players on expected aggregate effort $E_p^* = \bar{k} e_p^* = \bar{k} c'^{-1}(B_p(\theta))$. As in Section 3.2, this problem simplifies substantially when the cost function is quadratic, $c(e) = c_0 e^2$, in which case $E_p^* = \frac{\bar{k}}{2c_0} B_p(\theta)$. Using (13), it can be written as

$$E_p^* = \frac{\bar{k}}{2c_0} \sum_{k=1}^{n} \tilde{p}_k b_k = \frac{1}{2c_0} \sum_{k=0}^{n} p_k k b_k = \sum_{k=0}^{n} p_k E_k^* = \mathrm{E}_p(E_K^*).$$

Here, $E_k^* = \frac{k b_k}{2c_0}$ is the aggregate equilibrium effort in a tournament with deterministic size k, and the expectation is taken over the original group size distribution p. Lemma 4 and Proposition 3 then lead to the following results.

Proposition 6 *Suppose an increase in θ leads to an FOSD increase in the number of players K.*
(i) If $f(x)$ is IFR and $c(e)$ is more convex than quadratic (DFR and $c(e)$ is less convex than quadratic) then E_p^* is increasing (decreasing) in θ.
(ii) If $f(x)$ has a unimodal (U-shaped) failure rate, $-G_\theta(z, \theta)$ is log-supermodular and $c(e) = c_0 e^2$, then E_p^* is unimodal (U-shaped) in θ.

The proof of part (i) follows exactly the same steps as that of Proposition 3, while part (ii) follows from part (iii) of Lemma 4 and Lemmas 8 and 9. Two interesting special cases are the exponential distribution, which has a constant failure rate and generates aggregate effort E_k^* that is independent of k, and the uniform distribution, which generates individual effort that is independent of k (for $k \geq 2$, in both cases). When effort costs are quadratic, from part (ii) of Lemma 4, for the exponential distribution with parameter λ we have $E_k^* = \frac{\lambda}{2c_0}$ for $k \geq 2$, which gives $E_p^* = \frac{\lambda}{2c_0} \sum_{k=2}^{n} p_k = \frac{\lambda}{2c_0}(1 - P_1(\theta))$ where $P_1(\theta)$ is either constant or decreasing in θ; thus, E_p^* is either constant or increasing in θ. An important special case when E_p^* is constant is when $P_1(\theta) = 0$, i.e., there are always at least two participants in the tournament. For the uniform distribution, Eq. (14) gives $E_p^* = \frac{b_2(\bar{k} - p_1(\theta))}{2c_0}$, which is increasing in θ if $p_1'(\theta) = 0$ or $p_0'(\theta) = 0$ (in the latter case, $0 \geq P_1'(\theta) = p_0'(\theta) + p_1'(\theta) = p_1'(\theta)$).

Note that, unlike in Proposition 4, the condition for a stochastic increase in the number of players pertaining to Proposition 6 is formulated in terms of the original distribution of group size $p(\theta)$, and not the updated distribution $\tilde{p}(\theta)$. Condition $p_0'(\theta) = 0$ holds, in particular, in cases when $p_0(\theta) = 0$, i.e., the tournament is guaranteed to have at least one participant; more generally, it holds when the FOSD shift in K does not affect the probability of having no participants in the tournament.

Propositions 4 and 6(ii) can be generalized to the cases of multimodal density and failure rate, respectively, under the assumptions that $-\tilde{G}_\theta(z, \theta)$ and $-G_\theta(z, \theta)$, respectively, are TP_r for some $r \geq 2$, cf. the discussion at the end of Section 3.1.

4.3 Applications to tournament design

In this section, we investigate two optimal design questions for tournaments with stochastic participation: (i) how the level of uncertainty in the number of players affects aggregate effort, and (ii) whether it is optimal, from an *ex ante* perspective, to disclose the realized number of players.

4.3.1 The effect of uncertainty in the number of players

Is uncertainty in the number of players beneficial or detrimental for aggregate effort? The following proposition provides a general answer for any two distributions of the number of players ranked by second-order stochastic dominance (SOSD).

Proposition 7 *Consider two group size distributions, p and p', with the same mean $\bar{k} = \sum_{k=0}^{n} kp_k$ such that there are always at least two players ($p_0 = p_1 = p'_0 = p'_1 = 0$) and p' SOSD p. Suppose that $f'(x)$ is piecewise differentiable and continuous and $f'(\overline{x})$ is finite. Then, $E^*_{p'} \geq (\leq) E^*_p$ if $f(x)$ is log-concave (log-convex); moreover, the inequality is strict if $f(x)$ is strictly log-concave (log-convex).*

Proposition 7 says that higher uncertainty reduces expected aggregate effort if noise has a log-concave distribution. For a log-convex noise distribution the relationship is reversed. The comparison between aggregate efforts $E^*_{p'}$ and E^*_p is equivalent to a comparison between individual efforts $e^*_{p'}$ and e^*_p since the mean group size $\bar{k}$ is the same for the two distributions, which amounts to a comparison between $B_{p'}$ and B_p, cf. Eq. (11). From Eq. (13), B_p is proportional to the expectation of Kb_K conditional on $K \geq 2$, and the assumption that there are always at least two players makes this expectation unconditional. Jensen's inequality can then be applied if kb_k is concave (convex) in k for $k \geq 2$, which is the case when $f(x)$ is log-concave (log-convex).

As a special case, Proposition 7 allows for a comparison of aggregate effort between tournaments with deterministic and stochastic group sizes. It implies that the presence of uncertainty in the number of players, as opposed to a tournament where the number of players is fixed and equal to $\bar{k}$, reduces expected aggregate effort in the Tullock contest (since the Gumbel distribution is log-concave) and increases it for many heavy-tailed distributions such as Pareto (which is log-convex). This is in contrast to the two existing studies – restricted to Tullock contests – comparing aggregate effort in contests with deterministic and stochastic participation: Myerson and Wärneryd (2006) and Lim and Matros (2009). Both show that uncertainty in the number of players always reduces aggregate effort.[19]

However, Proposition 7 is more general than that and relates the ranking of expected aggregate effort to the SOSD order of group size distributions. It also shows that the presence of heavy tails in the distribution of noise reverses the prevailing "intuition" that uncertainty in the number of players is detrimental for aggregate effort.

[19] Proposition 7 is not a direct generalization of Myerson and Wärneryd (2006) and Lim and Matros (2009) because these two papers allow for the possibility of having fewer than two players in the contest, under additional restrictions. A version of Proposition 7 directly generalizing Myerson and Wärneryd (2006) and Lim and Matros (2009) is more nuanced and is available in the extended working paper version of this paper.

4.3.2 Optimal disclosure of the number of players

When the number of players K is stochastic, it might be possible for the tournament designer to reveal the realization of K to the players before they choose their efforts. Assuming commitment power, when does the tournament designer prefer to (commit to) disclose K? Lim and Matros (2009) show that in a standard Tullock contest with the binomial distribution of the number of players aggregate effort is independent of disclosure. Fu, Jiao and Lu (2011) generalize this result to contests with CSFs of the form $\frac{g(e_i)}{\sum_{j=1}^{k} g(e_j)}$. They show that full disclosure (no disclosure) is optimal if $\frac{g(e)}{g'(e)}$ is strictly convex (concave), while the indifference is recovered when $\frac{g(e)}{g'(e)}$ is linear. The following proposition generalizes these results to arbitrary tournaments and arbitrary distributions of the number of players.

Proposition 8 *Suppose b_k is non-constant for $k \geq 1$ in the support of p and $c'(\cdot)$ is nonlinear. Then it is optimal to disclose (not disclose) the number of participants in the tournament if $c''' \leq (\geq) 0$.*

Disclosure creates a mean-preserving variation in the marginal benefit of effort. Indeed, without disclosure the (expected) marginal benefit of effort is $c'(e_p^*) = B_p = \mathrm{E}_{\tilde{p}}(b_K)$, cf. (13), whereas with disclosure the realization of K is observed effort is chosen according to $c'(e_k^*) = b_k$. Such variation then increases (decreases) expected individual effort if the marginal cost function is concave (convex); that is, if $c''' \leq (\geq) 0$. For a quadratic cost function (i.e., when $c'(\cdot)$ is linear) disclosure is irrelevant. Note that the nature of coefficients b_k does not affect the optimality of disclosure. The only special case is when b_k is constant in the support of p for $k \geq 1$ (for example, noise is uniformly distributed and $p_1 = 0$), in which case disclosure does not matter.

The results of Fu, Jiao and Lu (2011) are recovered as a special case by introducing effective effort $y = g(e)$, which transforms their CSF into the lottery form and the cost of effort into $c(y) = g^{-1}(y)$. Following Section 2.2, the resulting cost of effort in the corresponding tournament with additive noise is $c_{\mathrm{m}}(y) = \int_0^y c'(t)t\,dt$, which gives the marginal cost $c'_{\mathrm{m}}(y) = c'(y)y = \frac{y}{g'(g^{-1}(y))} = \frac{g(e)}{g'(e)}$.

A similar effect of a (mean-preserving) variation in the marginal benefit of effort emerges in static biased contests (see Drugov and Ryvkin, 2017) and dynamic contests where revealing interim information is equivalent to biasing the next stage (see Lizzeri, Meyer and Persico, 1999, 2002; Aoyagi, 2010). Parallel results regarding the role of c''' hold in those settings as well.

5 Technical results

5.1 Equilibrium existence

Equilibrium existence and comparative statics are two separate issues, and in the rest of the paper we have focused on the latter. In this section, however, we address the equilibrium existence, which so far did not receive an adequate treatment in the literature on Lazear-Rosen tournaments. It is generally understood that a symmetric pure strategy equilibrium exists if the variance of shocks X_i is sufficiently large and/or the effort cost function $c(\cdot)$ is sufficiently convex (see, e.g., Nalebuff and Stiglitz, 1983), but general sufficient conditions for equilibrium existence have remained unknown.[20]

For e_k^* to be the unique symmetric equilibrium, it is sufficient to require that (i) Eq. (3) has a solution; (ii) payoff function $\pi_i(e_i, e^*)$, Eq. (2), is strictly concave in e_i.[21] The main difficulty is in the "revenue" part of the payoff function that may not be globally concave because, in general, $F(\cdot)$ is not concave; moreover, even if $F(\cdot)$ is concave, $F(\cdot)^{k-1}$ may not be, for a sufficiently large k. At the same time, $c(\cdot)$ is strictly convex, and hence a version of sufficient conditions can be obtained if the convexity of $c(\cdot)$ is restricted in some way. The simplest approach is to impose a uniform restriction on $c''(\cdot)$ on $[0, e_{\max}]$.[22]

Let $f_m = \sup\{f(x) : x \in U\}$, $f'_{\max} = \sup\{f'(x) : x \in U\}$ and $f'_{\min} = \inf\{f'(x) : x \in U\}$ denote the tight, possibly infinite, bounds of pdf $f(\cdot)$ and its derivative $f'(\cdot)$ on U. We impose the following restrictions on the pdf of noise.

Assumption 1 *(a) $f(\cdot)$ is uniformly bounded; that is, $f_m < \infty$.*
(b) $f'(\cdot)$ is uniformly bounded above or below or both; that is, either $f'_{\max} < \infty$ or $f'_{\min} > -\infty$ or both.

Proposition 9 *Suppose Assumption 1 is satisfied and*
(a) There exists a $c_0 > 0$ such that $c''(e) \geq c_0$ for all $e \in [0, e_{\max}]$.
(b) $c_0 > D \equiv \min\{D_+, D_-\}$, where

$$D_+ = (k-1)[(k-1)f_m^2 + f'_{\max}], \quad D_- = (k-1)(f_m^2 - f'_{\min}). \tag{16}$$

(c) $kc(c'^{-1}(f_m)) \leq 1$.

[20] For WTA Tullock contests, equilibrium existence and uniqueness are well-understood, see Szidarovszky and Okuguchi (1997).

[21] Since $\pi_i(0, e^*) \geq 0$, conditions (i) and (ii) automatically imply that the symmetric equilibrium payoff is positive, $\pi_i(e_k^*, e_k^*) > 0$.

[22] Any effort $e_i > e_{\max}$ is strictly dominated.

Then e_k^* is the unique equilibrium in the tournament

Conditions (a) and (b) in Proposition 9 guarantee the global strict concavity of payoff function (2) in e_i, while condition (c) ensures that Eq. (3) has a solution. The conditions are consistent with the intuition described above. For a given tournament model, they are easier to satisfy as noise becomes more dispersed (leading to a decrease in f_m, $f'_{\max}$ and $|f'_{\min}|$). Additionally, conditions (b) and (c) are harder to satisfy as k increases. Overall, the conditions of Proposition 9 are rather strong because the global strict concavity of the payoff function is not necessary. An alternative approach can be to impose a weaker restriction on $c(\cdot)$ but restrict attention to particular families of noise distributions. In contrast, for the purposes of this paper we have chosen to formulate conditions with maximum flexibility for the shape of the distribution of noise, at the expense of a rather restrictive positivity of $c''(\cdot)$ and substantial noise dispersion.

A quadratic cost function, $c(e) = \frac{c_0}{2}e^2$, satisfies condition (a). Generally, functions satisfying condition (a) have the form $c(e) = \frac{c_0}{2}e^2 + \kappa(e)$, where $\kappa : [0, e_{\max}] \to \mathbb{R}_+$ is convex. Note that a function can satisfy the condition even if it is less convex than quadratic. For example, function $c(e) = c_1 e^{\xi}$ has a positive second derivative bounded below by $c_0 = \xi(\xi - 1)c_1 e_{\max}^{\xi-2}$ when $\xi \in (1, 2]$.

5.2 Preservation of unimodality under uncertainty

Throughout this paper, we explore the comparative statics of individual and aggregate equilibrium effort in tournaments with respect to the number of players. First, in Section 3, we assume that this number, k, is fixed; then, in Section 4, we allow k to be a realization of a nonnegative integer random variable with some probability mass function (pmf). In the latter case, we explore the comparative statics with respect to changes in the parameters of the pmf ranked according to FOSD.

In both cases, we show that robust comparative statics for individual effort can be obtained for *unimodal* distributions of noise $f(x)$, whereas for aggregate effort the same holds for noise distributions with a unimodal failure rate $h(x) = \frac{f(x)}{1-F(x)}$. These comparative statics amount to preservation of unimodality under uncertainty. Indeed, coefficients b_k, Eq. (4), which determine the comparative statics of individual effort in the case of deterministic group size, can be written as expectations of inverse quantile density of the form $b_k = \int_0^1 m(z)dH(z,k)$, where $H(z,k) = z^{k-1}$ is a family of cdfs FOSD-ordered by parameter k. Our first lemma in this section provides a necessary and sufficient condition for such expectations, generally of the form $\gamma(\theta) = \int_0^1 u(z)dH(z,\theta)$, where $H(z,\theta)$ are

FOSD-ordered in θ, to be unimodal in θ for all unimodal functions $u(z)$. Turning to the case of stochastic group size, equilibrium effort is determined by discrete expectations of the form $\chi(\theta) = \sum_{k=1}^{n} p_k(\theta)u_k$, where $\{u_k\}_{k=1}^{n}$ is some sequence and $p(\theta) = \{p_k(\theta)\}_{k=1}^{n}$ is an FOSD-ordered family of pmfs. The second lemma in this section establishes a necessary and sufficient condition for such expectations to be unimodal in θ for all unimodal sequences u_k. We start with some definitions. All missing proofs are in Appendix A.

Definition 1 *A function (or sequence) $u : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}$, is unimodal if there exists $\hat{x} \in S$ such that $u(\cdot)$ is increasing for $x \leq \hat{x}$ and decreasing for $x \geq \hat{x}$. A function (or sequence) is interior unimodal if it is unimodal and nonmonotone.*

Definition 2 *A function $v : S_1 \times S_2 \to \mathbb{R}$, where $S_1, S_2 \subseteq \mathbb{R}$, is log-supermodular if for all $x_1, x_1' \in S_1$, $x_2, x_2' \in S_2$, such that $x_1' > x_1$ and $x_2' > x_2$,*

$$v(x_1, x_2')v(x_1', x_2) \leq v(x_1, x_2)v(x_1', x_2').$$

In other words, for all $x_2' > x_2$ the ratio $r(x_1, x_2, x_2') = \frac{v(x_1,x_2')}{v(x_1,x_2)}$ is increasing in x_1.

Consider integrals of the form $\gamma(\theta) = \int_0^1 u(z)dH(z, \theta)$, where $u(z) : [0, 1] \to \mathbb{R}$ is an integrable, continuous and piecewise differentiable function and $H(z, \theta)$ is a cdf of a random variable $Z|\theta$ defined on $[0, 1]$ and parameterized by $\theta \in \Theta \subseteq \mathbb{R}$.[23] We assume that an increase in θ leads to an upward probabilistic shift, in the FOSD sense, of $Z|\theta$; that is, $H(z, \theta)$ is decreasing in θ for all $z \in [0, 1]$ and $\theta \in \Theta$. Let $H_\theta(z, \theta) \leq 0$ denote the derivative of $H(z, \theta)$ with respect to θ if θ is a continuous parameter (in which case we assume that $H(z, \theta)$ is differentiable) or the first difference, $H(z, \theta + d) - H(z, \theta)$, if θ is a discrete index with step size $d > 0$.

Lemma 7 *$\gamma(\theta)$ is unimodal for all unimodal functions $u(z)$ if and only if $-H_\theta(z, \theta)$ is log-supermodular; that is, the ratio $r(z, \theta, \theta') = \frac{H_\theta(z,\theta')}{H_\theta(z,\theta)}$ is increasing in z for any $\theta' > \theta$.*

Consider now sums of the form $\chi(\theta) = \sum_{k=1}^{n} p_k(\theta)u_k$, where u_k is a nonnegative sequence and $p(\theta) = (p_1(\theta), \ldots, p_n(\theta))$ is a pmf parameterized by $\theta \in \Theta \subseteq \mathbb{R}$. We use $P_k(\theta) = \sum_{l=1}^{k} p_l(\theta)$ to denote the corresponding cumulative mass function (cmf), with $P_n(\theta) = 1$. The upper bound of the sum, $n \geq 2$, can be finite or infinite and applies

[23]Variables $Z|\theta$ do not have to have the same support; rather, we assume that $[0, 1]$ includes all of their supports, and $H(0, \theta) = 1 - H(1, \theta) = 0$ for all $\theta \in \Theta$.

uniformly for all values of θ.[24] We assume that an increase in θ shifts the distribution $p(\theta)$ upward in the FOSD sense. Let $P_k'(\theta) \leq 0$ denote the derivative or the first difference of the cmf with respect to θ.

Lemma 8 *$\chi(\theta)$ is unimodal for all unimodal sequences u_k if and only if $-P_k'(\theta)$ is log-supermodular; that is, the ratio $r(k, \theta, \theta') = \frac{P_k'(\theta')}{P_k'(\theta)}$ is increasing in k for any $\theta' > \theta$.*

In some cases, the log-supermodularity condition of Lemma 8 may be difficult to check directly because there is no closed-form expression for the cmf $P_k(\theta)$. The following lemma shows that a similar ratio condition can instead be checked for the probability-generating function (pgf) of distribution $p(\theta)$, defined as $G(z, \theta) = \sum_{k=1}^{n} p_k(\theta) z^{k-1}$. Probabilities $p_k(\theta)$ can be recovered from it as $p_k(\theta) = \frac{1}{(k-1)!} G^{(k-1)}(0, \theta)$. Moreover, the pgf can be related to the cmf $P(\theta)$ as

$$\sum_{k=1}^{n} P_k(\theta) z^{k-1} = \frac{G(z, \theta) - z^{n-1}}{1 - z}. \tag{17}$$

It follows from Eq. (17) that $G(z, \theta)$ is decreasing in θ whenever $P_k(\theta)$ is decreasing in θ for all k; that is, $G(z, \theta)$ behaves as an FOSD-ordered family of cdfs (except that $G(0, \theta) = p_1(\theta)$, which is, generally, nonzero). Let $G_\theta(z, \theta) \leq 0$ denote, similar to $H_\theta(z, \theta)$ in Lemma 7, either the derivative or the first difference of $G(z, \theta)$ with respect to θ.

Lemma 9 *$-G_\theta(z, \theta)$ is log-supermodular if and only if $P_k'(\theta)$ is log-supermodular; that is, the ratio $R(z, \theta, \theta') = \frac{G_\theta(z, \theta')}{G_\theta(z, \theta)}$ is increasing in z for any $\theta' > \theta$ if and only if the ratio $r(k, \theta, \theta')$ in Lemma 8 is increasing in k for any $\theta' > \theta$.*

The increasing ratio conditions in Lemmas 7, 8 and 9 are well-known in the literature on comparative statics under uncertainty (Athey, 2002). They are also known as total positivity of order 2 (Karlin, 1968), and increasing likelihood ratio properties when applied to parameterized probability density functions (see, e.g., Shaked and Shanthikumar, 2007). The results of this section are most closely related to those of Athey (2002) on the comparative statics of expectations of the form $\gamma(\theta) = \int_0^1 u(z) dH(z, \theta)$ for single-crossing functions $u(z)$. Lemma 7 is a straightforward corollary of these results applied to unimodal functions, i.e., functions with a single-crossing derivative. Indeed, assuming

[24] This is not to say that $p(\theta)$ has support independent of $\theta \in \Theta$; rather, $n = \sup_{\theta \in \Theta} n(\theta)$, where $n(\theta)$ is the upper bound of the support of $p(\theta)$. The definitions of $p(\theta)$ are extended to the uniform support so that $p_k(\theta) = 0$ and $P_k(\theta) = 1$ for $k > n(\theta)$.

$u(1)$ is finite (which is the case for interior unimodal functions) and integrating by parts, $\gamma(\theta) = u(1) - \int_0^1 u'(z)H(z,\theta)dz$, where $u'(z)$ is single-crossing and hence, following Athey (2002), $\gamma'(\theta) = -\int_0^1 u'(z)H_\theta(z,\theta)dz$ is single-crossing, i.e., $\gamma(\theta)$ is unimodal, if $-H_\theta(z,\theta)$ is log-supermodular. Lemma 8 is a discrete version of Lemma 7 and follows similarly via "summation by parts." Lemma 9, however, is less straightforward; the equivalence of log-supermodality of a discrete cmf and the corresponding pgf is a new result with potentially broader applications.

6 Conclusion

Tournament incentives are ubiquitous. Students applying to universities, researchers competing for grants, R&D firms competing for innovation, job candidates applying for an opening or employees competing for promotion, and numerous other examples, are situations where participants' outcomes are determined jointly by ability, effort and luck. Differences in ability stratify the playing field to some extent, but competition is the most fierce, and luck plays the biggest role, in tournaments among equally able contestants.

It is traditionally believed that competition increases productivity, fosters innovation, and promotes economic growth. However, it is also easy to imagine how competition may discourage effort in winner-take-all environments where luck plays a significant role. Our results demonstrate that there is a nontrivial interplay between the two effects, and the nature of shocks – the "shape of luck" – matters for the willingness to compete.

We show that individual effort reacts to an increase in competition, be it deterministic or stochastic, in a way that essentially follows the shape of the density of noise. As long as the density is unimodal, individual effort is also unimodal in the number of players, but it can be increasing, deceasing or nonmonotone when the distribution of noise is skewed. Aggregate effort behaves similarly, but following the shape of the failure rate of noise. Hence, the presence of heavy tails – a decreasing or interior unimodal failure rate – in the distribution of noise can lead to a *reduction* in aggregate effort with competition.

Heavy-tailed distributions, including the Pareto distribution (also known as the family of *power laws*), have been widely identified in economics, finance and other domains. For example, it has been known for a long time that economic variables such as income (Pareto, 1896), city sizes (Auerbach, 1913), firm sizes (Axtell, 2001), stock market movements (Mandelbrot, 1963) and CEO compensation (Roberts, 1956) follow power laws. More recently, power laws have been found to describe demand for books at Amazon (Chevalier and Goolsbee, 2003) and movie ratings in Netflix (Bimpikis and Markakis, 2016).[25] The nature of innovation as an unlikely breakthrough resulting from a large number of mostly unsuccessful attempts produces heavy tails in the value, quality and financial returns of inventions (Fleming, 2007).

Our results predict diverging effects of competition on aggregate effort (or investment) in tournaments characterized by different types of noise. Given the various contradictory findings and nonmonotonicities in the literature on the effects of competitive pressure on innovation (e.g., Aghion et al., 2005; Vives, 2008), our results provide an independent mechanism through which different reactions to competitive pressure may arise across industries, or even within the same industry across time.

To date, there is virtually no empirical research on the effects of variation in the shape of shocks on behavior in tournaments. The existing studies of tournaments using natural data (e.g., Ehrenberg and Bognanno, 1990; Knoeber and Thurman, 1994; Eriksson, 1999) treat noise as a nuisance and do not attempt to estimate its distribution. Similarly, laboratory experiments typically rely on a specific distribution of noise in their winner determination process – most often, a lottery contest or uniformly distributed additive shocks (for a review, see Dechenaux, Kovenock and Sheremeta, 2015) – and do not explore variation in its shape. We only know of one exception. List et al. (2014) study how effort depends on the number of players in tournaments with varying noise densities. They consider distributions with constant, increasing and decreasing densities and find, consistent with theory, that the comparative statics of individual effort follow similar patterns. However, all three distributions in their study are light-tailed with increasing failure rates, and, consistent with our results, they observe aggregate effort increasing in the number of players in all three cases.

Our last comment is methodological. The techniques developed in this paper can be extended to many applications of general tournament models, including optimal contract design and dynamic tournaments,[26] giving a new life to the literature that so far has been limited to considering a number of special cases.

[25]See Gabaix (2016) for a survey of many identified power laws and their underlying mechanisms. Many patterns outside economics are described by power laws as well, such as the frequency of words in natural languages (Zipf, 1949), the intensity of earthquakes (Christensen et al., 2002) or popularity in social networks (Barabási and Albert, 1999).

[26]For example, in a companion paper (Drugov and Ryvkin, 2018), we use them to obtain new general results on the optimal allocation of prizes.

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and Innovation: an Inverted-U Relationship*." *The Quarterly Journal of Economics*, 120(2): 701–728.

Akerlof, Robert J., and Richard T. Holden. 2012. "The nature of tournaments." *Economic Theory*, 51(2): 289–313.

Aoyagi, Masaki. 2010. "Information feedback in a dynamic tournament." *Games and Economic Behavior*, 70(2): 242–260.

Athey, Susan. 2002. "Monotone comparative statics under uncertainty." *Quarterly Journal of Economics*, 187–223.

Auerbach, Felix. 1913. "Das Gesetz Der Bevolkerungskonzentration." *Petermanns Geographische Mitteilungen*, 59: 74–76.

Axtell, Robert L. 2001. "Zipf Distribution of U.S. Firm Sizes." *Science*, 293(5536): 1818–1820.

Bagnoli, Mark, and Ted Bergstrom. 2005. "Log-concave probability and its applications." *Economic Theory*, 26(2): 445–469.

Balafoutas, Loukas, E. Glenn Dutcher, Florian Lindner, and Dmitry Ryvkin. 2017. "The optimal allocation of prizes in tournaments of heterogeneous agents." *Economic Inquiry*, 55(1): 461–478.

Barabási, Albert-László, and Réka Albert. 1999. "Emergence of scaling in random networks." *Science*, 286(5439): 509–512.

Baye, Michael R., Dan Kovenock, and Casper G. De Vries. 1996. "The all-pay auction with complete information." *Economic Theory*, 8(2): 291–305.

Bimpikis, Kostas, and Mihalis G. Markakis. 2016. "Inventory Pooling Under Heavy-Tailed Demand." *Management Science*, 62(6): 1800–1813.

Boosey, Luke, Philip Brookins, and Dmitry Ryvkin. 2018. "Contests between groups of unknown size." *Games and Economic Behavior*. https://doi.org/10.1016/j.geb.2018.09.001.

Chevalier, Judith, and Austan Goolsbee. 2003. "Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com." *Quantitative Marketing and Economics*, 1(2): 203–222.

Christensen, Kim, Leon Danon, Tim Scanlon, and Per Bak. 2002. "Unified scaling law for earthquakes." *Proceedings of the National Academy of Sciences*, 99(suppl 1): 2509–2513.

Connelly, Brian L., Laszlo Tihanyi, T. Russell Crook, and K. Ashley Gangloff. 2014. "Tournament Theory: Thirty Years of Contests and Competitions." *Journal of Management*, 40(1): 16–47.

Corchón, Luis C, and Marco Serena. 2018. "Contest Theory." *Handbook of Game Theory and Industrial Organization, Volume II*, , ed. Luis C Corchón and Marco A. Marini, 125–146. Edward Elgar Publishing, Inc.

Dechenaux, Emmanuel, Dan Kovenock, and Roman M. Sheremeta. 2015. "A survey of experimental research on contests, all-pay auctions and tournaments." *Experimental Economics*, 18(4): 609–669.

De Sinopoli, Francesco, and Carlos Pimienta. 2009. "Undominated (and) perfect equilibria in Poisson games." *Games and Economic Behavior*, 66(2): 775–784.

Drugov, Mikhail, and Dmitry Ryvkin. 2017. "Biased contests for symmetric players." *Games and Economic Behavior*, 103: 116–144.

Drugov, Mikhail, and Dmitry Ryvkin. 2018. "Tournament Rewards and Heavy Tails." *Working paper*. https://econpapers.repec.org/paper/fsuwpaper/wp2017_5f05_5f01.htm.

Ehrenberg, Ronald G., and Michael L. Bognanno. 1990. "The incentive effects of tournaments revisited: Evidence from the European PGA tour." *Industrial and Labor Relations Review*, 43(3): 74–S–88–S.

Eriksson, Tor. 1999. "Executive compensation and tournament theory: Empirical tests on Danish data." *Journal of Labor Economics*, 17(2): 262–280.

Fleming, Lee. 2007. "Breakthroughs and the "long tail" of innovation." *MIT Sloan Management Review*, 49(1): 69.

Frank, Robert H. 2016. *Success and Luck: Good Fortune and the Myth of Meritocracy.* Princeton University Press.

Fu, Qiang, Qian Jiao, and Jingfeng Lu. 2011. "On disclosure policy in contests with stochastic entry." *Public Choice*, 148(3-4): 419–434.

Gabaix, Xavier. 2016. "Power Laws in Economics: An Introduction." *Journal of Economic Perspectives*, 30(1): 185–206.

Gerchak, Yigal, and Qi-Ming He. 2003. "When will the range of prizes in tournaments increase in the noise or in the number of players?" *International Game Theory Review*, 5(02): 151–165.

Green, Jerry R., and Nancy L. Stokey. 1983. "A comparison of tournaments and contracts." *The Journal of Political Economy*, 349–364.

Harstad, Ronald M., John H. Kagel, and Dan Levin. 1990. "Equilibrium bid functions for auctions with an uncertain number of bidders." *Economics Letters*, 33(1): 35–40.

Hillman, Arye L., and John G. Riley. 1989. "Politically contestable rents and transfers." *Economics & Politics*, 1(1): 17–39.

Jia, Hao. 2008. "A stochastic derivation of the ratio form of contest success functions." *Public Choice*, 135(3-4): 125–130.

Jia, Hao, Stergios Skaperdas, and Samarth Vaidya. 2013. "Contest functions: Theoretical foundations and issues in estimation." *International Journal of Industrial Organization*, 31(3): 211 – 222.

Johnson, Norman L., Adrienne W. Kemp, and Samuel Kotz. 2005. *Univariate Discrete Distributions.* John Wiley & Sons, Inc.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan. 1995. *Continuous Univariate Distributions.* Vol. 2, John Wiley & Sons, Inc.

Kahana, Nava, and Doron Klunover. 2015. "A note on Poisson contests." *Public Choice*, 165(1-2): 97–102.

Kahana, Nava, and Doron Klunover. 2016. "Complete rent dissipation when the number of rent seekers is uncertain." *Economics Letters*, 141: 8–10.

Karlin, Samuel. 1968. *Total positivity.* Vol. 1, Stanford University Press.

Knoeber, Charles R., and Walter N. Thurman. 1994. "Testing the theory of tournaments: An empirical analysis of broiler production." *Journal of Labor Economics*, 12(2): 155–179.

Konrad, Kai A. 2009. *Strategy and dynamics in contests.* New York, NY:Oxford University Press.

Krishna, Vijay, and John Morgan. 1998. "The winner-take-all principle in small tournaments." *Advnaces in Applied Microeconomics*, 7: 61–74.

Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy*, 89(5): 841–864.

Levin, Dan, and Emre Ozdenoren. 2004. "Auctions with uncertain numbers of bidders." *Journal of Economic Theory*, 118(2): 229–251.

Lim, Wooyoung, and Alexander Matros. 2009. "Contests with a stochastic number of players." *Games and Economic Behavior*, 67: 584–597.

List, John A., Daan Van Soest, Jan Stoop, and Haiwen Zhou. 2014. "On the Role of Group Size in Tournaments: Theory and Evidence from Lab and Field Experiments." *NBER Working Paper 20008*. http://www.nber.org/papers/w20008.

Lizzeri, Alessandro, Margaret Meyer, and Nicola Persico. 1999. "Interim Evaluations in Dynamic Tournaments: The Effects of Midterm Exams." mimeo.

Lizzeri, Alessandro, Margaret Meyer, and Nicola Persico. 2002. "The Incentive Effects of Interim Performance Evaluations." CARESS Working Paper # 02-09, http://economics.sas.upenn.edu/sites/economics.sas.upenn.edu/files/02-09.pdf.

Makris, Miltiadis. 2008. "Complementarities and macroeconomics: Poisson games." *Games and Economic Behavior*, 62(1): 180–189.

Makris, Miltiadis. 2009. "Private provision of discrete public goods." *Games and Economic Behavior*, 67(1): 292–299.

Mandelbrot, Benoit. 1963. "The Variation of Certain Speculative Prices." *Journal of Business*, 36(4): 394–419.

Marshall, Albert W., Ingram Olkin, and Barry C. Arnold. 2011. *Inequalities: Theory of majorization and its applications.* Springer.

McAfee, R. Preston, and John McMillan. 1987*a*. "Auctions and bidding." *Journal of Economic Literature*, 25(2): 699–738.

McAfee, R. Preston, and John McMillan. 1987*b*. "Auctions with a stochastic number of bidders." *Journal of Economic Theory*, 43(1): 1–19.

McLaughlin, Kenneth J. 1988. "Aspects of tournament models: A survey." *Research in Labor Economics*, 9(1): 225–56.

Mlodinow, Leonard. 2009. *The drunkard's walk: How randomness rules our lives.* Pantheon Books, New York.

Mohlin, Erik, Robert Östling, and Joseph Tao-yi Wang. 2015. "Lowest unique bid auctions with population uncertainty." *Economics Letters*, 134: 53–57.

Moldovanu, Benny, and Aner Sela. 2001. "The optimal allocation of prizes in contests." *American Economic Review*, 542–558.

Münster, Johannes. 2006. "Contests with an unknown number of contestants." *Public Choice*, 129(3-4): 353–368.

Myerson, Roger B. 1998. "Population uncertainty and Poisson games." *International Journal of Game Theory*, 27(3): 375–392.

Myerson, Roger B. 2000. "Large Poisson games." *Journal of Economic Theory*, 94(1): 7–45.

Myerson, Roger B., and Karl Wärneryd. 2006. "Population uncertainty in contests." *Economic Theory*, 27(2): 469–474.

Nalebuff, Barry J., and Joseph E. Stiglitz. 1983. "Prizes and incentives: towards a general theory of compensation and competition." *The Bell Journal of Economics*, 21–43.

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

Pareto, Vilfredo. 1896. *Cours d'Économie Politique.* New Edition, G.-H. Bousquet and G. Busino, eds. Geneva: Librairie Droz, 1964.

Paris, R.B. 2010. "Incomplete beta functions." *NIST Handbook of Mathematical Functions*, , ed. Frank W. J. Olver, Daniel M. Lozier, Ronald F. Boisvert and Charles W. Clark. Cambridge University Press. http://dlmf.nist.gov/8.17.

Parzen, Emanuel. 1979. "Nonparametric statistical data modeling." *Journal of the American Statistical Association*, 74(365): 105–121.

Roberts, David R. 1956. "A general theory of executive compensation based on statistically tested propositions." *Quarterly Journal of Economics*, 70(2): 270–294.

Ruffin, Roy J. 1971. "Cournot oligopoly and competitive behaviour." *The Review of Economic Studies*, 38(4): 493–502.

Shaked, Moshe, and J. George Shanthikumar. 2007. *Stochastic orders.* Springer.

Siegel, Ron. 2009. "All-pay contests." *Econometrica*, 77(1): 71–92.

Skaperdas, Stergios. 1996. "Contest success functions." *Economic Theory*, 7(2): 283–290.

Szidarovszky, Ferenc, and Koji Okuguchi. 1997. "On the existence and uniqueness of pure Nash equilibrium in rent-seeking games." *Games and Economic Behavior*, 18(1): 135–140.

The New York Times. 2009. "A $1 Million Research Bargain for Netflix, and Maybe a Model for Others." 21 September, https://archive.nytimes.com/www.nytimes.com/2009/09/22/technology/internet/22netflix.html.

Tullock, Gordon. 1980. "Efficient rent seeking." *Toward a Theory of the Rent-Seeking Society*, , ed. James M. Buchanan, Robert D. Tollison and Gordon Tullock, 97–112. College Station:Texas A&M University Press.

Vives, Xavier. 2008. "Innovation and competitive pressure." *The Journal of Industrial Economics*, 56(3): 419–469.

Watts, Duncan J. 2011. *Everything is Obvious: Once You Know The Answer.* Crown Business.

Zipf, George Kingsley. 1949. *Human behaviour and the principle of least effort.* Addison-Wesley.

A Proofs

Proof of the "only if" part of part (ii) of Lemma 1 Let $m_k = \int_0^1 z^k m(z)dz$ denote the moments of the inverse quantile density. Suppose $b_k = b_2$ for all $k \geq 2$. This implies, from (4), $(k+1)m_k = b_2$ and hence $m_k = \frac{b_2}{k+1}$ for all $k = 0, 1, \ldots$. The moment-generating function of $m(z)$, defined as $\mu(x) = \mathrm{E}(\exp(xZ))$, can be written in the form of expansion over moments, $\mu(x) = \sum_{k=0}^{\infty} \frac{m_k}{k!} x^k$, which gives

$$\mu(x) = \sum_{k=0}^{\infty} \frac{b_2}{(k+1)!} x^k = \frac{b_2}{x}(\exp(x) - 1).$$

This is the moment-generating function of an (unnormalized) uniform distribution on $[0, 1]$, implying $m(z)$ is a constant and F is uniform. ■

Proof of Lemma 2 Recall from (4) that $b_k = \int_0^1 m(z)dz^{k-1}$; therefore, integrating by parts,

$$b_k - b_{k+1} = \int_0^1 m(z)d(z^{k-1} - z^k) = -\int_0^1 z^{k-1}(1-z)m'(z)dz.$$

Let $\hat{z} = F^{-1}(\hat{x})$ and suppose $m(z)$ is decreasing and nonconstant on $(\hat{z}, 1)$ (the case of an increasing and nonconstant $m(z)$ is proved similarly). Then

$$\begin{aligned}
&b_k - b_{k+1} = -\int_0^{\hat{z}} z^{k-1}(1-z)m'(z)dz + \int_{\hat{z}}^1 z^{k-1}(1-z)|m'(z)|dz \\
&\geq \int_{\hat{z}}^1 z^{k-1}(1-z)|m'(z)|dz - \int_0^{\hat{z}} z^{k-1}(1-z)|m'(z)|dz \\
&\geq \hat{z}^{k-1} \int_{\hat{z}}^1 (1-z)|m'(z)|dz - M_2 \int_0^{\hat{z}} z^{k-1}dz = M_1 \hat{z}^{k-1} - \frac{M_2 \hat{z}^k}{k} = \hat{z}^{k-1}\left(M_1 - \frac{M_2 \hat{z}}{k}\right).
\end{aligned}$$

Here, $M_1 = \int_{\hat{z}}^1 (1-z)|m'(z)|dz > 0$ is independent of k and $M_2 > 0$ is bounded (the existence of M_2 follows, for example, from the mean-value theorem for definite integrals). The last expression becomes positive for a sufficiently large k. ■

Proof of Proposition 2 Define

$$\Delta b_{k+3} = b_{k+3} - b_{k+2} = \int_0^1 \left[(k+2)z^{k+1} - (k+1)z^k\right] m(z)dz, \quad k = 0, 1, \ldots, n-3. \tag{18}$$

Integrating by parts, obtain

$$\Delta b_{k+3} = \int_0^1 m(z)d(z^{k+2} - z^{k+1}) = \int_0^1 z^{k+1}(1-z)m'(z)dz. \tag{19}$$

For part (ii), the symmetry of $f(x)$ around its mean μ implies $f(x) = f(2\mu - x)$ and $F(x) = 1 - F(2\mu - x)$ for all $x \in U$. Letting $z = F(x) = 1 - F(2\mu - x)$, obtain $1 - z = F(2\mu - x)$, $F^{-1}(1-z) = 2\mu - x$ and $m(1-z) = f(F^{-1}(1-z)) = f(2\mu - x) = f(x) = f(F^{-1}(z)) = m(z)$. Thus, the symmetry of the distribution of noise implies $m(z) = m(1-z)$ and $m'(z) = -m'(1-z)$ for all $z \in [0,1]$.

This gives, via a change of variable $z \to 1 - z$,

$$\Delta b_{k+3} = -\int_0^{\frac{1}{2}} z(1-z)[(1-z)^k - z^k]m'(z)dz,$$

which immediately implies that $\Delta b_3 = 0$ and $\Delta b_{k+3} < 0$ for $k > 0$.

For part (iii), note that $b_2 = \int_0^1 m(z)dz$ and, if $m(z) = m(1-z)$ (which only requires symmetry but not unimodality of f),

$$b_3 = 2\int_0^1 zm(z)dz = 2\int_0^1 (1-z)m(1-z)dz = 2\int_0^1 (1-z)m(z)dz = 2b_2 - b_3,$$

which implies $b_2 = b_3$. ■

Proof of Lemma 4 For a quadratic cost function, $E_k^* \propto \mathrm{E}(h_q(Z_{(k-1:k)}))$ and part (i) follows immediately from the the FOSD ordering of order statistics $Z_{(k-1:k)}$ in k. Part (ii) follows by direct computation. For part (iii), the result follows from Lemma 7 due to the log-supermodularity of $|F_k^B(z; k-1, 2)|$. Indeed, recall that $F^B(z; x, y)$ is the regularized incomplete beta function, and its properties include (Paris, 2010)

$$F^B(z; x+1, y) = F^B(z; x, y) - \frac{z^x(1-z)^y}{x\mathcal{B}(x,y)},$$

where $\mathcal{B}(x,y)$ is the beta function. This gives

$$F_k^B(z,k-1,2) = F^B(z;k,2) - F^B(z;k-1,2) = -\frac{z^{k-1}(1-z)^2}{(k-1)\mathcal{B}(k-1,2)},$$

and, for some $z' > z$,

$$\frac{F_k^B(z',k-1,2)}{F_k^B(z,k-1,2)} = \left(\frac{z'}{z}\right)^{k-1} \frac{(1-z')^2}{(1-z)^2}$$

is increasing in k. ■

Proof of Proposition 3 For concreteness, suppose $f(x)$ is IFR. Then, by Lemma 4, kb_k is increasing in k. Treating $k \geq 2$ as a continuous parameter, which is justified because b_k, Eq. (4), is differentiable in k, gives $(kb_k)' = b_k + kb'_k \geq 0$, where the prime denotes partial derivative with respect to k. In general, $E_k^* = ke_k^*$; therefore, $(E_k^*)' = e_k^* + k(e_k^*)'$. Differentiating the first-order condition $c'(e_k^*) = b_k$ with respect to k obtain $c''(e_k^*)(e_k^*)' = b'_k$, which gives $(e_k^*)' = \frac{b'_k}{c''(e_k^*)}$ and hence $(E_k^*)' = e_k^* + \frac{kb'_k}{c''(e_k^*)}$.

Suppose $c(e)$ is more convex than e^2. Then $c(\sqrt{t})$ is convex in t, which implies

$$\frac{\partial^2}{\partial t^2}c(\sqrt{t}) = \frac{\partial}{\partial t}\left[\frac{c'(\sqrt{t})}{2\sqrt{t}}\right] = \frac{c''(\sqrt{t})\sqrt{t} - c'(\sqrt{t})}{4t^{3/2}} \geq 0,$$

i.e., $c''(e)e \geq c'(e)$. Therefore, $c''(e_k^*)e_k^* \geq c'(e_k^*) = b_k$ and, using the condition $b_k + kb'_k \geq 0$,

$$(E_k^*)' = e_k^* + \frac{kb'_k}{c''(e_k^*)} \geq e_k^* - \frac{b_k}{c''(e_k^*)} \geq 0.$$

For the case when $f(x)$ is DFR and $c(e)$ is less convex than e^2 the derivation is similar. ■

Proof of Proposition 5 (i) From (9),

$$\tilde{p}_k = \frac{kp_k}{\bar{k}} = \frac{ka_k\theta^k}{\sum_{k=1}^{\infty} ka_k\theta^k} = \frac{\tilde{a}_k\theta^k}{\tilde{A}(\theta)},$$

where $\tilde{a}_k = ka_k$ and $\tilde{A}(\theta) = \sum_{k=1}^{\infty} \tilde{a}_k\theta^k$; that is, $\tilde{p}_k$ also has the PSD form.

(ii) Recall that $G(z,\theta) = \frac{A(\theta z)}{A(\theta)}$. This gives

$$\begin{aligned}
G_\theta(z,\theta) &= \frac{A'(\theta z)z}{A(\theta)} - \frac{A'(\theta)}{A(\theta)}\frac{A(\theta z)}{A(\theta)} \\
&= \frac{\sum_{k=0}^{\infty} k a_k \theta^{k-1} z^k}{A(\theta)} - \frac{\sum_{k=0}^{\infty} k a_k \theta^{k-1}}{A(\theta)}\frac{\sum_{k=0}^{\infty} a_k \theta^k z^k}{A(\theta)} \\
&= \frac{1}{\theta}\left(\mathrm{E}(Kz^K) - \mathrm{E}(K)\mathrm{E}(z^K)\right) = \frac{1}{\theta}\mathrm{Cov}(K, z^K) \le 0.
\end{aligned}$$

(iii) Let $A_k(\theta) = \frac{1}{A(\theta)}\sum_{l=0}^{k} a_l \theta^l$ denote the cmf of a PSD distribution. We will prove that $|A_k'(\theta)|$ is log-supermodular; the result then follows by Lemma 9. Note that

$$A_k'(\theta) = \frac{1}{A(\theta)^2}\sum_{l=0}^{k}\sum_{m\ge 0} a_l a_m \theta^{l+m-1}(l-m) = -\frac{1}{A(\theta)^2}\sum_{l=0}^{k}\sum_{m\ge k+1} a_l a_m \theta^{l+m-1}(m-l).$$

Consider some $\theta' > \theta$ and let $\beta = \frac{\theta'}{\theta} > 1$. For convenience, introduce the notation $\alpha_{lm} = a_l a_m \theta^{l+m-1}(m-l)$. The ratio $r(k,\theta,\theta')$ from Lemma 8 is $\frac{A_k'(\theta')}{A_k'(\theta)} = \frac{A(\theta)^2}{A(\theta')^2}\frac{N_k}{D_k}$, where

$$N_k = \sum_{l=0}^{k}\sum_{m\ge k+1} \beta^{l+m-1}\alpha_{lm}, \quad D_k = \sum_{l=0}^{k}\sum_{m\ge k+1} \alpha_{lm}.$$

We need to show that $\frac{N_k}{D_k}$ is increasing in k, or, equivalently, that $N_{k+1}D_k - N_k D_{k+1} \ge 0$. Notice that N_{k+1} can be expressed through N_k as follows:

$$N_{k+1} = N_k - \sum_{l=0}^{k} \beta^{l+k}\alpha_{l,k+1} + \sum_{m\ge k+2} \beta^{m+k}\alpha_{k+1,m}.$$

Similarly,

$$D_{k+1} = D_k - \sum_{l=0}^{k} \alpha_{l,k+1} + \sum_{m\ge k+2} \alpha_{k+1,m};$$

therefore,

$$N_{k+1}D_k - N_kD_{k+1} = \left(N_k - \sum_{l=0}^{k}\beta^{l+k}\alpha_{l,k+1} + \sum_{m\geq k+2}\beta^{m+k}\alpha_{k+1,m}\right)D_k$$

$$- N_k\left(D_k - \sum_{l=0}^{k}\alpha_{l,k+1} + \sum_{m\geq k+2}\alpha_{k+1,m}\right)$$

$$= \sum_{l=0}^{k}\alpha_{l,k+1}(N_k - \beta^{l+k}D_k) + \sum_{m\geq k+2}\alpha_{k+1,m}(\beta^{m+k}D_k - N_k).$$

It can be shown that each of the two terms in the last line is nonnegative. We demonstrate it explicitly for the first term; for the second term, the derivation is similar.

$$\sum_{l=0}^{k}\alpha_{l,k+1}(N_k - \beta^{l+k}D_k) = \sum_{l=0}^{k}\sum_{m\geq k+1}\sum_{l'=0}^{k}\left(\beta^{l'+m-1}\alpha_{l'm}\alpha_{l,k+1} - \beta^{l+k}\alpha_{l'm}\alpha_{l,k+1}\right)$$

$$= \sum_{l=0}^{k}\sum_{m\geq k+1}\sum_{l'=0}^{k}\left(\beta^{l+m-1}\alpha_{lm}\alpha_{l',k+1} - \beta^{l+k}\alpha_{l'm}\alpha_{l,k+1}\right)$$

$$\geq \sum_{l=0}^{k}\sum_{m\geq k+1}\sum_{l'=0}^{k}\beta^{l+k}\left(\alpha_{lm}\alpha_{l',k+1} - \alpha_{l'm}\alpha_{l,k+1}\right)$$

$$= \sum_{l=0}^{k}\sum_{m\geq k+1}\sum_{l'=0}^{k}\beta^{l+k}a_la_ma_{l'}a_{k+1}\theta^{l+m-1+l'+k}\left[(m-l)(k+1-l') - (m-l')(k+1-l)\right]$$

$$= \sum_{l=0}^{k}\sum_{m\geq k+1}\sum_{l'=0}^{k}\beta^{l+k}a_la_ma_{l'}a_{k+1}\theta^{l+m-1+l'+k}(m-k-1)(l-l')$$

$$= \sum_{m\geq k+1}\beta^k a_m a_{k+1}\theta^{m-1+k}(m-k-1)\sum_{l=0}^{k}\sum_{l'=0}^{k}\beta^l a_l a_{l'}\theta^{l+l'}(l-l').$$

The sum over l and l' can be rewritten as

$$\sum_{l=0}^{k}\sum_{l'=0}^{k}\beta^l a_l a_{l'}\theta^{l+l'}(l-l') = A_k(\theta)^2A(\theta)^2[\mathrm{E}(\beta^L L) - \mathrm{E}(\beta^L)\mathrm{E}(L)]$$

$$= A_k(\theta)^2A(\theta)^2\mathrm{Cov}(\beta^L, L) \geq 0.$$

Here, L is understood as a random variable with support $0, 1, \ldots, k$ and pmf $\frac{a_l\theta^l}{A_k(\theta)A(\theta)}$. The covariance is nonnegative because $\beta > 1$. ■

Proof of Proposition 7 First, we prove that if $f(x)$ is log-concave (log-convex) then kb_k is concave (convex) in k for $k \geq 2$. Integrating (4) by parts twice, obtain

$$E_k^* = kb_k = k(k-1)\int_0^1 z^{k-2}m(z)dz = k\left[m(z)z^{k-1}|_0^1 - \int_0^1 z^{k-1}m'(z)dz\right]$$
$$= km(1) - m'(z)z^k|_0^1 + \int_0^1 z^k m''(z)dz = km(1) - m'(1) + \int_0^1 z^k m''(z)dz.$$

This gives the second difference

$$\Delta^2 E_k^* = E_{k+2}^* - 2E_{k+1}^* + E_k^* = \int_0^1 z^k(1-z)^2 m''(z)dz \leq (\geq)0$$

for $m''(\cdot) \leq (\geq)0$, which holds when $f(x)$ is log-concave (log-convex), with strict inequality if $f(x)$ is strictly log-concave (log-convex).

Second, we compare $E_p^* = \bar{k}e_p^*$ to $E_{p'}^* = \bar{k}e_{p'}^*$. This is equivalent to comparing e_p^* and $e_{p'}^*$, i.e., it is sufficient to compare B_p to $B_{p'}$ or, from (13), $\sum_{k\geq 2} p_k kb_k$ to $\sum_{k\geq 2} p'_k kb_k$. Because $p_0 = p_1 = p'_0 = p'_1 = 0$, both sums represent unconditional expectations, and we are comparing $\mathrm{E}_p(Kb_K)$ to $\mathrm{E}_{p'}(Kb_K)$. The result then follows from the definition of second-order stochastic dominance. ■

Proof of Proposition 8 Without disclosure, the expected aggregate effort in the tournament is $E_p^* = \bar{k}e_p^* = \bar{k}c'^{-1}(B_p)$, where, from (13), $B_p = \mathrm{E}_{\tilde{p}}(b_K)$. With disclosure, the expected aggregate effort is $\mathrm{E}_p(Kc'^{-1}(b_K))$, which can be rewritten as

$$\mathrm{E}_p(Kc'^{-1}(b_K)) = \sum_{k=1}^n p_k k c'^{-1}(b_k) = \bar{k}\sum_{k=1}^n \tilde{p}_k c'^{-1}(b_k) = \bar{k}\mathrm{E}_{\tilde{p}}(c'^{-1}(b_K)).$$

Thus, comparing E_p^* and $\mathrm{E}_p(Kc'^{-1}(b_K))$ is equivalent to comparing $c'^{-1}(\mathrm{E}_{\tilde{p}}(b_K))$ and $\mathrm{E}_{\tilde{p}}(c'^{-1}(b_K))$.

It follows that when b_k is not constant in the support of $\tilde{p}$, and c'^{-1} is concave (convex) and nonlinear for at least some distinct values of b_k, disclosure is not optimal (optimal). The concavity (convexity) of c'^{-1} is equivalent to the convexity (concavity) of c', i.e., to the condition $c''' \geq (\leq)0$. ■

Proof of Proposition 9 We start by showing that condition (c) guarantees the existence of a unique e_k^* solving (3). Recall that $c'(\cdot)$ is strictly increasing and $c'(0) = 0$. It is, therefore, sufficient to show that $c'(e_{\max}) > b_k$. Condition (c) implies $c'^{-1}(f_m) \leq$

$c^{-1}(\frac{1}{k})$; therefore, $f_m \leq c'(c^{-1}(\frac{1}{k})) < c'(c^{-1}(1)) = c'(e_{\max})$. Representation (4) gives $b_k = \int_0^1 m(z)dz^{k-1} \leq f_m$, which produces the desired result.

Next, we use conditions (a) and (b) to show that payoff function (2) is strictly concave in e_i. Let $R(e) = \int_{\underline{x}}^{\overline{x}} F(e - e^* + x)^{k-1} dF(x)$ and suppose $c''(e) \geq c_0 > 0$ on $[0, e_{\max}]$. We need to show that $R''(e) < c_0$. For convenience, let $\Delta e = e - e^*$. Differentiating $R(e)$ once, obtain

$$R'(e) = (k-1) \int_{\underline{x}}^{\overline{x}} F(\Delta e + x)^{k-2} f(\Delta e + x) dF(x). \tag{20}$$

We need to evaluate the second derivative $R''(e)$. Note that the integrand in (20) is nonzero only for $x \in [\max\{\underline{x}, \underline{x} - \Delta e\}, \min\{\overline{x}, \overline{x} - \Delta e\}]$, and is continuous and piecewise differentiable in this interval under our assumptions; however, the integrand may be discontinuous on U. We, therefore, consider the cases when $\Delta e \geq 0$ and $\Delta e < 0$ separately.

(i) Suppose that $\Delta e \geq 0$. Then the interval of integration in (20) is $[\underline{x}, \overline{x} - \Delta e]$ and

$$\begin{aligned} R''(e) = (k-1) \bigg[&(k-2) \int_U F(\Delta e + x)^{k-3} f(\Delta e + x)^2 dF(x) \\ + &\int_U F(\Delta e + x)^{k-2} f'(\Delta e + x) dF(x) - f(\overline{x}) f(\overline{x} - \Delta e) \bigg] \leq (k-1)[(k-2) f_m^2 + f'_{\max}]. \end{aligned}$$

(ii) Suppose that $\Delta < 0$. Then the interval of integration in (20) is $[\underline{x} - \Delta e, \overline{x}]$ and

$$\begin{aligned} R''(e) &= (k-1) \bigg[(k-2) \int_U F(\Delta e + x)^{k-3} f(\Delta e + x)^2 dF(x) \\ &+ \int_U F(\Delta e + x)^{k-2} f'(\Delta e + x) dF(x) + F(\underline{x})^{k-2} f(\underline{x}) f(\underline{x} - \Delta e) \bigg] \\ &\leq (k-1)[(k-1) f_m^2 + f'_{\max}]. \end{aligned}$$

Thus, $D_+ = (k-1)[(k-1)f_m^2 + f'_{\max}]$ is a bound such that $c_0 > D_+$ ensures $R''(e) - c_0 < 0$.

An alternative bound on $R''(e)$ can be obtained by transforming (20) via a change of variables $x + \Delta e \to x$ into the form

$$R'(e) = (k-1) \int_{\underline{x}}^{\overline{x}} F(x)^{k-2} f(x - \Delta e) dF(x). \tag{21}$$

In this case the integrand is nonzero, continuous and piecewise differentiable for $x \in [\max\{\underline{x}, \underline{x} + \Delta e\}, \min\{\overline{x}, \overline{x} + \Delta e\}]$. We consider the same two cases as above.

(i) For $\Delta e \geq 0$, the interval of integration in (21) is $[\underline{x} + \Delta e, \overline{x}]$ and

$$R''(e) = (k-1)\left[-\int_U F(x)^{k-2} f'(x - \Delta e) dF(x) - F(\underline{x} + \Delta e)^{k-2} f(\underline{x}) f(\underline{x} + \Delta e)\right]$$
$$\leq -(k-1) f'_{\min}.$$

(ii) For $\Delta e < 0$, the interval of integration in (21) is $[\underline{x}, \overline{x} + \Delta e]$ and

$$R''(e) = (k-1)\left[-\int_U F(x)^{k-2} f'(x - \Delta e) dF(x) + F(\overline{x} + \Delta e)^{k-2} f(\overline{x}) f(\overline{x} + \Delta e)\right]$$
$$\leq (k-1)(f_m^2 - f'_{\min}).$$

This produces bound $D_- = (k-1)(f_m^2 - f'_{\min})$ such that $c_0 > D_-$ implies $R''(e) - c_0 < 0$. Since both bounds are valid, condition $c_0 > \min\{D_+, D_-\}$ is sufficient.

Note that condition (c) automatically implies that $\pi_i(e_k^*, e_k^*) = \frac{1}{k} - c(e_k^*) \geq 0$. Using the bound $b_k \leq f_m$ derived above, obtain $kc(e_k^*) = kc(c'^{-1}(b_k)) \leq kc(c'^{-1}(f_m)) \leq 1$. ■

Proof of Lemma 7 (i) Sufficiency: When $u(z)$ is monotone, it follows immediately that $\gamma(\theta)$ is monotone. Suppose that $u(z)$ is interior unimodal; in this case, $u(1)$ is finite. Integrating by parts, obtain

$$\gamma(\theta) = u(1) - \int_0^1 u'(z) H(z, \theta) dz. \tag{22}$$

Let $\hat{z} \in (0, 1)$ denote a mode of $u(z)$. Differentiating, or taking the first difference, with respect to θ, and splitting the integral in (22), obtain

$$\begin{aligned}\gamma'(\theta) &= -\int_0^{\hat{z}} u'(z) H_\theta(z, \theta) dz - \int_{\hat{z}}^1 u'(z) H_\theta(z, \theta) dz \\ &= \int_0^{\hat{z}} u'(z) |H_\theta(z, \theta)| dz - \int_{\hat{z}}^1 |u'(z)| |H_\theta(z, \theta)| dz.\end{aligned} \tag{23}$$

Suppose $\gamma'(\theta) \leq 0$ for some θ and consider a $\theta' > \theta$. Then (23) gives

$$\gamma'(\theta') = \int_0^{\hat{z}} u'(z)|H_\theta(z,\theta')|dz - \int_{\hat{z}}^1 |u'(z)||H_\theta(z,\theta')|dz$$
$$= \int_0^{\hat{z}} u'(z)r(z,\theta,\theta')|H_\theta(z,\theta)|dz - \int_{\hat{z}}^1 |u'(z)|r(z,\theta,\theta')|H_\theta(z,\theta')|dz$$
$$\leq r(\hat{z},\theta,\theta') \int_0^{\hat{z}} u'(z)|H_\theta(z,\theta)|dz - r(\hat{z},\theta,\theta') \int_{\hat{z}}^1 |u'(z)||H_\theta(z,\theta')|dz = r(\hat{z},\theta,\theta')\gamma'(\theta) \leq 0.$$

Here, the first inequality follows from the assumption that $r(z,\theta,\theta')$ is increasing in z. Thus, we showed that $\gamma(\theta)$ is unimodal.

(ii) Necessity: Suppose that there exist $\theta' > \theta$ and a $z \in [0,1]$ such that $r(z,\theta,\theta')$ is decreasing in z. The proof consists in showing that a unimodal function $u(z)$ can then be constructed such that $\gamma(\theta)$ is not unimodal. By continuity, there exists an interval of positive length $[z_1, z_2]$ where $r(z,\theta,\theta')$ is strictly decreasing. First, define a unimodal function $u(z)$ such that it is nonzero only withing this interval. Furthermore, $u(z)$ can be defined in a way that $\gamma'(\theta) = 0$. For example, it can be defined as a piecewise linear function such that $u'(z) = \int_{\hat{z}}^{z_2} |H_\theta(z,\theta)|dz$ for $z \in (z_1, \hat{z})$ and $|u'(z)| = \int_{z_1}^{\hat{z}} |H_\theta(z,\theta)|dz$ for $z \in (\hat{z}, z_2)$. In this case, it follows from (23) that $\gamma'(\theta) = 0$. Finally, we modify this $u(z)$ "slightly" to make $\gamma'(\theta)$ negative. For example, choose some $\epsilon > 0$ and set $u'(z) = \int_{\hat{z}}^{z_2} |H_\theta(z,\theta)|dz - \epsilon$ for $z \in (z_1, \hat{z})$. Then

$$\gamma'(\theta') = \int_{z_1}^{\hat{z}} u'(z)r(z,\theta,\theta')|H_\theta(z,\theta)|dz - \int_{\hat{z}}^{z_2} |u'(z)|r(z,\theta,\theta')|H_\theta(z,\theta')|dz$$
$$= r(z_1^*,\theta,\theta') \int_{z_1}^{\hat{z}} u'(z)|H_\theta(z,\theta)|dz - r(z_2^*,\theta,\theta') \int_{\hat{z}}^{z_2} |u'(z)||H_\theta(z,\theta')|dz$$
$$= r(z_1^*,\theta,\theta') \left[\int_{\hat{z}}^{z_2} |H_\theta(z,\theta)|dz - \epsilon\right] \int_{z_1}^{\hat{z}} |H_\theta(z,\theta)|dz$$
$$- r(z_2^*,\theta,\theta') \int_{\hat{z}}^{z_2} |H_\theta(z,\theta')|dz \int_{z_1}^{\hat{z}} |H_\theta(z,\theta)|dz$$
$$= (r(z_1^*,\theta,\theta') - r(z_2^*,\theta,\theta')) \int_{z_1}^{\hat{z}} |H_\theta(z,\theta)|dz \int_{\hat{z}}^{z_2} |H_\theta(z,\theta')|dz$$
$$- \epsilon r(z_1^*,\theta,\theta') \int_{\hat{z}}^{z_2} |H_\theta(z,\theta)|dz.$$

Here, $z_1^* \in (z_1, \hat{z})$ and $z_2^* \in (\hat{z}, z_2)$ exist due to the mean-value theorem for definite integrals. Note that $z_2^* > z_1^*$ and hence the first term in the last expression is positive, while the second term can be made arbitrarily small via the choice of ϵ; therefore, an $\epsilon > 0$ can be chosen such that $\gamma'(\theta') > 0$. Thus, $\gamma(\theta)$ is not unimodal. ∎

Proof of Lemma 8 (i) Sufficiency: Rewrite $\chi(\theta)$ as follows:

$$
\begin{aligned}
\chi(\theta) &= p_1(\theta)u_1 + p_2(\theta)u_2 + \ldots + p_{n-1}(\theta)u_{n-1} + p_n(\theta)u_n \\
&= P_1(\theta)u_1 + (P_2(\theta) - P_1(\theta))u_2 + \ldots + (P_{n-1}(\theta) - P_{n-2}(\theta))u_{n-1} + (P_n(\theta) - P_{n-1}(\theta))u_n \\
&= u_n + P_1(\theta)(u_1 - u_2) + P_2(\theta)(u_2 - u_3) + \ldots + P_{n-1}(\theta)(u_{n-1} - u_n) \\
&= u_n - \sum_{k=1}^{n-1} P_k(\theta)\Delta u_{k+1},
\end{aligned}
$$

where $\Delta u_{k+1} = u_{k+1} - u_k$. This "summation by parts" representation is similar to integration by parts and expresses the expectation $\chi(\theta)$ through the cmf $P(\theta)$ and the first difference of u_k. Taking the derivative, or the difference, with respect to θ, obtain

$$
\chi'(\theta) = -\sum_{k=1}^{n-1} P_k'(\theta)\Delta u_{k+1} = \sum_{k=1}^{n-1} |P_k'(\theta)|\Delta u_{k+1}.
$$

Let $\hat{k}$ denote a mode of u_k such that $\Delta u_{k+1} \geq (\leq) 0$ for $k < (\geq) \hat{k}$. This gives

$$
\chi'(\theta) = \sum_{k<\hat{k}} |P_k'(\theta)|\Delta u_{k+1} - \sum_{k\geq\hat{k}} |P_k'(\theta)||\Delta u_{k+1}|.
$$

Suppose that $\chi'(\theta) \leq 0$ for some θ and consider a $\theta' > \theta$. Then

$$
\begin{aligned}
\chi'(\theta') &= \sum_{k<\hat{k}} |P_k'(\theta')|\Delta u_{k+1} - \sum_{k\geq\hat{k}} |P_k'(\theta')||\Delta u_{k+1}| \\
&= \sum_{k<\hat{k}} |P_k'(\theta)| r(k,\theta,\theta')\Delta u_{k+1} - \sum_{k\geq\hat{k}} |P_k'(\theta)| r(k,\theta,\theta')|\Delta u_{k+1}| \\
&\leq r(\hat{k},\theta,\theta') \sum_{k<\hat{k}} |P_k'(\theta)|\Delta u_{k+1} - r(\hat{k},\theta,\theta') \sum_{k\geq\hat{k}} |P_k'(\theta)||\Delta u_{k+1}| = r(\hat{k},\theta,\theta')\chi'(\theta) \leq 0.
\end{aligned}
$$

Here, the first inequality follows from the assumption that $r(\hat{k},\theta,\theta')$ is increasing in k.

(ii) Necessity: Suppose that there exist $\theta' > \theta$ and k such that $r(k-1,\theta,\theta') > r(k,\theta,\theta')$. As in the proof of Lemma 7, we will show that it is possible to construct a unimodal sequence u_k such that $\chi(\theta)$ is not unimodal. Set $u_l = a$ for all $l \leq k-1$ and $u_l = b$ for all $l \geq k+1$; furthermore, set $u_k > \max\{a, b\}$. The resulting sequence u_k is interior unimodal with mode k and satisfies $\Delta u_k > 0$, $\Delta u_{k+1} < 0$, and $\Delta u_l = 0$ for all $l \neq k, k+1$. Then

$$\chi'(\theta) = |P'_{k-1}(\theta)|\Delta u_k - |P'_k(\theta)||\Delta u_{k+1}|.$$

Choosing a, u_k and b so that $\Delta u_k = |P'_k(\theta)| - \epsilon$ for some $\epsilon > 0$ and $|\Delta u_{k+1}| = |P'_{k-1}(\theta)|$, obtain $\chi'(\theta) = -\epsilon|P'_{k-1}(\theta)| < 0$. However,

$$\begin{aligned}
&\chi'(\theta') = |P'_{k-1}(\theta')|\Delta u_k - |P'_k(\theta')||\Delta u_{k+1}| \\
&= r(k-1,\theta,\theta')|P'_{k-1}(\theta)|(|P'_k(\theta)| - \epsilon) - r(k,\theta,\theta')|P'_k(\theta)||P'_{k-1}(\theta)| \\
&= (r(k-1,\theta,\theta') - r(k,\theta,\theta'))|P'_k(\theta)||P'_{k-1}(\theta)| - \epsilon r(k-1,\theta,\theta')|P'_{k-1}(\theta)|.
\end{aligned}$$

The first term on the last line is strictly positive, while the second term can be made arbitrarily small through the choice of ϵ; thus, an $\epsilon > 0$ can be chosen such that $\chi'(\theta') > 0$, i.e., $\chi(\theta)$ is not unimodal. ■

Proof of Lemma 9 (i) Sufficiency: By differentiating, or taking the first difference of, Eq. (17) with respect to θ, obtain

$$\sum_{k=1}^{n} P'_k(\theta) z^{k-1} = \frac{G_\theta(z,\theta)}{1-z},$$

which gives, for some $\theta' > \theta$,

$$R(z,\theta,\theta') = \frac{|G_\theta(z,\theta')|}{|G_\theta(z,\theta)|} = \frac{\sum_{k=1}^n |P'_k(\theta')|z^{k-1}}{\sum_{k=1}^n |P'_k(\theta)|z^{k-1}} = \frac{\sum_{k=1}^n |P'_k(\theta)|r(k,\theta,\theta')z^{k-1}}{\sum_{k=1}^n |P'_k(\theta)|z^{k-1}}. \tag{24}$$

Define a pmf $\alpha_k(z) = \frac{|P'_k(\theta)|z^{k-1}}{\sum_{l=1}^n |P'_l(\theta)|z^{l-1}}$ and the corresponding cmf $A_k(z) = \sum_{l=1}^k \alpha_k(z)$. Then (24) can be written as an expectation $R(z,\theta,\theta') = \sum_{k=1}^n \alpha_k(z) r(k,\theta,\theta')$ of an increasing random variable $r(K,\theta,\theta')$. This expectation is increasing in z provided an increase in z leads to an FOSD increase in distribution $\alpha(z)$, i.e., if $A_k(z)$ is decreasing in z. The derivative of $A_k(z)$ is

$$A'_k(z) = \frac{d}{dz}\left(\frac{\sum_{l=1}^k |P'_l(\theta)|z^{l-1}}{\sum_{l=1}^n |P'_l(\theta)|z^{l-1}}\right) = \frac{1}{(\sum_{l=1}^n |P'_l(\theta)|z^{l-1})^2} \sum_{l=1}^{k}\sum_{l'=1}^{n} |P'_l(\theta)||P'_{l'}(\theta)|z^{l+l'-3}(l-l')$$

$$= \frac{1}{(\sum_{l=1}^n |P'_l(\theta)|z^{l-1})^2} \sum_{l=1}^{k}\sum_{l'=k+1}^{n} |P'_l(\theta)||P'_{l'}(\theta)|z^{l+l'-3}(l-l') \leq 0. \tag{25}$$

(ii) Necessity: Define $\Delta r_{l+1} = r(l+1, \theta, \theta') - r(l, \theta, \theta')$, and suppose that $\Delta r_{k+1} < 0$ for some k and $\theta' > \theta$. Using the same "summation by parts" transformation as at the start of the proof of Lemma 8, write

$$R(z, \theta, \theta') = r(n, \theta, \theta') - \sum_{l=1}^{n-1} A_l(z) \Delta r_{l+1},$$

which gives, differentiating with respect to z,

$$R_z(z, \theta, \theta') = \sum_{l=1}^{n-1} |A_l'(z)| \Delta r_{l+1}.$$

Choose $P_l(\theta)$ so that $P_l'(\theta) = 0$ for all $l \neq k, k+1$ and $P_k'(\theta), P_{k+1}'(\theta) < 0$. Equation (25) then gives

$$A_k'(z) = \frac{-|P_k'(\theta)||P_{k+1}'(\theta)|z^{2k-2}}{(|P_k'(\theta)|z^{k-1} + |P_{k+1}'(\theta)|z^k)^2} < 0$$

and $A_l'(z) = 0$ for all $l \neq k$; therefore, we obtain $R_z(z, \theta, \theta') = |A_k'(z)| \Delta r_{k+1} < 0$, which is a contradiction. ■