DISCUSSION PAPER SERIES

DP12067

WINNER-TAKE-ALL TOURNAMENTS

Mikhail Drugov and Dmitry Ryvkin

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

WINNER-TAKE-ALL TOURNAMENTS

Mikhail Drugov and Dmitry Ryvkin

Discussion Paper DP12067
Published 29 May 2017
Submitted 29 May 2017

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

WINNER-TAKE-ALL TOURNAMENTS

Abstract

This paper provides new general results for winner-take-all rank-order tournaments with additive and multiplicative noise. We show that the comparative statics of the individual equilibrium effort with respect to the number of players follow the shape of the density the noise distribution. For aggregate effort, a similar relation holds for the failure (hazard) rate of the noise distribution. The equilibrium effort decreases as noise becomes more dispersed, in the sense of the dispersive order or appropriately defined entropy. These results are then extended to the case of a stochastic number of players, and new results on the effects of population uncertainty are obtained. All relevant results for the Tullock contest follow as a special case.

JEL Classification: C72, D72, D82

Keywords: tournament, stochastic number of players, unimodality, log-supermodularity, failure rate, entropy, dispersive order, Tullock contest

Mikhail Drugov - mdrugov@nes.ru
New Economic School and CEPR

Dmitry Ryvkin - dryvkin@fsu.edu
Florida State University

Acknowledgements
We are grateful to Johannes Hörner, Margaret Meyer, Marc Möller, Armin Schmutzler, Marco Serena and seminar participants at Universitat Autònoma de Barcelona, University of Bern and Toulouse School of Economics.

Winner-Take-All Tournaments*

Dmitry Ryvkin† Mikhail Drugov‡

This version: May 28, 2017

Abstract

This paper provides new general results for winner-take-all rank-order tournaments with additive and multiplicative noise. We show that the comparative statics of the individual equilibrium effort with respect to the number of players follow the shape of the density the noise distribution. For aggregate effort, a similar relation holds for the failure (hazard) rate of the noise distribution. The equilibrium effort decreases as noise becomes more dispersed, in the sense of the dispersive order or appropriately defined entropy. These results are then extended to the case of a stochastic number of players, and new results on the effects of population uncertainty are obtained. All relevant results for the Tullock contest follow as a special case.

Keywords: tournament, stochastic number of players, unimodality, log-supermodularity, failure rate, entropy, dispersive order

JEL codes: C72, D72, D82

*We are grateful to Johannes Hörner, Margaret Meyer, Marc Möller, Armin Schmutzler, Marco Serena and seminar participants at Universitat Autònoma de Barcelona, University of Bern and Toulouse School of Economics.

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

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

1 Introduction

Tournaments are environments in which participants compete for a valuable prize by spending effort or other resources. Examples include rent-seeking and lobbying, wars and conflicts, R&D races, litigation, sport competitions, political campaigns and tournaments in organizations where promotions or bonuses are based on the relative performance of workers. 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.[1]

In a generic model of a winner-take-all (WTA) tournament, a player's output is written as $y_i = \phi(e_i, X_i)$, where e_i is player i's costly effort (or another type of investment), X_i is an idiosyncratic random shock with some pdf $f(t)$, and ϕ is a "production function" increasing in both arguments. The player with the highest output wins the tournament and receives a fixed prize, and all players incur their effort costs. 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," X_i, in the winner determination process. The Lazear-Rosen model uses an additive production function, $\phi(e_i, X_i) = e_i + X_i$, while the contest success function (CSF) of Tullock (1980) can be obtained using a multiplicative production function, $\phi(e_i, X_i) = e_i X_i$, with the generalized inverse exponential distribution of noise. Jia (2008) and Jia, Skaperdas and Vaidya (2013) provide a unified framework for the two models showing that the Tullock contest can be represented as a special case of a Lazear-Rosen tournament.[2]

Yet, the existing analysis of *general* WTA 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), or the Lazear-Rosen tournament with two players.[3] Relatively little is known in

[1]For a recent summary, see, e.g., Konrad (2009), Congleton, Hillman and Konrad (2008), Corchón (2007), Connelly et al. (2014).

[2]Throughout this paper, we focus exclusively on models of "imperfectly discriminating" contests with noise and use "tournament" as a unifying term for such models. While it has been known in the demand estimation literature for a long time that the logit model can be derived from the random utility model (McFadden, 1974), in the tournament literature the Tullock contest and the Lazear-Rosen tournament models have been treated as two completely unrelated models, with the exception of Jia, Skaperdas and Vaidya (2013).

[3]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

general about the basic comparative statics of the WTA tournament model. Common wisdom suggests that as the number of players increases the individual probability of winning goes down and hence so does the marginal gain from increasing one's effort, leading to lower effort in equilibrium. This is indeed the case in the Tullock contest (see, for example, surveys by Nitzan, 1994; Corchón, 2007). However, the symmetric equilibrium effort is independent of the number of players in a Lazear-Rosen tournament when the distribution of individual noise is uniform. Gerchak and He (2003) further show that the equilibrium effort is decreasing in the number of players when $f(t)$ is decreasing or unimodal and symmetric, and increasing when $f(t)$ is increasing.[4] Are these disparate results part of a more general pattern? Can individual equilibrium effort be nonmonotone? Even less is known about the behavior of aggregate effort. In Tullock contests and Lazear-Rosen tournaments with a uniformly distributed noise it is increasing in the number of players, but can it be decreasing or nonmonotone?

Similar unanswered questions exist about the effects of the distribution of noise. Intuitively, as noise becomes more dispersed, the marginal gain from increasing one's effort declines, and hence equilibrium effort should go down. Indeed, when the distribution of noise is uniform with support $[-s, s]$, the equilibrium effort is proportional to $\frac{1}{2s}$, confirming the intuition. Consider, however, the distribution of noise with pdf $f(t) = \frac{|t|}{s^2}$ on the same support. Even though its variance is higher than that of the uniform distribution and, more generally, it is dominated by the uniform distribution in the sense of second-order stochastic dominance (SOSD), this distribution leads to a higher equilibrium effort than the uniform distribution in a two-player tournament. The reason is, as noted by Gerchak and He (2003), that this distribution has a lower *entropy*, and it is the Rényi entropy of order 2, and not the variance or SOSD ordering, that determines the effect of noise on the equilibrium effort. These results, however, are restricted to two-player tournaments because the entropy ordering of distributions is not, in general, preserved by order statistics. Is there a natural ordering of noise distributions in terms of their impact on the equilibrium effort for any number of players?

In this paper, we develop an innovative approach to the analysis of general rank-order tournament models, with far-reaching applications. We start with the analysis of the standard WTA Lazear-Rosen model[5] and show that the *unimodality* of the distribution

McLaughlin (1988).

[4]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.

[5]We use the formulation with additive noise. Models with multiplicative noise, such as the Tullock

of noise is a unifying characteristic that is preserved by the individual equilibrium effort. We provide a general characterization of the equilibrium comparative statics for unimodal noise distributions, from which all existing results follow as special cases. For aggregate effort, the failure (hazard) rate of the distribution of noise plays a similar role; that is, aggregate effort is unimodal in the number of players for unimodal failure rates, when the cost function of effort is quadratic. However, as expected, the behavior of aggregate effort depends critically on the curvature of the cost function. We show that for cost functions that are more convex than quadratic, in the sense of the likelihood ratio order, aggregate effort increases in the number of players for noise distributions with increasing failure rates. The opposite is true for cost functions that are less convex than quadratic, and decreasing failure rates.

We also analyze the effect of noise dispersion on the equilibrium effort. As we show, noise distributions can be ranked in terms of their effect on the equilibrium effort using the *dispersive order* (Lehmann, 1988). This order is preserved by order statistics and hence the ranking applies to tournaments with an arbitrary number of players.

We then extend the analysis to WTA tournaments with a stochastic number of players. Indeed, 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 would be the case, for example, in coding contests where an unknown and potentially very large number of coders submit their solutions; in hiring tournaments where a job seeker does 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 tradition of the literature on auctions with a stochastic number of bidders (e.g., McAfee and McMillan, 1987; 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 deterministic participation case, we show that 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

contest, are transformed into an appropriately defined additive noise form, and hence their comparative statics follow as a special case of a more general theory, see Section 2.2 (cf. also Jia, Skaperdas and Vaidya, 2013).

from arguments similar to those identified by Athey (2002) for the preservation of single-crossing under uncertainty. We show that this condition is rather weak and is satisfied, for example, by the family of power series distributions which includes the Poisson, binomial, negative binomial and logarithmic distributions. We also explore the effects of noise dispersion and show that, in addition to the dispersive order, they can be expressed through an appropriate entropy defined by a combination of the distribution of noise and the tournament size distribution. Finally, we generalize the results of Myerson and Wärneryd (2006) on the (dis)advantage of having uncertainty in the number of players, and of Fu, Jiao and Lu (2011) on the benefits of disclosing the realized number of players.

For general tournaments with a fixed number of players, the paper closest to ours is Gerchak and He (2003). These authors provide comparative statics results for the individual equilibrium effort when the density of the noise distribution is monotone or symmetric and unimodal. However, they have no results for general unimodal distributions or for aggregate effort. They also note that the entropy of the noise distribution determines the equilibrium effort, but this result is restricted to two-player tournaments, and they do not relate it to the dispersive order. Hence, Gerchak and He (2003) is a useful first step in the general study of tournaments, on which we vastly expand.[6]

The present paper is also the first one to study general rank-order tournaments with a stochastic number of players where the previous literature is restricted to the Tullock contest model and its lottery-form generalizations. 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. They show that aggregate expenditure is strictly lower in the former case if it is guaranteed that the contest has at least one participant. Lim and Matros (2009) show that, for the binomial distribution with parameters (n, q) of contest size, the equilibrium effort is nonmonotone and single-peaked in q when the number of potential players $n > 2$. They also show that, as long as q is not too high, effort is nonmonotone in n. Münster (2006) explores the effect of risk-aversion in the same setting. Fu, Jiao and Lu (2011) study the effect of disclosure of the number of participating players on aggregate effort. They show that disclosure or nondisclosure may be optimal depending on the properties of the "impact function" in the generalized lottery-form CSF; and in the special case of lottery CSF of Tullock (1980), the principal is indifferent between disclosure and nondisclosure. Finally, Fu and Lu (2010) study endogenous entry and the optimal allocation of winner's

[6]Their paper is not well-known, and we became aware of it only when this paper was almost finished.

prize and participation subsidy/fee. There is no contest size uncertainty in their model, however, because entry occurs sequentially and each player observes the number of prior entrants. Fu and Lu (2010) find that the optimal contract extracts all surplus from the contestants and restricts participation to two active players. 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 noise and shows how the case of multiplicative noise reduces to it as well. Section 3 provides general results on the preservation of unimodality under uncertainty that are used in the subsequent sections. Section 4 focuses on tournaments with deterministic number of players and presents the comparative statics with respect to the number of players and a ranking of noise distributions. Section 5 analyzes the case of a stochastic number of players, and Section 6 concludes. The proofs are contained in Appendix A.

2 Model setup

2.1 Additive noise

There are $k \geq 2$ identical, risk-neutral players indexed by $i \in \mathcal{K} = \{1, \ldots, k\}$. All players $i \in \mathcal{K}$ simultaneously and independently choose efforts $e_i \geq 0$. The cost of effort e_i to player i is $c(e_i)$, where $c(\cdot)$ is strictly increasing, strictly convex, and twice differentiable on $(0, c^{-1}(1)]$, with $c(0) = 0$. 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 and probability density function (pdf) f defined on interval support U. When necessary, we will use $\underline{x}$ and $\overline{x}$ to denote, respectively, the lower and upper bounds of U, which may be finite or infinite.[9] We assume that f is atomless, continuous and piece-wise differentiable in the interior of U, and has an inverse quantile density $m(z)$ (defined below) that is continuous and piece-wise differentiable on $(0, 1)$

[7]For a theoretical analysis of auctions with a stochastic number of bidders see, e.g., McAfee and McMillan (1987), Harstad, Kagel and Levin (1990) and Levin and Ozdenoren (2004). For a theoretical analysis of endogenous entry in auctions see, e.g., Levin and Smith (1994) and Pevnitskaya (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).

[9]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.

and integrable on $[0,1]$. The winner of the tournament is the player whose output is the highest.[10] The winner receives a prize normalized to one, whereas all other players receive zero.[11]

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

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

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

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

The first-order condition for payoff maximization evaluated at $e_i=e^*$, $\frac{\partial\pi_i(e_i,e^*)}{\partial e_i}\Big|_{e_i=e^*}=0$, gives

$$c'(e^*)=b_k\equiv(k-1)\int_U F(t)^{k-2}f(t)dF(t).\tag{3}$$

Note that $c'(e^*)$ is a strictly increasing function; therefore, if Eq. (3) has a solution, it is positive and unique for $k\geq 2$. In what follows we assume that such a solution, e_k^*, exists, and that it is a symmetric pure strategy equilibrium, i.e., $e_k^*\in\arg\max_{e_i\in[0,c^{-1}(1)]}\pi_i(e_i,e_k^*)$, where $\pi_i(e_i,e_k^*)$ is given by (2).[12]

Because the marginal cost of effort, $c'(\cdot)$, is strictly increasing, 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 $F^{-1}(z)=\inf\{t\in U:F(t)\geq z\}$ denote the quantile function of the distribution of noise. Introduce

[10] Ties are broken randomly but, under the assumption of atomless f, occur with probability zero.

[11] 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.

[12] Equilibrium existence and comparative statics are two separate issues, and here we focus on the latter, leaving the discussion of equilibrium existence (and uniqueness) outside the scope of this paper. In the Lazear-Rosen tournament model, these are still open questions. It is generally understood that the 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, cf. Nalebuff and Stiglitz (1983). Note that the second-order condition and the requirement that zero effort is not a best response are not sufficient for e_k^* to be a symmetric equilibrium because function $\pi_i(e_i,e_k^*)$ may have multiple local maxima in e_i. For completeness, we provide the second-order condition in Appendix A.

also an unnormalized density function $m(z) = f(F^{-1}(z))$, known as the inverse quantile density function (Parzen, 1979). Using the probability integral transformation $z = F(t)$, 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 $[0,1]$. The 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 $m(z)$, while the former are determined by a family of FOSD-ordered highest order statistics with cdfs z^{k-1}.

2.2 Multiplicative noise

Via simple transformations of the distribution of noise and the cost of effort, the 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

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

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 numbers. Defining $\hat{F}(t) = F(\exp(t))$ 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 is reduced to a tournament model with additive noise, and all the results above 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* $\hat{F}(t) = F(\exp(t))$ *and the cost of effort* $c_{\mathrm{m}}(e) = \int_0^e c'(x)x dx$.

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, $c(e) = e$. 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 $F(t) = \exp(-t^{-r})$.

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 type-I extreme value (or Gumbel) distribution with cdf $\hat{F}(t) = F(\exp(t)) = \exp[-\exp(-rt)]$ and pdf $\hat{f}(t) = r\exp[-rt - \exp(-rt)]$ (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 $c_{\mathrm{m}}(e) = \int_0^e x dx = \frac{e^2}{2}$. The first-order condition then takes the form $e_k^* = \hat{b}_k$, where $\hat{b}_k$ is given by Eq. (4) with $m(z) = \hat{f}(\hat{F}^{-1}(z)) = -rz\ln z$:

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

which is the equilibrium effort in the Tullock contest.

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

A word of caution is necessary regarding the correspondence between the tournament with multiplicative noise and cost $c(e)$ and the tournament with additive noise and cost $c_{\mathrm{m}}(e)$ described in Proposition 1. The two tournaments have the same symmetric equilibrium effort, e_k^* (provided the equilibrium exists in both cases), but otherwise these are two different games. That is, while the properties of equilibrium of the original model can be studied using the equilibrium of the additive noise model, the global properties of the two games, such as the equilibrium existence conditions and payoff levels, can be different. For illustration, consider the Tullock contest. The symmetric equilibrium payoff in the original model is $\pi^* = \frac{1}{k} - e_k^* = \frac{1}{k}\left[1 - \frac{r(k-1)}{k}\right]$, and hence the restriction $\pi^* \geq 0$ imposes the condition $r \leq \frac{k}{k-1}$, a well-known restriction on the discriminatory power. However, in the corresponding additive noise game, the symmetric equilibrium payoff is different, $\pi_{\mathrm{m}}^* = \frac{1}{k} - \frac{e_k^{*2}}{2} = \frac{1}{k}\left[1 - \frac{r^2(k-1)^2}{2k^3}\right]$, and is nonnegative for a much wider range of discriminatory powers, $r \leq \frac{\sqrt{2k^3}}{k-1}$.

3 Preservation of unimodality under uncertainty

In what follows, we explore the comparative statics of individual and aggregate equilibrium effort in tournaments with respect to the number of players, k. First, in Section 4, we assume that k is deterministically given; then, in Section 5, 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 leading to first-order stochastic dominance (FOSD).

In both cases, we show that robust comparative statics can be obtained for *unimodal* distributions of noise $f(t)$. These comparative statics amount to preservation of unimodality under uncertainty. Indeed, coefficients b_k, Eq. (4), which determine the comparative statics 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 cdfs $H(z,\theta)$ are FOSD-ordered in θ, to be unimodal in θ for all unimodal functions $u(z)$. When we turn to the case of stochastic group size, equilibrium effort will be determined by discrete expectations of the form $\chi(\theta) = \sum_{k=1}^n x_k p_k(\theta)$, where $x = \{x_k\}_{k=1}^n$ 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 x. 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 a $\hat{t} \in S$ such that $u(t)$ is increasing for $t \leq \hat{t}$ and decreasing for $t \geq \hat{t}$. 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* $t_1, t_1' \in S_1$*,* $t_2, t_2' \in S_2$*, such that* $t_1' > t_1$ *and* $t_2' > t_2$*,*

$$v(t_1, t_2')v(t_1', t_2) \leq v(t_1, t_2)v(t_1', t_2').$$

In other words, for all $t_2' > t_2$ *the ratio* $r(t_1, t_2, t_2') = \frac{v(t_1, t_2')}{v(t_1, t_2)}$ *is increasing in* t_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 piece-wise differentiable function and $H(z, \theta)$ is a cdf of a random variable $Z|\theta$ defined on $[0, 1]$ and parameterized by $\theta \in \Theta \subseteq \mathbb{R}$.[13] 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 1 $\gamma(\theta)$ *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 $\chi(\theta) = \sum_{k=1}^n x_k p_k(\theta)$, where x 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 will 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 uniformly for all values of θ.[14] 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 2 $\chi(\theta)$ *is unimodal for all unimodal sequences* x *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 2 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

[13]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 $\theta \in \Theta$.

[14]This is not to say that $p(\theta)$ have the same support for all $\theta \in \Theta$; 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 $k > n(\theta)$.

$p_k(\theta)$ 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

$$\sum_{k=1}^{n} P_k(\theta) z^{k-1} = \frac{G(z,\theta) - z^{n-1}}{1-z}. \tag{6}$$

It follows from Eq. (6) 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 $G_\theta(z,\theta) \le 0$ denote, similar to $H_\theta(z,\theta)$ in Lemma 1, either the derivative or the first difference of $G(z,\theta)$ with respect to θ.

Lemma 3 *$|G_\theta(z,\theta)|$ 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 2 is increasing in k for any $\theta' > \theta$.*

The increasing ratio conditions in Lemmas 1, 2 and 3 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). Our results 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 1 is a straightforward corollary of these results applied to unimodal functions, i.e., functions with a single-crossing derivative. Indeed, assuming $u(1)$ 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 2 is a discrete version of Lemma 1 and follows similarly via "summation by parts." Lemma 3, 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.

4 Tournaments with deterministic group size

4.1 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, Eq. (3). These coefficients represent the marginal benefit of effort in equilibrium, and can be written as

$$b_k = \int_U f(t) dF(t)^{k-1} = \int_U f(t) f_{(k-1:k-1)}(t) dt, \tag{7}$$

where $F(t)^{k-1}$ is the cdf of the $(k-1)$-th order statistic among $k-1$ i.i.d. draws from distribution F, and $f_{(k-1:k-1)}(t) = \frac{d}{dt}F(t)^{k-1}$ 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. (7). This representation immediately leads to comparative statics results for monotone pdfs $f(t)$.

Lemma 4 *(i) If $f(t)$ is increasing (decreasing) then b_k (and e_k^*) is increasing (decreasing) for $k \geq 2$.*
(ii) b_k (and e_k^*) is constant for $k \geq 2$ if and only if $f(t)$ 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 $f(t)$ become more important as k increases. For example, if $f(t)$ 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 4 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 implied by representation (7) allows us to also obtain large-k asymptotic results for an arbitrary $f(t)$. 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 as k increases; hence, the asymptotic behavior of b_k is determined by the shape of the upper tail of pdf $f(t)$. Specifically, a decreasing (increasing) upper tail of $f(t)$ will lead to a decreasing (increasing) b_k for large k. The following lemma states the result formally.

Lemma 5 *Define $\hat{t} = \inf\{t' \in U : f(t) \text{ is monotone for } t > \hat{t}\}$. If $f(t)$ is decreasing (increasing) for $t > \hat{t}$, then there exists a large enough $\hat{k}$ such that b_k is decreasing (increasing) for all $k > \hat{k}$.*

Point $\hat{t}$ defined in Lemma 5 determines the location of the "last" interior peak or dip of $f(t)$. If pdf $f(t)$ is monotone, $\hat{t} = \underline{x}$ and b_k is either decreasing or increasing for all

$k \geq 2$, by Lemma 4. If $f(t)$ is nonmonotone, b_k is asymptotically decreasing or increasing depending on the behavior of the "last" monotone part of $f(t)$. Lemma 5 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 1. Note that $m(z)$ has the same monotonicity as $f(t)$, 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 (7) applies.

Proposition 2 *(i) If $f(t)$ is unimodal then b_k (and e_k^*) is unimodal for $k \geq 2$.*
(ii) If $f(t)$ is unimodal and symmetric then $b_2 = b_3$ (and $e_2^* = e_3^*$), and b_k (and e_k^*) is decreasing for $k \geq 3$.
(iii) If $f(t)$ is symmetric (not necessarily unimodal) then $b_2 = b_3$ (and $e_2^* = e_3^*$).

Part (i) of Proposition 2 is the main result of this section, and it follows directly from Lemma 1. Indeed, considering the representation (4) with $H(z, k) = z^{k-1}$, 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 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 unifying 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 5). 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 $-f(t)$ is unimodal. Of interest is the case when $f(t)$ is U-shaped and nonmonotone (when f is monotone, Lemma 4 applies).

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

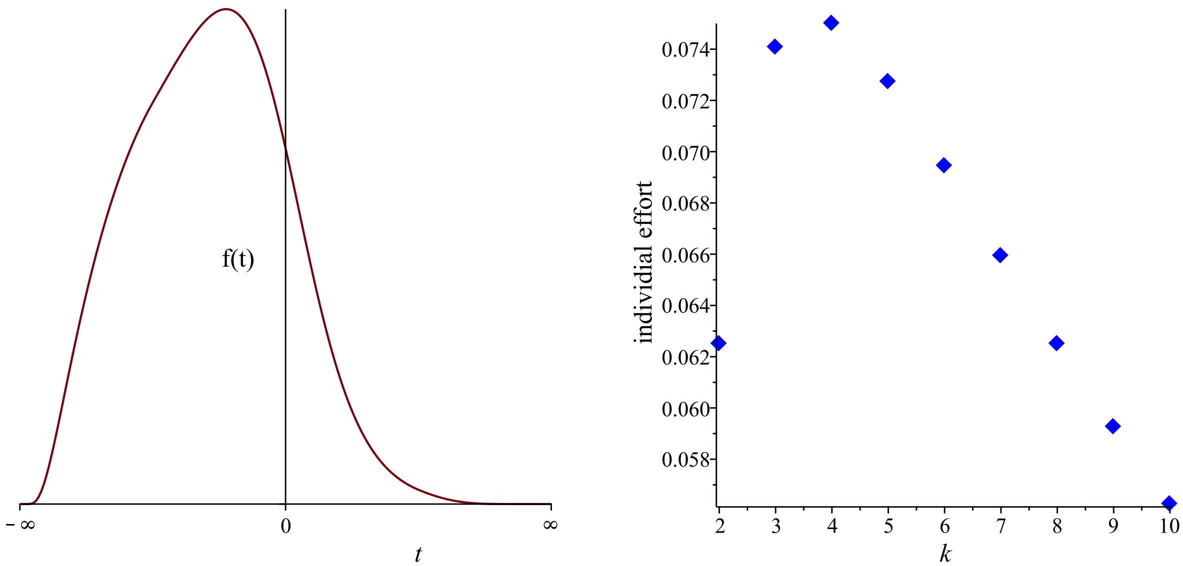

Figure 1: *Left*: The pdf $f(t)$ of the type I generalized logistic distribution with $a = \frac{1}{6}$. *Right*: Individual equilibrium effort e_k^* as a function of k for effort cost function $c(e) = \frac{1}{2}e^2$.

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 $F(t) = \frac{1}{(1+\exp(-t))^a}$ 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.[15] At the same time, a non-unimodal $f(t)$ does not necessarily lead to a non-unimodal sequence b_k. For example, a bimodal distribution with pdf $f(t) = \frac{1}{2}[f_{N(-12,4)}(t) + f_{N(12,4)}(t)]$, where $f_{N(\mu,\sigma^2)}(t)$ is the pdf of the Normal distribution with mean μ and variance σ^2, generates a decreasing sequence b_k. Thus, there is no "higher-order" universality of behavior of b_k for non-unimodal distributions.

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

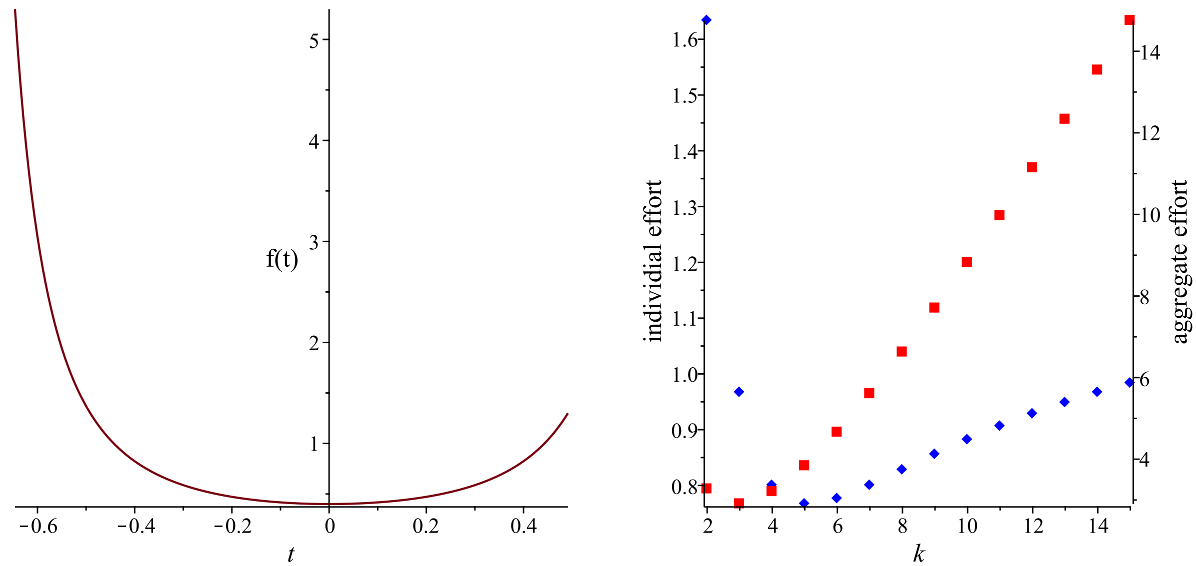

Figure 2: *Left*: The pdf $f(t)$ of a distribution with cdf $F(t) = 0.2\tan(2t) + 0.7$ defined on $[-0.646, 0.491]$. *Right*: Individual equilibrium effort e_k^* (blue diamonds, left scale) and aggregate equilibrium effort E_k^* (red squares, right scale) as a function of k for effort cost function $c(e)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

4.2 Aggregate equilibrium effort

Given the various possibilities for the dependence of individual equilibrium effort e_k^* on group size k, it is of interest to also explore how aggregate equilibrium effort $E_k^* = ke_k^*$ changes with the number of players. A natural question to ask is whether it can be established that E_k^* is unimodal for a unimodal $f(t)$. The answer is, in general, negative. Indeed, even when e_k^* is unimodal, the product of a strictly increasing and unimodal functions is not necessarily unimodal. Moreover, 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. Additional restrictions on $f(t)$ and/or $c(e)$ are needed to ensure the unimodality of E_k^*.

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 the aggregate equilibrium 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

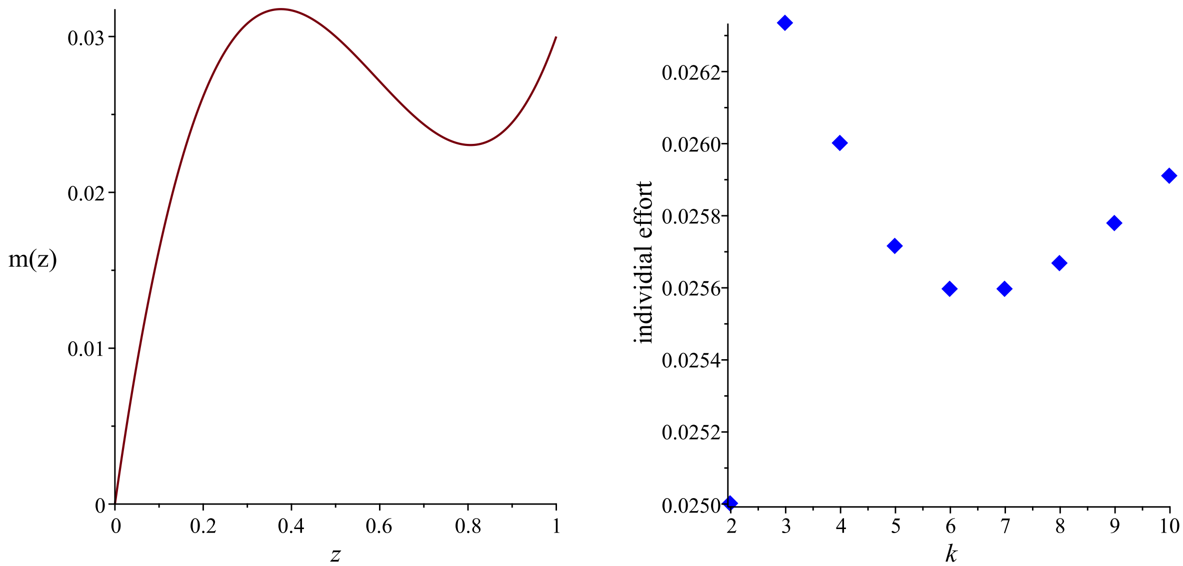

Figure 3: *Left*: $m(z) = 0.22z^3 - 0.39z^2 + 0.2z$. *Right*: Individual equilibrium effort e_k^* as a function of k for effort cost function $c(e) = \frac{1}{2}e^2$.

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.

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 the 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

$$\Delta E_k^* = (k-1)\left(\frac{b_{k-1}}{c_0\xi}\right)^{\frac{1}{\xi-1}}\left[\frac{k}{k-1}\left(\frac{b_k}{b_{k-1}}\right)^{\frac{1}{\xi-1}} - 1\right], \tag{8}$$

which immediately implies the following result.

Lemma 6 *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 6 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;[16] therefore,

[16] As ξ gets closer to 1, it becomes more difficult to satisfy the second-order condition for payoff

if $b_k < b_{k-1}$ for some k, it is always possible to find a $\xi > 1$ such that the condition of Lemma 6 does not hold and hence $E_k^* < E_{k-1}^*$.

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 $f(t)$ is symmetric or increasing. However, if $f(t)$ 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 $F(t) = t^\alpha$ and pdf $f(t) = \alpha t^{\alpha-1}$ on $[0,1]$, with $\alpha > \frac{1}{2}$.[17] 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., $f(t)$ 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 $f(t)$ and cdf $F(t)$, the failure (or hazard) rate is defined as $h(t) = \frac{f(t)}{1-F(t)}$. A distribution is characterized as having increasing failure rate (IFR) if $h(t)$ is increasing, and decreasing failure rate (DFR) if $h(t)$ is decreasing. IFR is implied by the log-concavity of pdf $f(t)$, while DFR is implied by the log-convexity of $f(t)$ provided $f(\overline{x}) = 0$. The exponential distribution, with $f(t) = \lambda \exp(-\lambda t)$, 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

$$b_k = \frac{1}{k}\int_0^1 \frac{m(z)}{1-z} dF^B(z; k-1, 2) = \frac{1}{k}\mathrm{E}(h_q(Z_{(k-1:k)})). \tag{9}$$

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 $h(t)$ 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. (9) 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

maximization at e_k^*, but for any given ξ it can always be satisfied for a sufficiently high c_0 and/or a sufficiently dispersed distribution of noise.

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

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 comdition, and hence Lemma 1 can be applied for unimodal and U-shaped failure rates. The results are summarized as follows.

Lemma 7 *Suppose the cost of effort is quadratic, $c(e) = c_0 e^2$. Then*
(i) If $f(t)$ is IFR (DRF) then E_k^* is increasing (decreasing) for $k \geq 2$.
(ii) If $f(t)$ is exponential, with $f(t) = \lambda \exp(-\lambda t)$, then $E_k^* = \frac{\lambda}{2c_0}$ is constant for $k \geq 2$.
(iii) If $f(t)$ 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 7, note that representation (9) 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 4.1. 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 $h(t) = \frac{f(t)}{1-F(t)} = f(t|X \geq t)$, i.e., the pdf of noise at $X = t$ conditional on $X \geq t$. 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. (9), which makes this representation suitable for aggregate effort as it conveniently subsumes the effect of multiplier k.

Part (i) of Lemma 7 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 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 the exponential. Hence, individual effort is decreasing (see Lemma 4(i)), and so fast that aggregate effort decreases too. For a simple example, consider the $F_{2,2}$-distribution whose pdf and cdf are $f(t) = \frac{1}{(1+t)^2}$ and $F(t) = \frac{t}{1+t}$ defined for $t \geq 0$. Then, $b_k = \frac{2}{k(k+1)}$ and, for $\xi = 2$, aggregate effort $E_k^* = \frac{2}{k+1}$ is strictly decreasing with the number of players.

For part (iii), examples of distributions with 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.

Lemma 6 shows that a "more convex" cost function is more likely to lead to aggregate effort increasing in the number of players. While in Lemma 6 this result is restricted to power cost functions, which are naturally ordered by parameters ξ, it is, in fact, very general. Specifically, Lemma 7 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 $x = c_2(e)$, obtain $u(x) = c_1(c_2^{-1}(x))$. 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 $\frac{f_Y(t)}{f_X(t)}$ is increasing in t. 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 $c(\sqrt{x})$ is convex in x 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. 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 $f(t)$ 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 7 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.

4.3 The effect of noise dispersion

In this section, we assume that the number of players, k, is fixed and focus on the effect of the distribution of noise on the individual equilibrium effort. Intuitively, when noise becomes more dispersed, the marginal gain from effort goes down and equilibrium effort should decrease. For example, when the distribution of noise is uniform on the interval $[-s, s]$, we have $b_k = \frac{1}{2s}$ for all $k \geq 2$; hence, the individual equilibrium effort decreases in s, a natural measure of dispersion for the uniform distribution. The variance of the uniform distribution is $\frac{s^2}{3}$, which implies that b_k (and e_k^*) decreases in the variance of noise as well. Similarly, in Tullock contests the dispersion of noise is determined by parameter r (see Section 2.2). As r goes down, noise becomes more dispersed and the equilibrium effort decreases. Here too, the variance of noise of the corresponding Gumbel distribution is $\frac{\pi^2}{6r^2}$, and b_k (and e_k^*) decreases in the variance.

Consider, however, a family of zero-mean, symmetrically distributed random variables X_α, parameterized by $\alpha \geq 0$, with pdfs $f(t, \alpha) = \frac{\alpha+1}{2}|t|^\alpha$ defined on support $[-1, 1]$. An increase in α leads to a higher variance, $\mathrm{Var}(X_\alpha) = \frac{\alpha+1}{\alpha+3}$, and, more generally, shifts the distribution in terms of second-order stochastic dominance (SOSD). At the same time, $b_2 = \frac{(\alpha+1)^2}{2(2\alpha+1)}$ increases with α. In other words, an increase in the level of *risk* associated with noise can lead to a higher equilibrium effort in a two-player tournament.

These examples show, perhaps surprisingly, that, in general, neither the variance nor SOSD ordering of noise distributions have a monotone effect on the equilibrium effort. To understand why this is the case, let X_1 and X_2 denote i.i.d. random variables with pdf f and recall that, from Eq. (3), $b_2 = \int_U f(t)^2 dt = f_{X_1-X_2}(0)$, where $f_{X_1-X_2}(\cdot)$ is the pdf of $X_1 - X_2$. In the example with variables X_α above, as α increases, the mass of the distribution is shifted away from the middle towards the edges of the support and, therefore, the density of $X_1 - X_2$ acquires a sharp peak at zero (and two additional, smaller peaks around -2 and $+2$) leading to an increase in b_2 even as the variance of X_α goes up.

For the rest of this section, we will use $b_k[f]$ and $e_k^*[f]$ to denote, respectively, the coefficient b_k and equilibrium effort e_k^* obtained from a noise distribution with pdf $f(t)$. As noted by Gerchak and He (2003), from Eq. (3), b_2 can be written in the form $b_2[f] = \int_U f(t)^2 dt = \exp(-H[f])$, where $H[f]$ is the Rényi entropy of order 2, also known as "collision entropy" (Rényi, 1961).[18] Thus, in two-player tournaments equilibrium effort

[18] The general expression for the Rényi entropy of order α is $H_\alpha[f] = \frac{1}{1-\alpha} \ln \left(\int_U f(t)^\alpha dt \right)$.

decreases in the *entropy* of the noise distribution. More generally, from Eq. (3),

$$b_k[f] = \frac{4(k-1)}{k^2} \int_U \left[\frac{k}{2} F(t)^{\frac{k}{2}-1} f(t) \right]^2 dt = \frac{4(k-1)}{k^2} b_2[f_{(k/2)}] = \frac{4(k-1)}{k^2} \exp(-H[f_{(k/2)}]), \tag{10}$$

where $f_{(k/2)}(t) = \frac{d}{dt} F(t)^{\frac{k}{2}}$ is the pdf corresponding to cdf $F_{(k/2)}(t) = F(t)^{\frac{k}{2}}$. Thus, coefficient b_k in a tournament of $k \geq 2$ players can be represented as an appropriately rescaled coefficient b_2 in a tournament of two symmetric players each having the cdf of noise $F_{(k/2)}(t)$. The latter coefficient can then be expressed through the entropy of pdf $f_{(k/2)}$.

Proposition 4 *In a tournament of k players, equilibrium effort decreases in the Rényi entropy of order 2 of a distribution with pdf $f_{(k/2)}$.*

The representation (10) and Proposition 4 have a straightforward interpretation when k is even: Consider a tournament of two players where each player has access to $\frac{k}{2}$ independent draws from the original noise distribution and selects the highest draw.[19] Another, though less precise, interpretation is that the k players are split arbitrarily into two equal subgroups with $\frac{k}{2}$ players each. Then $F_{(k/2)}(t)$ is the cdf of noise of the two players whose shocks are the largest in each subgroup, and the player with a larger shock between these two wins the tournament.

When support $[\underline{x}, \overline{x}]$ is finite, the entropy reaches its maximum for the uniform distribution. Proposition 4, therefore, also leads to the following corollary.

Corollary 2 *Of all noise distributions with a finite support $[\underline{x}, \overline{x}]$, the distribution that minimizes the symmetric equilibrium effort in the tournament of $k \geq 2$ players has cdf $F_{\min}(t) = \left(\frac{t-\underline{x}}{\overline{x}-\underline{x}}\right)^{\frac{2}{k}}$. The resulting minimized value of b_k is $b_k[f_{\min}] = \frac{4(k-1)}{k^2(\overline{x}-\underline{x})}$.*

As seen from the corollary, the effort-minimizing noise distribution in a k-player tournament is uniform for $k = 2$, but for $k > 2$ it has a concave cdf and monotonically decreasing pdf, more so the larger the number of players k, such that $F_{\min}(t)^{\frac{k}{2}}$ is uniform.

The representation (10) and Proposition 4 are not very convenient because function $f_{(k/2)}$ depends on k, and hence it may in general be difficult to compare entropy $H[f_{(k/2)}]$ across distributions. The reason is that, in general, entropy ordering is not preserved by order statistics (cf. the discussion below Corollary 3). A sufficient condition that allows

[19]This is the case in some Olympic sports where participants have several attempts and choose the best result, such as discus throw, shot put, javelin throw, long jump, triple jump, etc.

one to directly rank the entropy of order statistics across distributions, and hence to rank equilibrium efforts, is given by the *dispersive order* (Lehmann, 1988).[20]

Definition 3 *X is more dispersed than Y if for all $z, z' \in [0,1]$ such that $z' > z$*

$$F_X^{-1}(z') - F_X^{-1}(z) \geq F_Y^{-1}(z') - F_Y^{-1}(z).$$

and the inequality is strict in some open interval of z.

The definition is rather intuitive: X is more dispersed than Y if the distance between any two quantiles of X is at least as large as the distance between the same quantiles of Y. As discussed by Shaked and Shanthikumar (2007), whenever X is more dispersed than Y, $\mathrm{Var}(X) \geq \mathrm{Var}(Y)$; the converse, however, is not true. Similarly, the dispersive order for variables with equal means implies SOSD, but the converse is not true. Finally, whenever X is more dispersed than Y, it has a higher entropy. Moreover, the dispersive order is preserved for order statistics (Theorem 3.B.26 in Shaked and Shanthikumar, 2007), leading to the following result.

Lemma 8 *If X is more dispersed than Y then $H[f_{X(k/2)}] > H[f_{Y(k/2)}]$, and hence $e_k^*[f_X] < e_k^*[f_Y]$ for any $k \geq 2$.*

The proof of Lemma 8 is straightforward and based on Proposition 4 and the fact that X being more dispersed than Y is equivalent to $m_X(z) \leq m_Y(z)$ (see Appendix A).

An important special case which satisfies the dispersive order, allows for an explicit characterization of the equilibrium effort, and incorporates several important examples is when additional dispersion is generated by scaling: $X = sY$, where $s > 1$. A parameterized cdf $F(t,s)$ is said to have a scale parameter s if it satisfies $F(t,s) = F(\frac{t}{s}, 1)$. The corresponding scaled pdf is $f(t,s) = \frac{1}{s} f(\frac{t}{s}, 1)$. For example, the standard deviation of a zero-mean normal distribution, the length of the support of a uniform distribution, the expected value of an exponential distribution and the scale of the Gumbel distribution (and hence $\frac{1}{r}$, where r is the "discriminatory power" of the Tullock contest, see Section 2.2) are scale parameters. It is easy to see that an increase in s leads to a more dispersed distribution (Theorem 3.B.4 in Shaked and Shanthikumar, 2007) and hence to a lower equilibrium effort (Lemma 8). For an explicit characterization, note that the scale

[20]For recent applications of the dispersive order in the auction theory literature see, e.g., Ganuza and Penalva (2010) and Kirkegaard (2012).

transformation implies $F^{-1}(z,s) = sF^{-1}(z,1)$ and hence

$$m(z,s) = f(F^{-1}(z,s),s) = \frac{1}{s} f\left(\frac{sF^{-1}(z,1)}{s}, 1\right) = \frac{1}{s} m(z,1),$$

which immediately gives $b_k[f(t,s)] = \frac{1}{s} b_k[f(t,1)]$. Therefore, the individual and aggregate equilibrium efforts are decreasing in s.

Despite its generality, in many cases of interest the dispersive order does not rank distributions, and other methods of comparing the entropy of order statistics need to be developed. For example, a mean-preserving spread generated by adding an independent zero-mean random variable satisfies the dispersive order only under a special condition. Specifically, suppose $X = Y + W$, where $\mathrm{E}(W) = 0$ and W is independent of Y. In this case X is more dispersed than Y for any W (and hence Lemma 8 applies) if and only if the pdf of Y is log-concave (Theorem 3.B.7 in Shaked and Shanthikumar, 2007).

An important case when two (different) distributions cannot be ranked in the sense of dispersive order is when they have the same finite support (Theorem 3.B.14. in Shaked and Shanthikumar, 2007). The following lemma, due to Gerchak and He (2003) (also proved in Appendix A for completeness), provides a useful starting point for ranking of some distributions directly in terms of the entropy.

Lemma 9 *Consider random variables X and Y. Then $H[f_X] \geq H[f_Y]$ if*
(i) $f_X(t)$ and $f_Y(t)$ are increasing and Y FOSD X; or
(ii) $f_X(t)$ and $f_Y(t)$ are decreasing and X FOSD Y.

Condition (i) in Lemma 9 is satisfied, for example, when f_X and f_Y are both increasing and f_X crosses f_Y from above; that is, there exists a $\hat{t} \in [\underline{x}, \overline{x}]$ such that $f_X(t) \geq (\leq) f_Y(t)$ for $t \leq (\geq) \hat{t}$. Multiple crossings are also admissible as long as the FOSD relationships hold. Intuitively, $f_Y(t)$ is "steeper" than $f_X(t)$ and hence, it departs further away from the uniform distribution (which maximizes the entropy); therefore, it has a lower entropy. Condition (ii) has an similar intuitive explanation. Also note that the entropy is invariant with respect to a reflection of the probability distribution function around a vertical axis. Hence, condition (ii) is equivalent to condition (i) for reflected distributions.

Consider now another relationship between two distributions that we call *spread*.

Definition 4 *f_X is a spread of f_Y if*
(a) f_X and f_Y are unimodal;
(b) there exist $t_1 < t_2$ such that f_X crosses f_Y from above at t_1, then from below at t_2;
(c) the modes of both f_X and f_Y are between t_1 and t_2.

Condition (b) implies that the cdfs F_X and F_Y cross once. Hence, the spread is a special case of single-crossing cdfs that have been studied by Diamond and Stiglitz (1974), Hammond (1974) and Johnson and Myatt (2006), among others. While it may seem intuitive that the spread implies the ranking of entropies, further restrictions are needed for this to be the case in general. The following Lemma provides a sufficient condition.[21] Denote by c the intersection point of the two cdfs: $F_X(c) = F_Y(c)$. It is easy to see that $t_1 \leq c \leq t_2$.

Lemma 10 *Suppose that f_X is a spread of f_Y and*

$$f_X(c) + f_Y(c) \geq 2\max\{f_X(t_1), f_X(t_2)\}. \tag{11}$$

Then, $H[f_X] \geq H[f_Y]$.

To understand the intuition for condition (11) note that f_X and f_Y are FOSD-ranked on each side of c. However, unless both modes coincide with c, each of the two pdfs is not monotone on one side of c, and hence Lemma 9 does not apply. Condition (11) effectively guarantees that the nonmonotone parts do not contribute enough to reverse the entropy ordering. It is satisfied in several easy-to-check situations such as when f_X has its mode at c or when f_X and f_Y intersect at the same level, $f_X(t_1) = f_X(t_2)$. In the special case when both f_X and f_Y have modes at c, f_X and f_Y are monotone on each side of c and FOSD-ranked. Hence, Lemma 9 applies on each side of c implying the same entropy ranking overall for any number of crossings of f_X and f_Y.[22] For symmetric distributions this implies that the *peakedness order* (X is smaller than Y in the peakedness order if $|Y - c|$ FOSD $|X - c|$, Birnbaum (1948)) leads to the entropy ranking.

Generally, the ranking of entropies by Lemmas 9 and 10 only implies the ranking of efforts in two-player tournaments. However, since first-order stochastic dominance is preserved by order statistics, if condition (i) in Lemma 9 is satisfied for f_X and f_Y, the same condition is satisfied for $f_{X(k/2)}$ and $f_{Y(k/2)}$ for any $k \geq 2$. Indeed, if f_X and f_Y are both increasing then the cdfs F_X and F_Y are convex; therefore, the cdfs $F_{X(k/2)}(t)$ and $F_{Y(k/2)}(t)$ are convex as well, implying that the corresponding pdfs, $f_{X(k/2)}(t)$ and $f_{Y(k/2)}(t)$, still satisfy condition (i). This leads to the following result.

Corollary 3 *If f_X and f_Y satisfy condition (i) in Lemma 9 then $H[f_{X(k/2)}] \geq H[f_{Y(k/2)}]$ and hence $e^*_k[f_X] \leq e^*_k[f_Y]$ for any $k \geq 2$.*

[21] A counterexample showing that Lemma 10 does not hold without condition (11) is available from the authors upon request.

[22] This was shown first by Gerchak and He (2003).

Note that the same cannot be said about condition (ii) of Lemma 9 where F_X and F_Y are concave but $F_{X(k/2)}(t)$ and $F_{Y(k/2)}(t)$ can become convex for a sufficiently large k, and hence condition (ii) for $f_{X(k/2)}(t)$ and $f_{Y(k/2)}(t)$ will be reversed.

5 Tournaments with stochastic group size

5.1 Model setup

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 $\mathcal{K} \in 2^{\mathcal{N}}$ 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)

$$\tilde{p}_k = \Pr(\tilde{K} = k) = \frac{p_k k}{\bar{k}}, \quad k = 1, \ldots, n. \tag{12}$$

Equation (12) 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 (12).

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

$$\pi_i(e_i, e^*) = \sum_{k=1}^{n} \tilde{p}_k \int_U F(e_i - e^* + t)^{k-1} dF(t) - c(e_i). \tag{13}$$

The first-order condition for payoff maximization evaluated at $e_i = e^*$, $\frac{\partial \pi_i(e_i, e^*)}{\partial e_i}\Big|_{e_i = e^*} = 0$, gives

$$c'(e^*) = B_p \equiv \sum_{k=1}^{n} \tilde{p}_k (k-1) \int_U F(t)^{k-2} f(t) dF(t). \tag{14}$$

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

$$B_p = \sum_{k=1}^{n} \tilde{p}_k (k-1) \int_0^1 z^{k-2} m(z) dz = \int_0^1 m(z) d\tilde{G}(z). \tag{15}$$

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 (14), assuming that it exists and it is a symmetric pure strategy equilibrium.[23] When p is degenerate at some k, Eq. (14) 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. (15) and (12), and the definition of b_k, Eq. (3), coefficients B_p can also be written as

$$B_p = \sum_{k=1}^{n} \tilde{p}_k b_k = \mathrm{E}_{\tilde{p}}(b_K) = \frac{1}{\bar{k}} \sum_{k=2}^{n} p_k k b_k = \frac{1}{\bar{k}} \mathrm{E}_p(K b_K | K \geq 2) \mathrm{Pr}_p(K \geq 2). \tag{16}$$

Here, $\mathrm{E}_p(\cdot)$ and $\mathrm{Pr}_p(\cdot)$ denote expectation and probability with respect to distribution p. Note that the summation in (16) can start with $k = 2$ instead of $k = 1$ because $b_1 = 0$. Representation (16) 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

[23] As in Section 2, we leave the issues of equilibrium existence and uniqueness outside the scope of this paper.

noise is uniform. In this case, $b_k = b_2$ for any $k \geq 2$. Equation (15) then gives

$$B_p = b_2 \left(\tilde{G}(1) - \tilde{G}(0) \right) = b_2 \left(1 - \frac{p_1}{\bar{k}} \right), \tag{17}$$

leading to the following result.

Lemma 11 *Suppose F 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 11 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 than with deterministic group size, and is strictly lower if the probability for a player to be alone in the tournament is not zero. Indeed, if $p_1 = 0$, there are at least two players in the tournament (from the perspective of a player who has been selected), and the result follows because equilibrium effort is independent of tournament size for $k \geq 2$ when F is uniform (see Lemma 4(ii)).

5.2 Comparative statics for unimodal noise distributions

We are interested in the effects of changes in distribution p on coefficients B_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 $\tilde{P}(\theta)$, $\tilde{G}(z, \theta)$ and $B_p(\theta)$ denote, respectively, the corresponding cmf, pgf and B_p.

Suppose an increase in θ leads to a stochastic increase in the number of players in the sense of first-order stochastic dominance (FOSD); that is, assume that $\tilde{P}_k(\theta)$ is decreasing in θ for all $k = 1, 2, \ldots, n$. The simplest case that does not require any additional restrictions is when the sequence $\{b_k\}_{k=2}^n$ is increasing (which implies that $\{b_k\}_{k=1}^n$ is increasing because $b_1 = 0$). The following lemma and corollary follow immediately from (16) and Lemma 4.

Lemma 12 *Suppose an increase in θ leads to a stochastic increase in $\tilde{K}$ and $\{b_k\}_{k=2}^n$ is increasing. Then $B_p(\theta)$ (and e_p^*) is increasing in θ.*

Corollary 4 *Suppose an increase in θ leads to a stochastic increase in $\tilde{K}$ and $f(t)$ is increasing. Then $B_p(\theta)$ (and e_p^*) is increasing in θ.*

Note that a similar result cannot be established when $\{b_k\}_{k=2}^n$ is decreasing, because $b_1 = 0$ and hence $\{b_k\}_{k=1}^n$ would be nonmonotone, unless $p_1 = 0$. Tournaments with $p_1 = 0$, in which, from the perspective of a participating player, the number of players is known to be at least two, are rather common in applications. Indeed, organizers often include a provision that competition will be canceled if fewer than a pre-specified number of participants sign up. The following lemma is a direct consequence of Eq. (16).

Lemma 13 *Suppose an increase in θ leads to a stochastic increase in $\tilde{K}$, $p_1(\theta) = 0$ for all $\theta \in \Theta$, and $\{b_k\}_{k=2}^n$ is decreasing. Then $B_p(\theta)$ (and e_p^*) is decreasing in θ.*

It is then straightforward to relate the properties of the distribution of noise to the behavior of equilibrium effort.

Corollary 5 *Suppose an increase in θ leads to a stochastic increase in $\tilde{K}$ and $p_1(\theta) = 0$ for all $\theta \in \Theta$. Then,*
(i) if $f(t)$ is decreasing then e_p^* is decreasing in θ;
(ii) for $n \geq 4$, if $f(t)$ is interior unimodal and symmetric then e_p^* is decreasing in θ;
(iii) for $n = 3$, if $f(t)$ is symmetric then e_p^* is independent of θ.

Part (i) of the corollary follows from Lemma 4(i), while parts (ii) and (iii) follow from parts (ii) and (iii) of Proposition 2.

Lemma 13 has one other interesting implication. When $\{b_k\}_{k=2}^n$ is decreasing, *the only way e_p^** can be nonmonotone with respect to an upward probabilistic 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 $K \sim \text{Binomial}(n, q)$. Lemma 13 shows that this nonmonotonicity is a consequence entirely 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 will go 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.

We now turn to the general case and present our main result. Let $\tilde{G}_\theta(z, \theta) \leq 0$ denote the derivative or the first difference of the pgf with respect to θ. Combined with Proposition 2, Lemmas 2 and 3 produce the following.

Proposition 5 *Suppose an increase in θ leads to a stochastic increase in $\tilde{K}$ and*

(a) $f(t)$ 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 $B_p(\theta)$ (and e_p^*) is unimodal in θ.

The two distributions used most prominently in the literature to model population uncertainty – the Poisson and binomial distributions – satisfy the log-supermodularity condition (b) of Proposition 5. These distributions, along with the negative binomial and logarithmic distributions, belong to a family known as power series distributions (PSD). As we show in the following section, all PSD distributions satisfy condition (b).

5.3 Power series distributions of group size

Power series distributions (PSD) are characterized by pmfs of the form

$$p_k(\theta) = \frac{a_k \theta^k}{A(\theta)}, \tag{18}$$

where 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 5 is applicable to the whole PSD family due to the following three properties.

Proposition 6 *For any pmf p in the PSD family (18)*
(i) the updated pmf $\tilde{p}$ is also in the PSD family;
(ii) $G_\theta(z,\theta) \leq 0$;
(iii) $|G_\theta(z,\theta)|$ is log-supermodular.

Property (i) states that the PSD family is closed under the participation updating (12). 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 distribution are FOSD-ordered by parameter θ. Finally, property (iii) ensures that condition (b) of Proposition 5 is satisfied, and hence B_p (and e_p^*) is unimodal in θ for any PSD distribution provided $f(t)$ is unimodal.

5.4 Aggregate effort

In this section, we explore the effects of changes in the distribution of the number of players on expected aggregate effort $E_p^* = \bar{k}e_p^* = \bar{k}c'^{-1}(B_p(\theta))$. This problem simplifies substantially when the cost function is quadratic, $c(e) = c_0e^2$, in which case $E_p^* = \frac{\bar{k}}{2c_0}B_p(\theta)$. Using (16), 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, $E_k^* = \frac{kb_k}{2c_0}$ is the equilibrium aggregate effort in a tournament with deterministic size k, and the expectation is taken over the original group size distribution p. Lemma 7 and Proposition 3 then lead to the following results.

Proposition 7 *Suppose an increase in θ leads to an FOSD increase in the number of players K. Then*
(i) If $f(t)$ 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 $f(t)$ has a unimodal (U-shaped) failure rate, $|G_\theta(z,\theta)|$ is log-supermodular and $c(e) = c_0e^2$, 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 7 and Lemmas 2 and 3. 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 $k \geq 2$). From part (ii) of Lemma 7, for the exponential distribution with parameter λ we have $E_k^* = \frac{\lambda}{2c_0}$ for $k \geq 2$, which gives $E_p^* = \frac{\lambda}{2c_0}\sum_{k=2}^{n}p_k = \frac{\lambda}{2c_0}(1 - P_2(\theta))$, where $P_2(\theta)$ is decreasing in θ; thus, E_p^* is increasing in θ. For the uniform distribution, Eq. (17) 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 5, the condition for a stochastic increase in the number of players pertaining to Proposition 7 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.

5.5 The binomial distribution of group size

For illustration, we consider the binomial distribution of tournament size, with $K \sim \text{Binomial}(n, q)$, where $n \geq 2$ and $q \in [0, 1]$. The expected number of players is $\bar{k} = nq$ and the updated probability of group size k is

$$\tilde{p}_k = \frac{1}{nq}\binom{n}{k}q^k(1-q)^{n-k}k = \binom{n-1}{k-1}q^{k-1}(1-q)^{n-k};$$

that is, from the perspective of a participating player, the distribution of the number of *other players*, $\tilde{K} - 1$, is Binomial$(n-1, q)$. The pgf for the updated binomial distribution is

$$\tilde{G}(z, q) = \sum_{k=1}^{n}\binom{n-1}{k-1}q^{k-1}(1-q)^{n-k}z^{k-1} = (1-q+qz)^{n-1}. \tag{19}$$

Throughout this section, with a slight abuse of notation, we will write B_p as $B_n(q)$ and e_p^* as $e_n^*(q)$, in order to show explicitly the dependence on the parameters (n, q) of the binomial distribution. From Eqs. (15) and (19),

$$B_n(q) = (n-1)q\int_0^1 (1-q+qz)^{n-2}m(z)dz. \tag{20}$$

5.5.1 The effects of q and n

For a given n, an increase in q leads to an FOSD shift in the number of participants. It follows from (19) that $\tilde{G}_q(z, q) = -(n-1)(1-z)(1-q+qz)^{n-2} \leq 0$ and

$$R(z, q, q') = \frac{\tilde{G}_q(z, q')}{\tilde{G}_q(z, q)} = \left(\frac{1-q'(1-z)}{1-q(1-z)}\right)^{n-2}.$$

It is easy to see that $R(z, q, q')$ is increasing in z for any $q' > q$. Thus, $|\tilde{G}_q|$ is log-supermodular, and for a unimodal $f(t)$ all the assumptions of Proposition 5 are satisfied and $B_n(q)$ is unimodal in q.

Similarly, for a given q an increase in n leads to an FOSD shift in $\tilde{K}$. Equation (19) gives

$$\tilde{G}_n(z,n) = \tilde{G}(z,n+1) - \tilde{G}(z,n) = -q(1-z)(1-q+qz)^{n-1}$$

and $R(z,n,n') = (1-q+qz)^{n'-n}$, which is increasing in z for $n' > n$. Thus, for a unimodal $f(t)$, $B_n(q)$ is unimodal in n as well. As seen from Eq. (20), $B_n(q)$ is a polynomial in q; therefore, the unimodality implies that it is either monotonically increasing or has a unique interior maximum in q, which we denote q_n^*. These results confirm the findings of Lim and Matros (2009) for Tullock contests with the binomial distribution of group size. Indeed, such contests are equivalent to tournaments with the Gumbel distribution of noise, which is unimodal.

5.5.2 The Laplace distribution of noise

Consider the Laplace$(0,1)$ distribution of noise, whose pdf is $f(t) = \frac{1}{2}\exp(-|t|)$ and cdf is $F(t) = \frac{1}{2}\exp(t)$ for $t \le 0$ and $F(t) = 1 - \frac{1}{2}\exp(-t)$ for $t \ge 0$. Equation (20) then gives

$$B_n(q) = \frac{(1-q)^n - 2\left(1-\frac{q}{2}\right)^n + 1}{nq}. \tag{21}$$

Coefficients $b_k = \frac{1}{k}\left(1 - \frac{1}{2^{k-1}}\right)$ are decreasing for $k \ge 3$, with $b_2 = b_3$. Indeed, since the Laplace distribution is symmetric and unimodal, Proposition 2(iv) applies. Proposition 5 also applies, and $B_n(q)$ (and hence $e_n^*(q)$) is unimodal in q and n, as illustrated in the left panel of Figure 4.

We conclude this section by an example showing that, similar to the conditions of Proposition 2, the unimodality of $f(t)$ in Proposition 5 is a tight condition. Consider again the bimodal distribution shown in Figure 2, which produces a non-unimodal sequence $\{b_k\}$. This distribution generates a non-unimodal dependence of B_p (and e_p^*) on q shown in the right panel of Figure 4.[24]

5.5.3 Aggregate effort

The effect of q

In this section we explore how *aggregate* expected equilibrium effort $E_n^*(q) = nqe_n^*(q)$ changes with q. For Tullock contests, Lim and Matros (2009) showed that the aggregate

[24] Similar to Section 4.1, a bimodal distribution is not sufficient to generate a non-unimodal dependence of $B_n(q)$ on q. For example, the bimodal distribution with pdf $f(t) = \frac{1}{2}[f_{N(-12,4)}(t) + f_{N(12,4)}(t)]$ generates $B_n(q)$ which is strictly increasing in q for any n.

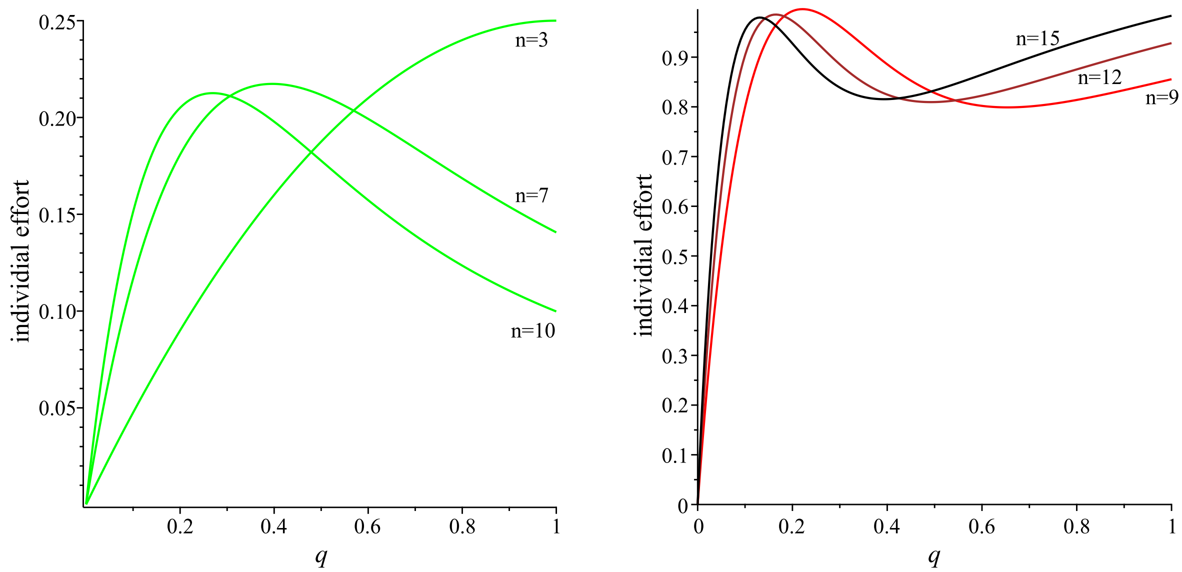

Figure 4: Individual effort as a function of q for different values of n for the binomial distribution of the number of players with parameters (n, q) and cost function $c(e) = \frac{1}{2}e^2$. *Left*: Noise is distributed according to the Laplace$(0,1)$ distribution. *Right*: Noise is distributed according to a distribution with cdf $F(t) = 0.2\tan(2t) + 0.7$ on $[-0.646, 0.491]$ (see Figure 2).

effort always increases in q. In our case, the situation is more complex due to the nonlinearity of the cost of effort and the effects of the shape of the distribution of noise. As we show, $E_n^*(q)$ may or may not be monotonically increasing in q, depending on parameters.

For the remainder of this section, we will restrict attention to the cases when $B_n(q)$ is unimodal in q, i.e., it is either increasing or there exists a $q_n^* \in (0,1)$ such that $B_n(q)$ is increasing for $q \le q_n^*$ and decreasing for $q \ge q_n^*$. A sufficient, but not necessary, condition for this is the unimodality of $f(t)$, as described in Proposition 5.

Differentiating both sides of the first-order condition $B_n(q) = c'(e_n^*(q))$, obtain

$$\frac{\partial E_n^*(q)}{\partial q} = n\left[e_n^*(q) + q\frac{\partial e_n^*(q)}{\partial q}\right] = \frac{nc'(e_n^*(q))}{c''(e_n^*(q))}\left[\eta(e_n^*(q)) + \beta_n(q)\right], \tag{22}$$

$$\eta(e) = \frac{c''(e)e}{c'(e)}, \qquad \beta_n(q) = \frac{B_n'(q)q}{B_n(q)}.$$

Thus, the sign of the derivative $\frac{\partial E_n^*(q)}{\partial q}$ is determined by the sum of two elasticities: $\eta(e_n^*(q))$ is the effort elasticity of the marginal cost $c'(e)$ in equilibrium, and $\beta_n(q)$ is the elasticity of $B_n(q)$ with respect to q. The former elasticity is always positive; the latter one is also positive when $B_n(q)$ is increasing for all q. Thus, the nontrivial case left to be considered is the one where $B_n(q)$ has an interior maximum q_n^* and, therefore, $\beta_n(q)$ is negative for

$q > q_n^*$.

It is difficult to proceed with the analysis of Eq. (22) in general since $\eta(e)$ may be an arbitrary function of effort. Therefore, from this point on we will consider the special case of constant elasticity, i.e., restrict attention to cost functions of the form $c(e) = c_0 e^{\xi}$, $\xi > 1$. In this case $\eta(e) = \xi - 1$ is independent of e and can be made arbitrarily small. It should, therefore, be possible to have $\xi - 1 + \beta_n(q) < 0$ for some q and a nonmonotonic dependence of aggregate effort on q. In other words, when individual effort is nonmonotonic in q, ξ should be large enough for the aggregate effort to be monotonically increasing in q. Formally, note that $\beta_n(q)$ is a continuous function on $[q_n^*, 1]$ and hence $\min_{q \in [q_n^*, 1]} \beta_n(q)$ is well-defined. This leads to the following proposition.

Proposition 8 *Suppose $c(e) = c_0 e^{\xi}$, $\xi > 1$. Then $E_n^*(q)$ is monotonically increasing in q if and only if*

$$\xi \geq 1 - \min_{q \in [q_n^*, 1]} \beta_n(q).$$

Under an additional restriction, the result of Proposition 8 can be written in terms of the primitives of the model. Recall that the nontrivial case arises only when $\beta_n(q)$ is negative for $q > q_n^*$. As shown in the following proposition, $\beta_n(q)$ is *decreasing* on $[q_n^*, 1]$, and hence the minimum of $\beta_n(q)$ is reached at $q = 1$, when the sequence $\{b_k\}_{k=2}^n$ is log-concave.

Definition 5 *A sequence of numbers $\{a_k\}_{k=1}^n$ is log-concave if $a_{k+1}^2 \geq a_k a_{k+2}$ for all $k = 1, \ldots, n-2$.*

Proposition 9 *Suppose $c(e) = c_0 e^{\xi}$, $\xi > 1$, and $B_n(q)$ has an interior maximum q_n^*. Suppose also that the sequence $\{b_k\}_{k=2}^n$ is log-concave. Then $E_n^*(q)$ is monotonically increasing in q if and only if*

$$\xi \geq (n-1)\frac{b_{n-1}}{b_n} - (n-2).$$

Otherwise, $E_n^*(q)$ has a unique interior maximum $Q_n^* \in (q_n^*, 1)$.

The effect of n

We now turn to the analysis of the dependence of aggregate effort $E_n^*(q)$ on n. As above, we restrict attention to the cases when $e_n^*(q)$ is single-peaked in q. Considering a change from $n-1$ to n players, obtain for the relative change in aggregate effort,

$$\delta E_n^*(q) = \frac{E_n^*(q) - E_{n-1}^*(q)}{E_{n-1}^*(q)} = \frac{n e_n^*(q) - (n-1)e_{n-1}^*(q)}{(n-1)e_{n-1}^*(q)} = \frac{n}{n-1}\frac{e_n^*(q)}{e_{n-1}^*(q)} - 1.$$

Clearly, this expression is positive for $q > 0$ when $e_n^*(q) \geq e_{n-1}^*(q)$, i.e., for $q \in (0, q_n^*]$. Thus, if $e_n^*(q)$ is monotonically increasing in q then $E_n^*(q) \geq E_{n-1}^*(q)$ for all q. The nontrivial case is when $e_n^*(q)$ has an interior maximum and $q > q_n^*$, i.e., $e_n^*(q) < e_{n-1}^*(q)$. In this case it is possible that $E_n^*(q) < E_{n-1}^*(q)$ for some $q > q_n^*$. As in the analysis above, it is difficult to obtain further results for a general effort cost function. Assuming $c(e) = c_0 e^\xi$, $\xi > 1$, we can write, from Eq. (14), $e_n^*(q) = \left(\frac{B_n(q)}{c_0\xi}\right)^{\frac{1}{\xi-1}}$ and

$$\frac{E_n^*(q) - E_{n-1}^*(q)}{E_{n-1}^*(q)} = \frac{n}{n-1}\left(\frac{B_n(q)}{B_{n-1}(q)}\right)^{\frac{1}{\xi-1}} - 1.$$

The function $\frac{B_n(q)}{B_{n-1}(q)}$ is continuous in q in $[q_n^*, 1]$, and hence its minimum on this interval is well-defined. This leads to the following proposition.

Proposition 10 *Suppose $c(e) = c_0 e^\xi$, $\xi > 1$. Then $E_n^*(q) \geq E_{n-1}^*(q)$ for all q if and only if*

$$\min_{q\in[q_n^*,1]} \frac{B_n(q)}{B_{n-1}(q)} \geq \left(\frac{n-1}{n}\right)^{\xi-1}. \tag{23}$$

It follows from Proposition 10 that for any n it is possible to find ξ high enough for the aggregate effort to increase in n. Indeed, the right-hand side of the inequality (23) can be made arbitrarily small by increasing ξ.

Similar to Proposition 9, condition (23) can be written more explicitly if $\frac{B_n(q)}{B_{n-1}(q)}$ is decreasing in q for $q \in [q_n^*, 1]$. It turns out that this condition is equivalent to $\beta_n(q)$ being a decreasing function of q and hence the log-concavity of $\{b_k\}_{k=2}^n$ is sufficient for it as well (cf. the proof of Lemma 15 in the Appendix).

Proposition 11 *Suppose $c(e) = c_0 e^\xi$, $\xi > 1$, and $B_n(q)$ has an interior maximum q_n^*. Suppose also that the sequence $\{b_k\}_{k=2}^n$ is log-concave. Then $E_n^*(q) \geq E_{n-1}^*(q)$ for all q if and only if*

$$\frac{b_n}{b_{n-1}} \geq \left(\frac{n-1}{n}\right)^{\xi-1}. \tag{24}$$

Note that inequality (24) coincides with condition (**??**) in Proposition **??** for the deterministic case. The two results are different, however. Given the binomial distribution of group size, condition (24) is necessary and sufficient for $E_n^*(q) \geq E_{n-1}^*(q)$ to hold for any q under the additional assumption of log-concavity of $\{b_k\}_{k=2}^n$, whereas condition (**??**) is necessary and sufficient for $E_n^*(1) \geq E_{n-1}^*(1)$ only, and does not require any additional assumptions.

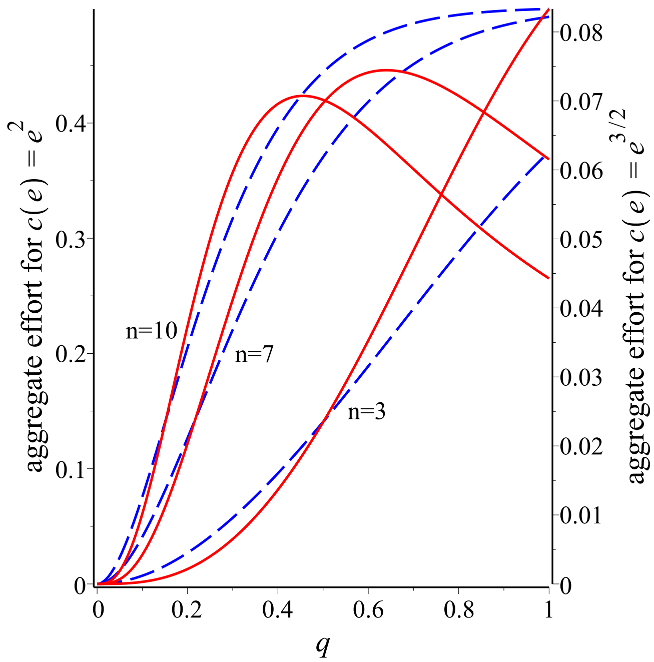

Figure 5: Aggregate effort as a function of q for cost functions $c(e) = e^2$ (dashed blue line, left scale) and $c(e) = e^{3/2}$ (solid red line, right scale), for noise distributed as Laplace$(0, 1)$.

For an illustration, continue with the example of the Laplace distribution considered in Section 5.5.2. Aggregate effort $E_n^*(q) = nqe_n^*(q)$ is monotonically increasing in q for $n = 3$ since individual effort $e_3^*(q)$ is monotonically increasing in q, due to the symmetry of the distribution. For $n \geq 4$, the behavior of $E_n^*(q)$ depends on the cost function. It can be shown that the minimum value of $\beta_n(q)$, the elasticity of $B_n(q)$ with respect to q, is above -1 and converges to this value when $q = 1$ and $n \to \infty$. Hence, if the cost function is $c(e) = c_0 e^{\xi}$, the aggregate effort is monotonically increasing in q if and only if $\xi \geq 2$ (see Proposition 8). This is illustrated in Figure 5.

5.6 The effect of noise dispersion

Similar to Section 4.3, suppose the distribution of group sizes, p, is fixed and consider the effect of changes in the dispersion of noise on the equilibrium effort. Throughout this section, we will use $B_p[f]$ and $e_p^*[f]$ to denote, respectively, coefficient B_p and the equilibrium effort e_p^* corresponding to the distribution of noise with pdf $f(t)$. Let $\tilde{g}(z) = \tilde{G}_z(z)$ denote the derivative of the pgf $\tilde{G}$ with respect to z. Changing the variable of integration to $z = F(t)$, rewrite (15) in the form

$$B_p[f] = \int_0^1 m(z)\tilde{g}(z)dz = \int_U \tilde{g}(F(t))f(t)^2 dt. \tag{25}$$

Consider a pdf $f_p(t)$ (with support U) defined as follows:

$$f_p(t) = \frac{1}{c_p} f(t)\sqrt{\tilde{g}(F(t))}, \quad c_p = \int_U f(t)\sqrt{\tilde{g}(F(t))}dt = \int_0^1 \sqrt{\tilde{g}(z)}dz, \tag{26}$$

where the normalization constant c_p is independent of f. Then Eq. (25) can be written in the form

$$B_p[f] = c_p^2 \int_U f_p(t)^2 dt = c_p^2 \exp(-H[f_p]), \tag{27}$$

where $H[\cdot]$ is the Rényi entropy. We arrive at the following results.

Proposition 12 *(i) In tournaments with stochastic participation, the equilibrium effort decreases in the Rényi entropy of a distribution with pdf* f_p.

(ii) Of all noise distributions with a finite support $[\underline{x}, \overline{x}]$*, the equilibrium effort is minimized by the distribution such that* $f_p(t) = \frac{1}{\overline{x}-\underline{x}}$*; that is, cdf* $F_{\min}$ *satisfies the differential equation*

$$F'(t) = \frac{c_p}{(\overline{x}-\underline{x})\sqrt{\tilde{g}(F(t))}}. \tag{28}$$

The minimized value of B_p *is* $B_p[f_{\min}] = \frac{c_p^2}{\overline{x}-\underline{x}}$.

It is easy to see that the results for deterministic participation can be recovered as a special case for a degenerate p. The right-hand side of Eq. (28) decreases in t; hence, similar to the deterministic participation case, the effort-minimizing cdf is concave, with a monotonically decreasing pdf.

For illustration, consider $K \sim \text{Binomial}(n, q)$. From (19), $\tilde{g}(z) = (n-1)q(1-q+qz)^{n-2}$, $c_p = \sqrt{\frac{4(n-1)}{qn^2}}[1-(1-q)^{\frac{n}{2}}]$, and

$$f_p(t) = \frac{nqf(t)[1-q+qF(t)]^{\frac{n}{2}-1}}{2[1-(1-q)^{\frac{n}{2}}]}.$$

The equilibrium effort is minimized when $f_p(t)$ is uniform on $[\underline{x}, \overline{x}]$, and the minimized value of B_p is $B_p[f_{\min}] = \frac{4(n-1)[1-(1-q)^{\frac{n}{2}}]^2}{qn^2(\overline{x}-\underline{x})}$.

Note that $\tilde{g}(z)$ is independent of the shape of the distribution of noise. Representation (25) then immediately implies that if X is more dispersed than Y then $B_p[f_X] \leq B_p[f_Y]$;

thus, the dispersive order of noise distributions has the same effect on the equilibrium effort as in the deterministic participation case (cf. Lemma 8).

Lemma 14 *If X is more dispersed than Y then $e_p^*[f_X] \leq e_p^*[f_Y]$.*

5.7 A comparison between stochastic and deterministic participation

It may be of interest to compare expected aggregate effort in a tournament with stochastic participation, $E_p^* = \bar{k} e_p^*$, to aggregate effort in the tournament with deterministic participation of size $\bar{k}$, $E_{\bar{k}}^* = \bar{k} e_{\bar{k}}^*$. The results are summarized in the following proposition.

Proposition 13 *(i) Suppose $\bar{k} = \sum_{k=0}^{n} k p_k$ is integer. Suppose also that $p_0 = 0$ and for all $k \geq 1$ in the support of p $k b_k$ is concave. Then $E_p^* \leq E_{\bar{k}}^*$; moreover, the inequality is strict if $k b_k$ is strictly concave.*

(ii) Suppose $\bar{k} \geq 2$ is integer. Suppose also that for all $k \geq 2$ in the support of p (a) $k b_k$ is concave and (b) b_k is decreasing. Then $E_p^* \leq E_{\bar{k}}^*$; moreover, the inequality is strict if $k b_k$ is strictly concave or $p_1 > 0$.

The comparison between aggregate efforts E_p^* and $E_{\bar{k}}^*$ for a given $\bar{k}$ is equivalent to the comparison of individual efforts e_p^* and $e_{\bar{k}}^*$. The general intuition behind Proposition 13 is that B_p, which determines e_p^*, is proportional to the expectation of $K b_K$, cf. Eq. (16), and the concavity of $k b_k$ gives the result by Jensen's inequality. However, since this expectation is conditional and also divided by the expected number of players $\bar{k}$, additional qualifiers are needed. For part (i), note that $\bar{k} = \mathrm{E}_p(K)$ is the unconditional expectation of K while B_p is proportional to the expectation of $K b_K$ conditional on $K \geq 1$. By setting $p_0 = 0$, this conditional expectation becomes unconditional and Jensen's inequality gives the result. For part (ii), as seen from (16), B_p can also be written as proportional to the expectation of $K b_K$ conditional on $K \geq 2$; while the expectation of K conditional on $K \geq 2$ is always (weakly) greater than the unconditional expectation of K. Then, the result is obtained using Jensen's inequality for conditional expectations (for concave $k b_k$) and the assumption that b_k is decreasing for $k \geq 2$. Part (i) of Proposition 13 generalizes the result of Myerson and Wärneryd (2006) who studied generalized Tullock contests with an arbitrary distribution of group size (subject to the restriction $p_0 = 0$). Part (ii) generalizes the result of Lim and Matros (2009) who analyzed Tullock contests with $K \sim \text{Binomial}(n, q)$.

For examples of violations of the conditions of Proposition 13, when stochastic participation can lead to a higher expected aggregate effort, consider the binomial distribution of tournament size, $K \sim \text{Binomial}(n, q)$. Let q_{opt} denote the *optimal* participation probability, that is, the probability q that maximizes expected aggregate effort $E_p^* = \bar{k} e_p^*$ subject to the constraint $\bar{k} = nq$. The deterministic contest generates a higher aggregate effort if $q_{\text{opt}} = 1$. The binomial distribution violates the conditions of part (i) of Proposition 13 since $p_0 = (1-q)^n > 0$. Also, for the bimodal distribution of noise in Figure 2 both assumptions (a) and (b) of part (ii) do not hold. Then, $q_{\text{opt}} \approx 0.9$ for $\bar{k} = 3$ and $q_{\text{opt}} \to 0$ (that is, a tournament with $n \to \infty$ potential players, each with zero probability of participation, is optimal) for $\bar{k} \geq 4$. For the $F_{2,2}$-distribution of noise (see the end of Section 4.2) assumption (a) of part (ii) is violated, and $q_{\text{opt}} \in (0, 1)$ for $3 \leq \bar{k} \leq 5$ while $q_{\text{opt}} \to 0$ for $\bar{k} \geq 6$.

Part (i) of Proposition 13 can be further generalized by considering two group size distributions, p and p', that have the same mean but are SOSD ranked. As shown in the following proposition, the aggregate equilibrium effort is ranked accordingly.

Proposition 14 *Consider two group size distributions, p and p', with the same integer mean $\bar{k}$ and $p_0 = p'_0 = 0$, such that p' SOSD p. Suppose also that for all $k \geq 1$ in the union of supports of p and p' kb_k is concave. Then $E_{p'}^* \geq E_p^*$; moreover, the inequality is strict if kb_k is strictly concave.*

5.8 Optimal disclosure of the number of players

Several authors investigated optimal disclosure policies under uncertainty, asking whether it makes sense for a principal whose goal is the maximization of aggregate effort, to (commit to) disclose to players how many participants there are in the tournament. 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 lottery-form contests with CSFs of the form $\frac{h(e_i)}{\sum_{j=1}^{k} h(e_j)}$. They show that full disclosure (no disclosure) is optimal if $\frac{h(e)}{h'(e)}$ is strictly convex (concave), while the indifference is recovered when $\frac{h(e)}{h'(e)}$ is linear.[25] In this section, we generalize these results to arbitrary tournaments and arbitrary distributions of the number of players.

[25] In asymmetric settings, the consequences of disclosure/nondisclosure become richer. For recent developments see, e.g., Denter, Morgan and Sisak (2014), Fu, Lu and Zhang (2016) and Zhang and Zhou (2016).

Without disclosure, the expected aggregate effort in the tournament is $E_p^* = \bar{k}e_p^* = \bar{k}c'^{-1}(B_p)$, where, from (16), $B_p = \mathrm{E}_{\tilde{p}}(b_K)$. 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 the optimality of disclosure depends entirely on the concavity/convexity of c'^{-1}, and not on the nature of coefficients b_k. One special case is when b_k is constant in the support of $\tilde{p}$ (for example, noise is uniformly distributed and $p_1 = 0$); in this case the two expressions are equal. 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). Note that the concavity (convexity) of c'^{-1} is equivalent to the convexity (concavity) of c', i.e., to the condition $c''' \geq (\leq) 0$.

Proposition 15 *Suppose b_k is non-constant for k in the support of $\tilde{p}$, and $c'(\cdot)$ is non-linear for at least some distinct values of b_k in the support of $\tilde{p}$. Then it is optimal to disclose (not disclose) the number of participants in the tournament if $c''' \leq (\geq) 0$.*

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 regrading the role of c''' hold in those settings as well.

6 Conclusion

In this paper we derive robust comparative statics results for WTA rank-order tournaments in which a player's effort is distorted by additive or multiplicative noise and the number of players is either deterministic or stochastic. The unimodality of the distribution of noise is critical for robust comparative statics, due to results on the preservation of unimodality under uncertainty. In the deterministic case, we show that the equilibrium effort is unimodal in the number of players when the distribution of noise is unimodal. In the stochastic case, the equilibrium effort is similarly unimodal in parameters shifting the distribution of the number of players in the sense of first-order stochastic dominance, albeit under an additional log-supermodularity restriction. The unimodality of the distribution of noise is a tight condition; we provide examples of non-unimodal noise distributions for which the comparative statics are no longer unimodal. For aggregate effort, the unimodlity of the failure rate of the distribution of noise plays a similar role when the cost of effort is quadratic. More generally, for effort costs that are more (less) convex than quadratic in the likelihood ratio order, we show that aggregate effort is increasing (decreasing) in the number of players whenthe distribution of noise is IFR (DFR).

The second dimension of our analysis is the effect of noise dispersion. We show that the equilibrium effort decreases in the appropriately defined Rényi entropy, as opposed to the often-cited variance or second-order stochastic dominance order. For the case of deterministic participation, it is the entropy of order statistics of the distribution of noise, while in the case of stochastic participation it is the entropy of a distribution that combines the distribution of noise with the distribution of tournament size. An important special case of entropy ordering that applies to both cases is the dispersive order of noise distributions.

References

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.

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.

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

Birnbaum, Z. W. 1948. "On Random Variables with Comparable Peakedness." *The Annals of Mathematical Statistics*, 19(1): 76–81.

Congleton, Roger D., Arye L. Hillman, and Kai A. Konrad. 2008. *40 Years of Research on Rent Seeking.* Vol. 1 and 2, Berlin:Springer.

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. 2007. "The theory of contests: a survey." *Review of Economic Design*, 11(2): 69–100.

Denter, Philipp, John Morgan, and Dana Sisak. 2014. "'Where ignorance is bliss,'tis folly to be wise': Transparency in contests." *Working paper.* https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=1836905.

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

Diamond, Peter A., and Joseph E. Stiglitz. 1974. "Increases in risk and in risk aversion." *Journal of Economic Theory*, 8(3): 337–360.

Drugov, Mikhail, and Dmitry Ryvkin. 2017. "Biased contests for symmetric players." *Games and Economic Behavior.*

Fu, Qiang, and Jingfeng Lu. 2010. "Contest design and optimal endogenous entry." *Economic Inquiry*, 48(1): 80–88.

Fu, Qiang, Jingfeng Lu, and Jun Zhang. 2016. "Disclosure policy in Tullock contests with asymmetric stochastic entry." *Canadian Journal of Economics*, 49(1): 52–75.

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

Ganuza, Juan-José, and José S. Penalva. 2010. "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions." *Econometrica*, 78(3): 1007–1030.

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.

Hammond, John S. 1974. "Simplifying the Choice between Uncertain Prospects Where Preference Is Nonlinear." *Management Science*, 20(7): 1047–1072.

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, Justin P., and David P. Myatt. 2006. "On the simple economics of advertising, marketing, and product design." *The American Economic Review*, 96(3): 756–784.

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.

Kirkegaard, René. 2012. "A Mechanism Design Approach to Ranking Asymmetric Auctions." *Econometrica*, 80(5): 2349–2364.

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.

Lehmann, Erich L. 1988. "Comparing location experiments." *The Annals of Statistics*, 16(2): 521–533.

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

Levin, Dan, and James L. Smith. 1994. "Equilibrium in auctions with entry." *American Economic Review*, 84(3): 585–599.

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

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.

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

McFadden, D.L. 1974. "Conditional logit analysis of qualitative choice behavior." *Frontiers in Econometrics*, , ed. P. Zarembka, 105–142. New York:Academic Press.

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

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.

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.

Pevnitskaya, Svetlana. 2004. "Endogenous entry in first-price private value auctions: The self-selection effect." *Unpublished Manuscript*.

Rényi, Alfred. 1961. "On Measures of Entropy and Information." *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 547–561. University of California Press.

Shaked, Moshe, and J. George Shanthikumar. 2007. *Stochastic orders.* Springer-Verlag New York.

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

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

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.

Zhang, Jun, and Junjie Zhou. 2016. "Information disclosure in contests: A Bayesian persuasion approach." *Economic Journal*, 126: 2197–2217.

A Proofs

Second-order condition Differentiating the payoff function (2) twice with respect to e_i and setting $e_i = e^*$, obtain $\frac{\partial^2 \pi_i(e_i, e^*)}{\partial e_i^2}\Big|_{e_i=e^*} = \eta_k - c''(e^*)$, where

$$\eta_k = (k-1)\left[(k-2)\int_U F(t)^{k-3} f(t)^2 dF(t) + \int_U F(t)^{k-2} f'(t) dF(t)\right].$$

Integrating the second term by parts, obtain

$$\eta_k = \frac{k-1}{2}\left[(k-2)\int_U F(t)^{k-3} f(t)^2 dF(t) + f(\overline{x})^2 - f(\underline{x})^2 I_{k=2}\right], \tag{29}$$

where $I_{k=2}$ is an indicator equal to one if $k = 2$ and zero otherwise. Thus, when $k = 2$ and the distribution of noise is symmetric the second-order condition is always satisfied. Otherwise, the restriction $\eta_k - c''(e^*) < 0$ has to be imposed.

For distributions admitting a scaling parameter s, such that $f(t, s) = \frac{1}{s} f(\frac{t}{s}, 1)$ and $F(t, s) = F(\frac{t}{s}, 1)$ (cf. Section 4.3), the second-order condition can be written explicitly as a restriction from below on the level of noise in the tournament. Let $\eta_k(s)$ denote coefficient η_k, Eq. (29), as a function of the scale parameter. Assuming the support of $f(t,1)$ is $U=[\underline{x},\overline{x}]$, Eq. (29) can be written as $\eta_k(s)=\frac{1}{s^2}\eta_k(1)$, and hence the second-order condition becomes $s^2>\frac{\eta_k(1)}{c''(e^*)}$.

Proof of Lemma 1 (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

$$\gamma(\theta)=u(1)-\int_0^1 u'(z)H(z,\theta)dz. \tag{30}$$

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 (30), obtain

$$\begin{aligned}\gamma'(\theta)&=-\int_0^{\hat{z}} u'(z)H_\theta(z,\theta)dz-\int_{\hat{z}}^1 u'(z)H_\theta(z,\theta)dz\\&=\int_0^{\hat{z}} u'(z)|H_\theta(z,\theta)|dz-\int_{\hat{z}}^1 |u'(z)||H_\theta(z,\theta)|dz.\end{aligned} \tag{31}$$

Suppose $\gamma'(\theta)\le 0$ for some θ and consider a $\theta'>\theta$. Then (31) gives

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

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 piece-wise 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 (31) 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

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

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 2 (i) Sufficiency: Rewrite $\chi(\theta)$ as follows:

$$
\begin{aligned}
\chi(\theta) &= p_1(\theta)x_1 + p_2(\theta)x_2 + \ldots + p_{n-1}(\theta)x_{n-1} + p_n(\theta)x_n \\
&= P_1(\theta)x_1 + (P_2(\theta) - P_1(\theta))x_2 + \ldots + (P_{n-1}(\theta) - P_{n-2}(\theta))x_{n-1} + (P_n(\theta) - P_{n-1}(\theta))x_n \\
&= x_n + P_1(\theta)(x_1 - x_2) + P_2(\theta)(x_2 - x_3) + \ldots + P_{n-1}(\theta)(x_{n-1} - x_n) \\
&= x_n - \sum_{k=1}^{n-1} P_k(\theta) \Delta x_{k+1},
\end{aligned}
$$

where $\Delta x_{k+1} = x_{k+1} - x_k$. 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 x_k. Taking the derivative, or the difference, with respect to θ, obtain

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

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

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

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

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

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 1, we will show that it is possible to construct a unimodal sequence x such that $\chi(\theta)$ is not unimodal. Set $x_l = a$ for all $l \leq k-1$ and $x_l = b$ for all $l \geq k+1$; furthermore, set $x_k > \max\{a, b\}$. The resulting sequence x is interior unimodal with mode k and satisfies $\Delta x_k > 0$, $\Delta x_{k+1} < 0$, and $\Delta x_l = 0$ for all $l \neq k, k+1$. Then

$$\chi'(\theta) = |P_{k-1}'(\theta)| \Delta x_k - |P_k'(\theta)| |\Delta x_{k+1}|.$$

Choosing a, x_k and b so that $\Delta x_k = |P_k'(\theta)| - \epsilon$ for some $\epsilon > 0$ and $|\Delta x_{k+1}| = |P_{k-1}'(\theta)|$, obtain $\chi'(\theta) = -\epsilon |P_{k-1}'(\theta)| < 0$. However,

$$\begin{aligned}
\chi'(\theta') &= |P_{k-1}'(\theta')| \Delta x_k - |P_k'(\theta')| |\Delta x_{k+1}| \\
&= r(k-1, \theta, \theta') |P_{k-1}'(\theta)| (|P_k'(\theta)| - \epsilon) - r(k, \theta, \theta') |P_k'(\theta)| |P_{k-1}'(\theta)| \\
&= (r(k-1, \theta, \theta') - r(k, \theta, \theta')) |P_k'(\theta)| |P_{k-1}'(\theta)| - \epsilon r(k-1, \theta, \theta') |P_{k-1}'(\theta)|.
\end{aligned}$$

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., $\chi(\theta)$ os not unimodal. ■

Proof of Lemma 3 (i) Sufficiency: By differentiating, or taking the first difference of, Eq. (6) 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 $\theta' > \theta$,

$$R(z,\theta,\theta') = \frac{|G_\theta(z,\theta')|}{|G_\theta(z,\theta)|} = \frac{\sum_{k=1}^{n} |P_k'(\theta')| z^{k-1}}{\sum_{k=1}^{n} |P_k'(\theta)| z^{k-1}} = \frac{\sum_{k=1}^{n} |P_k'(\theta)| r(k,\theta,\theta') z^{k-1}}{\sum_{k=1}^{n} |P_k'(\theta)| z^{k-1}}. \tag{32}$$

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 (32) 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

$$A_k'(z) = \frac{d}{dz}\left(\frac{\sum_{l=1}^{k} |P_l'(\theta)| z^{l-1}}{\sum_{l=1}^{n} |P_l'(\theta)| z^{l-1}}\right) = \frac{1}{(\sum_{l=1}^{n} |P_l'(\theta)| z^{l-1})^2} \sum_{l=1}^{k} \sum_{l'=1}^{n} |P_l'(\theta)||P_{l'}'(\theta)| z^{l+l'-3}(l-l')$$

$$= \frac{1}{(\sum_{l=1}^{n} |P_l'(\theta)| z^{l-1})^2} \sum_{l=1}^{k} \sum_{l'=k+1}^{n} |P_l'(\theta)||P_{l'}'(\theta)| z^{l+l'-3}(l-l') \le 0. \tag{33}$$

(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 2, 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 (33) 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. ■

Proof of the "only if" part of part (ii) of Lemma 4 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 $\mu(t) = \mathrm{E}(\exp(tZ))$, can be written in the form of expansion over moments, $\mu(t) = \sum_{k=0}^{\infty} \frac{m_k}{k!} t^k$, which gives

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

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 5 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 $\hat{z} = F^{-1}(\hat{t})$ 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

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

where M_1 and M_2 are positive constants (independent of k), the existence of which follows from the mean-value theorem for definite integrals. Evaluating the integrals, further obtain

$$b_k - b_{k+1} \geq \frac{1}{k}[M_1(1 - \hat{z}^k) - M_2\hat{z}^k] = \frac{1}{k}[M_1 - \hat{z}^k(M_1 + M_2)].$$

Since $\hat{z} < 1$, it is clear that the last expression becomes positive for a sufficiently large k. ■

Proof of Proposition 2 Define

$$\Delta b_{k+3} = b_{k+3} - b_{k+2} = \int_0^1 \left[(k+2)z^{k+1} - (k+1)z^k\right] m(z)dz, \quad k = 0, 1, \ldots, n-3. \tag{34}$$

Integrating by parts, obtain

$$\Delta b_{k+3} = \int_0^1 m(z)d(z^{k+2} - z^{k+1}) = \int_0^1 z^{k+1}(1-z)m'(z)dz. \tag{35}$$

For part (ii), the symmetry of $f(t)$ around its mean μ implies $f(t) = f(2\mu - t)$ and $F(t) = 1 - F(2\mu - t)$ for all $t \in U$. Letting $z = F(t) = 1 - F(2\mu - t)$, obtain $1 - z = F(2\mu - t)$, $F^{-1}(1-z) = 2\mu - t$ and $m(1-z) = f(F^{-1}(1-z)) = f(2\mu - t) = f(t) = f(F^{-1}(z)) = m(z)$. 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 7 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 1 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 $f(t)$ is IFR. Then, by Lemma 7, 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 $c(\sqrt{x})$ is convex in x, which implies

$$\frac{\partial^2}{\partial x^2}c(\sqrt{x}) = \frac{\partial}{\partial x}\left[\frac{c'(\sqrt{x})}{2\sqrt{x}}\right] = \frac{c''(\sqrt{x})\sqrt{x} - c'(\sqrt{x})}{4x^{3/2}} \geq 0,$$

i.e., $c''(e)e \geq c'(e)$. Therefore, $c''(e_k^*)e_k^* \geq c'(e_k^*) = b_k$ and

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

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

Proof of Lemma 8 Definition 3 is equivalent to the requirement that $F_X^{-1}(z) - F_Y^{-1}(z)$ is increasing in z. Differentiating with respect to z, obtain $\frac{1}{f_X(F_X^{-1}(z))} - \frac{1}{f_Y(F_Y^{-1}(z))} \geq 0$, or, using the definition of inverse quantile density, $m_X(z) \leq m_Y(z)$ (with a strict inequality in some open interval). Equation (4) then gives the result. ■

Proof of Lemma 9 For part (i), note that since f_X and f_Y are increasing and Y FOSD X, for any increasing function $u(t)$ we have $\int_{\underline{x}}^{\overline{x}} f_Y(t)u(t)dt \geq \int_{\underline{x}}^{\overline{x}} f_X(t)u(t)dt$. Using $u(t) = f_Y(t)$, obtain $\int_{\underline{x}}^{\overline{x}} f_Y(t)^2dt \geq \int_{\underline{x}}^{\overline{x}} f_X(t)f_Y(t)dt$; using $u(t) = f_X(t)$, obtain $\int_{\underline{x}}^{\overline{x}} f_Y(t)f_X(t)dt \geq \int_{\underline{x}}^{\overline{x}} f_X(t)^2dt$. Combining the two inequalities, obtain the result.

For part (ii), similarly, note that X FOSD Y and hence for any decreasing function $u(t)$ we have $\int_{\underline{x}}^{\overline{x}} f_X(t)u(t)dt \leq \int_{\underline{x}}^{\overline{x}} f_Y(t)u(t)dt$. Using $u(t) = f_Y(t)$ and $u(t) = f_X(t)$

consecutively, obtain the result. ■

Proof of Lemma 10 Let $[\underline{x}, \overline{x}]$ denote the union of supports of X and Y. Define

$$\Delta b_2 = \int_{\underline{x}}^{\overline{x}} [f_Y^2(t) - f_X^2(t)]dt = \int_{\underline{x}}^{\overline{x}} f_+(t) f_-(t)dt,$$

where $f_\pm(t) = f_Y(t) \pm f_X(t)$. Note that $f_-(t) \leq 0$ for $t \in [\underline{x}, t_1] \cup [t_2, \overline{x}]$ and $f_-(t) \geq 0$ for $t \in [t_1, t_2]$. Thus, we can write

$$\Delta b_2 = -\int_{\underline{x}}^{t_1} f_+(t)|f_-(t)|dt + \int_{t_1}^{c} f_+(t) f_-(t)dt + \int_{c}^{t_2} f_+(t) f_-(t)dt - \int_{t_2}^{\overline{x}} f_+(t)|f_-(t)|dt.$$

By the mean-value theorem for definite integrals, there exist $t_1^* \in (\underline{x}, t_1)$, $t_2^* \in (t_1, c)$, $t_3^* \in (c, t_2)$ and $t_4^* \in (t_2, \overline{x})$ such that

$$\Delta b_2 = -f_+(t_1^*) \int_{\underline{x}}^{t_1} |f_-(t)|dt + f_+(t_2^*) \int_{t_1}^{c} f_-(t)dt + f_+(t_3^*) \int_{c}^{t_2} f_-(t)dt - f_+(t_4^*) \int_{t_2}^{\overline{x}} |f_-(t)|dt.$$

Recall that $F_X(c) = F_Y(c)$, which implies $\int_{\underline{x}}^{c} f_X(t)dt = \int_{\underline{x}}^{c} f_Y(t)dt$, and hence $\int_{\underline{x}}^{c} f_-(t)dt = 0$ and $\int_{t_1}^{c} f_-(t)dt = \int_{\underline{x}}^{t_1} |f_-(t)|dt$. Similarly, $\int_{c}^{t_2} f_-(t)dt = \int_{t_2}^{\overline{x}} |f_-(t)|dt$. This gives

$$\Delta b_2 = [f_+(t_2^*) - f_+(t_1^*)] \int_{\underline{x}}^{t_1} |f_-(t)|dt + [f_+(t_3^*) - f_+(t_4^*)] \int_{t_2}^{\overline{x}} |f_-(t)|dt.$$

It follows from the condition $f_+(c) \geq 2\max\{f_X(t_1), f_X(t_2)\}$ that $f_+(c) \geq f_+(t_1)$, which implies $f_+(t_2^*) \geq f_+(t_1)$ and hence $f_+(t_2^*) \geq f_+(t_1^*)$. Similarly, $f_+(t_3^*) \geq f_+(t_4^*)$, which implies $\Delta b_2 \geq 0$. ■

Proof of Proposition 6 (i) From (12),

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

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 3. 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 2 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,

$$\begin{aligned}
N_{k+1}D_k - N_k D_{k+1} &= \left(N_k - \sum_{l=0}^{k} \beta^{l+k}\alpha_{l,k+1} + \sum_{m\geq k+2} \beta^{m+k}\alpha_{k+1,m} \right) D_k \\
&- N_k \left(D_k - \sum_{l=0}^{k} \alpha_{l,k+1} + \sum_{m\geq k+2} \alpha_{k+1,m} \right) \\
&= \sum_{l=0}^{k} \alpha_{l,k+1}(N_k - \beta^{l+k}D_k) + \sum_{m\geq k+2} \alpha_{k+1,m}(\beta^{m+k}D_k - N_k).
\end{aligned}$$

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.

$$\begin{aligned}
&\sum_{l=0}^{k} \alpha_{l,k+1}(N_k - \beta^{l+k}D_k) = \sum_{l=0}^{k} \sum_{m\geq k+1} \sum_{l'=0}^{k} \left(\beta^{l'+m-1}\alpha_{l'm}\alpha_{l,k+1} - \beta^{l+k}\alpha_{l'm}\alpha_{l,k+1} \right) \\
&= \sum_{l=0}^{k} \sum_{m\geq k+1} \sum_{l'=0}^{k} \left(\beta^{l+m-1}\alpha_{lm}\alpha_{l',k+1} - \beta^{l+k}\alpha_{l'm}\alpha_{l,k+1} \right) \\
&\geq \sum_{l=0}^{k} \sum_{m\geq k+1} \sum_{l'=0}^{k} \beta^{l+k} \left(\alpha_{lm}\alpha_{l',k+1} - \alpha_{l'm}\alpha_{l,k+1} \right) \\
&= \sum_{l=0}^{k} \sum_{m\geq k+1} \sum_{l'=0}^{k} \beta^{l+k} a_l a_m a_{l'} a_{k+1} \theta^{l+m-1+l'+k} \left[(m-l)(k+1-l') - (m-l')(k+1-l) \right] \\
&= \sum_{l=0}^{k} \sum_{m\geq k+1} \sum_{l'=0}^{k} \beta^{l+k} a_l a_m a_{l'} a_{k+1} \theta^{l+m-1+l'+k} (m-k-1)(l-l') \\
&= \sum_{m\geq k+1} \beta^k a_m a_{k+1} \theta^{m-1+k}(m-k-1) \sum_{l=0}^{k} \sum_{l'=0}^{k} \beta^l a_l a_{l'} \theta^{l+l'}(l-l').
\end{aligned}$$

The sum over l and l' can be rewritten as

$$\begin{aligned}
&\sum_{l=0}^{k} \sum_{l'=0}^{k} \beta^l a_l a_{l'} \theta^{l+l'}(l-l') = A_k(\theta)^2 A(\theta)^2 [\mathrm{E}(\beta^L L) - \mathrm{E}(\beta^L)\mathrm{E}(L)] \\
&= A_k(\theta)^2 A(\theta)^2 \mathrm{Cov}(\beta^L, L) \geq 0.
\end{aligned}$$

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 $\beta > 1$. ■

Lemma 15 *Suppose $B_n(q)$ has a unique interior maximum $q_n^* \in (0,1)$ and the sequence $\{b_k\} = \{B_k(1)\}$ is log-concave. Then function $\beta_n(q) = \frac{qB_n'(q)}{B_n(q)}$ is decreasing in q for $q \in (q_n^*, 1]$.*

Proof From 20, it is easy to see that $B_n'(q) = \frac{n-1}{q}[B_n(q) - B_{n-1}(q)]$, which gives

$$\beta_n(q) = \frac{(n-1)[B_n(q) - B_{n-1}(q)]}{B_n(q)} = (n-1)\left[1 - \frac{B_{n-1}(q)}{B_n(q)}\right].$$

Thus, in order to show that $\beta_n(q)$ is decreasing in q, we will show that $\frac{B_{n-1}(q)}{B_n(q)}$ is increasing in q. This is the case if and only if $B_{n-1}'(q)B_n(q) > B_{n-1}(q)B_n'(q)$ or, using the representation for $B_n'(q)$ again,

$$(n-2)[B_{n-1}(q) - B_{n-2}(q)]B_n(q) > (n-1)[B_n(q) - B_{n-1}(q)]B_{n-1}(q),$$

which can be transformed into

$$(n-1)B_{n-1}(q)^2 > (n-2)B_n(q)B_{n-2}(q) + B_n(q)B_{n-1}(q). \tag{36}$$

Recall that we are considering $q \in (q_n^*, 1]$ and hence $B_{n-1}(q) > B_n(q)$. Therefore, it is sufficient to show that $B_{n-1}(q)^2 > B_n(q)B_{n-2}(q)$. Introducing variable $s = \frac{1-q}{q}$, rewrite $B_n(q)$ as

$$B_n(s) = \frac{n-1}{(s+1)^{n-1}} \int_0^1 (s+z)^{n-2} m(z)dz.$$

The inequality $B_{n-1}(q)^2 > B_n(q)B_{n-2}(q)$ then can be written as

$$(n-2)^2 \left[\int (s+z)^{n-3} m(z)dz\right]^2 > (n-1)(n-3) \int (s+z)^{n-2} m(z)dz \int (s+z)^{n-4} m(z)dz. \tag{37}$$

Introduce coefficients $m_k = \int z^k m(z)dz$. For convenience, we also set $m_k = 0$ if $k < 0$. Then, representing the square of the integral on the left-hand side as a product of integrals and using binomial expansions, the inequality becomes

$$\begin{aligned}(n-2)^2 &\sum_{k=0}^{n-3} \binom{n-3}{k} s^k m_{n-3-k} \sum_{l=0}^{n-3} \binom{n-3}{l} s^l m_{n-3-l} \\ &> (n-1)(n-3) \sum_{k=0}^{n-2} \binom{n-2}{k} s^k m_{n-2-k} \sum_{l=0}^{n-4} \binom{n-4}{l} s^l m_{n-4-l}.\end{aligned}$$

On both sides of the inequality we have strictly increasing, positive polynomials of degree $2n-6$ in s. The inequality will hold for all $s \geq 0$ if for any $t = 0, 1, \ldots, 2n-6$ the coefficient on s^t on the left-hand side exceeds the coefficient on s^t on the right-hand side. Thus, we require that for each $t = 0, 1, \ldots, 2n-6$,

$$\begin{aligned}
&(n-2)^2 \sum_{k=0}^{t} \binom{n-3}{k} \binom{n-3}{t-k} m_{n-3-k} m_{n-3-t+k} \\
&> (n-1)(n-3) \sum_{k=0}^{t} \binom{n-2}{k} \binom{n-4}{t-k} m_{n-2-k} m_{n-4-t+k}.
\end{aligned}$$

This inequality can be rewritten in terms of coefficients $b_n = B_n(1) = (n-1) \int_0^1 z^{n-2} m(z) dz = (n-1) m_{n-2}$. The left-hand side becomes

$$\begin{aligned}
&(n-2)^2 \sum_{k=0}^{t} \binom{n-3}{k} \binom{n-3}{t-k} \frac{b_{n-1-k} b_{n-1-t+k}}{(n-2-k)(n-2-t+k)} \\
&= \sum_{k=0}^{t} \frac{(n-2)^2 (n-3)!(n-3)!}{k!(n-3-k)!(t-k)!(n-3-t+k)!} \frac{b_{n-1-k} b_{n-1-t+k}}{(n-2-k)(n-2-t+k)} \\
&= \sum_{k=0}^{t} \binom{n-2}{k} \binom{n-2}{t-k} b_{n-1-k} b_{n-1-t+k}.
\end{aligned}$$

Similarly, the right-hand side becomes

$$\begin{aligned}
&(n-1)(n-3) \sum_{k=0}^{t} \binom{n-2}{k} \binom{n-4}{t-k} \frac{b_{n-k} b_{n-2-t+k}}{(n-1-k)(n-3-t+k)} \\
&= \sum_{k=0}^{t} \frac{(n-1)(n-3)(n-2)!(n-4)!}{k!(n-2-k)!(t-k)!(n-4-t+k)!} \frac{b_{n-k} b_{n-2-t+k}}{(n-1-k)(n-3-t+k)} \\
&= \sum_{k=0}^{t} \binom{n-1}{k} \binom{n-3}{t-k} b_{n-k} b_{n-2-t+k}.
\end{aligned}$$

Thus, we need to show that for all $t = 0, \ldots, 2n-6$,

$$\sum_{k=0}^{t} \binom{n-2}{k} \binom{n-2}{t-k} b_{n-1-k} b_{n-1-t+k} > \sum_{k=0}^{t} \binom{n-1}{k} \binom{n-3}{t-k} b_{n-k} b_{n-2-t+k}. \tag{38}$$

Observe that the sums of the products of binomial coefficients on the two sides of (38) are equal to each other. This follows from the identity $(s+1)^{n-2}(s+1)^{n-2} = (s+1)^{n-1}(s+1)^{n-3}$

by equating the coefficients on the same powers of s.

The right-hand side of (38) can be transformed as

$$\begin{aligned}
&\sum_{k=0}^{t}\binom{n-1}{k}\binom{n-3}{t-k}b_{n-k}b_{n-2-t+k}=\binom{n-3}{t}b_nb_{n-2-t}\\
&+\sum_{k=1}^{t}\binom{n-1}{k}\binom{n-3}{t-k}b_{n-k}b_{n-2-t+k}=\binom{n-3}{t}b_nb_{n-2-t}\\
&+\sum_{k=0}^{t-1}\binom{n-1}{k+1}\binom{n-3}{t-k-1}b_{n-1-k}b_{n-1-t+k}.
\end{aligned}$$

Thus, inequality (38) becomes

$$\begin{aligned}
&\binom{n-2}{t}b_{n-1}b_{n-1-t}-\binom{n-3}{t}b_nb_{n-2-t}\\
&+\sum_{k=0}^{t-1}\left[\binom{n-2}{k}\binom{n-2}{t-k}-\binom{n-1}{k+1}\binom{n-3}{t-k-1}\right]b_{n-1-k}b_{n-1-t+k}>0.
\end{aligned}$$

Note that the sum can be extended to $k=t$ and the first term can be included back into it using the convention that $\binom{n-3}{-1}=0$. This gives

$$\sum_{k=0}^{t}\left[\binom{n-2}{k}\binom{n-2}{t-k}-\binom{n-1}{k+1}\binom{n-3}{t-k-1}\right]b_{n-1-k}b_{n-1-t+k}>\binom{n-3}{t}b_nb_{n-2-t}. \tag{39}$$

To gain some intuition, consider several special cases. For $t=0$, the inequality reduces to $b_{n-1}^2>b_nb_{n-2}$, which is true due to the log-concavity of b_k. For $t=1$, we obtain

$$[(n-2)-(n-1)]b_{n-1}b_{n-2}+(n-2)b_{n-2}b_{n-1}>(n-3)b_nb_{n-3},$$

which reduces to $b_{n-1}b_{n-2}>b_nb_{n-3}$, which is also true due to the log-concavity of b_k. For $t=2$, the inequality becomes

$$\begin{aligned}
&\left[\binom{n-2}{2}-(n-1)(n-3)\right]b_{n-1}b_{n-3}+\left[(n-2)^2-\binom{n-1}{2}\right]b_{n-2}^2\\
&+\binom{n-2}{2}b_{n-3}b_{n-1}>\binom{n-3}{2}b_nb_{n-4}.
\end{aligned}$$

Simplifying, obtain

$$\frac{(n-2)(n-3)}{2}b_{n-2}^2 > (n-3)b_{n-1}b_{n-3} + \frac{(n-3)(n-4)}{2}b_n b_{n-4},$$

or

$$(n-2)b_{n-2}^2 > 2b_{n-1}b_{n-3} + (n-4)b_n b_{n-4}.$$

This inequality holds because $b_{n-2}^2 > b_{n-1}b_{n-3}$ and $b_{n-2}^2 > b_n b_{n-4}$ due to the log-concavity of b_k.

Building on these examples, we proceed as follows. Suppose first that t is odd, i.e., $t = 2p+1$ where $p \geq 1$. There are $2p+2$ terms in the sum in the left-hand side of (39), which come in pairs multiplying $b_{n-1-k}b_{n-1-t+k}$ for $k = 0, \dots, p$. Therefore, (39) can be written as

$$\sum_{k=0}^{p} D_{tk} b_{n-1-k} b_{n-1-t+k} > \binom{n-3}{t} b_n b_{n-2-t}, \tag{40}$$
$$D_{tk} = 2\binom{n-2}{k}\binom{n-2}{t-k} - \binom{n-1}{k+1}\binom{n-3}{t-k-1} - \binom{n-1}{t-k+1}\binom{n-3}{k-1}.$$

Note that

$$\binom{n-2}{k}\binom{n-2}{t-k} - \binom{n-1}{k+1}\binom{n-3}{t-k-1} = \frac{(n-2)!^2}{k!(n-2-k)!(t-k)!(n-2-t+k)!}$$
$$- \frac{(n-1)!(n-3)!}{(k+1)!(n-2-k)!(t-k-1)!(n-2-t+k)!}$$
$$= \frac{(n-2)!(n-3)![(n-2)(k+1)-(n-1)(t-k)]}{(k+1)!(t-k)!(n-2-k)!(n-2+t-k)!},$$

$$\binom{n-2}{k}\binom{n-2}{t-k} - \binom{n-1}{t-k+1}\binom{n-3}{k-1} = \frac{(n-2)!^2}{k!(n-2-k)!(t-k)!(n-2-t+k)!}$$
$$- \frac{(n-1)!(n-3)!}{(t-k+1)!(n-2-t+)!(k-1)!(n-2-k)!}$$
$$= \frac{(n-2)!(n-3)![(n-2)(t-k+1)-(n-1)k]}{k!(t-k+1)!(n-2-k)!(n-2+t-k)!}.$$

Combining the two differences, obtain for the coefficient on $b_{n-1-k}b_{n-1-t+k}$ in (40):

$$D_{tk} = \frac{(n-2)!(n-3)!T_{tk}}{(k+1)!(t-k+1)!(n-2-k)!(n-2-t+k)!},$$

where

$$\begin{aligned}T_{tk} &= [(n-2)(k+1)-(n-1)(t-k)](t-k+1)+[(n-2)(t-k+1)-(n-1)k](k+1)\\ &= 2(n-2)(k+1)(t-k+1)-(n-1)[(t-k)^2+t-k+k^2+k]\\ &= -(n-1)t^2+2(2n-3)kt-2(2n-3)k^2+(n-3)t+2(n-2).\end{aligned}$$

For $k \leq p$, T_{tk} is increasing in k. It can also be shown that $T_{tp} > 0$. Indeed,

$$\begin{aligned}T_{tp} &= -(n-1)(4p^2+4p+1)+2(2n-3)(p^2+p)+(n-3)(2p+1)+2(n-2)\\ &= 2[-p^2+(n-4)p+(n-3)] = 2(n-3-p)(p+1).\end{aligned}$$

Recall that $t = 2p+1 \leq 2n-6$, and hence $p < n-3$, which implies $T_{tp} > 0$. Next, we show that $T_{t0} < 0$ for a sufficiently large t. Indeed,

$$T_{t0} = -(n-1)t^2+(n-3)t+2(n-2) = -[(n-1)t-2(n-2)](t+1);$$

therefore, $T_{t0} < 0$ for $t > 2$. The cases with $t \leq 2$ have been considered separately above, and from this point on we assume that $t > 2$. This implies that coefficients D_{tk} in (40) are negative for low k and positive for high k, and there exists a $\bar{k}$ such that $0 < \bar{k} < p$ and $D_{tk} \leq 0$ for $k \leq \bar{k}$ (where the inequality is strict at least for $k = 0$) and $D_{tk} > 0$ for $k > \bar{k}$.

Separating the terms with $k \leq \bar{k}$ and $k > \bar{k}$ in the sum in (40), rewrite it as

$$\sum_{k=\bar{k}+1}^{p} D_{tk}b_{n-1-k}b_{n-1-t+k} > -\sum_{k=0}^{\bar{k}} D_{tk}b_{n-1-k}b_{n-1-t+k} + \binom{n-3}{t}b_nb_{n-2-t}. \tag{41}$$

Now all the coefficients on both sides are positive, and because the sums of coefficients on both sides of (38) are equal, so are the sums of coefficients on both sides of (41). In addition, it follows from the log-concavity of $\{b_k\}$ that each product $b_{n-1-k}b_{n-1-t+k}$ on the left-hand side of (41) is greater than each product on the right-hand side because the values of k in the sum on the left are all larger and $t-k > k$. These observations imply that the inequality holds.

Suppose now that t is even, $t = 2p$, $p > 1$. In this case there are $2p+1$ terms in the sum on the left-hand side of (39). There is one term proportional to b_{n-1-p}^2, and all other $2p$ terms come in pairs multiplying $b_{n-1-k}b_{n-1-t+k}$ for $k = 0, \dots, p-1$. Therefore, (39) can be written as

$$\left[\binom{n-2}{p}^2 - \binom{n-1}{p+1}\binom{n-3}{p-1}\right] b_{n-1-p}^2 + \sum_{k=0}^{p-1} D_{tk} b_{n-1-k} b_{n-1-t+k} > \binom{n-3}{t} b_n b_{n-2-t}, \tag{42}$$

where D_{tk} is defined in (40). The coefficient on b_{n-1-p}^2 can be simplified as

$$\binom{n-2}{p}^2 - \binom{n-1}{p+1}\binom{n-3}{p-1} = \frac{(n-2)!^2}{p!^2(n-2-p)!^2} - \frac{(n-1)!(n-3)!}{(p+1)!(p-1)!(n-2-p)!^2}$$
$$= \frac{(n-2)!(n-3)![(n-2)(p+1)-(n-1)p]}{p!(p+1)!(n-2-p)!^2} = \frac{1}{p+1}\binom{n-2}{p}\binom{n-3}{p}.$$

Thus, the coefficient on b_{n-1-p}^2 is always positive. The remaining coefficients D_tk, $k = 0, \dots, p-1$, are nonpositive for $k \le \bar{k}$ (and strictly negative at least for $k = 0$) and positive for $k > \bar{k}$ as shown above. We can, therefore, separate the sum into two parts and move the negative part to the right-hand side in the same way as above. The result again will be an inequality where all coefficients are positive, the sums of coefficients on the left and on the right are equal, and all terms $b_{n-1-k}b_{n-1-t+k}$ on the left have a higher value of k than on the right and hence, due to the log-concavity of $\{b_k\}$, they are larger. The inequality, therefore, holds.

We have shown that if the sequence $\{b_k\}$ is log-concave then inequality (37) holds for all s and, therefore, $B_{n-1}(q)^2 > B_n(q)B_{n-2}(q)$ for all q, which implies that (36) holds for $q \in (q_n^*, 1]$. This, in turn, implies that $\beta_n(q)$ is decreasing in q for $q \in (q_n^*, 1]$. ■

Proof of Proposition 9 For $c(e) = c_0 e^{\xi}$, it follows from (22) that

$$\frac{\partial E_n^*(q)}{\partial q} = \frac{nc'(e_n^*(q))}{c''(e_n^*(q))}[\xi - 1 + \beta_n(q)]. \tag{43}$$

This expression is positive if and only if $\xi > 1 - \beta_n(q)$. This is always true for $\beta_n(q) \ge 0$, i.e., for $q \le q_n^*$. For $q > q_n^*$, it follows from Lemma 15 that $\beta_n(q)$ is decreasing in q, and hence $\xi > 1 - \beta_n(q)$ if and only if

$$\xi > 1 - \beta_n(1) = 1 - (n-1)\left[1 - \frac{B_{n-1}(1)}{B_n(1)}\right],$$

and the result follows. ■

Proof of Proposition 13 In order to compare $E_p^* = \bar{k}e_p^*$ to $E_{\bar{k}}^* = \bar{k}e_{\bar{k}}^*$, we need to compare e_p^* and $e_{\bar{k}}^*$, i.e., it is sufficient to compare B_p given by (16) and $b_{\bar{k}}$.

(i) Suppose $p_0 = 0$ and kb_k is concave for $k \geq 1$. Then

$$B_p = \frac{1}{\bar{k}}\sum_{k=1}^{n} p_k k b_k = \frac{1}{\bar{k}}\mathrm{E}_p(Kb_K) \leq \frac{1}{\bar{k}}\bar{k}b_{\bar{k}} = b_{\bar{k}},$$

where the inequality follows from Jensen's inequality, which will be strict if kb_k is strictly concave.

(ii) From Jensen's inequality for conditional expectations, and assumptions (a) and (b),

$$\mathrm{E}_p(Kb_K|K \geq 2) \leq \mathrm{E}_p(K|K \geq 2)b_{\mathrm{E}_p(K|K\geq 2)} \leq \mathrm{E}_p(K|K \geq 2)b_{\bar{k}}.$$

The first inequality will be strict if kb_k is strictly concave. Multiplying both sides by $\mathrm{Pr}_p(K \geq 2)$,

$$\mathrm{E}_p(Kb_K|K \geq 2)\mathrm{Pr}_p(K \geq 2) \leq \mathrm{E}_p(K|K \geq 2)\mathrm{Pr}_p(K \geq 2)b_{\bar{k}},$$

or

$$\bar{k}B_p \leq \sum_{k=2}^{n} kp_k b_{\bar{k}} \leq \sum_{k=0}^{n} kp_k b_{\bar{k}} = \bar{k}b_{\bar{k}}.$$

The last inequality will be strict if $p_1 > 0$. Thus, we showed that $B_p \leq b_{\bar{k}}$, with strict inequality if kb_k is strictly concave or $p_1 > 0$. ■

Proof of Proposition 14 In order to compare $E_p^* = \bar{k}e_p^*$ to $E_{p'}^* = \bar{k}e_{p'}^*$, we need to compare e_p^* and $e_{p'}^*$, i.e., it is sufficient to compare B_p to $B_{p'}$. From (16), assuming $p_0 = p_0' = 0$, $B_p = \frac{1}{\bar{k}}\sum_{k=1}^{n} p_k k b_k = \frac{1}{\bar{k}}\mathrm{E}_p(Kb_K)$ and, similarly, $B_{p'} = \frac{1}{\bar{k}}\mathrm{E}_{p'}(Kb_K)$. The result follows by concavity of kb_k directly from the definition of second-order stochastic dominance. ■