Optimal Contest Design: A General Approach

Igor Letina, Shuo Liu and Nick Netzer*

May 2020

Abstract

We consider the design of contests for n agents when the principal can choose both the prize profile and the contest success function. Our framework includes Tullock contests, Lazear-Rosen tournaments and all-pay contests as special cases, among others. We show that the optimal contest has an intermediate degree of competitiveness in the contest success function, and a minimally competitive prize profile with $n-1$ identical prizes. The optimum can be achieved with a nested Tullock contest. We extend the model to allow for imperfect performance measurement and for heterogeneous agents. We relate our results to a recent literature which has asked similar questions but has typically focused on the design of either the prize profile or the contest success function.

Keywords: contest design, optimal contests, tournaments.

JEL: D02, D82, M52

*Letina: Department of Economics and CEPR, University of Bern. Liu: Guanghua School of Management, Peking University. Netzer: Department of Economics, University of Zurich. Email: igor.letina@vwi.unibe.ch, shuo.liu@gsm.pku.edu.cn, nick.netzer@econ.uzh.ch. We would like to thank Philipp Denter, Mikhail Drugov, Dana Sisak and Cédric Wasser for helpful comments. Large parts of the material in this paper are from the older discussion paper Letina, Liu, and Netzer (2018). Other parts of that older paper have appeared as Letina, Liu, and Netzer (2020).

1 Introduction

Many economic interactions can be summarized as situations where a group of agents compete for a set of prizes. Examples of such *contests* are: (i) competition for promotions or bonuses among employees, (ii) elections where candidates campaign in an effort to win political office, (iii) entrance exams where students compete for a limited number of places in schools and universities, (iv) scientists competing for grants and prizes, and (v) sporting events. What all of these contests have in common is that they are *designed*. In other words, some principal chooses the rules of the contest as well as the prizes that can be won. While the equilibrium behavior of agents in standard contests (Tullock contests, Lazear-Rosen tournaments, all-pay contests) is by now well understood,[1] the question of how the principal should optimally design a contest has received only partial answers.

Several recent articles have analyzed the optimal allocation of prizes in specific classes of contests. Examples include Schweinzer and Segev (2012) and Fu, Wang, and Wu (2019) for Tullock contests, Drugov and Ryvkin (2020b) and Morgan, Tumlinson, and Várdy (2019) for Lazear-Rosen tournaments, and Fang, Noe, and Strack (2020) and Olszewski and Siegel (2020) for all-pay contests. While these papers have produced important insights, sometimes the intuition obtained from one contest class does not translate well to a different class. For example, from Clark and Riis (1998) and Schweinzer and Segev (2012) we learn that in a nested Tullock contest with risk-neutral agents a winner-take-all (WTA) prize structure is optimal if a pure-strategy equilibrium exists, while Fang et al. (2020) show that in an all-pay contest the exact opposite is optimal, with all agents but one receiving an equal positive prize. Furthermore, it is not clear if the principal should use a Tullock contest or an all-pay contest, or even some other contest format which has not been studied yet.[2] Our paper proposes a general framework in which these contest design questions can be analyzed, and which can provide an intuition for the different results in the literature.

In our model, the principal can choose any prize profile and any contest success function (CSF), which includes the standard contest models as special cases. The objective of the principal is to maximize the expected aggregate effort net of the sum of prizes. The agents can be risk-neutral or risk-averse, and they have convex effort cost functions. Our main result is that, even though the principal can choose from a large set of contests, the optimum can be implemented by a nested Tullock contest (Clark and Riis, 1996). The optimal prize profile has $n-1$ equal positive prizes and one zero prize. The optimal Tullock CSF is characterized by a precision parameter $r^*(n)$ which is the largest r such that a symmetric pure-strategy equilibrium still exists. Note that we do not restrict attention to

[1] For an excellent textbook treatment of the standard contests, see Konrad (2009).

[2] A broader question is whether the principal should use a contest at all, rather than some other incentive mechanism. One setting in which contests are optimal within a larger set of mechanisms is provided in Letina et al. (2020). In their model, contests are optimal because they give a lenient reviewer the commitment to punish shirking agents.

pure-strategy equilibria, but they emerge in the optimum. The precision parameter $r^*(n)$ increases in n and approaches infinity in the limit. In other words, the optimal Tullock CSF approximates the all-pay CSF when the number of agents is large, but is less competitive for smaller numbers of agents.

Our results provide a unifying perspective on the seemingly contradictory findings in the literature. In line with Fang et al. (2020), we show that the minimally competitive prize profile, with $n-1$ identical positive prizes and one prize of zero, is optimal (see also Glazer and Hassin, 1988; Letina et al., 2020). More generally, the message of Fang et al. (2020) is that "turning up the heat" in an all-pay contest, by increasing the dispersion in prizes, decreases the effort that agents exert. Our results show that it is optimal to turn down the heat even more, by moving from the perfectly discriminating and very competitive all-pay CSF towards a smoother and less competitive Tullock CSF. Our results are also in line with the seemingly contradictory intuition of Schweinzer and Segev (2012), who argue that optimal Tullock contests should concentrate prizes on the top. This holds subject to the constraint that a pure-strategy equilibrium exists in the contest. Our optimal Tullock contest is indeed as competitive as possible without destroying the pure-strategy equilibrium. In other words, "turning down the heat" should stop exactly at the point where a pure-strategy equilibrium emerges. Such an insight can only be obtained in a setting like ours, where both the prize profile and the CSF are endogenous and can be chosen without functional-form constraints.

In our baseline model, we assume that the principal perfectly observes the efforts by the agents and that the agents are symmetric. To illustrate the flexibility of our approach, we relax these assumptions in turn. We first provide a general result about optimal contests with imperfect performance measurement. We illustrate this result with two examples, one where the principal observes a noisy measure of efforts, and one where the principal observes only the difference in efforts of two agents. Next, we consider agents with heterogeneous effort cost functions. We derive the optimal contest for $n = 2$, and for $n > 2$ we provide results for the case when heterogeneity is sufficiently small. In all our extensions, the optimal contest has only one prize of zero, no dispersion among the positive prizes, and a CSF that is less competitive than all-pay.

The model is introduced in Section 2. The optimal contest is derived in Section 3. The extensions can be found in Section 4. Section 5 provides a more detailed overview of the related literature, and Section 6 concludes. All proofs are in the Appendix.

2 The Model

2.1 Environment

The basic model setup is the same as in Letina et al. (2020). There is a principal and a set of agents $I = \{1, \ldots, n\}$, where $n \geq 2$. Each agent $i \in I$ chooses an effort level $e_i \geq 0$, incurs a cost of effort equal to $c(e_i)$, and obtains a monetary transfer $t_i \geq 0$. The payoff of agent i is

$$\pi_i(e_i, t_i) = u(t_i) - c(e_i).$$

The utility function $u : \mathbb{R}_+ \to \mathbb{R}$ is twice differentiable, strictly increasing, weakly concave, and satisfies $u(0) = 0$. The cost function $c : \mathbb{R}_+ \to \mathbb{R}_+$ is twice differentiable, strictly increasing, strictly convex, and satisfies $c(0) = 0$, $c'(0) = 0$, and $\lim_{e_i \to \infty} c'(e_i) = \infty$. The outside option of agents is zero.

Let effort profiles be denoted by $e = (e_1, \ldots, e_n) \in E = \mathbb{R}^n_+$ and transfer vectors by $t = (t_1, \ldots, t_n) \in T = \mathbb{R}^n_+$. The payoff of the principal is

$$\pi_P(e, t) = \sum_{i=1}^{n} e_i - \sum_{i=1}^{n} t_i.$$

2.2 Contests

As in Letina et al. (2020), a *contest* (y, μ) is defined by a prize profile y and a CSF μ. The prize profile $y = (y_1, \ldots, y_n)$ is w.l.o.g. assumed to satisfy $y_1 \geq \ldots \geq y_n \geq 0$. Each effort profile e results in some, possibly random, allocation of the prizes to the agents. Formally, let $T(y)$ be the set of all permutations of y. That is, $t = (t_1, \ldots, t_n) \in T(y)$ if and only if there exists a bijective mapping $s : I \to I$ such that $t_i = y_{s(i)}$, $\forall i \in I$. Then, t represents the specific allocation where agent i obtains the $s(i)$-th prize. The probability that prize allocation t is realized is governed by the CSF $\mu = (\mu^e)_{e \in E}$, which maps each effort profile $e \in E$ into a probability measure $\mu^e \in \Delta T(y)$. Hence $\mu^e(t)$ is the probability of obtaining prize allocation t when the agents' efforts are given by e.[3]

Example. When there are two agents, the prize profile is $y = (y_1, y_2)$ with $y_1 \geq y_2 \geq 0$, where y_1 is the prize that the winner obtains and y_2 is the prize that the loser obtains. There are only two permutations of y. One is $t' = (y_1, y_2)$, which is the transfer vector when agent 1 wins the contest, and the other is $t'' = (y_2, y_1)$, which is the transfer vector when agent 2 wins. For any given effort profile $e = (e_1, e_2)$, the probability that agent 1 wins is given by $\mu^e(t')$, and the probability that agent 2 wins is $\mu^e(t'') = 1 - \mu^e(t')$. In an

[3] We assume that $\mu^e(t)$ is a measurable function of e, for each t, which implies that all expected values in the following are well-defined.

all-pay contest

$$\mu^e(t') = \begin{cases} 1 & \text{if } e_1 > e_2, \\ 1/2 & \text{if } e_1 = e_2, \\ 0 & \text{if } e_1 < e_2, \end{cases}$$

while in a Tullock contest with impact function f we have

$$\mu^e(t') = \begin{cases} \dfrac{f(e_1)}{f(e_1) + f(e_2)} & \text{if } \max\{e_1, e_2\} > 0, \\ 1/2 & \text{else.} \end{cases}$$

With more than 2 agents and thus multiple prizes, the function μ is a convenient way to represent the probabilities with which prizes are allocated to agents, which encompasses all conventional CSFs that the literature has studied. □

Given a contest (y, μ), the agents choose their efforts simultaneously, anticipating that the prizes y will be distributed according to μ. Let $\sigma_i \in \Delta\mathbb{R}_+$ be agent i's mixed strategy and let $e_i \in \mathbb{R}_+$ represent pure strategies. Strategy profiles are given by $\sigma = (\sigma_1, ..., \sigma_n) \in (\Delta\mathbb{R}_+)^n$. We also use σ to denote the induced product measure in ΔE. We say that a contest (y, μ) *implements* a strategy profile σ if it satisfies

$$\Pi_i(\sigma_i, \sigma_{-i} \mid (y, \mu)) \geq \Pi_i(\sigma'_i, \sigma_{-i} \mid (y, \mu)) \quad \forall \sigma'_i \in \Delta\mathbb{R}_+, \forall i \in I, \tag{IC-A}$$

where $\Pi_i(\sigma \mid (y, \mu)) = \mathbb{E}_\sigma[\mathbb{E}_{\mu^e}[u(t_i)]] - \mathbb{E}_{\sigma_i}[c(e_i)]$. Each agent can always deviate to zero effort and thus guarantee himself a payoff of at least zero. Therefore we can ignore the agents' participation constraints.

The principal chooses a contest (y, μ) which implements a strategy profile σ in order to maximize her expected payoff. Formally, the principal's problem is given by

$$\max_{\sigma, y, \mu} \Pi_P(\sigma \mid (y, \mu)) \quad \text{s.t.} \quad \text{(IC-A)}, \tag{P}$$

where $\Pi_P(\sigma \mid (y, \mu)) = \mathbb{E}_\sigma[\sum_{i=1}^n e_i] - \sum_{i=1}^n y_i$. A contest (y^*, μ^*) is *optimal* if there exists σ^* such that $(\sigma^*, (y^*, \mu^*))$ solves (P).[4]

[4] Letina et al. (2020) study a related but different problem. They are interested in the design of performance evaluation schemes when a reviewer observes the agents' efforts and reports to the principal. Since the preferences of the reviewer and the principal are misaligned, an additional incentive constraint is required. Furthermore, Letina et al. (2020) optimize over a substantially larger class of mechanisms, but show that a contest is optimal within that class.

3 Optimal Contest

In this section, we will show that a specific Tullock contest solves the principal's problem. Tullock contests with n agents and a single positive prize are typically characterized by a contest success function of the form

$$p_i(e) = \frac{f(e_i)}{\sum_{j \in I} f(e_j)}, \tag{1}$$

which determines the probability that agent i wins the prize as a function of the effort profile. The impact function f is continuous, strictly increasing and satisfies $f(0) = 0$ (Skaperdas, 1996).[5] If all agents exert zero effort, each of them wins with equal probability. With more than one positive prize, the contest success function can be applied in a nested fashion (see Clark and Riis, 1996). The first prize is allocated according to (1) among all n agents, the second prize is allocated according to (1) restricted to those $n-1$ agents who have not received the first prize, and so on. In terms of our notation based on μ introduced above, when all efforts are strictly positive a nested Tullock contest gives rise to the allocation probabilities

$$\mu^e(t_1, \ldots, t_n) = \mu^e(y_{s(1)}, \ldots, y_{s(n)}) = \prod_{k=1}^{n} \left[\frac{f(e_{s^{-1}(k)})}{\sum_{j=k}^{n} f(e_{s^{-1}(j)})} \right], \tag{2}$$

where s is the permutation generating t from y. The extension to the case where some efforts are zero is straightforward.

Proposition 1 *The following contest (y^*, μ^*) is optimal:*

(i) The prize profile is $y^* = (x^*/(n-1), \ldots, x^*/(n-1), 0)$, where x^* is given by

$$u'\left(\frac{x^*}{n-1}\right) = c'\left(c^{-1}\left(\frac{n-1}{n} u\left(\frac{x^*}{n-1}\right)\right)\right).$$

(ii) The CSF μ^* is of the nested Tullock type (2) with

$$f(e_i) = c(e_i)^{r^*(n)} \quad \text{and} \quad r^*(n) = \frac{n-1}{H_n - 1},$$

where $H_n = \sum_{k=1}^{n} 1/k$ is the n-th harmonic number.

To obtain an intuition for this result, start from condition *(i)*, which characterizes the optimal prize sum x^* by a simple cost-benefit condition. The optimal prize profile y^* has

[5] In Section 4.2, we will study an extension with individual-specific impact functions f_i (Cornes and Hartley, 2005).

one zero prize and splits x^* equally among the other $n-1$ prizes.[6] The optimality of this prize structure is related to the reason why, in Fang et al. (2020), reducing inequality in the prize profile is beneficial to the principal if the CSF is an all-pay contest. Reducing prize inequality in the all-pay contest reduces dispersion of efforts chosen in the mixed strategy equilibrium, and random effort choice is inefficient due to convex effort costs. Fang et al. (2020) therefore conclude that it is optimal to move towards the least unequal prize profile y^* in all-pay contests (for both risk-neutral and risk-averse agents). With the optimal CSF characterized in condition (ii) of our proposition, the contest actually implements a symmetric effort profile in pure strategies $(e^*, \ldots, e^*)$, where

$$e^* = c^{-1}\left(\frac{n-1}{n} u\left(\frac{x^*}{n-1}\right)\right)$$

is such that the agents' expected payoff is zero. This is achieved by using a CSF that is less competitive than the all-pay contest. Put differently, it is optimal to turn down the heat of the contest not only in the prize profile but also in the allocation rule.

The optimum is a nested Tullock contest with a finite randomness parameter $r^*(n)$.[7] If there are two agents and hence one positive prize, we obtain $r^*(2) = 2$. It is well-known that this is the largest value of the parameter r for which the Tullock contest still has a pure-strategy equilibrium. For the case of n risk-neutral agents and cost functions of the monomial form, Schweinzer and Segev (2012) show that there is a continuum of nested Tullock contests that all generate the first-best pure-strategy efforts, for a given prize sum. That continuum is parametrized by the precision parameter $r \in [n/(n-1), (n-1)/(H_n-1)]$, and the prizes are concentrated on the top as much as possible so that the pure-strategy equilibrium still exists. Considering their special case, where the first-best is achievable due to risk-neutrality, this multiplicity of optimal contests of course carries over to our setting. In the general case with risk-averse agents, where the first-best is not achievable, the optimal contest described in Proposition 1 has the highest possible precision parameter $r^*(n) = (n-1)/(H_n - 1)$ from along the continuum.[8] A higher r would make the contest too competitive and induce wasteful mixing in equilibrium. A lower r would induce less effort, and, in contrast to Schweinzer and Segev (2012), the weaker incentives cannot be

[6]Part (i) of Proposition 1 is a corollary of Theorem 2 in Letina et al. (2020). This optimal prize profile is actually uniquely optimal whenever the agents are risk-averse. If the agents are risk-neutral, then x^* can be split arbitrarily among the first $n-1$ prizes, but the CSF also has to be adjusted in that case.

[7]Readers familiar with the Tullock form may be surprised that f depends on the cost function c. However, standard formulations of the Tullock contest assume linear cost functions, which is then equivalent to a reformulation of our model where agents choose expenditure levels $c(e_i)$ directly. Note also that the nested Tullock CSF is not the unique solution to the problem. As Letina et al. (2020) have shown, another way to reduce competitiveness is to put an effort cap on an all-pay contest.

[8]This is also similar to the finding in Morgan et al. (2019), who show that in a large Lazear-Rosen tournament, the optimal level of precision of the CSF is such that the agents are indifferent between dropping out of the contest and participating.

compensated by a more unequal prize profile when the agents are risk-averse.

To prove Proposition 1, we employ a novel approach that is of independent interest and could prove useful in other settings. Instead of showing directly that no profitable deviation exists, we fix an arbitrary deviation and ask for which levels of r this deviation is not profitable. Using this approach, we can show that when $r \geq r^*(n)$, there are no profitable deviations from the equilibrium effort to lower effort levels, and when $r \leq r^*(n)$, there are no profitable deviations to higher effort levels. This is intuitive, because a higher r implies a more precise contest in which shirking is less profitable, but not trivial to demonstrate.

As already mentioned by Schweinzer and Segev (2012), the randomness parameter $r^*(n)$ is strictly increasing in n and satisfies $\lim_{n\to\infty} r^*(n) = \infty$.[9] In other words, the optimal contest becomes more precise and more competitive as n grows, and it approximates an all-pay contest in the limit when the contest becomes *large*.[10]

4 Extensions

4.1 Imperfect Performance Measurement

In this section, we show how our framework can be enriched to study the optimal contest design problem in settings where the agents' efforts are measured imperfectly. We do not want to impose one specific measurement constraint, so we first introduce a general information structure. We then derive conditions under which an optimal contest can be described, and we illustrate the approach with two specific examples.

Suppose that, after the agents have chosen their efforts e, a signal $s \in S$ is drawn according to an effort-dependent probability measure $\eta^e \in \Delta S$. Only this signal can be used to evaluate the performance of the agents, i.e., the CSF must condition the prize allocation on s rather than on e directly. For instance, in a firm the allocation of bonuses among the sales force will typically depend on measures such as realized sales, which are only noisy proxies for the agents' actual efforts.

We denote $\eta = (\eta^e)_{e\in E}$ and call (S, η) the *observational structure* of the model. We do not impose any assumptions on the set of signals S or the stochastic signal-generating process η.[11] Hence a large range of applications and examples can be modelled by different observational structures. Our previous setting with perfect performance measurement is a special case where $S = E$ and η^e is the Dirac measure on e. A second example is

[9] Our Appendix A.2 contains a formal proof of that claim.

[10] See Siegel (2009) for a general treatment of all-pay contests and Olszewski and Siegel (2016) for large contests.

[11] We only need the regularity condition that $\eta^e(A)$ is a measurable function of e for each measurable subset $A \subseteq S$, to ensure that expected payoffs remain well-defined.

the classical moral-hazard setting where each agent's effort e_i produces a random output s_i such that $\mathbb{E}_{\eta^e}[s_i] = e_i$. The CSF can condition only on the observable output vector $s \in S \subseteq \mathbb{R}^n$. It is possible to assume that the principal cares about output rather than effort in this application, because her expected payoffs are unaffected by the noise with zero mean. A third example is a setting where only an aggregate statistic of the effort profile becomes observable. For instance, suppose there are two agents and only the difference between their efforts but not the levels can be observed. This amounts to an observational structure where $S = \mathbb{R}$ and η^e is the Dirac measure on $e_1 - e_2$. One could also model the observation of ordinal performance ranks, or a blind review process where the individual efforts are anonymized. Finally, the observational structure allows for stochastic signals which are correlated across the agents like in Green and Stokey (1983) or Nalebuff and Stiglitz (1983).

Given an observational structure (S, η), a contest is defined by (y, τ), where y is the prize profile as before, and $\tau = (\tau^s)_{s \in S}$ is the CSF describing how prizes are allocated to the agents depending on the realized signal. Similar to before, $\tau^s \in \Delta T(y)$ denotes the prize allocation associated with a specific signal realization $s \in S$. Given a contest (y, τ), the payoff of agent i with strategy profile σ is

$$\Pi_i(\sigma \mid (y, \tau)) = \mathbb{E}_\sigma \left[\mathbb{E}_{\eta^e} \left[\mathbb{E}_{\tau^s} \left[u(t_i) \right] \right] \right] - \mathbb{E}_{\sigma_i} \left[c(e_i) \right].$$

The contest implements σ if $\Pi_i((\sigma_i, \sigma_{-i}) \mid (y, \tau)) \geq \Pi_i((\sigma'_i, \sigma_{-i}) \mid (y, \tau))$ for all $\sigma'_i \in \Delta\mathbb{R}_+$ and $i \in I$. The principal maximizes

$$\Pi_P(\sigma \mid (y, \tau)) = \mathbb{E}_\sigma \left[\sum_{i=1}^n e_i \right] - \sum_{i=1}^n y_i$$

by choosing a contest (y, τ) and a strategy profile σ to be implemented.

As the following result shows, optimal contests can sometimes be described despite the generality of this framework.

Proposition 2 *Fix an arbitrary observational structure* (S, η). *A contest with prize profile* $y = (x^*/(n-1), \ldots, x^*/(n-1), 0)$ *is optimal if it implements* $(e^*, \ldots, e^*)$.

A contest with a prize profile as characterized in Proposition 1 is optimal for any observational structure, provided that it still implements the optimal effort profile $(e^*, \ldots, e^*)$ and therefore achieves the same maximal payoff for the principal as in the case of perfect observation. Intuitively, a coarser observational structure can always be replicated when observation is perfect, by emulating the observational friction in the CSF. Hence, the maximal payoff with perfect observation is an upper bound on the principal's payoff with imperfect observation.

Despite its simplicity, the result is a powerful tool for the design of optimal contests. Whenever we can find a signal-contingent CSF that ensures implementation of $(e^*, \ldots, e^*)$ with prize profile $(x^*/(n-1), \ldots, x^*/(n-1), 0)$, we can be sure to have constructed an optimal contest for the given observational structure. We illustrate the applicability of this tool in two simple examples.

Example. Consider a classical moral-hazard setting with two agents. The effort cost function is given by $c(e_i) = \gamma e_i^\beta$ for some $\gamma > 0$ and $\beta > 1$. The noise in output takes a multiplicative (or log-additive) form: the output of agent i who exerts effort e_i is given by $\tilde{s}_i = e_i \tilde{r}_i$, where the pair of random variables $(\tilde{r}_1, \tilde{r}_2)$ follows a bivariate log-normal distribution,

$$(\tilde{r}_1, \tilde{r}_2) \sim \ln \mathcal{N} \left[\begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right].$$

We show in Appendix A.4 that the optimal effort profile (e^*, e^*) can be implemented by a contest with prize profile $(x^*, 0)$ whenever the inequality

$$\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} \leq 2/(\pi \beta^2)$$

is satisfied, which just requires that the noise in output is not too strong. The contest that achieves implementation of (e^*, e^*) – and is therefore optimal by Proposition 2 – allocates the positive prize x^* to agent 1 with a probability that is increasing in the realized ratio of outputs s_1/s_2. More precisely, agent 1 receives the prize whenever s_1/s_2 is larger than a log-normally distributed random number. Similar *contests with multiplicative noise* have been studied in the literature.[12] With this construction, the overall randomness in the prize allocation can be adjusted to a level that guarantees implementation of (e^*, e^*).

Example. Consider a setting with two agents in which only the difference $s = e_1 - e_2$ but not the entire profile (e_1, e_2) can be observed. We show in Appendix A.4 that, despite this constraint, the optimal effort profile (e^*, e^*) can always be implemented by a contest with prize profile $(x^*, 0)$. The contest that achieves implementation of (e^*, e^*) – and is therefore optimal – allocates the positive prize to agent 1 with a probability that is increasing in the observed difference s. More precisely, agent 1 receives the prize whenever s is larger than a uniformly distributed random number. Such *contests with additive noise* have also been studied in the literature.[13] An appropriate level of randomness in the allocation rule again

[12] See, for instance, Jia, Skaperdas, and Vaidya (2013). We are not aware of an explicit treatment of the multiplicative log-normal noise model in the literature, but of course it can be transformed into a specific probit model with additive normal noise (Dixit, 1987).

[13] See, for instance, Lazear and Rosen (1981) and Hirshleifer (1989). Che and Gale (2000) provide a general treatment of contests with additive uniform noise. They show that these contests often do not

ensures that unilateral deviations from (e^*, e^*) are not profitable.

Of course, Proposition 2 is not always applicable. For instance, it is clear that only zero effort can be implemented if the signals are completely uninformative ($\eta^e = \eta^{e'}$ for all $e, e' \in E$). More generally, it will be impossible to implement the effort profile $(e^*, \dots, e^*)$ in a contest when the signals on which the prize allocation can be conditioned are too coarse or too noisy. We leave a characterization of optimal contests for such environments to future research.

4.2 Heterogeneous Contestants

Our framework can also incorporate heterogeneity in the abilities of the agents. To illustrate this, we consider a variation of the basic model in which the payoff of agent i is given by

$$\pi_i(e_i, t_i) = u(t_i) - c_i(e_i),$$

where the cost functions c_i satisfy our previous assumptions but can be different across agents. For the case of two agents, we provide a result that generalizes Proposition 1 to arbitrary cost functions.

Proposition 3 *Suppose $n = 2$. For any profile of cost functions (c_1, c_2), the following contest (y^*, μ^*) is optimal:*

(i) The prize profile is $y^* = (x^*, 0)$, where x^* is given by

$$(x^*, e_1^*, e_2^*) \in \underset{x, e_1, e_2 \geq 0}{\operatorname{argmax}} \; e_1 + e_2 - x \quad s.t. \quad c_1(e_1) + c_2(e_2) = u(x).$$

(ii) The CSF μ^* is of the Tullock type (1) with individual-specific impact functions

$$f_i(e_i) = \frac{c_i(e_i)^{r_i^*}}{c_i(e_i^*)^{r_i^* - 1}} \quad and \quad r_i^* = 1 + \frac{c_i(e_i^*)}{c_j(e_j^*)}, \;\; \forall i = 1, 2, j \neq i.$$

For the special case where $c_1 = c_2 = c$, we obtain $e_1^* = e_2^* = e^*$ and $r_i^* = 2$, so that the optimal impact functions are (up to an irrelevant multiplicative constant) given by $f_i(e_i) = c(e_i)^2$, exactly like in Proposition 1 for $n = 2$. With asymmetric cost functions, by contrast, the implemented effort levels will typically not be identical for the two agents. Consequently, the winning probabilities can also not be identical in equilibrium, because the agents have to be compensated for different effort costs. That this kind of biasing of a contest is beneficial when agents are heterogeneous is well-known (see e.g. Ewerhart, 2017;

have a symmetric pure-strategy equilibrium. The uniform distribution used in our construction is chosen precisely to avoid this problem.

Franke, Leininger, and Wasser, 2018). Our result establishes the form of biasing that is optimal when the principal is not restricted to a specific class of CSFs. To the best of our knowledge, the asymmetric Tullock contest described in Proposition 3 has not been studied before.

That the principal would optimally choose a zero prize $y_n = 0$ continues to hold with $n > 2$ asymmetric agents (see Lemma 7 in the Appendix). Generalizing the optimality of $n-1$ equal positive prizes faces the difficulty that some agents may have substantially higher effort costs in equilibrium than others, and cannot be compensated for their costs even if they win one of the identical prizes for sure. Our next result rests on the insight that effort profiles for which the agents' costs are so strongly heterogeneous cannot be optimal if their cost functions are not strongly heterogeneous. To formalize this idea, we fix any sequence of cost function profiles $(c_1^m, \dots, c_n^m)_{m \in \mathbb{N}}$ such that, for each $i \in I$, the sequence $(c_i^m)_{m \in \mathbb{N}}$ converges uniformly to a common cost function c as $m \to \infty$.

Proposition 4 *Let $(c_1^m, \dots, c_n^m) \to (c, \dots, c)$ uniformly. Then, there exists $\underline{m} \in \mathbb{N}$ such that for all $m \geq \underline{m}$, a contest with $n-1$ equal positive prizes and one zero prize is optimal.*

The optimality of a minimally competitive prize profile is robust to heterogeneity even with $n > 2$ agents, as long as the heterogeneity is not too large. Again, an optimal contest will typically ask for different effort levels from different agents, and allocates the zero prize with non-identical probabilities across the agents in equilibrium. While Proposition 4 only states the existence of an optimal contest with $n-1$ identical prizes, it is easy to show that those prizes and the optimal effort levels are characterized by a generalized version of the optimization problem in part (i) of Proposition 3, namely

$$(x^*, e^*) \in \operatorname*{argmax}_{x,e} \sum_{i=1}^{n} e_i - x \quad \text{s.t.} \quad \sum_{i=1}^{n} c_i(e_i) = (n-1)u\left(\frac{x}{n-1}\right).$$

Given the complexity of the problem, we leave the question whether a suitably defined asymmetric nested Tullock CSF can achieve the optimum to future research.

5 Related Literature

A contest is described by two dimensions: the prize profile and the CSF. The contest design literature has typically treated the design of these two dimensions separately. We will first discuss existing results on the optimal prize profile,[14] an then existing results on the optimal CSF.

[14] For a survey on the optimal allocation of prizes in contests see Sisak (2009).

For the class of Tullock CSFs, Clark and Riis (1998) show that, if a symmetric pure-strategy equilibrium exists for a WTA prize structure, then WTA is optimal. More generally, Schweinzer and Segev (2012) argue that prizes should be concentrated on the top as much as possible so that a pure-strategy equilibrium still exists, always under the assumption of risk-neutral agents. Fu, Jiao, and Lu (2015) focus on entry into Tullock contests and also show that a single prize can be optimal. Feng and Lu (2018) study a multi-battle Tullock contest and show that the optimal prize profile depends on the randomness of the CSF. In particular, when randomness is significant, WTA is optimal.

For Lazear-Rosen tournaments, Drugov and Ryvkin (2020b) characterize the optimal prize profile and show that the distribution of noise plays a crucial role. For light-tailed shocks, WTA is optimal, while with heavy-tailed shocks, more equal prize-sharing becomes optimal. For large tournaments, Morgan et al. (2019) show that when the distribution of noise is optimally chosen (see below), any number of equal positive prizes is optimal.

For the class of all-pay CSFs, Fang et al. (2020) show that it is optimal to give equal positive prizes to all agents but one, who receives a zero prize. More generally, their message is that making an all-pay contest less competitive, by decreasing the dispersion in prizes, increases the effort that agents exert. When agents are heterogeneous in an all-pay contest, finding the optimal prize vector becomes difficult. Xiao (2016) shows that a WTA prize profile is in general not optimal. By studying large all-pay contests, Olszewski and Siegel (2020) are able to characterize the optimal prize profile under very general conditions and show that prize sharing is optimal in general. When agents have heterogeneous private types, Moldovanu and Sela (2001) show that WTA is optimal for weakly concave cost functions, but that multiple prizes can be optimal for convex cost functions.

In some settings, the principal can also assign punishments in addition to prizes. Punishments can be effective tools for incentivizing effort in all-pay contests, as shown by Moldovanu, Sela, and Shi (2012) and Liu, Lu, Wang, and Zhang (2018). Similar results for Tullock contests and Lazear-Rosen tournaments can be found in Amiad and Sela (2016) and Akerlof and Holden (2012), respectively.

Most of the papers in this literature assume risk-neutral agents. Risk-aversion makes more equal prize sharing better from the principal's perspective, because it reduces the amount of risk to which the agents are exposed. This was shown by Glazer and Hassin (1988) for all-pay contests, Fu et al. (2019) for Tullock contests, and Drugov and Ryvkin (2019) for Lazear-Rosen tournaments.

Instead of characterizing the optimal prize profile, several papers consider how changes in the CSF affect equilibrium effort, for given prizes. For Tullock contests, Fu et al. (2015) show that increasing randomness leads to more entry into the contest, at the cost of potentially lower effort by the agents who enter. The optimal level of randomness trades off these effects. For two agents in a Tullock contest, Wang (2010) shows that increasing randomness can be an optimal response to more heterogeneous agents. Drugov and Ryvkin (2020a) show that, as a Lazear-Rosen tournament becomes more noisy, equilibrium effort decreases. Morgan et al. (2019) analyze large Lazear-Rosen tournaments where noise is a random variable from the location-scale family. They vary the scale parameter (the randomness of the contest) and find that intermediate levels of randomness are optimal. Olszewski and Siegel (2019) model college admissions as a large all-pay contest and show how treating students with similar results equally, in essence making the all-pay contest more random, can improve outcomes.

The contribution of our paper is to study jointly optimal choice of the prize profile and the CSF. The contest theory literature has developed foundations for various functional forms of the CSF (for a comprehensive survey, see Jia et al., 2013). Our main result contributes to this literature by showing that, even when the principal can freely choose among all conceivable contest success functions, it is optimal to use a Tullock contest.

6 Conclusion

In this paper, we provide a framework which enables us to study the optimal design of contests without being restricted to a single class of contests. We show that when the principal can choose any contest success function, the optimum can be achieved by an appropriately chosen nested Tullock contest. The optimal prize vector features a single zero prize and $n-1$ equal positive prizes. We also show how our framework can be used in cases where the measurement of the agents' performance is imperfect, and when the agents are heterogeneous in their abilities.

Our general message is that optimal contests exhibit a relatively small degree of competitiveness, embodied by a minimally competitive prize profile and an imperfectly discriminating CSF. The optimal degree of competition is achieved when a pure-strategy equilibrium emerges. Reducing competitiveness beyond that point would decrease the efforts that the principal can elicit, and increasing competitiveness would induce wasteful mixing in equilibrium.

We conclude with a discussion of two important questions for future research. First, we have focused on the optimal design of contests when the principal's objective is the maximization of total effort. However, contest mechanisms are also used for other purposes. One important application is to incentivize development of innovations. *Innovation contests* have been used both by governments (for example the 2012 EU Vaccine Prize) and by private firms (such as the 2006 Netflix Prize). In innovation contests, the principal is usually only interested in the best innovation and not in the total effort that the agents have exerted. For this reason, the literature studying innovation contests usually assumes that the objective of the principal is to maximize the highest realization of the agents' outputs.[15] In future work, our framework could be extended to this setting by adjusting the principal's payoff function $\Pi_P(\sigma \mid (y, \mu))$ in problem (P) appropriately.

Second, we have focused on the case where the agents have publicly known types. Future work could extend the framework to the case with private types, as in Moldovanu and Sela (2001, 2006). When agents have private types, contests can also be used as mechanisms for the selection of the best agent. The usual application of such *selection contests* is to promotions within firms.[16] This could be incorporated into our framework by allowing for private types of the agents and again adjusting the objective function of the principal appropriately.

[15] Classical references are Taylor (1995) and Che and Gale (2003), while more recent examples are Erkal and Xiao (2019), Lemus and Temnyalov (2019) and Benkert and Letina (2020). For a similar objective in *prediction contests* see Lemus and Marshall (2019).

[16] For examples of selection contests see Meyer (1991), and Fang and Noe (2019) for a more recent contribution.

References

Akerlof, R. J. and R. T. Holden (2012): "The Nature of Tournaments," *Economic Theory*, 51, 289–313.

Amiad, M. and A. Sela (2016): "The Optimal Allocation of Punishments in Tullock Contests," Mimeo.

Barvinek, E., I. Daler, and J. Francu (1991): "Convergence of Sequences of Inverse Functions," *Archivum Mathematicum*, 27, 201–204.

Benkert, J.-M. and I. Letina (2020): "Designing Dynamic Research Contests," *American Economic Journal: Microeconomics*, forthcoming.

Chang, S.-H., P. C. Cosman, and L. B. Milstein (2011): "Chernoff-Type Bounds for the Gaussian Error Function," *IEEE Transactions on Communications*, 59, 2939–2944.

Che, Y.-K. and I. Gale (2000): "Difference-Form Contests and the Robustness of All-Pay Auctions," *Games and Economic Behavior*, 30, 22–43.

——— (2003): "Optimal Design of Research Contests," *American Economic Review*, 93, 646–671.

Clark, D. J. and C. Riis (1996): "A Multi-Winner Nested Rent-Seeking Contest," *Public Choice*, 87, 177–184.

——— (1998): "Influence and the Discretionary Allocation of Several Prizes," *European Journal of Political Economy*, 14, 605–625.

Cornes, R. and R. Hartley (2005): "Asymmetric Contests with General Technologies," *Economic Theory*, 26, 923–946.

Dixit, A. (1987): "Strategic Behavior in Contests," *American Economic Review*, 77, 891–898.

Drugov, M. and D. Ryvkin (2019): "Optimal Prizes in Tournaments with Risk-Averse Agents," Mimeo.

——— (2020a): "How Noise Affects Effort in Tournaments," *Journal of Economic Theory*, forthcoming.

——— (2020b): "Tournament Rewards and Heavy Tails," CEPR Discussion Paper No. 12368.

Erkal, N. and J. Xiao (2019): "Scarcity of Ideas and Optimal Prizes in Innovation Contests," Mimeo.

Ewerhart, C. (2017): "Revenue Ranking of Optimally Biased Contests: The Case of Two Players," *Economics Letters*, 157, 167–170.

Fang, D. and T. Noe (2019): "Less Competition, More Meritocracy?" Mimeo.

Fang, D., T. Noe, and P. Strack (2020): "Turning Up the Heat: The Discouraging Effect of Competition in Contests," *Journal of Political Economy*, 128, 1940–1975.

Feng, X. and J. Lu (2018): "How to Split the Pie: Optimal Rewards in Dynamic Multi-Battle Competitions," *Journal of Public Economics*, 160, 82–95.

Franke, J., W. Leininger, and C. Wasser (2018): "Optimal Favoritism in All-Pay Auctions and Lottery Contests," *European Economic Review*, 104, 22–37.

Fu, Q., Q. Jiao, and J. Lu (2015): "Contests with Endogenous Entry," *International Journal of Game Theory*, 44, 387–424.

Fu, Q., X. Wang, and Z. Wu (2019): "Multi-Prize Contests with Risk-Averse Players," Mimeo.

Glazer, A. and R. Hassin (1988): "Optimal Contests," *Economic Inquiry*, 26, 133–143.

Green, J. R. and N. L. Stokey (1983): "A Comparison of Tournaments and Contracts," *Journal of Political Economy*, 91, 349–364.

Hirshleifer, J. (1989): "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success," *Public Choice*, 63, 101–112.

Jia, H., S. Skaperdas, and S. Vaidya (2013): "Contest Functions: Theoretical Foundations and Issues in Estimation," *International Journal of Industrial Organization*, 31, 211–222.

Konrad, K. A. (2009): *Strategy and Dynamics in Contests*, Oxford University Press.

Lazear, E. P. and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841–864.

Lemus, J. and G. Marshall (2019): "Dynamic Tournament Design: An Application to Prediction Contests," Mimeo.

Lemus, J. and E. Temnyalov (2019): "Diversification and Information in Contests," Mimeo.

Letina, I., S. Liu, and N. Netzer (2018): "Delegating Performance Evaluation," University of Zurich, Department of Economics, Working Paper No. 266.

——— (2020): "Delegating Performance Evaluation," *Theoretical Economics*, 15, 477–509.

Liu, B., J. Lu, R. Wang, and J. Zhang (2018): "Optimal Prize Allocation in Contests: The Role of Negative Prizes," *Journal of Economic Theory*, 175, 291–317.

Meyer, M. A. (1991): "Learning from Coarse Information: Biased Contests and Career Profiles," *The Review of Economic Studies*, 58, 15–41.

Moldovanu, B. and A. Sela (2001): "The Optimal Allocation of Prizes in Contests," *American Economic Review*, 91, 542–558.

——— (2006): "Contest Architecture," *Journal of Economic Theory*, 1, 70–96.

Moldovanu, B., A. Sela, and X. Shi (2012): "Carrots and Sticks: Prizes and Punishments in Contests," *Economic Inquiry*, 50, 453–462.

Morgan, J., J. Tumlinson, and F. Várdy (2019): "The Limits of Meritocracy," Mimeo.

Nalebuff, B. J. and J. E. Stiglitz (1983): "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *The Bell Journal of Economics*, 14, 21–43.

Olszewski, W. and R. Siegel (2016): "Large Contests," *Econometrica*, 84, 835–854.

——— (2019): "Pareto Improvements in the Contest for College Admissions," Mimeo.

——— (2020): "Performance-Maximizing Large Contests," *Theoretical Economics*, 15, 57–88.

Pinelis, I. (2002): "L'Hospital Type Rules for Monotonicity, with Applications," *Journal of Inequalities in Pure and Applied Mathematics*, 3, article 5, 5 pp. (electronic).

Schweinzer, P. and E. Segev (2012): "The Optimal Prize Structure of Symmetric Tullock Contests," *Public Choice*, 153, 69–82.

Siegel, R. (2009): "All-Pay Contests," *Econometrica*, 77, 71–92.

Sisak, D. (2009): "Multiple-prize Contests – The Optimal Allocation of Prizes," *Journal of Economic Surveys*, 23, 82–114.

Skaperdas, S. (1996): "Contest Success Functions," *Economic Theory*, 7, 283–290.

Taylor, C. R. (1995): "Digging for Golden Carrots: An Analysis of Research Tournaments," *American Economic Review*, 85, 872–890.

Wang, Z. (2010): "The Optimal Accuracy Level in Asymmetric Contests," *The B.E. Journal of Theoretical Economics*, 10, 1–18.

Xiao, J. (2016): "Asymmetric All-Pay Contests with Heterogeneous Prizes," *Journal of Economic Theory*, 163, 178–221.

A Proofs

A.1 Proof of Proposition 1

We first state without proof a result which is due to Letina et al. (2020), and which characterizes the optimal effort profile as well as all optimal prize profiles.[17] Statement (i) of Proposition 1 follows immediately from this lemma.

Lemma 1 (Letina et al., 2020) *A contest is optimal if and only if it satisfies conditions (i) and (ii):*

(i) The prizes satisfy $y_n^* = 0$ *and* $\sum_{k=1}^{n} y_k^* = x^*$*, where* x^* *is given by*

$$u'\left(\frac{x^*}{n-1}\right) = c'\left(c^{-1}\left(\frac{n-1}{n}u\left(\frac{x^*}{n-1}\right)\right)\right).$$

If the agents are risk-averse, then the prize profile is unique and given by

$$y^* = (x^*/(n-1), \dots, x^*/(n-1), 0).$$

(ii) The contest implements $(e^*, \dots, e^*)$*, where* e^* *is given by*

$$e^* = c^{-1}\left(\frac{n-1}{n}u\left(\frac{x^*}{n-1}\right)\right).$$

Consider now a contest with prize profile $y^* = (x^*/(n-1), ..., x^*/(n-1), 0)$ and CSF μ^* of the nested Tullock form (2). We will show that, for an appropriate choice of f, the effort profile $(e^*, \dots, e^*)$ is an equilibrium. The proof proceeds in three steps. In Step 1, we derive the agents' payoff function in the nested contest. Step 2 introduces the specific value $r^*(n)$ stated in the proposition. In Step 3, we then complete the proof that the resulting contest indeed implements the desired effort profile.

Step 1. Let $p(e_i)$ denote the probability that agent i wins none of the $n-1$ positive prizes, given that all other agents exert effort e^*. Furthermore, let u^* be the utility derived from a positive prize. Then, the expected payoff of agent i, when all other agents exert e^*, is given by

$$\begin{aligned}\Pi_i(e_i) &= [1 - p(e_i)]\, u^* - c(e_i) \\ &= \left[1 - \frac{(n-1)!f(e^*)^{n-1}}{\prod_{k=1}^{n-1}[f(e_i) + (n-k)f(e^*)]}\right] u^* - c(e_i)\end{aligned}$$

[17]The result in Letina et al. (2020) is more general as it allows for a possibly binding budget constraint of the principal.

$$= \left[1 - \prod_{k=1}^{n-1} \frac{(n-k)f(e^*)}{[f(e_i) + (n-k)f(e^*)]}\right] u^* - c(e_i)$$

$$= \left[1 - \prod_{k=1}^{n-1} \frac{(n-k)f(e^*)}{[f(e_i) + (n-k)f(e^*)]}\right] \left(\frac{n}{n-1}\right) c(e^*) - c(e_i).$$

Now suppose $f(e_i) = c(e_i)^r$ for some $r \geq 0$. It is easy to see that $\Pi_i(0) = \Pi_i(e^*) = 0$ for any r. We will show in the next two steps that $\Pi_i(e_i) \leq 0$ for all e_i when $r = r^*(n) = (n-1)/(H_n - 1)$, where $H_n = \sum_{k=1}^{n} 1/k$ is the n-th harmonic number. This implies that $(e^*, \dots, e^*)$ is an equilibrium.

Step 2. Consider any $e_i > 0$ (we already know the value of Π_i for $e_i = 0$). To determine the sign of $\Pi_i(e_i)$, we can equivalently examine the sign of

$$\Pi_i(e_i) \left[\frac{n-1}{nc(e^*)}\right] = \left[1 - \prod_{k=1}^{n-1} \frac{(n-k)c(e^*)^r}{[c(e_i)^r + (n-k)c(e^*)^r]}\right] - \left(\frac{n-1}{n}\right) \frac{c(e_i)}{c(e^*)}.$$

Make the change of variables $y^* = c(e^*)^r$ and $y = c(e_i)^r$ to obtain

$$F(y|r) := \left[1 - \prod_{k=1}^{n-1} \frac{(n-k)y^*}{[y + (n-k)y^*]}\right] - \frac{n-1}{n} \left(\frac{y}{y^*}\right)^{\frac{1}{r}}.$$

After the additional variable substitution $x = y^*/y$ we obtain

$$F(x|r) := \left[1 - \prod_{k=1}^{n-1} \frac{(n-k)x}{[1 + (n-k)x]}\right] - \frac{n-1}{n} \left(\frac{1}{x}\right)^{\frac{1}{r}}.$$

Showing that $F(x|r) \leq 0$ for all $x > 0$, $x \neq 1$, is then sufficient to ensure that the contest with parameter r implements the optimum.

Fix any x and let us look for $r(x)$ such that $F(x|r(x)) = 0$. Since F is strictly increasing in r whenever $x \in (0, 1)$, we obtain that $F(x|r) \leq 0$ for any fixed $x \in (0, 1)$ whenever $r \leq r(x)$, so $r(x)$ gives an upper bound on the possible values of r. Similarly, since F is strictly decreasing in r whenever $x \in (1, \infty)$, we obtain that $F(x|r) \leq 0$ for any fixed $x \in (1, \infty)$ whenever $r \geq r(x)$, so $r(x)$ gives a lower bound on the possible values of r. Thus it is sufficient to find a value r^* such that $r(x) \geq r^*$ for all $x \in (0, 1)$ and $r(x) \leq r^*$ for all $x \in (1, \infty)$.

Rewriting the equation $F(x|r(x)) = 0$, we have

$$\left[1 - \prod_{k=1}^{n-1} \frac{(n-k)x}{[1 + (n-k)x]}\right] = \frac{n-1}{n} \left(\frac{1}{x}\right)^{\frac{1}{r(x)}}$$

$$\log\left[1-\prod_{k=1}^{n-1}\frac{(n-k)x}{[1+(n-k)x]}\right]=\log\left(\frac{n-1}{n}\right)-\frac{1}{r(x)}\log(x)$$

$$\frac{1}{r(x)}\log(x)=\log\left(\frac{n-1}{n}\right)-\log\left[1-\frac{(n-1)!x^{n-1}}{\prod_{k=1}^{n-1}[1+(n-k)x]}\right]$$

$$\frac{1}{r(x)}\log(x)=\log\left[\frac{n-1}{n}\frac{\prod_{k=1}^{n-1}[1+(n-k)x]}{\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}}\right]$$

$$r(x)=\frac{\log(x)}{\log\left[\frac{n-1}{n}\frac{\prod_{k=1}^{n-1}[1+(n-k)x]}{\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}}\right]}.$$

Denote

$$g(x)=\frac{n-1}{n}\frac{\prod_{k=1}^{n-1}[1+(n-k)x]}{\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}}$$

so that

$$r(x)=\frac{\log(x)}{\log(g(x))}.$$

Note that $g(x)>0$ for any $x>0$. We will first show that $\lim_{x\nearrow 1}r(x)=\lim_{x\searrow 1}r(x)=r^*(n)=(n-1)/(H_n-1)$. Note that for $x=1$ both the denominator and the numerator of $r(x)$ equal zero. Hence we use l'Hôpital's rule. Observe that

$$\begin{aligned}
(\log(g(x)))' &= \frac{g'(x)}{g(x)}\\
&= \frac{\left(\frac{\partial}{\partial x}\prod_{k=1}^{n-1}[1+(n-k)x]\right)\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]}\\
&\quad-\frac{\left(\prod_{k=1}^{n-1}[1+(n-k)x]\right)\frac{\partial}{\partial x}\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]}\\
&= \frac{\left(\prod_{k=1}^{n-1}[1+(n-k)x]\right)\frac{\partial}{\partial x}\left((n-1)!x^{n-1}\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]}\\
&\quad-\frac{\left((n-1)!x^{n-1}\right)\left(\frac{\partial}{\partial x}\prod_{k=1}^{n-1}[1+(n-k)x]\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]}\\
&= \frac{\left(\prod_{k=1}^{n-1}[1+(n-k)x]\right)(n-1)\left((n-1)!x^{n-2}\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]}
\end{aligned}$$

$$-\frac{((n-1)!x^{n-1})\left(\sum_{k=1}^{n-1}(n-k)\prod_{j\neq k}[1+(n-j)x]\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]}.$$

We evaluate this at $x=1$, that is,

$$\begin{aligned}
(\log(g(x)))'|_{x=1} &= \frac{\left(\prod_{k=1}^{n-1}[1+(n-k)]\right)(n-1)(n-1)!}{\left(\prod_{k=1}^{n-1}[1+(n-k)]-(n-1)!\right)\prod_{k=1}^{n-1}[1+(n-k)]} \\
&\quad -\frac{(n-1)!\left(\sum_{k=1}^{n-1}(n-k)\prod_{j\neq k}[1+(n-j)]\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)]-(n-1)!\right)\prod_{k=1}^{n-1}[1+(n-k)]} \\
&= \frac{n!(n-1)(n-1)!}{(n!-(n-1)!)\,n!} \\
&\quad -\frac{(n-1)!\left(\sum_{k=1}^{n-1}(n-k)\prod_{j\neq k}[1+(n-j)]\right)}{(n!-(n-1)!)\,n!} \\
&= 1-\frac{\left(\sum_{k=1}^{n-1}(n-k)\prod_{j\neq k}[1+(n-j)]\right)}{(n-1)\,n!} \\
&= 1-\frac{n!\left(\sum_{k=1}^{n-1}\frac{n-k}{n-k+1}\right)}{(n-1)\,n!} \\
&= \frac{n-1-\left(\sum_{k=1}^{n-1}\frac{n-k}{n-k+1}\right)}{n-1} \\
&= \frac{1+\sum_{k=1}^{n-1}\frac{n-k+1}{n-k+1}-\sum_{k=1}^{n-1}\frac{n-k}{n-k+1}-1}{n-1} \\
&= \frac{1+\sum_{k=1}^{n-1}\frac{1}{n-k+1}-1}{n-1} \\
&= \frac{H_n-1}{n-1}.
\end{aligned}$$

Thus we have

$$\lim_{x\nearrow 1} r(x) = \lim_{x\searrow 1} r(x) = \left.\frac{1/x}{(\log(g(x)))'}\right|_{x=1} = \frac{n-1}{H_n-1}.$$

To complete the proof of Proposition 1 , it is now sufficient to show that $r(x)$ is weakly monotonically decreasing on $(0,1)$ and on $(1,\infty)$. We will do this in the next step.

Step 3. To show monotonicity of $r(x)$, we will apply a suitable version of the l'Hôpital monotone rule. Proposition 1.1 in Pinelis (2002) (together with Corollary 1.2 and Remark

1.3) implies that $r(x) = \log(x)/\log(g(x))$ is weakly decreasing on $(0, 1)$ and $(1, \infty)$ if

$$\frac{(\log(x))'}{(\log(g(x)))'} = \frac{g(x)}{xg'(x)}$$

is weakly decreasing.[18] We will thus show that

$$\left(\frac{g(x)}{xg'(x)}\right)' = \frac{[g'(x)x - g(x)]g'(x) - xg(x)g''(x)}{(xg'(x))^2} \leq 0.$$

For this, it is sufficient to show the following three conditions:

(*a*) $g'(x) > 0$,

(*b*) $g''(x) \geq 0$,

(*c*) $g'(x)x - g(x) \leq 0$.

We will verify these conditions in the following three lemmas. To do this, consider the function g. We can write

$$\begin{aligned}\prod_{k=1}^{n-1}[1 + (n-k)x] &= (n-1)!x^{n-1} + a_{n-2}x^{n-2} + a_{n-3}x^{n-3} + \cdots + a_1 x + 1 \\ &= (n-1)!x^{n-1} + \gamma(x),\end{aligned}$$

where $a_1, \ldots, a_{n-2}$ are strictly positive coefficients (that depend on n), so that γ is a polynomial of degree $n-2$ which is strictly positive for all $x > 0$.[19] We can then rewrite

$$g(x) = \frac{n-1}{n}\frac{(n-1)!x^{n-1} + \gamma(x)}{\gamma(x)}.$$

Lemma 2 *Condition $g'(x) > 0$ is satisfied.*

Proof. Observe that

$$\begin{aligned}g'(x) &= \frac{n-1}{n}\frac{(n-1)(n-1)!x^{n-2}\gamma(x) - (n-1)!x^{n-1}\gamma'(x)}{\gamma(x)^2} \\ &= \frac{n-1}{n}\frac{(n-1)!x^{n-2}[(n-1)\gamma(x) - x\gamma'(x)]}{\gamma(x)^2},\end{aligned}$$

[18] Proposition 1.1 is applicable because $\log(x)$ and $\log(g(x))$ are differentiable on the respective intervals and $\lim_{x\to 1}\log(x) = \lim_{x\to 1}\log(g(x)) = 0$ holds. The remaining prerequisite $(\log(g(x)))' = g'(x)/g(x) > 0$ also holds, because $g(x) > 0$ and $g'(x) > 0$ according to Lemma 2 below.

[19] To avoid confusion, the formula should be read as $\gamma(x) = 1$ for $n = 2$ and as $\gamma(x) = a_1 x$ for $n = 3$.

and, since

$$(n-1)\gamma(x) = (n-1)a_{n-2}x^{n-2} + (n-1)a_{n-3}x^{n-3} + \ldots + (n-1)a_1 x + n - 1 \text{ and}$$
$$x\gamma'(x) = (n-2)a_{n-2}x^{n-2} + (n-3)a_{n-3}x^{n-3} + \ldots + a_1 x,$$

it follows that $(n-1)\gamma(x) - x\gamma'(x) > 0$, which implies that $g'(x) > 0$. □

Lemma 3 *Condition $g''(x) \geq 0$ is satisfied.*

Proof. Observe that

$$g''(x) = \frac{(n-1)(n-1)!}{n}\left[\frac{(n-1)x^{n-2}\gamma(x) - x^{n-1}\gamma'(x)}{\gamma(x)^2}\right]',$$

so that $g''(x) \geq 0$ is equivalent to

$$\begin{aligned}
0 \leq &\left[\frac{(n-1)x^{n-2}\gamma(x) - x^{n-1}\gamma'(x)}{\gamma(x)^2}\right]' \\
= &\frac{[(n-2)(n-1)x^{n-3}\gamma(x) + (n-1)x^{n-2}\gamma'(x) - (n-1)x^{n-2}\gamma'(x) - x^{n-1}\gamma''(x)]\gamma(x)^2}{\gamma(x)^4} \\
&- \frac{[(n-1)x^{n-2}\gamma(x) - x^{n-1}\gamma'(x)]2\gamma(x)\gamma'(x)}{\gamma(x)^4} \\
= &\frac{[(n-2)(n-1)x^{n-3}\gamma(x) - x^{n-1}\gamma''(x)]\gamma(x)^2}{\gamma(x)^4} \\
&- \frac{[(n-1)x^{n-2}\gamma(x) - x^{n-1}\gamma'(x)]2\gamma(x)\gamma'(x)}{\gamma(x)^4} \\
= &\frac{\gamma(x)x^{n-3}}{\gamma(x)^4}\left[(n-2)(n-1)\gamma(x)^2 - x^2\gamma''(x)\gamma(x) - 2(n-1)x\gamma(x)\gamma'(x) + 2x^2\gamma'(x)^2\right].
\end{aligned}$$

The expression in the square bracket is a polynomial of degree $(2n-4)$. We will show that all coefficients of this polynomial are positive, which implies that the polynomial, and hence also $g''(x)$, is non-negative.

Using the auxiliary definitions $a_0 = 1$ and $a_\kappa = 0$ for $\kappa < 0$, the coefficient multiplying x^{2n-j} in this polynomial, for any $4 \leq j \leq 2n$, is given by

$$\begin{aligned}
&\sum_{k=2}^{j-2}(n-2)(n-1)a_{n-k}a_{n-j+k} - \sum_{k=2}^{j-2}(n-k)(n-k-1)a_{n-k}a_{n-j+k} \\
&\quad - \sum_{k=2}^{j-2}2(n-1)(n-k)a_{n-k}a_{n-j+k} + \sum_{k=2}^{j-2}2(n-k)(n-j+k)a_{n-k}a_{n-j+k} \\
= &\sum_{k=2}^{j-2}(n^2-3n+2)a_{n-k}a_{n-j+k} - \sum_{k=2}^{j-2}(n^2-2nk-n+k^2+k)a_{n-k}a_{n-j+k}
\end{aligned}$$

$$
- \sum_{k=2}^{j-2} 2(n^2 - nk - n + k) a_{n-k} a_{n-j+k} + \sum_{k=2}^{j-2} 2(n^2 - nj + jk - k^2) a_{n-k} a_{n-j+k}
$$

$$
= \sum_{k=2}^{j-2} (2 + 4nk - 3k^2 - 3k - 2nj + 2jk) a_{n-k} a_{n-j+k}.
$$

Let $\varphi(n, j, k) = 2 + 4nk - 3k^2 - 3k - 2nj + 2jk$. We will show in several steps that $\sum_{k=2}^{j-2} \varphi(n, j, k) a_{n-k} a_{n-j+k} \geq 0$. For $n = 2$ and $n = 3$, this condition can easily be verified directly. Hence we suppose that $n > 3$ from now on.

Observe that for any k there is $k' = j - k$ such that $a_{n-k} a_{n-j+k} = a_{n-k'} a_{n-j+k'}$. Hence we first consider the case where j is odd, so that we can write

$$
\sum_{k=2}^{j-2} \varphi(n, j, k) a_{n-k} a_{n-j+k} = \sum_{k=2}^{\frac{j-1}{2}} [\varphi(n, j, k) + \varphi(n, j, j - k)] a_{n-k} a_{n-j+k}.
$$

Since $\varphi(n, j, k) + \varphi(n, j, j - k)$ is an integer, we can think of this expression as a long sum where each of the terms $a_{n-k} a_{n-j+k}$ appears exactly $|\varphi(n, j, k) + \varphi(n, j, j - k)|$ times, added or subtracted depending on the sign of $\varphi(n, j, k) + \varphi(n, j, j - k)$. Now note that $\sum_{k=2}^{(j-1)/2} [\varphi(n, j, k) + \varphi(n, j, j - k)] = 0$ holds. This follows because we can write

$$
\begin{aligned}
&\sum_{k=2}^{\frac{j-1}{2}} [\varphi(n, j, k) + \varphi(n, j, j - k)] \\
&= \sum_{k=2}^{j-2} \varphi(n, j, k) \\
&= \sum_{k=2}^{j-2} (2 - 2nj) + (4n - 3 + 2j) \sum_{k=2}^{j-2} k - 3 \sum_{k=2}^{j-2} k^2 \\
&= (j - 3)(2 - 2nj) + (4n - 3 + 2j) \frac{j(j - 3)}{2} - 3 \frac{(j - 3)(2j^2 - 3j + 4)}{6} \\
&= (j - 3) \left(2 - 2nj + 2nj - \frac{3j}{2} + j^2 - j^2 + \frac{3j}{2} - 2 \right) \\
&= 0.
\end{aligned}
$$

Thus, for each instance where a term $a_{n-k'} a_{n-j+k'}$ is subtracted in the long sum, we can find a term $a_{n-k''} a_{n-j+k''}$ which is added. We claim that the respective terms which are added are weakly larger than the terms which are subtracted. This claim follows once we show that both $\varphi(n, j, k) + \varphi(n, j, j - k)$ and $a_{n-k} a_{n-j+k}$ are weakly increasing in k within the range of the sum. In that case, the terms which are subtracted are those for small k and the terms which are added are those for large k, and the latter are weakly larger. The same argument in fact applies when j is even, so that we can write

$$\sum_{k=2}^{j-2} \varphi(n,j,k) a_{n-k} a_{n-j+k}$$
$$= \sum_{k=2}^{\frac{j-2}{2}} [\varphi(n,j,k) + \varphi(n,j,j-k)] a_{n-k} a_{n-j+k} + \varphi(n,j,j/2) a_{n-j/2}^2.$$

Importantly, for the last term we have

$$\begin{aligned}
\varphi(n,j,j/2) &= 2 - 2nj - 3\left(\frac{j}{2}\right)^2 + \frac{j}{2}(4n - 3 + 2j) \\
&= 2 - j^2\frac{3}{4} - j\frac{3}{2} + j^2 \\
&= 2 + j\left(\frac{j}{4} - \frac{3}{2}\right) \\
&> 0,
\end{aligned}$$

so that the last and largest term $a_{n-j/2}^2 = a_{n-j/2} a_{n-j/2}$ is indeed also added.

We first show that $\varphi(n,j,k) + \varphi(n,j,j-k)$ is weakly increasing in k in the relevant range. We have

$$\begin{aligned}
&\varphi(n,j,k) + \varphi(n,j,j-k) \\
&= (2 - 2nj - 3k^2 + k(4n - 3 + 2j)) + (2 - 2nj - 3(j-k)^2 + (j-k)(4n - 3 + 2j)) \\
&= 4 - 4nj - 3(2k^2 + j^2 - 2jk) + j(4n - 3 + 2j).
\end{aligned}$$

Treating k as a real variable, we obtain

$$\begin{aligned}
\frac{\partial}{\partial k}\left[\varphi(n,j,k) + \varphi(n,j,j-k)\right] &= -3(4k - 2j) \\
&= -6(2k - j) > 0
\end{aligned}$$

for all $k < j/2$, so the claim follows.

We now show that $a_{n-k} a_{n-j+k}$ is weakly increasing in k in the relevant range. Formally, we show that $a_{n-k} a_{n-j+k} \leq a_{n-k-1} a_{n-j+k+1}$ for any $k < j/2$. Observe that we can write

$$\begin{aligned}
a_1 &= \sum_{k_1=1}^{n-1} (n - k_1), \\
a_2 &= \sum_{k_2=1}^{n-2} \sum_{k_1=k_2+1}^{n-1} (n - k_2)(n - k_1),
\end{aligned}$$

$$\vdots$$

$$a_j = \sum_{k_j=1}^{n-j} \sum_{k_{j-1}=k_j+1}^{n-j+1} \cdots \sum_{k_1=k_2+1}^{n-1} (n-k_j)(n-k_{j-1})\dots(n-k_1).$$

Intuitively, each summand in the definition of a_j is the product of j different elements chosen from the set $\{(n-1),(n-2),\dots,1\}$, and the nested summation goes over all the different possibilities in which these j elements can be chosen. Using simplified notation for the nested summation, we can thus write (where α, β, λ, and η take the role of the indices of summation, like k in the expression above):

$$\begin{aligned} a_{n-k} &= \sum (n-\alpha_{n-k})(n-\alpha_{n-k-1})\dots(n-\alpha_1), \\ a_{n-j+k} &= \sum (n-\beta_{n-j+k})(n-\beta_{n-j+k-1})\dots(n-\beta_1), \\ a_{n-k-1} &= \sum (n-\lambda_{n-k-1})(n-\lambda_{n-k-2})\dots(n-\lambda_1), \\ a_{n-j+k+1} &= \sum (n-\eta_{n-j+k+1})(n-\eta_{n-j+k})\dots(n-\eta_1). \end{aligned}$$

Rewriting the inequality $a_{n-k}a_{n-j+k} \leq a_{n-k-1}a_{n-j+k+1}$ using this notation, we obtain

$$\begin{aligned} &\sum (n-\alpha_{n-k})(n-\alpha_{n-k-1})\dots(n-\alpha_1)(n-\beta_{n-j+k})(n-\beta_{n-j+k-1})\dots(n-\beta_1) \\ &\leq \sum (n-\lambda_{n-k-1})(n-\lambda_{n-k-2})\dots(n-\lambda_1)(n-\eta_{n-j+k+1})(n-\eta_{n-j+k})\dots(n-\eta_1). \end{aligned}$$

Observe that each summand of the LHS sum is the product of $(n-k)+(n-j+k) = 2n-j$ elements, all of them chosen from the set $\{(n-1),(n-2),\dots,1\}$. The first $n-k$ elements are all different from each other, and the last $n-j+k$ elements are all different from each other. Thus, since $n-k > n-j+k$ when $k < j/2$, in each summand at most $n-j+k$ elements can appear twice. Furthermore, the LHS sum goes over all the different combinations that satisfy this property. Similarly, each summand of the RHS sum is the product of $(n-k-1)+(n-j+k+1) = 2n-j$ elements, all of them chosen from the same set $\{(n-1),(n-2),\dots,1\}$. The first $n-k-1$ elements are all different from each other, and the last $n-j+k+1$ elements are all different from each other. Thus, (weakly) more than $n-j+k$ elements can appear twice in these summands.[20] Since the RHS sum goes over all the different combinations that satisfy this property, for each summand on the LHS there exists an equal summand on the RHS. This shows that the inequality indeed holds. □

Lemma 4 *Condition $g'(x)x - g(x) \leq 0$ is satisfied.*

[20]The inequality $n-k-1 \geq n-j+k+1$ can be rearranged to $k \leq j/2-1$, which follows from $k < j/2$, except if j is odd and $k = (j-1)/2$. Thus, typically, up to $n-j+k+1$ elements can appear twice. If j is odd and $k = (j-1)/2$, up to $n-k-1$ elements can appear twice, which is identical to $n-j+k$ in that case.

Proof. We have

$$g'(x)x - g(x) = \frac{n-1}{n}\left[\frac{(n-1)!x^{n-1}[(n-1)\gamma(x) - x\gamma'(x)]}{\gamma(x)^2} - \frac{(n-1)!x^{n-1} + \gamma(x)}{\gamma(x)}\right],$$

and therefore $g'(x)x - g(x) \leq 0$ if and only if

$$\begin{aligned}
0 &\geq (n-1)!x^{n-1}[(n-1)\gamma(x) - x\gamma'(x)] - (n-1)!x^{n-1}\gamma(x) - \gamma(x)^2 \\
&= (n-1)!x^{n-1}(n-2)\gamma(x) - (n-1)!x^n\gamma'(x) - \gamma(x)^2 \\
&= (n-1)![(n-2)a_{n-2}x^{2n-3} + (n-2)a_{n-3}x^{2n-4} + \cdots + (n-2)a_1x^n + (n-2)x^{n-1} \\
&\quad - (n-2)a_{n-2}x^{2n-3} - (n-3)a_{n-3}x^{2n-4} - \cdots - a_1x^n] - \gamma(x)^2 \\
&= (n-1)![a_{n-3}x^{2n-4} + 2a_{n-4}x^{2n-5} + \cdots + (n-3)a_1x^n + (n-2)x^{n-1}] - \gamma(x)^2 \\
&= (n-1)![a_{n-3}x^{2n-4} + 2a_{n-4}x^{2n-5} + \cdots + (n-3)a_1x^n + (n-2)x^{n-1}] \\
&\quad - \sum_{j=4}^{n+1}\sum_{k=2}^{j-2} a_{n-k}a_{n-j+k}x^{2n-j} - \rho,
\end{aligned}$$

where $\rho \geq 0$ is some positive remainder of $\gamma(x)^2$. To show $g'(x)x - g(x) \leq 0$, it is therefore sufficient to ignore ρ and show that the overall coefficient on x^{2n-j} in the last expression is not positive. That is, it is sufficient to show that, for all $j \in \{4, \ldots, n+1\}$,

$$(n-1)!(j-3)a_{n-j+1} - \sum_{k=2}^{j-2} a_{n-k}a_{n-j+k} \leq 0.$$

Observe that the sum has exactly $(j-3)$ elements. Then, it is sufficient to show that, for all $k \in \{2, \ldots, j-2\}$,

$$(n-1)!a_{n-j+1} \leq a_{n-k}a_{n-j+k}. \tag{3}$$

To demonstrate condition (3), we will first write the values of the coefficients a_j in a different way. Instead of summing over all possibilities in which j different elements from the set $\{(n-1), (n-2), \ldots, 1\}$ can be chosen, we can sum over the $n-j-1$ elements not chosen, and divide the factorial $(n-1)!$ by the product of these elements. This yields

$$\begin{aligned}
a_{n-2} &= \sum_{k_1=1}^{n-1} \frac{(n-1)!}{n-k_1}, \\
a_{n-3} &= \sum_{k_2=1}^{n-2}\sum_{k_1=k_2+1}^{n-1} \frac{(n-1)!}{(n-k_2)(n-k_1)}, \\
&\vdots
\end{aligned}$$

$$a_{n-j} = \sum_{k_{j-1}=1}^{n-j+1} \sum_{k_{j-2}=k_{j-1}+1}^{n-j+2} \cdots \sum_{k_1=k_2+1}^{n-1} \frac{(n-1)!}{(n-k_{j-1})(n-k_{j-2})\dots(n-k_1)},$$

$$\vdots$$

$$a_1 = \sum_{k_{n-2}=1}^{2} \sum_{k_{n-3}=k_{n-2}+1}^{3} \cdots \sum_{k_1=k_2+1}^{n-1} \frac{(n-1)!}{(n-k_{n-2})(n-k_{n-3})\dots(n-k_1)}.$$

Rewriting condition (3), we then have

$$\sum_{\lambda_{j-2}=1}^{n-j+2} \sum_{\lambda_{j-3}=\lambda_{j-2}+1}^{n-j+3} \cdots \sum_{\lambda_1=\lambda_2+1}^{n-1} \frac{((n-1)!)^2}{(n-\lambda_{j-2})(n-\lambda_{j-3})\dots(n-\lambda_1)}$$
$$\leq \left[\sum_{\alpha_{k-1}=1}^{n-k+1} \sum_{\alpha_{k-2}=\alpha_{k-1}+1}^{n-k+2} \cdots \sum_{\alpha_1=\alpha_2+1}^{n-1} \frac{(n-1)!}{(n-\alpha_{k-1})(n-\alpha_{k-2})\dots(n-\alpha_1)}\right]$$
$$\times \left[\sum_{\beta_{j-k-1}=1}^{n-j+k+1} \sum_{\beta_{j-k-2}=\beta_{j-k-1}+1}^{n-j+k+2} \cdots \sum_{\beta_1=\beta_2+1}^{n-1} \frac{(n-1)!}{(n-\beta_{j-k-1})(n-\beta_{j-k-2})\dots(n-\beta_1)}\right].$$

Observe that for each summand on the LHS, the denominator is a product of $j-2$ different elements from the set $\{(n-1), (n-2), \dots, 1\}$. In fact, the LHS sum goes over all the different possibilities in which these $j-2$ elements can be chosen. On the RHS, after multiplication, the denominator of each summand is a product of $(k-1)+(j-k-1) = j-2$ elements from the same set, where replication of some elements may be possible (but is not necessary). Since the RHS sum goes over all these different possibilities, for each summand on the LHS there exists an equal summand on the RHS. This shows that the inequality holds. □ ■

A.2 Comparative Statics of $r^*(n)$

Claim 1: $r^*(n+1) > r^*(n)$. Consider any $n \geq 2$. By definition of $r^*(n)$ we have

$$r^*(n) = \frac{n-1}{H_n - 1} = \frac{(n-1)(H_{n+1}-1)}{(H_n-1)(H_{n+1}-1)},$$
$$r^*(n+1) = \frac{n}{H_{n+1}-1} = \frac{n(H_n-1)}{(H_{n+1}-1)(H_n-1)}.$$

Since $H_n - 1 > 0$ for any $n \geq 2$, $r^*(n+1) > r^*(n)$ holds if and only if

$$n(H_n - 1) - (n-1)(H_{n+1} - 1) > 0.$$

We have

$$n(H_n - 1) - (n-1)(H_{n+1} - 1) = n(H_n - H_{n+1}) + H_{n+1} - 1$$

$$
\begin{aligned}
&= -\frac{n}{n+1} + \sum_{k=2}^{n+1} \frac{1}{k} \\
&= \sum_{k=2}^{n+1} \left(\frac{1}{k} - \frac{1}{n+1} \right) > 0.
\end{aligned}
$$

Claim 2: $\lim_{n\to\infty} r^*(n) = \infty$.

$$
\begin{aligned}
\lim_{n\to\infty} r^*(n) &= \lim_{n\to\infty} \frac{n-1}{H_n - 1} \\
&= \lim_{n\to\infty} \frac{n-(n-1)}{(H_{n+1}-1)-(H_n-1)} \\
&= \lim_{n\to\infty} \frac{1}{1/(n+1)} \\
&= \infty,
\end{aligned}
$$

where the second equality follows from the Stolz-Cesàro Theorem.

A.3 Proof of Proposition 2

Fix an arbitrary observational structure (S, η) and suppose that a contest (y, τ) with prize profile $y = (x^*/(n-1), \ldots, x^*/(n-1), 0)$ implements $(e^*, ..., e^*)$. By contradiction, assume that (y, τ) is not optimal, i.e., there exists a contest $(\tilde{y}, \tilde{\tau})$ that implements some strategy profile σ and

$$
\begin{aligned}
\Pi_P(\sigma \mid (\tilde{y}, \tilde{\tau})) &= \mathbb{E}_\sigma \left[\sum_{i=1}^n e_i \right] - \sum_{i=1}^n \tilde{y}_i \\
&> \Pi_P((e^*, \ldots, e^*) \mid (y, \tau)) = ne^* - x^*.
\end{aligned}
$$

Construct a contest $(\tilde{y}, \mu)$ for the setting with perfect observation by defining

$$
\mu^e(t) = \mathbb{E}_{\eta^e} \left[\tilde{\tau}^s(t) \right]
$$

for all $t \in T(\tilde{y})$ and all $e \in E$. We then obtain

$$
\begin{aligned}
\Pi_i(\sigma' \mid (\tilde{y}, \mu)) &= \mathbb{E}_{\sigma'} \left[\mathbb{E}_{\mu^e} \left[u(t_i) \right] \right] - \mathbb{E}_{\sigma_i'} \left[c(e_i) \right] \\
&= \mathbb{E}_{\sigma'} \left[\mathbb{E}_{\eta^e} \left[\mathbb{E}_{\tilde{\tau}^s} \left[u(t_i) \right] \right] \right] - \mathbb{E}_{\sigma_i'} \left[c(e_i) \right] \\
&= \Pi_i(\sigma' \mid (\tilde{y}, \tilde{\tau}))
\end{aligned}
$$

for all profiles σ' and all $i \in I$. Hence $(\tilde{y}, \mu)$ implements σ with perfect observation because $(\tilde{y}, \tilde{\tau})$ implements σ with imperfect observation. We obtain

$$\begin{aligned}\Pi_P(\sigma \mid (\tilde{y}, \mu)) &= \mathbb{E}_\sigma \left[\sum_{i=1}^n e_i\right] - \sum_{i=1}^n \tilde{y}_i \\ &= \Pi_P(\sigma \mid (\tilde{y}, \tilde{\tau})) > ne^* - x^*,\end{aligned}$$

in contradiction to Lemma 1. ■

A.4 Examples of Imperfect Performance Measurement

A.4.1 First Example

Consider the moral-hazard example with multiplicative log-normal noise. Suppose that the condition $\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} \leq 2/(\pi\beta^2)$ is satisfied. Consider a contest with prize profile $y^* = (x^*, 0)$ in which x^* is given to agent 1 if and only if $\tilde{r}\tilde{s}_1/\tilde{s}_2 \geq 1$, where $\tilde{r} \sim \ln \mathcal{N}[\nu_r, \sigma_r^2]$ is a log-normal random variable with parameters

$$\nu_r = \nu_2 - \nu_1 \quad \text{and} \quad \sigma_r^2 = \frac{2}{\pi\beta^2} - (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}).$$

This allows for $\sigma_r^2 = 0$, by which we mean that $\tilde{r}$ is degenerate and takes the value e^{ν_r} with probability one. We now proceed in two steps. Step 1 derives an expression for agent i's expected payoff as a function of the effort profile e. Step 2 shows that $e_i = e^*$ is a best response when agent $j \neq i$ chooses $e_j = e^*$.

Step 1. Given an effort profile e, the probability that agent 1 wins the prize is

$$p(e) = \Pr\left[\frac{\tilde{r}\tilde{s}_1}{\tilde{s}_2} \geq 1\right] = \Pr\left[\frac{\tilde{r}\tilde{r}_1 e_1}{\tilde{r}_2 e_2} \geq 1\right] = \Pr\left[\frac{\tilde{r}_2}{\tilde{r}\tilde{r}_1} \leq \frac{e_1}{e_2}\right].$$

Since the variables $\tilde{r}_1$, $\tilde{r}_2$ and $\tilde{r}$ are log-normally distributed, it follows that the compound variable $\tilde{r}_2/(\tilde{r}\tilde{r}_1)$ is also log-normal, with location parameter $\nu = \nu_2 - \nu_1 - \nu_r = 0$ and scale parameter $\sigma^2 = \sigma_1^2 + \sigma_2^2 - \sigma_{12} + \sigma_r^2 = 2/(\pi\beta^2)$. The cdf of the log-normal distribution is given by $F(x) = \Phi\left((\log x - \nu)/\sigma\right)$, where Φ is the cdf of the standard normal distribution. Thus we can write

$$p(e) = \Phi\left(\log(e_1/e_2)\beta\sqrt{\frac{\pi}{2}}\right).$$

For the probability that agent 2 wins the prize we obtain

$$1 - p(e) = 1 - \Phi\left(\log(e_1/e_2)\beta\sqrt{\frac{\pi}{2}}\right)$$

$$= \Phi\left(-\log(e_1/e_2)\beta\sqrt{\frac{\pi}{2}}\right)$$
$$= \Phi\left(\log(e_2/e_1)\beta\sqrt{\frac{\pi}{2}}\right).$$

Hence the expected payoff of agent $i = 1, 2$ is

$$\Pi_i(e) = \Phi\left(\log(e_i/e_j)\beta\sqrt{\frac{\pi}{2}}\right) u(x^*) - c(e_i)$$
$$= \Phi\left(\log(e_i/e_j)\beta\sqrt{\frac{\pi}{2}}\right) 2\gamma e^{*\beta} - \gamma e_i^\beta,$$

where we have used that $u(x^*) = 2c(e^*)$ by Lemma 1.

Step 2. Suppose $e_j = e^*$ and consider the choice of agent $i \neq j$. We immediately obtain $\Pi_i(e^*, e^*) = 0$. We will now show that $\Pi_i(e_i, e^*) \leq 0$ always holds, i.e.,

$$\Phi\left(\log(e_i/e^*)\beta\sqrt{\frac{\pi}{2}}\right) \leq \frac{1}{2}\left(\frac{e_i}{e^*}\right)^\beta$$

for all $e_i \in \mathbb{R}_+$. After the change of variables $x = \log(e_i/e^*)\beta\sqrt{\pi/2}$ this becomes the requirement that

$$\Phi(x) \leq \frac{1}{2}e^{x\sqrt{2/\pi}} \tag{4}$$

for all $x \in \mathbb{R}$. Inequality (4) is satisfied for $x = 0$, where LHS and RHS both take a value of $1/2$. Furthermore, the LHS function and the RHS function are tangent at $x = 0$, because their derivatives are both equal to $1/\sqrt{2\pi}$ at this point. It then follows immediately that inequality (4) is also satisfied for all $x > 0$, because the LHS is strictly concave in x in this range, while the RHS is strictly convex. We now consider the remaining case where $x < 0$. We use the fact that $\Phi(x) = \text{erfc}(-x/\sqrt{2})/2$, where

$$\text{erfc}(y) = \frac{2}{\sqrt{\pi}}\int_y^\infty e^{-t^2}dt$$

is the complementary error function (see e.g. Chang, Cosman, and Milstein, 2011). After the change of variables $y = -x/\sqrt{2}$ we thus need to verify

$$\text{erfc}(y) \leq e^{-2y/\sqrt{\pi}} \tag{5}$$

for all $y > 0$. Inequality (5) is satisfied for $y = 0$, where LHS and RHS both take a value of 1. Now observe that the derivative of the LHS with respect to y is given by $-2e^{-y^2}/\sqrt{\pi}$, while the derivative of the RHS is $-2e^{-2y/\sqrt{\pi}}/\sqrt{\pi}$. The condition that the former is weakly smaller than the latter can be rearranged to $y \leq 2/\sqrt{\pi}$, which implies that (5) is satisfied for $0 < y \leq 2/\sqrt{\pi}$. For larger values of y, we can use a Chernoff bound for the complementary error function. Theorem 1 in Chang et al. (2011) implies that

$$\text{erfc}(y) \leq e^{-y^2}$$

for all $y \geq 0$. The inequality $e^{-y^2} \leq e^{-2y/\sqrt{\pi}}$ can be rearranged to $y \geq 2/\sqrt{\pi}$. This implies that (5) is satisfied also for $y > 2/\sqrt{\pi}$.

A.4.2 Second Example

Consider the example where only the effort difference $s = e_1 - e_2$ can be observed. Consider a contest with prize profile $y^* = (x^*, 0)$ in which x^* is given to agent 1 if and only if $s + \tilde{r} \geq 0$, where $\tilde{r} \sim \mathcal{U}[-c(e^*)/c'(e^*), c(e^*)/c'(e^*)]$ is a uniform random variable.

Observe that $c(e^*)/c'(e^*) < e^*$ holds due to strict convexity of c and $c(0) = 0$. We can therefore write the probability that agent 1 wins the prize, holding the effort $e_2 = e^*$ fixed, as a piecewise function

$$p(e_1) = \begin{cases} 1 & \text{if } e_1 > e^* + \frac{c(e^*)}{c'(e^*)}, \\ \frac{1}{2} + \frac{1}{2}\frac{c'(e^*)}{c(e^*)}(e_1 - e^*) & \text{if } e^* - \frac{c(e^*)}{c'(e^*)} \leq e_1 \leq e^* + \frac{c(e^*)}{c'(e^*)}, \\ 0 & \text{if } e_1 < e^* - \frac{c(e^*)}{c'(e^*)}. \end{cases}$$

Then, the expected payoff of agent 1 is given by

$$\Pi_1(e_1) = p(e_1)u(x^*) - c(e_1) = p(e_1)2c(e^*) - c(e_1).$$

It follows that $\Pi_1(e^*) = 0$. We now consider the three types of deviations from e^*.

Case 1: $e_1 < e^* - c(e^*)/c'(e^*)$. It follows immediately that $\Pi_1(e_1) \leq 0$ in this range, which implies that these deviations are not profitable.

Case 2: $e^* - c(e^*)/c'(e^*) \leq e_1 \leq e^* + c(e^*)/c'(e^*)$. Observe that $\Pi_1'(e_1) = c'(e^*) - c'(e_1)$ in this range. Hence the first-order condition yields the unique solution $e_1 = e^*$. Since $\Pi_1''(e_1) = -c''(e_1) < 0$, this is indeed the maximum over this range.

Case 3: $e_1 > e^* + c(e^*)/c'(e^*)$. We have $\Pi_1(e_1) < \Pi_1(e^* + c(e^*)/c'(e^*))$ for this range. Hence, by the arguments for the previous case, these deviations are not profitable either.

We conclude that $e_1 = e^*$ is a best response to $e_2 = e^*$. The argument for agent 2 is symmetric, which implies that the contest implements (e^*, e^*).

A.5 Proof of Proposition 3

We will first derive three lemmas which hold under cost heterogeneity for any number n of agents.

Lemma 5 *For any contest (y,μ) that implements a strategy profile σ, there exists a contest $(y,\tilde{\mu})$ that implements the pure-strategy profile $\bar{e}=(\bar{e}_1,\ldots,\bar{e}_n)$ where $\bar{e}_i=\mathbb{E}_\sigma[e_i]\ \forall i\in I$.*

Proof. Suppose (y,μ) implements σ. Define a probability measure $\eta\in\Delta T(y)$ by $\eta(t)=\mathbb{E}_\sigma[\mu^e(t)]$ for all $t\in T(y)$. Further, for each $i\in I$, define a probability measure $\eta^{(i)}\in\Delta T(y)$ by $\eta^{(i)}(t)=\mathbb{E}_\sigma\left[\mu^{(0,e_{-i})}(t)\right]$ for all $t\in T(y)$. Now construct $(y,\tilde{\mu})$ as follows. For $e=\bar{e}$, let $\tilde{\mu}^e=\eta$. For $e=(e_i,\bar{e}_{-i})$ with $e_i\neq\bar{e}_i$, let $\tilde{\mu}^e=\eta^{(i)}$. For all other e, let $\tilde{\mu}^e=\mu^e$. We will show that $(y,\tilde{\mu})$ implements $\bar{e}$, because $\bar{e}_i$ is a best response to $\bar{e}_{-i}$ for each $i\in I$. Indeed,

$$\begin{aligned}\Pi_i(\bar{e}\mid(y,\tilde{\mu}))&=\mathbb{E}_\eta[u(t_i)]-c_i(\bar{e}_i)\\&=\mathbb{E}_\sigma[\mathbb{E}_{\mu^e}[u(t_i)]]-c_i(\mathbb{E}_\sigma[e_i])\\&\geq\mathbb{E}_\sigma[\mathbb{E}_{\mu^e}[u(t_i)]]-\mathbb{E}_\sigma[c_i(e_i)]\\&\geq\mathbb{E}_\sigma[\mathbb{E}_{\mu^{(0,e_{-i})}}[u(t_i)]]\\&\geq\mathbb{E}_\sigma[\mathbb{E}_{\mu^{(0,e_{-i})}}[u(t_i)]]-c_i(e_i')\\&=\mathbb{E}_{\eta^{(i)}}[u(t_i)]-c_i(e_i')\\&=\Pi_i((e_i',\bar{e}_{-i})\mid(y,\tilde{\mu})),\end{aligned}$$

for all $e_i'\neq\bar{e}_i$, where the first inequality follows from convexity of c_i, and the second inequality follows because (y,μ) implements σ. □

Since the principal is indifferent between the mixed-strategy effort profile σ and its pure-strategy expectation $\bar{e}$, holding fixed the prize profile y, we can without loss of generality restrict attention to contests which implement a, possibly asymmetric, pure effort profile.[21] For any such contest, we obtain the following result.

Lemma 6 *If a contest (y,μ) implements a pure-strategy effort profile $\bar{e}$, it holds that*

$$\frac{1}{n-1}\sum_{i=1}^{n}c_i(\bar{e}_i)\leq u\left(\frac{x}{n-1}\right),$$

where $x=\sum_{k=1}^{n}y_k$.

Proof. Since (y,μ) implements $\bar{e}$, for each $i\in I$ it must hold that

$$c_i(\bar{e}_i)\leq\mathbb{E}_{\mu^{\bar{e}}}[u(t_i)]-\mathbb{E}_{\mu^{(0,\bar{e}_{-i})}}[u(t_i)].$$

[21] Lemma 5 generalizes Lemma 4 in Letina et al. (2020) to arbitrary costs functions, but restricted to the class of contests, while Letina et al. (2020) consider arbitrary incentive contracts.

Summing over all $i \in I$, we obtain

$$\begin{aligned}
\sum_{i=1}^{n} c_i(\bar{e}_i) &\leq \sum_{i=1}^{n} \mathbb{E}_{\mu^{\bar{e}}}\left[u(t_i)\right] - \sum_{i=1}^{n} \mathbb{E}_{\mu^{(0,\bar{e}_{-i})}}\left[u(t_i)\right] \\
&= \sum_{i=1}^{n} \mathbb{E}_{\mu^{(0,\bar{e}_{-1})}}\left[u(t_i)\right] - \sum_{i=1}^{n} \mathbb{E}_{\mu^{(0,\bar{e}_{-i})}}\left[u(t_i)\right] \\
&= \sum_{i=2}^{n} \mathbb{E}_{\mu^{(0,\bar{e}_{-1})}}[u(t_i)] - \sum_{i=2}^{n} \mathbb{E}_{\mu^{(0,\bar{e}_{-i})}}[u(t_i)] \\
&\leq \sum_{i=2}^{n} \mathbb{E}_{\mu^{(0,\bar{e}_{-1})}}[u(t_i)] \\
&\leq \sum_{i=2}^{n} u\left(\mathbb{E}_{\mu^{(0,\bar{e}_{-1})}}[t_i]\right) \\
&\leq \sum_{i=2}^{n} u\left(\frac{x}{n-1}\right) \\
&= (n-1)u\left(\frac{x}{n-1}\right),
\end{aligned}$$

where the first equality holds because the sum of expected utilities from money is the same for all effort profiles in a contest, the third inequality follows from concavity of u, and the fourth inequality follows from concavity of u together with the fact that $\sum_{i=2}^{n} \mathbb{E}_{\mu^{(0,\bar{e}_{-1})}}\left[t_i\right] \leq \sum_{i=1}^{n} y_i = x$. □

Our next result shows that we can restrict attention to contests in which the smallest prize is zero.

Lemma 7 *For any contest (y, μ) that implements a pure-strategy effort profile $\bar{e}$, there exists a contest $(\tilde{y}, \tilde{\mu})$ with $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_{n-1}, \tilde{y}_n) = (y_1, \ldots, y_{n-1}, 0)$ that also implements $\bar{e}$.*

Proof. Suppose that (y, μ) implements $\bar{e}$. Denote by $p_i^k(e)$ the probability that agent i obtains prize y_k when the effort profile is e, as induced by μ. The fact that the contest implements $\bar{e}$ implies that

$$\sum_{k=1}^{n} p_i^k(\bar{e})u(y_k) - c_i(\bar{e}_i) \geq \sum_{k=1}^{n} p_i^k(0, \bar{e}_{-i})u(y_k) \geq u(y_n) \;\forall i \in I. \tag{6}$$

Now consider another contest $(\tilde{y}, \tilde{\mu})$ with $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_{n-1}, \tilde{y}_n) = (y_1, \ldots, y_{n-1}, 0)$ and any CSF $\tilde{\mu}$ that induces $\tilde{p}_i^k(\bar{e}) = p_i^k(\bar{e})$ for all $i, k \in I$, and $\tilde{p}_i^n(e_i, \bar{e}_{-i}) = 1$ whenever $e_i \neq \bar{e}_i$. By construction, we have for all $i \in I$,

$$\sum_{k=1}^{n} \tilde{p}_i^k(\bar{e})u(\tilde{y}_k) - c_i(\bar{e}_i) = \sum_{k=1}^{n-1} p_i^k(\bar{e})u(y_k) - c_i(\bar{e}_i)$$

$$
\begin{aligned}
&\geq [1 - p_i^n(\bar{e})]u(y_n) \\
&\geq 0 \\
&= u(\tilde{y}_n),
\end{aligned}
$$

where the first inequality follows from (6). Hence, $(\tilde{y}, \tilde{\mu})$ also implements $\bar{e}$. $\square$

From now on we consider the special case of $n = 2$. It follows from Lemma 6 that any contest (y, μ) with $y = (x, 0)$ that implements a pure-strategy effort profile $\bar{e}$ must satisfy $c_1(\bar{e}_1) + c_2(\bar{e}_2) \leq u(x)$. Since restricting attention to such contests is without loss of generality by Lemmas 5 and 7, the problem

$$
\max_{x, e_1, e_2 \geq 0} \; e_1 + e_2 - x \quad \text{s.t.} \quad c_1(e_1) + c_2(e_2) \leq u(x)
$$

describes an upper bound on the payoff that the principal can achieve. Obviously, any solution (x^*, e_1^*, e_2^*) to this problem must satisfy the constraint with equality, and it must be strictly positive. We complete the proof by showing that the contest described in the proposition achieves that bound, by implementing the effort profile (e_1^*, e_2^*) using prize x^*.

Lemma 8 *Suppose $n = 2$. The contest (y^*, μ^*) implements the effort profile (e_1^*, e_2^*).*

Proof. Consider a tuple (x^*, e_1^*, e_2^*) as described in the proposition. Using a Tullock CSF with individual-specific impact functions $f_i(e_i) = c_i(e_i)^{r_i}/c_i(e_i^*)^{r_i - 1}$ for any $r_i > 1$, it follows that the probability that agent i wins the prize x^* with effort profile $e = (e_i, e_j)$ is

$$
\begin{aligned}
p_i(e_i, e_j) &= \frac{c_i(e_i)^{r_i}/c_i(e_i^*)^{r_i-1}}{c_i(e_i)^{r_i}/c_i(e_i^*)^{r_i-1} + c_j(e_j)^{r_j}/c_j(e_j^*)^{r_j-1}} \\
&= \frac{c_i(e_i)^{r_i} c_j(e_j^*)^{r_j-1}}{c_i(e_i)^{r_i} c_j(e_j^*)^{r_j-1} + c_j(e_j)^{r_j} c_i(e_i^*)^{r_i-1}} \\
&= 1 - \frac{c_j(e_j)^{r_j} c_i(e_i^*)^{r_i-1}}{c_i(e_i)^{r_i} c_j(e_j^*)^{r_j-1} + c_j(e_j)^{r_j} c_i(e_i^*)^{r_i-1}}.
\end{aligned}
$$

To simplify notation, let $c_i = c_i(e_i)$ and $c_i^* = c_i(e_i^*)$. Then we can write agent i's optimization problem as $\max_{c_i \geq 0} U(c_i, c_j^*)$, where $U_i(c_i, c_j^*) = p_i(c_i, c_j^*)u(x^*) - c_i$. We obtain after some simplifications

$$
\frac{\partial U_i(c_i, c_j^*)}{\partial c_i} = r_i \left[\frac{(c_i^*)^{r_i-1} c_j^* c_i^{r_i-1}}{(c_i^{r_i} + (c_i^*)^{r_i-1} c_j^*)^2} \right] u(x^*) - 1 \tag{7}
$$

and

$$
\frac{\partial^2 U_i(c_i, c_j^*)}{\partial c_i^2} = \frac{r_i u(x^*)(c_i^*)^{r_i-1} c_j^*}{(c_i^{r_i} + (c_i^*)^{r_i-1} c_j^*)^3} \left[(r_i - 1)c_i^{r_i-2}(c_i^{r_i} + (c_i^*)^{r_i-1} c_j^*) - 2r_i c_i^{2(r_i-1)} \right]. \tag{8}
$$

We immediately obtain $U_i(0, c_j^*) = 0$ and $\partial U_i(0, c_j^*)/\partial c_i < 0$, so that $c_i = 0$ is a local maximum. Using that $c_i^* + c_j^* = u(x^*)$, it also follows immediately that $U_i(c_i^*, c_j^*) = 0$. Now let $r_i = r_i^* = 1 + c_i^*/c_j^*$. From (7) we obtain

$$\frac{\partial U_i(c_i^*, c_j^*)}{\partial c_i} = \left(\frac{c_i^* + c_j^*}{c_j^*}\right) \left[\frac{(c_i^*)^{2(r_i^*-1)} c_j^*}{(c_i^*)^{2(r_i^*-1)}(c_i^* + c_j^*)^2}\right] (c_i^* + c_j^*) - 1 = 0, \tag{9}$$

so that the first-order condition is satisfied at $c_i = c_i^*$. By (8), the sign of $\partial^2 U_i/\partial c_i^2$ is equal to the sign of $(r_i - 1)c_i^{r_i-2}(c_i^{r_i} + (c_i^*)^{r_i-1}c_j^*) - 2r_i c_i^{2(r_i-1)}$, which for $r_i = r_i^* = 1 + c_i^*/c_j^*$ can be rearranged to

$$c_i^{r_i^*-2}(c_i^*)^{r_i^*} - (r_i^* + 1)c_i^{2(r_i^*-1)}. \tag{10}$$

Using (10) we thus obtain that $\partial^2 U_i/\partial c_i^2 \leq 0$ if and only if

$$c_i^{r_i^*} \geq \left(\frac{1}{r_i^* + 1}\right) (c_i^*)^{r_i^*}.$$

It follows that $c_i = c_i^*$ is also a local maximum. Furthermore, the sign of $\partial^2 U_i/\partial c_i^2$ changes only once as c_i increases from 0 to ∞, and hence both $c_i = 0$ and $c_i = c_i^*$ are global maxima of the function $U_i(c_i, c_j^*)$. Therefore, c_i^* is a best response of agent i to c_j^*, which implies that the contest implements (e_1^*, e_2^*). □ ■

A.6 Proof of Proposition 4

We first state some additional properties that hold for any given profile of effort cost functions $(c_1, \ldots, c_n)$. By Lemmas 5 and 7, we can restrict attention to the implementation of pure-strategy effort profiles by contests with $y_n = 0$. This allows us to show that the principal's optimization problem has a solution.

Lemma 9 *An optimal contest exists.*

Proof. When optimizing over contests that implement a pure-strategy effort profile e and have $y_n = 0$, it is without loss of generality to assume that an agent who deviates unilaterally from e obtains $y_n = 0$ with probability one, which is the harshest possible punishment. Thus constraint (IC-A) can be written as

$$\sum_{k=1}^{n} p_i^k u(y_k) - c_i(e_i) \geq 0 \;\; \forall i \in I, \tag{11}$$

where p_i^k denotes the probability that agent i obtains y_k when the effort profile is e. The principal therefore maximizes $\sum_{i=1}^n e_i - \sum_{i=1}^n y_i$ by choosing $e = (e_1, \ldots, e_n) \in \mathbb{R}_+^n$, $y =$

$(y_1, \ldots, y_n) \in \mathbb{R}^n_+$ and $p = (p_i^k)_{i,k} \in [0,1]^{n^2}$, subject to (11) and the constraints that $y_n = 0$ and p is a well-defined probability system. The allocation probabilities after multilateral deviations from e can be chosen arbitrarily. Using notation $x = \sum_{k=1}^n y_k$, constraint (11) implies $e_i \leq c_i^{-1}(u(x))$ for all $i \in I$. This implies $\sum_{i=1}^n e_i - \sum_{i=1}^n y_i \leq \sum_{i=1}^n c_i^{-1}(u(x)) - x$. Since u is weakly concave and each c_i is strictly convex with $\lim_{e_i \to \infty} c_i'(e_i) = \infty$, there exists $X > 0$ such that $\sum_{i=1}^n c_i^{-1}(u(x)) - x < 0$ whenever $x > X$, so that a contest with $x > X$ cannot be optimal. It is therefore without loss to impose $y_i \in [0, X]$ and $e_i \in [0, c_i^{-1}(u(X))]$ for all $i \in I$. Continuity of u and each c_i then implies that the constraint set is compact. Since the principal's objective is continuous, a solution exists. □

The next result provides a lower bound on maximal profits. Fix any $T > 0$ and define

$$\underline{\Pi} = \max_{x \in [0,T]} \left[c_1^{-1}(u(x)) - x \right],$$

which exists and satisfies $\underline{\Pi} > 0$ due to our assumptions on c_1 and u.

Lemma 10 *There exists a contest (y, μ) that implements a pure-strategy effort profile e such that $\Pi_P(e \mid (y, \mu)) = \underline{\Pi}$.*

Proof. Let $x^* = \arg\max_{x \in [0,T]} \left[c_1^{-1}(u(x)) - x \right]$ and $e_1^* = c_1^{-1}(u(x^*))$. Consider a contest with prize profile $y = (x^*, 0, \ldots, 0)$. If the effort profile e is such that $e_1 = e_1^*$, then agent 1 receives the prize x^* while all other agents receive a zero prize. For any other effort profile, agent 2 receives x^* and all other agents receive a zero prize. It follows that this contest implements $(e_1^*, 0, \ldots, 0)$ and yields the payoff $e_1^* - x^* = \underline{\Pi}$ to the principal. □

The next result states that it is without loss to focus on the implementation of effort profiles that are not too heterogeneous relative to the cost functions. The proof proceeds like the proof of Lemma 5 in Letina et al. (2020) and is therefore omitted.

Lemma 11 *For any contest (y, μ) that implements a pure-strategy effort profile $\bar{e}$ such that*

$$\frac{1}{n} \sum_{i=1}^n c_i(\bar{e}_i) > c_k \left(\frac{1}{n} \sum_{i=1}^n \bar{e}_i \right) \quad \forall k \in I,$$

there exists a contest (y', μ') that implements the pure-strategy effort profile $\hat{e}$ given by $\hat{e}_1 = \ldots = \hat{e}_n = \frac{1}{n} \sum_{i=1}^n \bar{e}_i$, and yields the same expected payoff to the principal.

Now consider a sequence $(c_1^m, \ldots, c_n^m)_{m \in \mathbb{N}}$ such that $(c_1^m, \ldots, c_n^m) \to (c, \ldots, c)$ uniformly. Let $(\bar{e}^m, (y^m, \mu^m))_{m \in \mathbb{N}}$ be a corresponding sequence of optimal solutions, i.e., (y^m, μ^m) implements $\bar{e}^m = (\bar{e}_1^m, \ldots, \bar{e}_n^m)$ and solves the principal's problem when the cost functions are $(c_1^m, \ldots, c_n^m)$. Given the above results, we can assume that $\Pi_P(\bar{e}^m \mid (y^m, \mu^m)) \geq \underline{\Pi}^m > 0$, where $\underline{\Pi}^m = \max_{x\in[0,T]} [(c_1^m)^{-1}(u(x)) - x]$. We can also assume that

$$\frac{1}{n}\sum_{i=1}^{n} c_i^m(\bar{e}_i^m) \leq \max_{k\in I} c_k^m \left(\frac{1}{n}\sum_{i=1}^{n} \bar{e}_i^m\right). \tag{12}$$

We will write $\hat{e}^m = (1/n)\sum_{i=1}^n \bar{e}_i^m$ for the average effort and $x^m = \sum_{i=1}^n y_i^m$ for the total budget of the contest at step m in the sequence. We first show that the total budget must be bounded.

Lemma 12 *There exists $B \in \mathbb{R}$ such that $x^m \leq B$ for all m.*

Proof. Since (y^m, μ^m) implements $\bar{e}^m$, we must have

$$\Pi_P(\bar{e}^m \mid (y^m, \mu^m)) \leq \left[\sum_{i=1}^{n}(c_i^m)^{-1}(u(x^m))\right] - x^m.$$

Using Theorem 2 in Barvinek, Daler, and Francu (1991), it can be shown that $(c_i^m)^{-1}$ converges uniformly to c^{-1} for all i.[22] Thus, for every $\epsilon > 0$ there exists $\underline{m}' \in \mathbb{N}$ such that for all $m \geq \underline{m}'$ and all i,

$$|(c_i^m)^{-1}(u(x^m)) - c^{-1}(u(x^m))| < \epsilon/n,$$

which implies $\sum_{i=1}^n |(c_i^m)^{-1}(u(x^m)) - c^{-1}(u(x^m))| < \epsilon$, and therefore

$$\left|\left(\sum_{i=1}^{n}(c_i^m)^{-1}(u(x^m))\right) - x^m - \left(nc^{-1}(u(x^m)) - x^m\right)\right| < \epsilon. \tag{13}$$

Since u is weakly concave and c is strictly convex with $\lim_{e_i\to\infty} c'(e_i) = \infty$, there exists $\tilde{B} > 0$ such that $nc^{-1}(u(x)) - x < -\epsilon$ for all $x > \tilde{B}$. Therefore, if for any $m \geq \underline{m}'$ it was the case that $x^m > \tilde{B}$, inequality (13) would imply that $(\sum_{i=1}^n (c_i^m)^{-1}(u(x^m))) - x^m < 0$, which in turn implies $\Pi_P(\bar{e}^m \mid (y^m, \mu^m)) < 0$. This is in contradiction to the assumption that $\Pi_P(\bar{e}^m \mid (y^m, \mu^m)) \geq \underline{\Pi}^m > 0$. Hence we know that $x^m \leq \tilde{B}$ for all $m \geq \underline{m}'$. Now simply let $B = \max\{x^1, \ldots, x^{\underline{m}'-1}, \tilde{B}\}$. □

For the remainder of the proof, we fix any $B \in \mathbb{R}$ such that $x^m \leq B$ for all m.

[22] The theorem is directly applicable and implies our claim after we extend the functions c and c_i^m to $\mathbb{R}$ by defining $c_i^m(e) = c(e) = e$ for all $e < 0$.

Lemma 13 *The sequence*

$$\kappa^m = \max_{k\in I} c_k^m(\bar{e}_k^m) - \frac{1}{n}\sum_{i=1}^{n} c_i^m(\bar{e}_i^m)$$

converges to zero as $m \to \infty$.

Proof. For every $m \in \mathbb{N}$, let

$$\delta^m = \max_{k\in I} c_k^m(\hat{e}^m) - \frac{1}{n}\sum_{i=1}^{n} c_i^m(\bar{e}_i^m) \quad \text{and} \quad \psi^m = \max_{i\in I} c_i^m(\bar{e}_i^m) - \max_{k\in I} c_k^m(\hat{e}^m),$$

and hence $\kappa^m = \delta^m + \psi^m$. We will show that $\lim_{m\to\infty}\delta^m = \lim_{m\to\infty}\psi^m = 0$, which immediately implies that $\lim_{m\to\infty}\kappa^m = 0$. For the sequence δ^m, note that

$$\delta^m = \left[\max_{k\in I} c_k^m(\hat{e}^m) - c(\hat{e}^m)\right] + \left[\frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i^m) - \frac{1}{n}\sum_{i=1}^{n} c_i^m(\bar{e}_i^m)\right] + \left[c(\hat{e}^m) - \frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i^m)\right].$$

By uniform convergence of c_i^m to c, $\forall i \in I$, we have

$$\lim_{m\to\infty}(c_i^m(\hat{e}^m) - c(\hat{e}^m)) = 0 \quad \text{and} \quad \lim_{m\to\infty}(c_i^m(\bar{e}_i^m) - c(\bar{e}_i^m)) = 0 \;\; \forall i \in I, \tag{14}$$

and thus

$$\lim_{m\to\infty}\max_{k\in I}(c_k^m(\hat{e}^m) - c(\hat{e}^m)) = 0 \quad \text{and} \quad \lim_{m\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n} c_i^m(\bar{e}_i^m) - \frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i^m)\right) = 0.$$

In addition, by convexity of c we have $c(\hat{e}^m) - \frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i^m) \le 0$ for all $m \in \mathbb{N}$, and by condition (12) we have $\delta^m \ge 0$ for all $m \in \mathbb{N}$. Hence, we must also have

$$\lim_{m\to\infty}\left(c(\hat{e}^m) - \frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i^m)\right) = 0, \tag{15}$$

as otherwise for some large m we would have $\delta^m < 0$, a contradiction. This concludes that $\lim_{m\to\infty}\delta^m = 0$. For the sequence ψ^m, we have

$$\psi^m = \max_{k\in I}(c(\hat{e}^m) - c_k^m(\hat{e}^m)) + \max_{i\in I}\left[c_i^m(\bar{e}_i^m) - c(\bar{e}_i^m) + c(\bar{e}_i^m) - c(\hat{e}^m)\right].$$

Hence, by (14), a sufficient condition for $\lim_{m\to\infty}\psi^m = 0$ is

$$\lim_{m\to\infty}(c(\bar{e}_i^m) - c(\hat{e}^m)) = 0 \;\; \forall i \in I. \tag{16}$$

To establish (16), we first claim that there exists $\tilde{e} > 0$ such that $\bar{e}_i^m \in [0, \tilde{e}]$ for all $i \in I$

and all $m \in \mathbb{N}$. The fact that (y^m, μ^m) implements $\bar{e}^m$ implies $c_i^m(\bar{e}_i^m) \leq u(B)$ for all $i \in I$. Now fix any $\tilde{u} > u(B)$. By uniform convergence of each c_i^m to c it follows that there exists $\underline{m}' \in \mathbb{N}$ such that for all $m \geq \underline{m}'$,

$$|c_i^m(\bar{e}_i^m) - c(\bar{e}_i^m)| \leq \tilde{u} - u(B) \;\; \forall i \in I,$$

which then implies $c(\bar{e}_i^m) \leq \tilde{u}$ and therefore $\bar{e}_i^m \leq c^{-1}(\tilde{u})$. Now just define $\tilde{e}$ as the maximum among $c^{-1}(\tilde{u})$ and the finite number of values $\bar{e}_i^m$ for all $i \in I$ and $m < \underline{m}'$. We next claim that $\lim_{m\to\infty}(\bar{e}_i^m - \hat{e}^m) = 0$ holds for all $i \in I$. By contradiction, assume there exists $i \in I$ and $\epsilon > 0$ such that for all $\underline{m}' \in \mathbb{N}$ there exists $m \geq \underline{m}'$ so that $|\bar{e}_i^m - \hat{e}^m| \geq \epsilon$. Define $E_i = \{(e_1, \ldots, e_n) \in [0, \tilde{e}]^n \mid |e_i - \frac{1}{n}\sum_{j=1}^n e_j| \geq \epsilon\}$. The set E_i is compact and the function $\chi(e) = \frac{1}{n}\sum_{j=1}^n c(e_j) - c\left(\frac{1}{n}\sum_{j=1}^n e_j\right)$ is continuous on E_i, with $\chi(e) > 0$ due to strict convexity of c and $\epsilon > 0$. Hence $\tilde{\epsilon} = \min_{e\in E_i} \chi(e)$ exists and satisfies $\tilde{\epsilon} > 0$. We have thus shown that there exists $\tilde{\epsilon} > 0$ such that for all $\underline{m}' \in \mathbb{N}$ there exists $m \geq \underline{m}'$ so that $\chi(\bar{e}^m) = -(c(\hat{e}^m) - \frac{1}{n}\sum_{i=1}^n c(\bar{e}_i^m)) \geq \tilde{\epsilon}$, contradicting (15). Finally, (16) now follows immediately because $\bar{e}_i^m \in [0, \tilde{e}]$ and $\hat{e}^m \in [0, \tilde{e}]$ and c is continuous on $[0, \tilde{e}]$. □

Next we show that the sum of effort costs is bounded away from zero for large m.

Lemma 14 *There exist $\underline{m}' \in \mathbb{N}$ and $\underline{c} > 0$ such that $\sum_{i=1}^n c_i^m(\bar{e}_i^m) \geq \underline{c}$ for all $m \geq \underline{m}'$.*

Proof. Let $\underline{\Pi}^m = \max_{x\in[0,T]} \underline{\Pi}_1^m(x)$ with $\underline{\Pi}_1^m(x) = (c_1^m)^{-1}(u(x)) - x$ be the lower profit bound for the cost functions $(c_1^m, \ldots, c_n^m)$ as defined earlier. Hence $\Pi_P(\bar{e}^m \mid (y^m, \mu^m)) \geq \underline{\Pi}^m$ holds for all $m \in \mathbb{N}$. Similarly, let $\underline{\Pi}^\infty = \max_{x\in[0,T]} \underline{\Pi}_1(x)$ with $\underline{\Pi}_1(x) = c^{-1}(u(x)) - x$ be the bound when the cost functions are $(c, \ldots, c)$. We first claim that $\lim_{m\to\infty} \underline{\Pi}^m = \underline{\Pi}^\infty$. The claim follows immediately once we show that $\underline{\Pi}_1^m$ converges uniformly to $\underline{\Pi}_1$ on $[0, T]$. Again using Theorem 2 in Barvinek et al. (1991), it can be shown that $(c_1^m)^{-1}$ converges uniformly to c^{-1} on $[0, u(T)]$. Thus for every $\epsilon > 0$ there exists $\underline{m}'' \in \mathbb{N}$ such that for all $m \geq \underline{m}''$,

$$|\underline{\Pi}_1^m(x) - \underline{\Pi}_1(x)| = |(c_1^m)^{-1}(u(x)) - c^{-1}(u(x))| < \epsilon$$

for all $x \in [0, T]$, which establishes uniform convergence. Now fix any ϵ with $0 < \epsilon < \underline{\Pi}^\infty$ and define $\tilde{\Pi} = \underline{\Pi}^\infty - \epsilon > 0$. Hence there exists $\underline{m}''' \in \mathbb{N}$ such that for all $m \geq \underline{m}'''$,

$$\sum_{i=1}^n \bar{e}_i^m \geq \Pi_P(\bar{e}^m \mid (y^m, \mu^m)) \geq \underline{\Pi}^m \geq \tilde{\Pi} > 0.$$

Define

$$\underline{c}^m = \min_{e\in E} \sum_{i=1}^n c_i^m(e_i) \quad \text{s.t.} \quad \sum_{i=1}^n e_i = \tilde{\Pi}.$$

We then obtain that $\sum_{i=1}^n c_i^m(\bar{e}_i^m) \geq \underline{c}^m$ for all $m \geq \underline{m}'''$. Similarly, define

$$\underline{c}^\infty = \min_{e \in E} \sum_{i=1}^n c(e_i) \quad \text{s.t.} \quad \sum_{i=1}^n e_i = \tilde{\Pi},$$

noting that $\underline{c}^\infty > 0$. It again follows from uniform convergence of c_i^m to c for each $i \in I$ that $\lim_{m \to \infty} \underline{c}^m = \underline{c}^\infty$. Fix any ϵ' such that $0 < \epsilon' < \underline{c}^\infty$ and define $\underline{c} = \underline{c}^\infty - \epsilon' > 0$. It follows that there exists $\underline{m}' \in \mathbb{N}$ such that for all $m \geq \underline{m}'$,

$$\sum_{i=1}^n c_i^m(\bar{e}_i^m) \geq \underline{c}^m \geq \underline{c},$$

which completes the proof. □

We can now combine Lemmas 13 and 14 to obtain the following result.

Lemma 15 *There exists $\underline{m} \in \mathbb{N}$ such that for all $m \geq \underline{m}$,*

$$\max_{k \in I} c_k^m(\bar{e}_k^m) \leq \frac{1}{n-1} \sum_{i=1}^n c_i^m(\bar{e}_i^m).$$

Proof. By Lemma 14, there exist $\underline{m}' \in \mathbb{N}$ and $\underline{c} > 0$ such that $\sum_{i=1}^n c_i^m(\bar{e}_i^m) \geq \underline{c}$ for all $m \geq \underline{m}'$. In addition, from the limiting statement about κ^m in Lemma 13 we can conclude that there exists $\underline{m}'' \in \mathbb{N}$ such that for all $m \geq \underline{m}''$,

$$\max_{k \in I} c_k^m(\bar{e}_k^m) - \frac{1}{n} \sum_{i=1}^n c_i^m(\bar{e}_i^m) \leq \frac{\underline{c}}{n(n-1)}.$$

Thus for all $m \geq \underline{m} = \max\{\underline{m}', \underline{m}''\}$ we obtain

$$\begin{aligned} \max_{k \in I} c_k^m(\bar{e}_k^m) - \frac{1}{n-1} \sum_{i=1}^n c_i^m(\bar{e}_i^m) &= \max_{k \in I} c_k^m(\bar{e}_k^m) - \frac{1}{n} \sum_{i=1}^n c_i^m(\bar{e}_i^m) - \frac{1}{n(n-1)} \sum_{i=1}^n c_i^m(\bar{e}_i^m) \\ &\leq \frac{\underline{c}}{n(n-1)} - \frac{1}{n(n-1)} \sum_{i=1}^n c_i^m(\bar{e}_i^m) \\ &\leq 0. \end{aligned}$$

□

Now consider any fixed $m \geq \underline{m}$, with $\underline{m}$ from Lemma 15. Combined with Lemma 6 we can conclude that the contest (y^m, μ^m) and the effort profile $\bar{e}^m$ satisfy

$$\max_{k \in I} c_k^m(\bar{e}_k^m) \leq u\left(\frac{x^m}{n-1}\right). \tag{17}$$

We now show that $\bar{e}^m$ can also be implemented in a contest with the same budget and $n-1$ identical prizes, given the cost functions $(c_1^m, \ldots, c_n^m)$.

Lemma 16 *Fix any* $m \geq \underline{m}$*. There exists a contest* (y, μ) *which implements* $\bar{e}^m$ *and has the prize profile* $y = (x^m/(n-1), \ldots, x^m/(n-1), 0)$.

Proof. We construct the allocation rule μ as follows. If $e = \bar{e}^m$, the zero prize is given to agent i with probability $p_i \geq 0$, while all other agents obtain one of the identical positive prizes. Below we will determine the values p_i such that $\sum_{i=1}^n p_i = 1$. If $e = (e_i, \bar{e}_{-i}^m)$ with $e_i \neq \bar{e}_i^m$ for some $i \in I$, the deviating agent i obtains the zero prize for sure and all other agents obtain one of the identical positive prizes. For all other effort profiles e, the allocation of the prizes can be chosen arbitrarily. First define $\tilde{p}_i$ implicitly by

$$(1 - \tilde{p}_i) u \left(\frac{x^m}{n-1} \right) = c_i^m(\bar{e}_i^m).$$

Since the LHS of this equation describes the expected payoff of agent i who expects to obtain the zero prize with probability $\tilde{p}_i$, it follows that the contest (y, μ) indeed implements $\bar{e}^m$ if $p_i \leq \tilde{p}_i$ holds for all $i \in I$. The fact that $c_i^m(\bar{e}_i^m) \leq u(x^m/(n-1))$ for all $i \in I$ due to (17) guarantees $\tilde{p}_i \geq 0$. Lemma 6 also implies that

$$\sum_{i=1}^n c_i^m(\bar{e}_i^m) = \sum_{i=1}^n (1 - \tilde{p}_i) u \left(\frac{x^m}{n-1} \right) = \left(n - \sum_{i=1}^n \tilde{p}_i \right) u \left(\frac{x^m}{n-1} \right) \leq (n-1) u \left(\frac{x^m}{n-1} \right),$$

which guarantees that $\sum_{i=1}^n \tilde{p}_i \geq 1$. It is therefore possible to find equilibrium punishment probabilities p_i such that $0 \leq p_i \leq \tilde{p}_i \; \forall i \in I$ and $\sum_{i=1}^n p_i = 1$. □

In sum, whenever $m \geq \underline{m}$, we can replace the optimal contest (y^m, μ^m) by a contest with $n-1$ identical prizes that implements the same effort profile and generates the same payoff for the principal. ■