Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 213 (2023) 105616

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Optimal contest design: Tuning the heat ☆

Igor Letina [a,b,*], Shuo Liu [c], Nick Netzer [d]

[a] *Department of Economics, University of Bern, Switzerland*
[b] *CEPR, UK*
[c] *Guanghua School of Management, Peking University, China*
[d] *Department of Economics, University of Zurich, Switzerland*

Received 13 June 2020; final version received 24 November 2022; accepted 23 January 2023
Available online 2 February 2023

Abstract

We consider the design of contests when the principal can choose both the prize profile and how the prizes are allocated as a function of a possibly noisy signal about the agents' efforts. We provide sufficient conditions that guarantee optimality of a contest. Optimal contests have a minimally competitive prize profile and an intermediate degree of competitiveness in the contest success function. Whenever observation is not too noisy, the optimum can be achieved by an all-pay contest with a cap. When observation is perfect, the optimum can also be achieved by a nested Tullock contest. 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.

JEL classification: D02; D82; M52

Keywords: Contest design; Optimal contests; Tournaments

☆ This paper was previously circulated under the title "Optimal Contest Design: A General Approach". We would like to thank the editor Tilman Börgers, an anonymous referee, Philipp Denter, Mikhail Drugov, Jingfeng Lu, Oleg Muratov, Dana Sisak, Nicolas Schutz, Cédric Wasser, and Zenan Wu for helpful comments. Parts of the material in this paper are from the older discussion paper Letina et al. (2018). Other parts of that older paper have appeared as Letina et al. (2020). Shuo Liu acknowledges financial support from the National Natural Science Foundation of China (Grant No. 72103006 and 72192844).

* Corresponding author.
E-mail addresses: igor.letina@vwi.unibe.ch (I. Letina), shuo.liu@gsm.pku.edu.cn (S. Liu), nick.netzer@econ.uzh.ch (N. Netzer).

https://doi.org/10.1016/j.jet.2023.105616

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*. 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) has been extensively studied, the question of optimal contest design has received significantly less attention.[1]

Several recent articles have analyzed the optimal allocation of prizes in specific classes of contests. Examples include Schweinzer and Segev (2012) and Fu et al. (2021a) for Tullock contests, Drugov and Ryvkin (2020b) and Morgan et al. (2022) for Lazear-Rosen tournaments, and Fang et al. (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 Schweinzer and Segev (2012) we learn that in a nested Tullock contest with risk-neutral agents, a winner-take-all prize structure is optimal, 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 explains the different results in the literature. In particular, we establish an upper payoff bound for the principal who optimally designs a contest, and we provide sufficient conditions under which there exists a contest that achieves this upper bound.

In our model, the principal can choose any prize profile and any rule specifying how the prizes are allocated to the agents as a function of a possibly noisy signal about their efforts. When the principal perfectly observes the efforts, then our framework includes all the standard contest success functions (CSFs) as special cases. When the observation of efforts is imperfect, then the observational noise puts constraints on the set of CSFs that the principal can induce. The objective of the principal is to maximize the expected aggregate effort minus the sum of prizes. The agents can be risk-neutral or risk-averse, and they have convex effort cost functions.

Our first main result provides sufficient conditions for a contest to achieve our upper payoff bound and thus be optimal, given an arbitrary observational structure. These conditions are that the prize profile is minimally competitive, with $n-1$ equal positive prizes and a single zero prize, and that the CSF has an intermediate degree of competitiveness, so that the off-equilibrium probability of winning a positive prize remains below an explicitly given bound for any effort deviation. The conditions can be easily verified and are a powerful tool for the design of optimal contests under diverse observational assumptions. The result builds on the previous work in Letina et al. (2020), who consider perfect effort observation only. For that case, they show that the optimal prize profile has $n-1$ equal positive prizes and that the optimum can be achieved by an all-pay contest with a cap. Our results here rely on the insight that there are many other CSFs

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.

that can also achieve the optimum, and some of them are feasible for the principal even if effort observation is imperfect.

For the case when efforts are indeed perfectly observable, our second main result shows that, in addition to an all-pay contest with a cap, the optimum can also be achieved by a nested Tullock contest (Clark and Riis, 1996). 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 ourselves to pure-strategy equilibria, but they emerge as part of 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. This result provides a unifying perspective on the findings in the literature mentioned above. The message of Fang et al. (2020) is that "turning up the heat" in a standard all-pay contest, by making the prize profile more unequal, increases the dispersion of the equilibrium effort distributions and decreases the expected equilibrium effort that agents exert. It follows that the principal should select the most equal prize profile with $n-1$ identical prizes, but the agents are still mixing in equilibrium with this prize profile. 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 CSF, exactly to the point where a pure-strategy equilibrium emerges. Our results are also in line with the seemingly contradictory intuition of Schweinzer and Segev (2012), who argue that optimal nested Tullock contests should turn up the heat by concentrating 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 in that sense; more concentration of the prizes on the top would destroy the pure-strategy equilibrium. Such insights 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.

Perhaps the most commonly studied model with imperfect observation of efforts is one where the principal observes effort plus an i.i.d. noise term. Our third main result focuses on this setting. We provide a condition on the distribution of noise such that the principal can still generate a CSF that satisfies our general optimality conditions. Recall that the principal aims for an intermediate degree of precision in the CSF anyway. Random observational noise makes a contest less competitive for any given allocation rule. As long as observation is not too noisy, the principal can tune the heat by combining the observational noise and the allocation rule in a way that implements the optimum. In particular, we show that an all-pay contest with a cap (applied to the stochastic signals about individual agents' efforts) is optimal. This generalizes the result of Letina et al. (2020) mentioned above to the case of imperfect effort observability. We then extend the analysis to one specific environment where the observation of effort is so imprecise that the general optimality conditions cannot be satisfied. We characterize the optimal contests in this setting and show that they still feature $n-1$ equal positive prizes, a single zero prize, and a CSF with an intermediate degree of competitiveness.

Although the focus of our analysis is on the optimal design of contests, we also compare the principal's payoff from an optimal contest to the payoff of a principal who instead can use an arbitrary incentive mechanism. We show that the comparison crucially depends on the observational structure. There exist observational structures for which a contest is unconstrained optimal, others for which the first-best can be achieved by a mechanism that is not a contest, and yet others for which the optimal mechanism achieves a payoff in between the contest and the first-best. Using an arbitrary incentive mechanism either results in a payoff that is equal to that of the optimal contest, equal to the first-best payoff, or anything in between. We also show that the optimal contest payoff and the first-best payoff converge as the number of agents increases, so that any potential gain from using arbitrary incentive mechanisms disappears in large contests.

In our baseline model, we assume that the agents are symmetric, entry into the contest is costless, and that the agents are (weakly) risk-averse. To illustrate the flexibility of our approach, we relax these assumptions in turn. We first 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. When agents have to incur a cost to enter the contest, we show that the optimal prize profile still features $n - 1$ equal top prizes but the lowest prize is potentially positive, to give rents to the agents and incentivize entry. Finally, for risk-loving agents, the optimal prize structure becomes winner-take-all but the CSF still exhibits an intermediate degree of competitiveness.

The paper is organized as follows. The model is introduced in Section 2. Our main results are 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

There is a principal and a set of agents $I = \{1, \dots, n\}$, 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 $\lim_{e_i \to \infty} c'(e_i) = \infty$.

Denote the effort profile of all agents by $e = (e_1, \dots, e_n) \in E = \mathbb{R}^n_+$ and the transfer profile by $t = (t_1, \dots, t_n) \in T = \mathbb{R}^n_+$. The payoff of the principal is

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

That is, the principal maximizes the sum of efforts net of transfers.

After the agents have chosen their efforts, a signal $s \in S$ is drawn according to an effort-dependent probability measure $\eta^e \in \Delta S$. The principal observes s but not e. 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.[3] The observational structure is assumed to be common knowledge.

A large range of applications and examples can be modeled by different observational structures. Perfect observability of effort is the special case where $S = E$ and η^e is the Dirac measure on e. A second example is the classical moral-hazard setting where each agent's effort e_i produces a random output s_i such that $\mathbb{E}_{\eta^e}[s_i] = e_i$ (and therefore it does not matter whether a

[3] We do require 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 are well-defined.

risk-neutral principal cares about effort e_i, as we assume, or about output s_i). Our general observational structure also allows for stochastic outputs which are correlated across the agents like in Green and Stokey (1983) or Nalebuff and Stiglitz (1983). 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.

2.2. Contests

A *contest* (y,π) is defined by a prize profile y and an allocation rule π. The prize profile $y=(y_1,\ldots,y_n)$ is w.l.o.g. assumed to satisfy $y_1\geq\ldots\geq y_n\geq 0$. The allocation rule π determines the possibly random allocation of the prizes to the agents as a function of the realized signal s. Formally, $\pi=(\pi_i^k(s))_{i,k\in I,s\in S}$ is a collection of allocation probabilities, where $\pi_i^k(s)$ is the probability that agent i gets prize y_k when the realized signal is s. For any fixed signal s, these allocation probabilities satisfy $\sum_{i=1}^n \pi_i^k(s)=1$ for all $k=1,\ldots,n$ (each prize is allocated with probability one) and $\sum_{k=1}^n \pi_i^k(s)=1$ for all $i=1,\ldots,n$ (each agent obtains a prize with probability one). In other words, the probabilities form a doubly stochastic matrix for any given s. By the Birkhoff-von Neumann theorem, each doubly stochastic matrix can be decomposed as a probability distribution over permutation matrices, which in our setting describe deterministic allocations of the prizes to the agents.[4] Conversely, each probability distribution over permutation matrices generates a doubly stochastic matrix. This allows us to identify the allocation rule with a collection of doubly stochastic matrices, one for each signal realization.[5]

Given the fixed observational structure (S,η) and a contest (y,π), the probability that agent i wins prize y_k when the effort profile is e can be calculated as

$$p_i^k(e)=\mathbb{E}_{\eta^e}[\pi_i^k(s)]. \tag{1}$$

It follows that the winning probabilities in (1) also form a doubly stochastic matrix for any given effort profile e, because they are an average of doubly stochastic matrices. We refer to the collection of probabilities $p=(p_i^k(e))_{i,k\in I,e\in E}$ as the *contest success function* (CSF). The incentives of the agents depend exclusively on the probabilities of winning the different prizes as a function of their efforts, and these are jointly determined by the allocation rule π chosen by the principal and the exogenously given distribution of the signals η. With perfect observability of effort, the distinction between π and p is not important. However, as we will see later, what the principal can implement will depend on the extent to which she can influence the CSF p by choice of the allocation rule π.

Example 1. Suppose efforts can be perfectly observed ($s=e$ with probability one). Assume that there are two agents, so that the prize profile is $y=(y_1,y_2)$ with $y_1\geq y_2\geq 0$. If the principal designs an all-pay contest, we obtain for $i,j=1,2$ with $i\neq j$,

[4] The decomposition is not necessarily unique, but this is not relevant in our setting since the agents care only about their own probabilities of winning each prize.

[5] Similar to before, we require the regularity condition that each $\pi_i^k(s)$ is a measurable function of s.

$$\pi_i^1(e) = p_i^1(e) = \begin{cases} 1 & \text{if } e_i > e_j, \\ 1/2 & \text{if } e_i = e_j, \\ 0 & \text{if } e_i < e_j. \end{cases}$$

If the principal designs a Tullock contest with impact function f, we have

$$\pi_i^1(e) = p_i^1(e) = \begin{cases} \dfrac{f(e_i)}{f(e_i) + f(e_j)} & \text{if } \max\{e_i, e_j\} > 0, \\ 1/2 & \text{else.} \end{cases}$$

In this interpretation, the noise in the prize allocation is deliberately designed by the principal and not the consequence of imperfect observation of efforts. □

Example 2. Now suppose the efforts of the two agents are imperfectly observed, and $s = (s_1, s_2) \in \mathbb{R}^2$. Assume the principal designs an all-pay allocation rule as a function of the signals,

$$\pi_i^1(s) = \begin{cases} 1 & \text{if } s_i > s_j, \\ 1/2 & \text{if } s_i = s_j, \\ 0 & \text{if } s_i < s_j. \end{cases}$$

The induced CSF then depends on the shape of the observational noise. If, for example, $s_i = e_i + \epsilon_i$ and the noise terms ϵ_i are i.i.d. Gumbel with mean zero, then it follows like in the well-known logit model (McFadden, 1974) that

$$p_i^1(e) = \frac{\exp(e_i/\beta)}{\exp(e_i/\beta) + \exp(e_j/\beta)}$$

for some $\beta > 0$. In this interpretation, a specific Tullock CSF arises as a consequence of imperfect observation of efforts. See Jia et al. (2013) for similar results for various other noise structures. □

Given a contest, the agents choose their efforts simultaneously, anticipating that the prizes y will be distributed according to the CSF induced by the observational structure and the allocation rule. 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, π) *implements* a strategy profile σ if it satisfies

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

where $\Pi_i(\sigma \mid (y, \pi)) = \mathbb{E}_\sigma\left[\sum_{k=1}^n p_i^k(e) u(y_k)\right] - \mathbb{E}_{\sigma_i}[c(e_i)]$ and the winning probabilities are given by (1). The principal chooses a contest (y, π) which implements a strategy profile σ in order to maximize her expected payoff. Formally, the principal's problem is given by

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

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

3. Optimal contests

3.1. General properties

The level of effort that the principal can induce will depend on the observational structure. As a simple illustration of this point, consider the case when the signals are independent of the effort profile. In this case, since effort is costly and does not affect the distribution of signals and prizes, the agents would always choose zero effort and the principal optimally decides not to award any positive prizes. However, when signals are sufficiently informative about efforts, the principal will be able to design contests which generate strictly positive payoffs.

In this subsection, we derive general conditions which are sufficient for a contest to be optimal. These conditions are stated in terms of the prize profile and the induced CSF. Whether it is possible for the principal to generate a CSF satisfying these conditions will depend on the exogenously given observational structure. In the subsequent subsections, we will study various observational structures, check whether the sufficient conditions can be attained, and explore the shape of the required allocation rule π.

Consider the prize profile $y^* = (x^*/(n-1), \dots, x^*/(n-1), 0)$ where the total sum x^* is characterized by

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

Profile y^* features $n-1$ equal positive prizes and a single zero prize. An agent participating in any contest (y^*, π) only cares about the aggregate probability of obtaining any one of the identical positive prizes. Denote this probability by $p_i^{-n}(e) = 1 - p_i^n(e)$. Furthermore, let the effort level e^* be given by

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

We will denote by (e_i, e^*_{-i}) the effort profile where agent i chooses effort e_i and all agents $j \neq i$ choose effort $e_j = e^*$. Letina et al. (2020) have shown that – in a setting with perfect observability of effort – optimal contests feature the prize profile y^* and implement the symmetric effort profile $(e^*, \dots, e^*)$. The resulting maximal payoff $\Pi^* = ne^* - x^*$ under perfect information is clearly an upper bound on the principal's payoff with an arbitrary observational structure. Our following proposition gives conditions under which a contest achieves this upper bound even with imperfect observability of effort and is thus optimal.

Proposition 1. *Fix an arbitrary observational structure (S, η). If a contest (y, π) has the prize profile $y = y^*$ and the CSF satisfies, for each $i \in I$,*

(i) $p_i^{-n}(e^*, e^*_{-i}) = \frac{n-1}{n}$, *and*
(ii) $p_i^{-n}(e_i, e^*_{-i}) \leq \frac{c(e_i)}{u(x^*/(n-1))}$, $\forall e_i \neq e^*$,

then it is optimal, because it implements $(e^*, \dots, e^*)$ and achieves the payoff bound Π^*.

To achieve optimality, property (i) of the proposition requires that in equilibrium all agents must receive a positive prize with equal probability $(n-1)/n$. Note that, if the observational

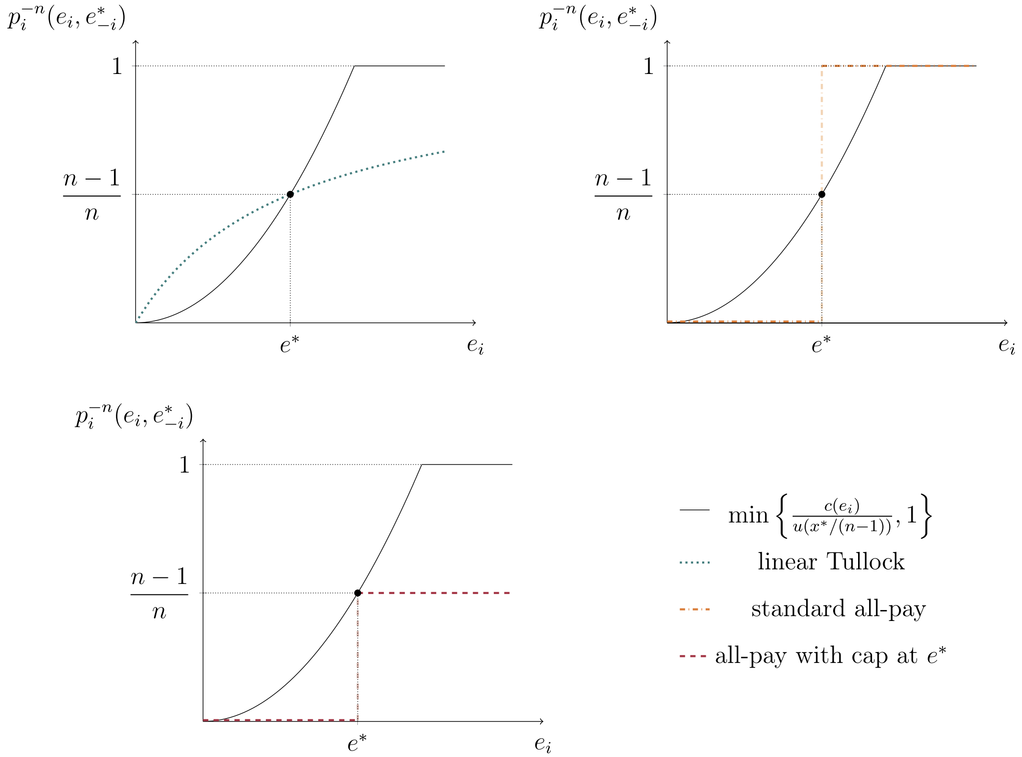

Fig. 1. Probability that agent i wins a positive prize when deviating from e^*, calculated for $n = 2$ and $u(t) = \sqrt{t}$ and $c(e) = e^2$.

structure (S, η) is not symmetric across agents, the allocation rule π must compensate that asymmetry. Property (*ii*) specifies how precisely the CSF has to discriminate between different levels of effort in order to achieve the optimum. If an agent deviates to some effort $e_i \neq e^*$, the probability of winning a positive prize has to remain below a certain boundary. This boundary is a continuous and strictly increasing function of the deviation effort, and therefore optimal CSFs must have an *intermediate level of precision*. They must be precise enough to detect and punish downward deviations sufficiently strongly, but they are not allowed to be too precise in detecting and rewarding upward deviations.

This insight is illustrated in Fig. 1 for the case of two agents. The top left panel shows the probability that the deviating agent i wins the positive prize with a linear Tullock CSF where $p_i^1(e) = e_i/(e_i + e_j)$. As noted before, this probability could be due to deliberate randomization of the principal and/or noise in the observation of efforts. The linear Tullock CSF is *not sufficiently precise* in punishment because downward deviations from e^* do not lead to a sufficient decrease in the winning probability which would deter the deviation. The top right panel shows a standard all-pay CSF, which the principal can induce if she observes at least the ordinal ranking of the agents' efforts. This all-pay CSF has the opposite problem; it is *too precise* in rewarding upward deviations because small upward deviations lead to a too large increase in the probability of winning, making such deviations profitable. The bottom left panel shows an all-pay CSF with a cap at e^*. It perfectly discriminates any downward deviation but does not discriminate upward deviations, and hence is feasible whenever the principal can detect downward deviations from the equilibrium effort with probability one. In a setting with perfect observability of efforts, Letina et al. (2020, Theorem 3) have shown that an all-pay contest with a cap at e^* is optimal. However, as Proposition 1 and Fig. 1 suggest, there are other CSFs which can achieve the optimum. We will demonstrate that some of these CSFs are feasible even with quite imprecise or coarse observation of efforts.

Fang et al. (2020) have shown that reducing inequality in the prize profile is beneficial for the principal in a contest with an all-pay CSF, which only admits mixed-strategy equilibria. Reducing prize inequality reduces dispersion of efforts chosen in the mixed equilibrium, and random effort choice is inefficient due to convex effort costs. Fang et al. (2020) therefore conclude that it is optimal in all-pay contests to move towards the least unequal prize profile with $n-1$ identical positive prizes and one prize of zero (see also Glazer and Hassin, 1988; Letina et al., 2020). Generalizing this insight by Fang et al. (2020), our Proposition 1 shows that a minimally competitive prize profile y^* is optimal also when the CSF is not exogenously fixed to be all-pay. The general message of Fang et al. (2020) is that "turning up the heat" in an all-pay contest increases the dispersion of the equilibrium effort distributions and decreases the expected equilibrium effort that agents exert. As a consequence, the principal should turn down the heat by using the minimally competitive prize profile. Our result in Proposition 1 shows that it is optimal to turn down the heat of the contest even more, by moving away from the perfectly discriminatory and thus very competitive all-pay CSF towards less discriminatory and hence less competitive CSFs. The optimal precision of the CSF is such that dispersion of equilibrium efforts vanishes entirely and a pure-strategy equilibrium emerges.

3.2. *Perfect observability of effort*

Perfect observability of effort is a special case of the observational structure (S, η) where $S = E$ and η^e is the Dirac measure on e. When effort is perfectly observable, we now show that, in addition to the all-pay contest with a cap at e^*, a properly designed Tullock contest can also achieve the optimum. This result is interesting for at least three reasons: (i) it shows that the optimum can be achieved by a smooth and strictly increasing contest success function, (ii) Tullock CSFs can be naturally ordered by a precision parameter which will provide insights about the optimal intensity of competition in contests, and (iii) it shows that a commonly studied contest format is optimal.

Tullock contests with n agents and a single positive prize are typically characterized by an allocation function of the form

$$\pi_i^1(e) = p_i^1(e) = \frac{f(e_i)}{\sum_{j\in I} f(e_j)}. \tag{2}$$

The impact function f is continuous, strictly increasing and satisfies $f(0) = 0$ (Skaperdas, 1996). 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).[6] The first prize is allocated according to (2) among all n agents, the second prize is allocated according to (2) restricted to those $n-1$ agents who have not received the first prize, and so on. To write this in our notation, let P_i^k denote the set of all permutations $\tau : I \to I$ which

[6] For sufficient conditions guaranteeing existence and uniqueness of pure strategy equilibria in nested Tullock contests see Fu et al. (2021b, 2022).

satisfy $\tau(i) = k$. When all efforts are strictly positive, a nested Tullock contest then gives rise to the allocation probabilities

$$\pi_i^k(e) = p_i^k(e) = \sum_{\tau \in P_i^k} \prod_{\ell=1}^{n} \left[\frac{f(e_{\tau^{-1}(\ell)})}{\sum_{j=\ell}^{n} f(e_{\tau^{-1}(j)})} \right]. \tag{3}$$

The extension to the case where some efforts are zero is straightforward.

Proposition 2. *Suppose efforts are perfectly observed. Then, the nested Tullock contest is optimal if the prize profile is $y = y^*$ and the CSF is given by (3) with*

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

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

To prove Proposition 2, 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 the Tullock exponent 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. The inequality $r \geq r^*(n)$ reflects that the CSF must be sufficiently precise to deter downward deviations. When $r \leq r^*(n)$, there are no profitable deviations to higher effort levels. The inequality $r \leq r^*(n)$ reflects that the CSF cannot be too precise, as otherwise upward deviations would become attractive. Altogether, the optimal Tullock contest features the intermediate precision parameter $r^*(n)$ and therefore an intermediate intensity of competition.[7] This is illustrated in Fig. 2, again for the case of two agents. The optimal Tullock CSF is smooth and tangential to our upper bound from Proposition 1 at the equilibrium effort.

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 two-agent Tullock contest still has a pure-strategy equilibrium. This property of the optimal contest carries over to $n > 2$. The precision parameter $r^*(n)$ is such that any increase in r would destroy the pure-strategy equilibrium.[8]

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 of the agents, 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 2 always has

[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. It may also appear that an all-pay contest with a cap is more "detail-free" than the optimal Tullock contest, since the function f depends on c. However, this is not the case, as the level of the cap in the optimal all-pay contest also depends on c (as does the optimal prize profile y^*).

[8] This is also similar to the finding in Morgan et al. (2022), 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.

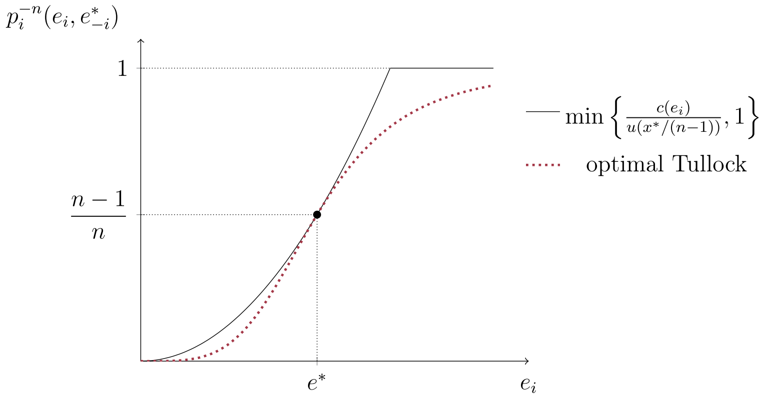

Fig. 2. Probability that agent i wins a positive prize when deviating from e^* in the optimal Tullock contest, calculated for $n = 2$ and $u(t) = \sqrt{t}$ and $c(e) = e^2$.

the highest possible precision parameter $r^*(n) = (n-1)/(H_n - 1)$ from along the continuum. 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), these weaker incentives cannot be compensated by a more unequal prize profile when the agents are risk-averse.

As already mentioned by Schweinzer and Segev (2012), the randomness parameter $r^*(n)$ is strictly increasing in n and satisfies $\lim_{n\to\infty} r^*(n) = \infty$. 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.[9]

3.3. *Imperfect observability of effort: some instructive cases*

In this section, we look for optimal contests when observability of effort is imperfect. We first consider several cases where our upper payoff bound can be achieved despite the imperfect observation. When that is the case, Proposition 1 can be used to greatly simplify the identification of optimal contests. Then, we consider one case in which observation is so imperfect that the upper payoff bound can no longer be achieved, and use it to show how our approach can be fruitful even in such environments.

3.3.1. *Upper payoff bound achievable*

The most common way in which imperfect observation is modeled in the literature is to assume that a random shock is added to the effort that an individual agent exerts. Formally, we say that the observational structure (S, η) features *i.i.d. additive noise* if, for each agent $i \in I$, the principal observes a signal $s_i = e_i + \varepsilon_i$, where the noise terms ε_i are i.i.d. draws from a distribution with cumulative distribution function F and support contained in an interval $[\underline{\varepsilon}, \overline{\varepsilon}]$.

Given this noise structure, we can derive a condition on the distribution F under which the principal can choose an allocation function which generates a CSF satisfying the optimality

[9] Our Appendix A.3 contains a formal proof of that claim. See Siegel (2009) for a general treatment of all-pay contests and Olszewski and Siegel (2016) for large contests.

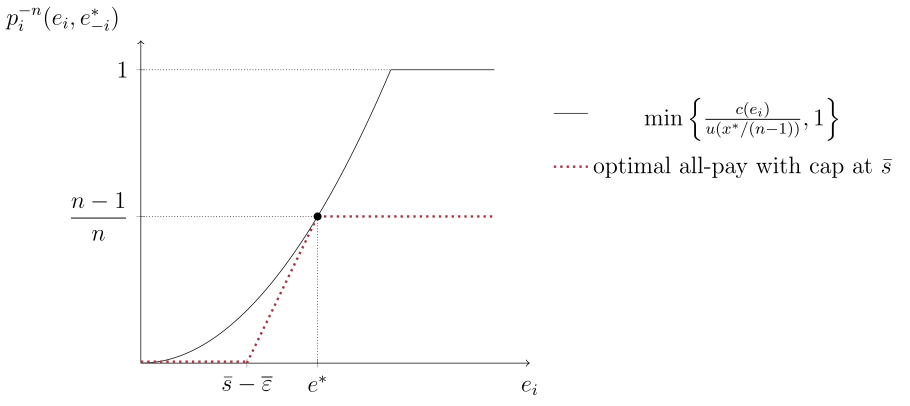

Fig. 3. Probability that agent i wins a positive prize when deviating from e^* in the optimal all-pay contest with a cap for i.i.d. additive noise $\varepsilon_i \sim U[-0.1, 0.1]$, calculated for $n = 2$ and $u(t) = \sqrt{t}$ and $c(e) = e^2$.

conditions in Proposition 1, and we can study properties of the optimal CSF. To this end, denote by F^- the left-continuous limit of F, i.e., $F^-(x) = \lim_{y \nearrow x} F(y)$ for all x.

Proposition 3. *Suppose efforts are observed with i.i.d. additive noise. If*

$$F^-(\underline{\varepsilon} + e^* - e) \geq 1 - \frac{c(e)}{c(e^*)}, \quad \forall e \in [0, e^*], \tag{4}$$

then a contest with prize profile $y = y^*$ and an all-pay allocation rule with a cap at $\bar{s} = e^* + \underline{\varepsilon}$ is optimal.

Any agent who exerts the equilibrium effort e^* will generate a signal $s_i \geq \bar{s}$, which is at or above the cap. An agent who unilaterally deviates downwards to $e_i \in [\bar{s} - \bar{\varepsilon}, e^*)$ may still generate a signal at or above the cap with positive probability and thus win a prize with positive probability. In order for $(e^*, \ldots, e^*)$ to be an equilibrium, the observational structure must be precise enough so that such downwards deviations are detected and punished with sufficiently high probability. Proposition 3 provides exactly this condition on the distribution of observational noise. Upwards deviations are never rewarded because of the cap, like in the case of an all-pay contest with a cap under perfect observability. Fig. 3 illustrates this for the case of two agents. The optimal CSF is a continuous modification of the previous all-pay contest with a cap, and it is feasible despite imperfect observation of downwards deviations from equilibrium.

Perfect observation of efforts is still a special case to which Proposition 3 applies, by setting $\underline{\varepsilon} = \bar{\varepsilon} = 0$. Thus, the optimality of an all-pay contest with a cap for perfectly observable efforts by Letina et al. (2020) is a corollary of Proposition 3.

The following corollary demonstrates a straightforward application of the condition in Proposition 3.

Corollary 1. *Suppose ε_i is uniformly distributed and $\Delta \equiv \bar{\varepsilon} - \underline{\varepsilon} < e^*$. Then, the condition in Proposition 3 is satisfied if $\Delta \leq c(e^*)/c'(e^*)$.*

The i.i.d. additive noise setting of Proposition 3 is by far not the only case in which optimal contests can be derived using our Proposition 1. We illustrate this with two additional examples.

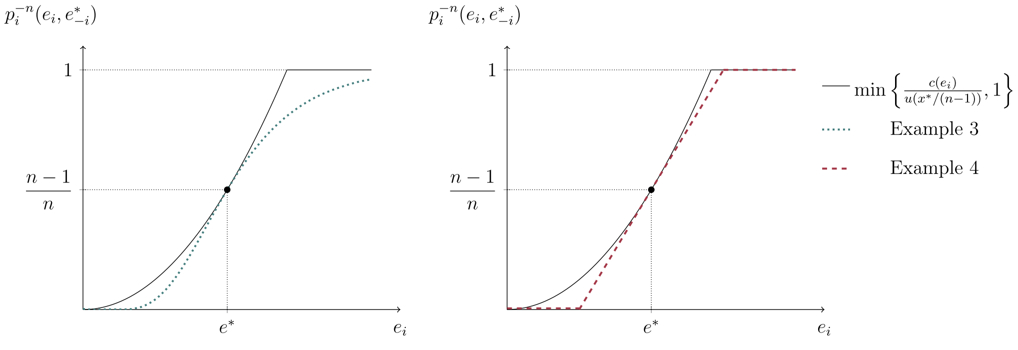

Fig. 4. Probability that agent i wins a positive prize when deviating from e^* in Examples 3 and 4, calculated for $u(t) = \sqrt{t}$ and $c(e) = e^2$.

In the first example, the observational noise has an unbounded support and can be correlated across the agents. In the second example, observation is not noisy, but only a coarse aggregate statistic of the effort profile can be observed.

Example 3. Consider a setting with two agents and effort cost function $c(e_i) = \gamma e_i^\beta$ for some $\gamma > 0$ and $\beta > 1$. The observational noise takes a multiplicative or log-additive form: when agent i exerts effort e_i, then a signal $s_i = e_i r_i$ is generated, where the pair (r_1, r_2) follows a bivariate log-normal distribution,

$$(r_1, r_2) \sim \ln \mathcal{N}\left[\begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right].$$

We show in Appendix A.6 that the principal can choose an allocation rule which generates a CSF satisfying the conditions in Proposition 1 whenever the inequality

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

holds, which again just requires that the observational noise is not too large. The optimal contest allocates the positive prize to agent 1 with a probability that is increasing in the ratio of observed signals s_1/s_2. More precisely, agent 1 receives the prize when s_1/s_2 is larger than a log-normally distributed random number, and conversely for agent 2. Similar *contests with multiplicative noise* have been studied in the literature.[10] With this construction, the overall randomness in the prize allocation can be adjusted to the appropriate interior level which guarantees optimality. This is illustrated in the left panel of Fig. 4.

Example 4. Consider a setting with two agents in which the difference $s = e_1 - e_2$ can be observed (and perfectly so), but no additional information about the effort profile. We show in Appendix A.6 that, despite this strong constraint, the principal can always choose an allocation rule which generates a CSF satisfying the conditions in Proposition 1. The optimal contest allocates the positive prize to agent 1 with a probability that is increasing in the observed difference

[10] See, for instance, Jia et al. (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).

s. More precisely, agent 1 receives the prize whenever s is larger than a uniformly distributed random number, and conversely for agent 2. Such *contests with additive noise* have also been studied in the literature.[11] An appropriate level of randomness in the allocation rule again ensures that the contest has the optimal intermediate level of precision. This situation is illustrated in the right panel of Fig. 4.

3.3.2. *Upper payoff bound not achievable*

Of course, Proposition 1 is not always applicable. One example is the extreme case of uninformative signals discussed in Section 3.1. More generally, it will be impossible to implement the effort profile $(e^*, \ldots, e^*)$ when the signals on which the prize allocation can be conditioned are too noisy or too coarse.

To illustrate that our approach can still be fruitful in such environments, consider the following example of an observational structure. Given the agents' effort profile $e \in E$, the principal observes a signal $s \in S = E$ which fully reveals the truth ($s = e$) with probability $\omega \in [0, 1]$; with the remaining probability, the signal is pure noise, generated by a fixed probability measure $\hat{\eta} \in \Delta E$ that is independent of effort. This "truth-or-noise" signal structure has been studied before in the information economics literature (e.g., Johnson and Myatt, 2006; Lewis and Sappington, 1994; Shi, 2012). It has the advantage of providing a measure of the informational friction of the environment. By varying the value of ω, we can cover both the cases of perfect observability ($\omega = 1$) and zero observability ($\omega = 0$).

The possibility of pure noise prohibits the principal from achieving the same payoff as under perfect observability. To see why, note that the agents receive zero rents in optimal contests with perfect observability. With pure noise, the principal has to leave some rents to the agents, because the noise implies that some agents must win a prize with positive probability even when they deviate to zero effort and have zero costs. The optimal contest can still be derived using a modified approach. Let us define the prize profile $y^\omega = (x^\omega/(n-1), ..., x^\omega/(n-1), 0)$ $y^\omega = (x^\omega, \ldots, x^\omega, 0)$ for every $\omega \in [0, 1]$, where the total sum x^ω is uniquely pinned down by

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

When $\omega > 0$, the prize profile y^ω features $n-1$ equal positive prizes and a single zero prize; for $\omega = 0$, all prizes are zero. It again suffices to describe any contest (y^ω, π) by its probability of assigning one of the identical positive prizes to each agent i given the signal realization $s \in E$, which we write as $\pi_i^{-n}(s) = 1 - \pi_i^n(s)$. Further, let the effort level e^ω be given by

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

We will use (e_i, e_{-i}^ω) to denote the effort profile where agent i chooses effort e_i and all agents $j \neq i$ choose effort $e_j = e^\omega$.

Proposition 4. *Suppose the observational structure is truth-or-noise. A contest (y, π) is optimal if the prize profile is $y = y^\omega$ and the allocation rule π satisfies, for each $i \in I$,*

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

(*i*) $\pi_i^{-n}(e^\omega, e_{-i}^\omega) = \frac{n-1}{n}$, *and*
(*ii*) $\pi_i^{-n}(e_i, e_{-i}^\omega) \leq \frac{c(e_i)}{u(x^\omega/(n-1))}$, $\forall e_i \neq e^\omega$.

Proposition 4 implies the existence of an optimal contest – one can always construct an allocation rule π that satisfies both conditions (i) and (ii). The proposition also subsumes the optimal contest in Letina et al. (2020) as a special case, because $y^\omega = y^*$ and $e^\omega = e^*$ when $\omega = 1$. However, when $\omega < 1$, the principal can never be sure whether an agent is working or shirking, and has to leave a strictly positive rent to the agents if she wants to incentivize them to work. Although a contest with $n-1$ equal positive prizes and a single zero prize remains optimal, the principal finds it desirable to limit the agents' information rents by providing a lower incentive power ($x^\omega < x^*$) and asking them to exert less effort than before ($e^\omega < e^*$).

3.4. *Second-best payoffs*

In this subsection, we compare the principal's payoffs when using an optimal contest with the second-best payoffs, where second-best is defined as what the principal can achieve by using an arbitrary incentive mechanism without being constrained to the class of contests, but still under a possibly imperfect observational structure. Can the principal do better by not using a contest? The answer to this question depends crucially on the observational structure, as we show by the following three examples.

Example 5. There are observational structures for which an optimal contest is second-best, so that the restriction to the class of contest mechanisms comes without loss for the principal. Consider an example with two agents and two signals, $I = S = \{1, 2\}$. The signals are generated by the probabilities

$$\eta^e(\{i\}) = \begin{cases} 1 & \text{if } e_i < e_j = e^*, \\ 0 & \text{if } e_j < e_i = e^*, \\ 1/2 & \text{otherwise,} \end{cases}$$

for $i, j = 1, 2$ with $i \neq j$. In words, if one agent deviates downward from (e^*, e^*), this agent is announced. In all other cases, one of the two agents is announced randomly.

Consider the contest (y, π) with $y = (x^*, 0)$ and

$$\pi_i^1(s) = \begin{cases} 1 & \text{if } s = j, \\ 0 & \text{if } s = i, \end{cases}$$

so agent i wins the positive prize if and only if the signal announces the other agent. It is easy to see that this contest implements (e^*, e^*). Hence it is an optimal contest and generates an average payoff per agent of size $e^* - x^*/2$ for the principal.

We claim that this is also the highest payoff that the principal can achieve when using more general incentive mechanisms. For the given observational structure, a general mechanism is described by $\Phi = \{(\bar{t}_i, \underline{t}_i)\}_{i=1,2}$, where $\bar{t}_i$ ($\underline{t}_i$) is transfer that the principal pays to agent i when the signal is $s = j$ ($s = i$).[12] Note that an effort level $e' > e^*$ can never be a best response for an

[12] It is without loss to restrict attention to deterministic transfers given each signal, because the agents are weakly risk-averse and the principal is risk-neutral.

agent, irrespective of what strategy the other agent is playing, because reducing the effort slightly saves effort costs without changing the distribution of the signal. By the same argument, an effort level $0 < e' < e^*$ can never be a best response. Hence we can restrict attention to effort strategies which are binary distributions over $\{0, e^*\}$ and denote by σ_i the probability that agent i chooses e^*. Given σ_j, agent i weakly prefers $e = e^*$ over $e = 0$ if and only if

$$\left(1 - \frac{\sigma_j}{2}\right) \cdot u(\bar{t}_i) + \frac{\sigma_j}{2} \cdot u(\underline{t}_i) - c(e^*) \geq \frac{1 - \sigma_j}{2} \cdot u(\bar{t}_i) + \frac{1 + \sigma_j}{2} \cdot u(\underline{t}_i),$$

or, equivalently, $u(\bar{t}_i) - u(\underline{t}_i) \geq c(e^*)/2 = u(x^*)$, which is thus a necessary condition to generate any positive effort in equilibrium. The most cost-effective way to meet this inequality is to set $\bar{t}_i = x^*$ and $\underline{t}_i = 0$. But then it is clear that the principal cannot do better with an arbitrary mechanism than with the optimal contest. □

Example 6. There are other observational structures for which the principal can do better by not using a contest. Consider the case of n agents and perfect observability of effort. In that case, the principal can treat each agent separately and pay a monetary transfer if and only if the agent exerts a desired effort level. Transfers only have to compensate agents for their cost. It is thus possible to achieve the first-best solution. The first-best effort is

$$e^{FB} = c^{-1}\left(u\left(x^{FB}\right)\right),$$

and the first-best transfer x^{FB} to an agent is defined by the first-order condition

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

Denote by $\Pi^{FB} = e^{FB} - x^{FB}$ the first-best payoff per agent for the principal. It is easy to show that $e^* \leq e^{FB}$ and $e^* - x^*/n \leq \Pi^{FB}$, and the inequalities are strict whenever u is strictly concave. □

Example 7. There are also examples where the principal can do better than with a contest even though the first-best is not achievable. Suppose that each agent's effort is perfectly observable up to a cap $\bar{e}$ with $e^* < \bar{e} < e^{FB}$, but the individual signal remains capped at $\bar{e}$ for all higher effort levels. It is immediate to see that the first-best is not achievable. However, it is possible to elicit the effort $\bar{e}$ from each agent by paying the transfer $u^{-1}(c(\bar{e}))$ if the signal indicates that the effort was at least $\bar{e}$ and zero otherwise. This generates a payoff for the principal that is strictly larger than with a contest. □

Giving a general answer to the question which share of second-best payoffs the principal can achieve by using a contest is impeded by the fact that we (and decades of literature) do not know the second-best for all conceivable observational structures, many of which are untractable. We now show, however, that the problem becomes less pressing when the number of agents is large, because the contest payoffs converge to the first-best as n grows.

Let $(S_n, \eta_n)_{n \in \mathbb{N}}$ be a sequence of observational structures such that there exists a contest achieving our payoff bound when the number of agents is n and the observational structure is given by (S_n, η_n), for every $n \in \mathbb{N}$. Examples include perfect observation of each agent's effort or i.i.d. additive noise as characterized in Proposition 3. Denote by x_n^* the optimal sum of prizes in the contest when there are n agents, defined as before, and by e_n^* the corresponding individual effort elicited by an optimal contest. Let $\Pi_n^* = e_n^* - x_n^*/n$ denote the average payoff per agent for the principal.

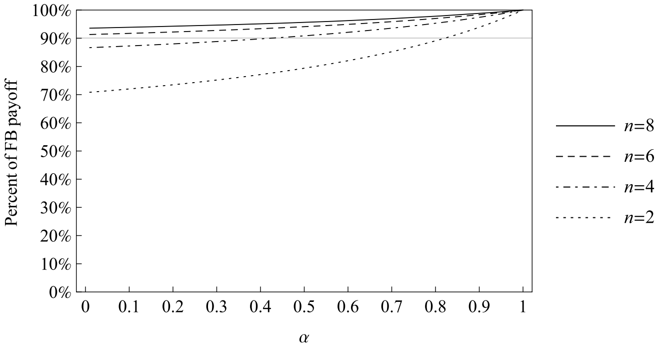

Fig. 5. Share of first-best payoffs with an optimal contest.

Proposition 5. $\lim_{n\to\infty} e_n^* = e^{FB}$ *and* $\lim_{n\to\infty} \Pi_n^* = \Pi^{FB}$.

Similar arguments can be made about risk-aversion. For a parameterized example where the agents' payoffs are $\Pi_i(e_i, t_i) = t_i^{\alpha} - e_i^2$, Fig. 5 depicts the percentage of first-best payoffs that the principal can achieve with an optimal contest as a function of the risk-aversion parameter α and for several values of n. As $\alpha \to 1$ the share of first-best payoffs that the principal can capture converges to one. Even for a modest number of agents, the principal obtains a substantial share of the first-best by running an optimal contest. For instance, already for $n = 6$ the principal captures more than 90% of the first-best payoffs for any $\alpha \in (0, 1)$.

4. Extensions

4.1. Heterogeneous agents

Our framework can also incorporate heterogeneity in the abilities of the agents. Consider a variation of the 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 2 (which assumes perfect effort observability) to arbitrary cost functions.

Proposition 6. *Suppose efforts are perfectly observed and* $n = 2$. *Then, for any profile of cost functions* (c_1, c_2)*, the following contest 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 is of the Tullock type (2) with individual-specific impact functions*

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

For the special case where $c_1(\cdot) = c_2(\cdot) = c(\cdot)$, 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 2 for $n = 2$. With asymmetric cost functions, by contrast, the optimal impact functions must be individual-specific.[13] The implemented effort levels will typically also not be identical for the two agents. Consequently, the winning probabilities cannot 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; Franke et al., 2018). Our result establishes the form of biasing that is optimal when the principal is not restricted to a specific class of CSFs.

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 $(c_1^m, \ldots, c_n^m)_{m\in\mathbb{N}}$ 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 7. *Suppose efforts are perfectly observed and let $(c_1^m, \ldots, c_n^m) \to (c, \ldots, c)$ 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 7 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 6, namely

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

Given the complexity of the problem, we leave the question whether a suitably defined asymmetric nested Tullock contest can achieve the optimum, and the possible extensions to imperfect effort observation, to future research.

4.2. Costly entry

Often, simply participating in a contest is costly. For example, applying for a research grant requires an investment of time and effort to understand the rules and requirements. Furthermore, these costs are likely to be agent-specific and, from the principal's point of view, uncertain. The

[13] See, for example, Cornes and Hartley (2005) for an equilibrium analysis of Tullock contests with individual-specific impact functions. Sahm (2022) studies the optimal choice of the Tullock exponent r with heterogeneous agents under the constraint that the contest must be symmetric.

literature has studied such situations but typically restricts the analysis to a specific class of CSFs (e.g., Fu et al., 2015; Liu and Lu, 2019; Morgan et al., 2012). This section shows how our generalized approach to contest design can be extended to settings with costly entry.

We again restrict attention to the case of perfectly observable efforts. Suppose that each of the n agents has a private entry cost $z_i \geq 0$ which is independently drawn according to some common cumulative distribution function G. This cost is assumed to be additive to the agents' previous payoff functions. Each agent observes his entry cost and decides whether to enter the contest or not. The set of entrants then becomes observable to everyone. The sunk entry cost z_i is irrelevant after the entry decision. Furthermore, we require contests to be anonymous, so that an agent's name i is also irrelevant. This allows us to denote by $I^m = \{1, \ldots, m\}$ the set of agents in the contest for each positive number of entrants $m \leq n$ and to drop the index i from the active agents' payoff functions. The corresponding set of possible effort profiles is $E^m = \mathbb{R}^m_+$.

The principal specifies the rules of the contest in advance, for each positive number of entrants $m \leq n$. Thus, a contest $((y^m, \pi^m))_{m=1,\ldots,n}$ describes for each m

(i) a prize vector $y^m = (y^m_1, ..., y^m_m) \in \mathbb{R}^m_+$ with $y^m_1 \geq \ldots \geq y^m_m$, and
(ii) anonymous allocation probabilities $\pi^m = (\pi_i^{m,k}(e))_{i,k \in I^m, e \in E^m}$.

We restrict attention to the implementation of symmetric strategy profiles characterized by $(\bar{z}, e^1, \ldots, e^n)$: each agent chooses to participate in the contest if and only if $z_i \leq \bar{z}$ for some common cutoff $\bar{z} \geq 0$, and when m agents join the contest, each active agent $i \in I^m$ chooses the same effort level $e^m \geq 0$.

If an agent is active and chooses effort e in a contest with m participants while all other $m-1$ participants are choosing the effort level e^m, then he will obtain the expected payoff

$$\Pi^m \left(e, e^m_{-i} \mid (y^m, \pi^m)\right) = \sum_{k=1}^{m} \pi_i^{m,k} \left(e, e^m_{-i}\right) u(y^m_k) - c(e).$$

We say that a contest $((y^m, \pi^m))_{m=1,\ldots,n}$ implements $(\bar{z}, e^1, \ldots, e^n)$ if both of the following conditions are satisfied:

(i) $\bar{\Pi}(m) \equiv \Pi^m \left(e^m, e^m_{-i} \mid (y^m, \pi^m)\right) \geq \Pi^m \left(e, e^m_{-i} \mid (y^m, \pi^m)\right)$, $\forall e \in \mathbb{R}_+$ and $m \in \{1, ..., n\}$,
(ii) $\sum_{m=1}^{n} \binom{n-1}{m-1} G(\bar{z})^{m-1}(1-G(\bar{z}))^{n-m}\bar{\Pi}(m) = \bar{z}$.

Condition (i) is simply the previous constraint (IC-A) applied to each possible number of entrants separately. Condition (ii) determines the cutoff $\bar{z}$ at which an agent is indifferent between participating or not, anticipating that all other agents apply the same cutoff for their entry decision. The expected payoff of the principal is given by

$$\Pi_P(\bar{z}, e^1, ..., e^n \mid ((y^m, \pi^m))_{m=1,\ldots,n}) = \sum_{m=1}^{n} \binom{n}{m} G(\bar{z})^m (1-G(\bar{z}))^{n-m} \left[m e^m - \sum_{k=1}^{m} y^m_k \right].$$

Our next result identifies two important characteristics of optimal contests.

Proposition 8. *Suppose efforts are perfectly observed. Take any contest* $((y^m, \pi^m))_{m=1,\ldots,n}$ *that implements some* $(\bar{z}, e^1, ..., e^n)$*. Then, there exists a contest* $((\hat{y}^m, \hat{\pi}^m))_{m=1,\ldots,n}$ *that yields a weakly higher expected payoff to the principal and implements* $(\bar{z}, \hat{e}^1, ..., \hat{e}^n)$ *such that, for each* $m = 1, ..., n$*,*

(i) *the prize profile satisfies* $\hat{y}_1^m = \ldots = \hat{y}_{m-1}^m \geq \hat{y}_m^m$, *and*
(ii) *the agents are indifferent between choosing the effort level* $\hat{e}^m$ *and zero.*

The proposition implies that the principal can without loss of generality restrict attention to contests which satisfy the conditions (i) and (ii). Such contests have $m - 1$ identical prizes, for each number of entrants m, which shows the robustness of our previous result. There is a difference to the case of a fixed number of agents, though. In the model without endogenous entry, the lowest prize in the optimal contest is always zero, the agents are indifferent between the equilibrium effort and an effort of zero, and consequently their expected payoff is zero. With endogenous entry this is not optimal – the principal has to leave some rents to the agents in order to incentivize costly entry. There are two ways in which the principal could do this, either by increasing the identical positive prizes but leaving the zero prize unchanged, or by increasing all prizes simultaneously, so that the agents are still indifferent between the equilibrium effort and an effort of zero. Condition (ii) reveals that the principal optimally uses the latter approach. This provides more insurance to the risk-averse agents who are harmed by the possibility of receiving the low prize.

4.3. Risk-loving agents

In our main model, we assumed that agents are either risk-neutral or risk-averse. This is in line with most of the contest theory literature. However, our approach can be equally well applied to contests with risk-loving agents.

Suppose that u is strictly convex (while keeping all other assumptions). For any prize sum $x > 0$, we define the effort level $e^x = c^{-1}(u(x)/n)$. As the next result shows, assuming risk-loving agents changes our result on the optimal prize profile, which becomes winner-take-all, but leaves our result on the intermediate level of precision of the CSF unchanged.

Proposition 9. *Fix an arbitrary observational structure* (S, η). *For any fixed prize sum* $x > 0$, *a contest* (y, π) *maximizes the principal's payoff if the prize profile is* $y = (x, 0, ..., 0)$ *and the CSF satisfies, for each* $i \in I$,

(i) $p_i^1\left(e^x, e_{-i}^x\right) = \frac{1}{n}$, *and*
(ii) $p_i^1(e_i, e_{-i}^x) \leq \frac{c(e_i)}{u(x)}$, $\forall e_i \neq e^x$.

When the agents are risk-averse, the principal wants to provide the maximal degree of insurance that is still compatible with incentives to provide effort, and this is achieved with $n - 1$ equal prizes. With risk-loving preferences, by contrast, the principal wants to provide the maximal degree of risk, as risk creates value for the agents, and this is achieved by allocating the entire budget to a single positive prize. The shape of the optimal CSF is driven by a different consideration. The competitiveness of the CSF must be chosen with the goal of inducing the highest possible pure strategy effort equilibrium, and this gives rise to an intermediate degree of competitiveness – or heat as Fang et al. (2020) call it – just like before. It can be achieved by appropriately choosing a cap in an all-pay contest, but other CSFs are also optimal.

Note that Proposition 9 holds for any exogenously fixed prize sum $x > 0$. The principal can then try to find the optimal prize sum by solving

$$\max_{x \in \mathbb{R}_+} nc^{-1}\left(\frac{u(x)}{n}\right) - x.$$

Note, however, that a solution to this problem may not exist, because both c and u are strictly convex. Depending on their relative convexity, the objective function of the principal may not have a maximum. Of course, if the principal was budget-constrained, the optimal prize sum would then be equal to the maximal budget.

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] and then existing results on the optimal CSF.

For the class of Tullock CSFs, Clark and Riis (1998) show that, if a symmetric pure-strategy equilibrium exists for a winner-take-all (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 et al. (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 Lazear-Rosen tournaments, Morgan et al. (2022) 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 et al. (2012), Liu et al. (2018) and Liu and Lu (2021). Similar results for Tullock contests and Lazear-Rosen tournaments can be found in Sela (2020) 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. (2021a) for Tullock contests, and Drugov and Ryvkin (2021) 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

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

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. (2022) 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 contest theory literature has also developed foundations for various functional forms of the CSF (for a comprehensive survey, see Jia et al., 2013). Our results contribute to this literature by characterizing the family of CSFs that can implement the optimal outcome, with perfect or imperfect observability of effort.

The main contribution of our paper is to study jointly optimal choice of the prize profile and the CSF. More recently, Zhang (2021) has revisited the contest design problem with incomplete information. Like us, she also allows the principal to choose both the prize profile and the CSF. She provides a necessary and sufficient condition for WTA to be optimal and characterizes the optimal prize profile when this condition fails. The main difference between our papers is that Zhang (2021) focuses on risk-neutral agents with perfect observability of effort and private types, while we consider general risk attitudes and imperfect observability of effort, but without private types.

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 or a particular observational structure. We provide easily verifiable sufficient conditions for a contest, described by a prize profile and a contest success function, to be optimal given an arbitrary observational structure. We apply these conditions to various settings. With perfect observability of effort, an appropriately chosen nested Tullock contest is optimal. When efforts are imperfectly observed, we derive an upper bound on the level of noise such that an all-pay contest with a cap is optimal. We also show how our tools can be used in cases where the agents are heterogeneous in their abilities, entry into the contest is costly, and when the agents have risk-loving preferences.

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 three 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 appropriately.

Second, we do not examine the number of equilibria in optimal contests. If other equilibria exist, there is a danger that the agents will coordinate on suboptimal equilibria, especially if they generate higher payoffs to the agents. We do know that the equilibrium is unique for some optimal contests. For example, with perfect observability of effort, both a two-agent Tullock contest and an n-agent all-pay contest with a cap have a unique equilibrium.[16] We do not know whether the equilibrium is unique in the optimal Tullock contest when there are more than two agents. To the best of our knowledge, the only paper investigating uniqueness in multi-prize Tullock contests is Fu et al. (2021b), but they restrict attention to contests that are less precise than what would be optimal, so we cannot rely on their results.

Third, we have mostly focused on observational structures for which the principal can implement the same outcome as in the optimal contest with perfect observability of efforts. As we have discussed earlier, when the observability of efforts is very limited, implementing this outcome will no longer be possible. The general characterization of optimal contests in those circumstances remains an open question. We provide the characterization for a truth-or-noise observational structure. In this setting, we show that optimal contests still feature $n-1$ equal positive prizes, a single zero prize, and a CSF with an intermediate degree of competitiveness. However, based on the intuition gained from our results, we conjecture that there are also observational structures where optimal contests feature more top-heavy prize structures. This could be the case if the optimal "heat" cannot be generated via the CSF due to observational noise, but competitiveness can be increased via the other channel – the prize vector. Understanding which observational structures lead to flat prizes and which to top-heavy ones is an interesting avenue for future research.

Data availability

No data was used for the research described in the article.

Appendix A. Proofs

A.1. *Proof of Proposition 1*

Fix an arbitrary observational structure (S, η) and suppose that a contest (y, π) with prize profile $y = y^*$ satisfies conditions (i) and (ii) in the proposition. We first claim that it implements the effort profile $(e^*, ..., e^*)$. If agent i exerts effort e_i when all other agents exert e^*, his payoff his

$$\Pi_i(e_i, e^*_{-i} \mid (y, \pi)) = p_i^{-n}(e_i, e^*_{-i}) u\left(\frac{x^*}{n-1}\right) - c(e_i).$$

From condition (i) and the definition of e^* it follows that $\Pi_i(e^*, e^*_{-i} \mid (y, \pi)) = 0$. From condition (ii) it follows that

[15] Classical references are Taylor (1995) and Che and Gale (2003), while more recent examples are Erkal and Xiao (2019), Lemus and Temnyalov (2021) and Benkert and Letina (2020). For a similar objective in *prediction contests* see Lemus and Marshall (2021). Another possible objective of contest design is the selection of best agents. For examples of such *selection contests* see Meyer (1991), and Fang and Noe (2022) for a more recent contribution.

[16] See Ewerhart (2017) and Letina et al. (2020).

$$\Pi_i(e_i, e^*_{-i} \mid (y,\pi)) \leq \frac{c(e_i)}{u(x^*/(n-1))} u\left(\frac{x^*}{n-1}\right) - c(e_i) = 0$$

for all $e_i \neq e^*$, which proves the claim.

Now suppose by contradiction that (y,π) is not optimal, i.e., there exists a contest $(\tilde{y},\tilde{\pi})$ that implements some strategy profile σ and

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

Construct a contest $(\tilde{y},\hat{\pi})$ for a setting with perfect observation of efforts by defining

$$\hat{\pi}_i^k(e) = \mathbb{E}_{\eta^e}\left[\tilde{\pi}_i^k(s)\right]$$

for all $i,k \in I$ and all $e \in E$. It follows that the induced CSF $\hat{p}$ of the contest $(\tilde{y},\hat{\pi})$ with perfect observation is identical to the induced CSF $\tilde{p}$ of the contest $(\tilde{y},\tilde{\pi})$ with the original observational structure (S,η). Since the prize profiles are also identical, it follows that $(\tilde{y},\hat{\pi})$ implements σ under perfect observation and achieves a payoff for the principal strictly larger than $ne^* - x^*$. This is a contradiction to the following Lemma 1 for the setting with perfect observability of efforts, which is due to Letina et al. (2020) and which we state without proof.[17]

Lemma 1 *(Letina et al., 2020). Suppose efforts are perfectly observed. Then, 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

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

We conclude that (y,π) must be optimal for the observational structure (S,η). ■

A.2. Proof of Proposition 2

Consider a contest with prize profile $y^* = (x^*/(n-1), ..., x^*/(n-1), 0)$ and allocation rule π of the nested Tullock form (3). 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

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

in the proposition. In Step 3, we then complete the proof that the 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

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

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

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

$$
\begin{aligned}
&= \frac{\left(\prod_{k=1}^{n-1}[1+(n-k)x]\right)(n-1)\left((n-1)!x^{n-2}\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]} \\
&\quad - \frac{\left((n-1)!x^{n-1}\right)\left(\sum_{k=1}^{n-1}(n-k)\prod_{j\neq k}[1+(n-j)x]\right)}{\left(\prod_{k=1}^{n-1}[1+(n-k)x]-(n-1)!x^{n-1}\right)\prod_{k=1}^{n-1}[1+(n-k)x]}.
\end{aligned}
$$

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

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

Thus we have

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

To complete the proof of Proposition 2, 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

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

and, since

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

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

[18] Proposition 1.1 in Pinelis (2002) 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 $n = 3$.

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

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

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

$$\begin{aligned}
&\sum_{k=2}^{j-2}(n-2)(n-1)a_{n-k}a_{n-j+k} - \sum_{k=2}^{j-2}(n-k)(n-k-1)a_{n-k}a_{n-j+k} \\
&\quad - \sum_{k=2}^{j-2}2(n-1)(n-k)a_{n-k}a_{n-j+k} + \sum_{k=2}^{j-2}2(n-k)(n-j+k)a_{n-k}a_{n-j+k} \\
&= \sum_{k=2}^{j-2}(n^2-3n+2)a_{n-k}a_{n-j+k} - \sum_{k=2}^{j-2}(n^2-2nk-n+k^2+k)a_{n-k}a_{n-j+k} \\
&\quad - \sum_{k=2}^{j-2}2(n^2-nk-n+k)a_{n-k}a_{n-j+k} + \sum_{k=2}^{j-2}2(n^2-nj+jk-k^2)a_{n-k}a_{n-j+k} \\
&= \sum_{k=2}^{j-2}(2+4nk-3k^2-3k-2nj+2jk)a_{n-k}a_{n-j+k}.
\end{aligned}$$

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

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

Importantly, for the last term we have

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

$$= 2 + j\left(\frac{j}{4} - \frac{3}{2}\right)$$
$$> 0,$$

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

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

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

$$\begin{aligned}
a_1 &= \sum_{k_1=1}^{n-1}(n - k_1), \\
a_2 &= \sum_{k_2=1}^{n-2}\sum_{k_1=k_2+1}^{n-1}(n - k_2)(n - k_1), \\
&\vdots \\
a_j &= \sum_{k_j=1}^{n-j}\sum_{k_{j-1}=k_j+1}^{n-j+1}\cdots\sum_{k_1=k_2+1}^{n-1}(n - k_j)(n - k_{j-1})\ldots(n - k_1).
\end{aligned}$$

Intuitively, each summand in the definition of a_j is the product of j different elements chosen from the set $\{(n-1), (n-2), \ldots, 1\}$, 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):

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

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

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

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.*

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

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

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 $j\in\{4,\dots,n+1\}$,

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

[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.

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

$$(n-1)!a_{n-j+1} \le a_{n-k}a_{n-j+k}. \tag{5}$$

To demonstrate condition (5), 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

$$a_{n-2} = \sum_{k_1=1}^{n-1} \frac{(n-1)!}{n-k_1},$$

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

$$\vdots$$

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

$$\vdots$$

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

Rewriting condition (5), we then have

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

Observe that for each summand on the LHS, the denominator is a product of $j-2$ different elements from the set $\{(n-1), (n-2), \ldots, 1\}$. 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.3. 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

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

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.4. Proof of Proposition 3

Consider a contest with prize profile $y = y^*$ and an all-pay allocation rule π with a cap at $\bar{s} = e^* + \underline{\varepsilon}$. Note that when an agent chooses any effort level $e_i \geq e^*$, the value of his signal will be at or above the cap $\bar{s}$ with probability one. Therefore, when agent i exerts effort e_i while all other agents $j \neq i$ choose effort e^*, the probability that agent i will get one of the identical positive prizes is given by

$$\begin{aligned} p_i^{-n}(e_i, e_{-i}^*) &= \left(1 - \Pr\left(e_i + \varepsilon_i < e^* + \underline{\varepsilon}\right)\right) \cdot \frac{n-1}{n} \\ &= \left(1 - F^-(\underline{\varepsilon} + e^* - e_i)\right) \cdot \frac{n-1}{n}. \end{aligned}$$

Since $F^-(x) = 0$ for all $x \leq \underline{\varepsilon}$, we have, for all $e_i \geq e^*$,

$$p_i^{-n}(e_i, e_{-i}^*) = \frac{n-1}{n} = \frac{c(e^*)}{u\,(x^*/(n-1))} \leq \frac{c(e_i)}{u\,(x^*/(n-1))}.$$

Hence condition (i) in Proposition 1 is satisfied, and condition (ii) is satisfied for all $e_i > e^*$. Furthermore, the inequality condition (4) implies that, for all $e_i \in [0, e^*)$,

$$p_i^{-n}(e_i, e_{-i}^*) \leq \frac{n-1}{n} \cdot \frac{c(e_i)}{c(e^*)} = \frac{c(e^*)}{u\,(x^*/(n-1))} \cdot \frac{c(e_i)}{c(e^*)} = \frac{c(e_i)}{u\,(x^*/(n-1))}.$$

Therefore, condition (ii) is also satisfied for all $e_i < e^*$. It follows from Proposition 1 that the proposed contest is optimal. ■

A.5. Proof of Corollary 1

With the uniform distribution given in the corollary, we have $F^- = F$ and condition (4) becomes

$$Q(e) \equiv \frac{\min\{e^* - e, \Delta\}}{\Delta} - 1 + \frac{c(e)}{c(e^*)} \geq 0 \;\; \forall e \in [0, e^*].$$

$Q(e)$ is continuous, satisfies $Q(0) = Q(e^*) = 0$, and $Q(e) > 0$ for all $e \in (0, e^* - \Delta)$. Moreover, we have for all $e \in (e^* - \Delta, e^*)$,

$$Q'(e) = -\frac{1}{\Delta} + \frac{c'(e)}{c(e^*)} \leq -\frac{1}{\Delta} + \frac{c'(e^*)}{c(e^*)} \leq 0,$$

where the first inequality follows from convexity of c and the second inequality follows from the assumption $\Delta \leq c(e^*)/c'(e^*)$. Therefore, on the interval $[0, e^*]$, the function $Q(e)$ is first positive and then decreases (weakly) towards zero. It follows that $Q(e) \geq 0$ holds for all $e \in [0, e^*]$. ■

A.6. Examples of imperfect effort observation

A.6.1. Example 3

Suppose 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 the positive prize x^* is given to agent 1 if and only if $rs_1/s_2 \geq 1$, where $r \sim \ln\mathcal{N}[\nu_r, \sigma_r^2]$ is distributed log-normally with parameters

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

This allows for $\sigma_r^2 = 0$, by which we mean that r is degenerate and takes the value e^{ν_r} with probability one. Formally, the principal sets

$$\pi_1^1(s) = \Pr\left[\frac{rs_1}{s_2} \geq 1\right],$$

and $\pi_2^1(s) = 1 - \pi_1^1(s)$. Given any effort profile e, the probability that agent 1 wins the positive prize is then given by

$$p_1^1(e) = \Pr\left[\frac{rr_1e_1}{r_2e_2} \geq 1\right] = \Pr\left[\frac{r_2}{rr_1} \leq \frac{e_1}{e_2}\right].$$

Since r_1, r_2 and r are all log-normally distributed, it follows that $r_2/(rr_1)$ is also log-normal, with location $\nu = \nu_2 - \nu_1 - \nu_r = 0$ and scale $\sigma^2 = \sigma_1^2 + \sigma_2^2 - \sigma_{12} + \sigma_r^2 = 2/(\pi\beta^2)$. The cdf of the log-normal distribution is given by $F(x) = \Phi((\log x - \nu)/\sigma)$, where Φ is the cdf of the standard normal distribution. Thus we can write

$$p_1^1(e) = \Phi\left(\log(e_1/e_2)\beta\sqrt{\frac{\pi}{2}}\right).$$

For the probability that agent 2 wins the prize we obtain

$$\begin{aligned} p_2^1(e) = 1 - p_1^1(e) &= 1 - \Phi\left(\log(e_1/e_2)\beta\sqrt{\frac{\pi}{2}}\right) \\ &= \Phi\left(-\log(e_1/e_2)\beta\sqrt{\frac{\pi}{2}}\right) \\ &= \Phi\left(\log(e_2/e_1)\beta\sqrt{\frac{\pi}{2}}\right). \end{aligned}$$

It follows immediately that $p_i^1(e^*, e^*) = 1/2$, which is condition (i) in Proposition 1. We will now establish that

$$p_i^1(e_i, e^*) \leq \frac{c(e_i)}{u(x^*)} = \frac{c(e_i)}{2c(e^*)} = \frac{1}{2}\left(\frac{e_i}{e^*}\right)^{\beta},$$

for all $e_i \neq e^*$, which is condition (ii) in Proposition 1, and where the first equality follows by definition of e^*. After the change of variables $x = \log(e_i/e^*)\beta\sqrt{\pi/2}$, this inequality becomes the requirement that

$$\Phi(x) \leq \frac{1}{2}e^{x\sqrt{2/\pi}} \tag{6}$$

for all $x \in \mathbb{R}$. Inequality (6) 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 (6) 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 $\Phi(x) = \mathrm{erfc}(-x/\sqrt{2})/2$, where

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

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

$$\mathrm{erfc}(y) \leq e^{-2y/\sqrt{\pi}} \tag{7}$$

for all $y > 0$. Inequality (7) 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 (7) 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

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

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 (7) is satisfied also for $y > 2/\sqrt{\pi}$.

A.6.2. Example 4

Consider a contest with prize profile $y^* = (x^*, 0)$ in which the positive prize x^* is given to agent 1 if and only if $s + r \geq 0$, where $r \sim \mathcal{U}[-c(e^*)/c'(e^*), c(e^*)/c'(e^*)]$ is distributed uniformly. Formally, the principal sets

$$\pi_1^1(s) = \Pr[r + s \geq 0],$$

and $\pi_2^1(s) = 1 - \pi_1^1(s)$. Observe that $c(e^*)/c'(e^*) < e^*$ holds due to strict convexity of c and $c(0) = 0$. We can then write the probability that agent 1 wins the positive prize, holding the effort $e_2 = e^*$ fixed, as a piecewise function

$$p_1^1(e_1, e^*) = \begin{cases} 1 & \text{if } e_1 > e^* + \frac{c(e^*)}{c'(e^*)}, \\ \frac{1}{2} + \frac{1}{2}\frac{c'(e^*)}{c(e^*)}(e_1 - e^*) & \text{if } e^* - \frac{c(e^*)}{c'(e^*)} \leq e_1 \leq e^* + \frac{c(e^*)}{c'(e^*)}, \\ 0 & \text{if } e_1 < e^* - \frac{c(e^*)}{c'(e^*)}. \end{cases}$$

It follows immediately that $p_1^1(e^*, e^*) = p_2^1(e^*, e^*) = 1/2$, which is condition (i) in Proposition 1. We will now establish condition (ii) in Proposition 1, first for agent $i = 1$. It is trivially satisfied for any $e_1 < e^* - c(e^*)/c'(e^*)$. Next, consider any e_1 with $e^* - c(e^*)/c'(e^*) \leq e_1 \leq e^* + c(e^*)/c'(e^*)$. Condition (ii) can then be rearranged to

$$e_1 - \frac{c(e_1)}{c'(e^*)} \leq e^* - \frac{c(e^*)}{c'(e^*)}.$$

The LHS is strictly concave and reaches its unique maximum at $e_1 = e^*$, where it equals the constant RHS. Hence the inequality holds. The fact that condition (ii) also holds for any $e_1 > e^* + c(e^*)/c'(e^*)$ follows because it holds for $e_1 = e^* + c(e^*)/c'(e^*)$, from the previous argument, where we already have $p_1^1(e_1, e^*) = 1$. The argument for agent 2 is symmetric.

A.7. Proof of Proposition 4

Fix a total prize sum x and define the effort level $e(\omega, x) = c^{-1}(\omega u(x/(n-1))(n-1)/n)$. If the principal could use a contest (y, π) with $\sum_{k=1}^n y_k = x$ to implement the pure strategy profile $(e(\omega, x), ..., e(\omega, x))$, her expected payoff would be $ne(\omega, x) - x$. We will first prove that, under the truth-or-noise observational structure, $ne(\omega, x) - x$ constitutes an upper bound on the principal's payoff in any contest with a total prize sum x.

Take any contest (y, π) with $\sum_{k=1}^n y_k = x$. Suppose that this contest implements some strategy profile σ. Then, implementing σ requires that

$$\mathbb{E}_\sigma\left[\sum_{k=1}^n p_i^k(e)u(y_k) - c(e_i)\right] \geq \mathbb{E}_{\sigma_{-i}}\left[\sum_{k=1}^n p_i^k(0, e_{-i})\right]u(y_k) \tag{8}$$

for all $i \in I$. Now, recall that with probability $1-\omega$, the signal s will be generated by a probability measure $\hat{\eta} \in \Delta E$ that is independent of effort. Hence, given an effort profile e, the probability that agent i wins prize y_k can be calculated as

$$p_i^k(e) = \omega\pi_i^k(e) + (1-\omega)\mathbb{E}_{\hat{\eta}}\left[\pi_i^k(s)\right]. \tag{9}$$

Using (9) and that $y_1 \geq \ldots \geq y_n$, we obtain the following implication of condition (8):

$$\omega \left[\mathbb{E}_\sigma \left[\sum_{k=1}^{n} \pi_i^k(e) u(y_k) \right] - u(y_n) \right] \geq \mathbb{E}_{\sigma_i}[c(e_i)], \tag{10}$$

for all $i \in I$. Summing up (10) over all $i \in I$, we obtain

$$\omega \left[\sum_{k=1}^{n-1} u(y_k) - (n-1)u(y_n) \right] \geq \sum_{i=1}^{n} \mathbb{E}_{\sigma_i}[c(e_i)]. \tag{11}$$

Since $\sum_{k=1}^{n} y_k = x$, u is concave and c is convex, (11) further implies

$$c\left(\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_{\sigma_i}[e_i]\right) \leq \omega \frac{n-1}{n} u\left(\frac{x}{n-1}\right) = c(e(\omega, x)). \tag{12}$$

Because c is strictly increasing, we finally have

$$\sum_{i=1}^{n} \mathbb{E}_{\sigma_i}[e_i] - x \leq ne(\omega, x) - x.$$

That is, for a fixed total prize sum x, the expected payoff of the principal can never exceed $ne(\omega, x) - x$, which is the upper bound that we mentioned before.

By construction, x^ω is the unique solution to the following payoff-maximization problem of the principal:

$$\max_{x \geq 0} \; ne(\omega, x) - x.$$

Hence, $ne^\omega - x^\omega$ constitutes an upper bound for the principal's payoff in any contest. It is clear that a contest satisfying the conditions of the proposition will be able to implement the pure strategy effort profile $(e^\omega, \ldots, e^\omega)$ and yields the expected payoff $ne^\omega - x^\omega$ to the principal. Therefore, such a contest must be optimal. ■

A.8. *Proof of Proposition 5*

The optimal prize sum x_n^* is defined by the first-order condition

$$u'\left(\frac{x_n^*}{n-1}\right) = c'\left(c^{-1}\left(\frac{n-1}{n} u\left(\frac{x_n^*}{n-1}\right)\right)\right). \tag{13}$$

Holding $x_n^*/(n-1)$ fixed, the LHS of (13) is constant and the RHS is strictly increasing in n. Furthermore, the LHS is weakly decreasing and the RHS is strictly increasing in $x_n^*/(n-1)$. It thus follows that $x_n^*/(n-1)$ must be strictly decreasing in n. It also holds that $x_n^*/(n-1) > x^{FB}$. By contradiction, from $x_n^*/(n-1) \leq x^{FB}$ we would obtain

$$\begin{aligned} u'(x^{FB}) &\leq u'\left(\frac{x_n^*}{n-1}\right) \\ &= c'\left(c^{-1}\left(\frac{n-1}{n} u\left(\frac{x_n^*}{n-1}\right)\right)\right) \\ &< c'\left(c^{-1}\left(u\left(\frac{x_n^*}{n-1}\right)\right)\right) \\ &\leq c'\left(c^{-1}\left(u\left(x^{FB}\right)\right)\right) = u'(x^{FB}). \end{aligned}$$

It thus follows that $\lim_{n\to\infty} x_n^*/(n-1)$ exists. We claim that $\lim_{n\to\infty} x_n^*/(n-1) = x^{FB}$. By contradiction, if $\lim_{n\to\infty} x_n^*/(n-1) = \bar{x} > x^{FB}$, then the LHS of (13) converges to $u'(\bar{x})$ and the RHS converges to $c'(c^{-1}(u(\bar{x})))$. From $u'\left(x^{FB}\right) = c'\left(c^{-1}\left(u\left(x^{FB}\right)\right)\right)$ together with $\bar{x} > x^{FB}$ we conclude that $u'(\bar{x}) < c'\left(c^{-1}(u(\bar{x}))\right)$, and hence (13) must be violated for sufficiently large n. It now follows that

$$\lim_{n\to\infty} e_n^* = \lim_{n\to\infty} c^{-1}\left(\frac{n-1}{n}u\left(\frac{x_n^*}{n-1}\right)\right) = c^{-1}\left(u\left(x^{FB}\right)\right) = e^{FB}.$$

It also follows that

$$\lim_{n\to\infty} \Pi_n^* = \lim_{n\to\infty}\left(e_n^* - \frac{x_n^*}{n}\right) = \lim_{n\to\infty}\left(e_n^* - \frac{x_n^*}{n-1}\right) = e^{FB} - x^{FB} = \Pi^{FB},$$

which completes the proof. ■

A.9. *Proof of Proposition 6*

We first derive three lemmas which hold under cost heterogeneity for any number n of agents. Since we assume perfect observability of efforts, we do not make a distinction between the allocation rule π and the CSF p. We use the notation π throughout.

Lemma 5. *For any contest (y, π) that implements a strategy profile σ, there exists a contest $(y, \hat{\pi})$ 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, π) implements σ. Define an allocation rule $\hat{\pi}$ as follows:

$$\hat{\pi}_i^k(\tilde{e}) = \begin{cases} \mathbb{E}_\sigma\left[\pi_i^k(e)\right] & \text{if } \tilde{e} = \bar{e}, \\ \mathbb{E}_\sigma\left[\pi_i^k(0, e_{-j})\right] & \text{if } \tilde{e}_j \neq \bar{e}_j \text{ and } \tilde{e}_\ell = \bar{e}_\ell \ \forall \ell \neq j, \\ \pi_i^k(\tilde{e}) & \text{otherwise,} \end{cases}$$

for all $i, k \in I$. We now show that, in the contest $(y, \hat{\pi})$, for each agent $i \in I$ it is a best response to play $\bar{e}_i$ when the remaining agents are playing $\bar{e}_{-i}$, which implies that $(y, \hat{\pi})$ implements $\bar{e}$. This claim holds because, for any $i \in I$ and $\forall e_i' \neq \bar{e}_i$,

$$\begin{aligned}
\Pi_i(\bar{e} \mid (y, \hat{\pi})) &= \sum_{k=1}^{n} \hat{\pi}_i^k(\bar{e}) u(y_k) - c_i(\bar{e}_i) \\
&= \sum_{k=1}^{n} \mathbb{E}_\sigma\left[\pi_i^k(e) u(y_k)\right] - c_i\left(\mathbb{E}_\sigma[e_i]\right) \\
&\geq \sum_{k=1}^{n} \mathbb{E}_\sigma\left[\pi_i^k(e) u(y_k)\right] - \mathbb{E}_\sigma\left[c_i(e_i)\right] \\
&\geq \sum_{k=1}^{n} \mathbb{E}_\sigma\left[\pi_i^k(0, e_{-i}) u(y_k)\right] \\
&\geq \sum_{k=1}^{n} \mathbb{E}_\sigma\left[\pi_i^k(0, e_{-i}) u(y_k)\right] - c_i(e_i')
\end{aligned}$$

$$= \sum_{k=1}^{n} \hat{\pi}_i^k(e_i', \bar{e}_{-i})u(y_k) - c_i(e_i')$$
$$= \Pi_i(e_i', \bar{e}_{-i} \mid (y, \hat{\pi})),$$

where the first inequality follows the convexity of c_i and the second inequality follows from the fact that (y, π) 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, π) 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, π) implements $\bar{e}$, for each $i \in I$ it must hold that

$$c_i(\bar{e}_i) \leq \sum_{k=1}^{n} \pi_i^k(\bar{e}_i, \bar{e}_{-i})u(y_k) - \sum_{k=1}^{n} \pi_i^k(0, \bar{e}_{-i})u(y_k).$$

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

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

where the first equality holds because the sum of all agents' expected utilities from the prizes is the same for all effort profiles (due to π being a doubly stochastic matrix for all e), the third

[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.

inequality follows from concavity of u, and the fourth inequality follows from concavity of u together with the fact that $\sum_{i=2}^{n}\sum_{k=1}^{n}\pi_i^k(0,\bar{e}_{-1})y_k \leq \sum_{k=1}^{n} y_k = x$. □

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

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

Proof. Suppose that (y,π) implements $\bar{e}$. Then, in particular,

$$\sum_{k=1}^{n}\pi_i^k(\bar{e})u(y_k)-c_i(\bar{e}_i)\geq\sum_{k=1}^{n}\pi_i^k(0,\bar{e}_{-i})u(y_k)\geq u(y_n)\ \forall i\in I. \tag{14}$$

Now consider another contest $(\tilde{y},\tilde{\pi})$ with $\tilde{y}=(\tilde{y}_1,\dots,\tilde{y}_{n-1},\tilde{y}_n)=(y_1,\dots,y_{n-1},0)$ and any allocation rule $\tilde{\pi}$ such that $\tilde{\pi}_i^k(\bar{e})=\pi_i^k(\bar{e})$ for all $i,k\in I$, and $\tilde{\pi}_i^n(e_i,\bar{e}_{-i})=1$ whenever $e_i\neq\bar{e}_i$. By construction, we have for all $i\in I$,

$$\begin{aligned}\sum_{k=1}^{n}\tilde{\pi}_i^k(\bar{e})u(\tilde{y}_k)-c_i(\bar{e}_i)&=\sum_{k=1}^{n-1}\pi_i^k(\bar{e})u(y_k)-c_i(\bar{e}_i)\\&\geq[1-\pi_i^n(\bar{e})]u(y_n)\\&\geq 0\\&=u(\tilde{y}_n),\end{aligned}$$

where the first inequality follows from (14). Hence, $(\tilde{y},\tilde{\pi})$ also implements $\bar{e}$. □

From now on we consider the special case of $n=2$. It follows from Lemma 6 that any contest (y,π) 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^*,π^*) 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^* when the effort profile is $e=(e_i,e_j)$ is

$$\pi_i^1(e_i,e_j)=\frac{c_i(e_i)^{r_i}/c_i(e_i^*)^{r_i-1}}{c_i(e_i)^{r_i}/c_i(e_i^*)^{r_i-1}+c_j(e_j)^{r_j}/c_j(e_j^*)^{r_j-1}}$$

$$= \frac{c_i(e_i)^{r_i} c_j(e_j^*)^{r_j-1}}{c_i(e_i)^{r_i} c_j(e_j^*)^{r_j-1} + c_j(e_j)^{r_j} c_i(e_i^*)^{r_i-1}}$$
$$= 1 - \frac{c_j(e_j)^{r_j} c_i(e_i^*)^{r_i-1}}{c_i(e_i)^{r_i} c_j(e_j^*)^{r_j-1} + c_j(e_j)^{r_j} c_i(e_i^*)^{r_i-1}}.$$

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 $U_i(c_i, c_j^*) = \pi_i^1(c_i, c_j^*)u(x^*) - c_i$. We obtain after some simplifications

$$\frac{\partial U_i(c_i, c_j^*)}{\partial c_i} = r_i \left[\frac{(c_i^*)^{r_i-1} c_j^* c_i^{r_i-1}}{(c_i^{r_i} + (c_i^*)^{r_i-1} c_j^*)^2} \right] u(x^*) - 1 \tag{15}$$

and

$$\frac{\partial^2 U_i(c_i, c_j^*)}{\partial c_i^2} = \frac{r_i u(x^*)(c_i^*)^{r_i-1} c_j^*}{(c_i^{r_i} + (c_i^*)^{r_i-1} c_j^*)^3} \left[(r_i - 1) c_i^{r_i-2} (c_i^{r_i} + (c_i^*)^{r_i-1} c_j^*) - 2 r_i c_i^{2(r_i-1)} \right]. \tag{16}$$

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 $r_i = 1 + c_i^*/c_j^* \equiv r_i^*$. From (15) we obtain

$$\frac{\partial U_i(c_i^*, c_j^*)}{\partial c_i} = \left(\frac{c_i^* + c_j^*}{c_j^*} \right) \left[\frac{(c_i^*)^{2(r_i^*-1)} c_j^*}{(c_i^*)^{2(r_i^*-1)} (c_i^* + c_j^*)^2} \right] (c_i^* + c_j^*) - 1 = 0, \tag{17}$$

so that the first-order condition is satisfied at $c_i = c_i^*$. By (16), 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 $r_i = r_i^*$ can be rearranged to

$$c_i^{r_i^*-2} (c_i^*)^{r_i^*} - (r_i^* + 1) c_i^{2(r_i^*-1)}. \tag{18}$$

Using (18) 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.10. *Proof of Proposition 7*

Since we assume perfect observability of efforts, we do not make a distinction between the allocation rule π and the CSF p. We use the notation π throughout. 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

$$\sum_{k=1}^{n} \pi_i^k(e)u(y_k) - c_i(e_i) \geq 0 \;\; \forall i \in I. \tag{19}$$

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 $\pi(e) = (\pi_i^k(e))_{i,k \in I} \in [0,1]^{n^2}$ for the given e, subject to (19) and the constraints that $y_n = 0$ and $\pi(e)$ is a doubly stochastic matrix. The allocation probabilities after multilateral deviations from e can be chosen arbitrarily. Using notation $x = \sum_{k=1}^n y_k$, constraint (19) 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, π) that implements a pure-strategy effort profile e such that $\Pi_P(e \mid (y, \pi)) = \underline{\Pi}$.*

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, π) that implements a pure-strategy effort profile $\bar{e}$ such that*

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

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

$$\frac{1}{n}\sum_{i=1}^{n} c_i^m(\bar{e}_i^m) \leq \max_{k\in I} c_k^m \left(\frac{1}{n}\sum_{i=1}^{n} \bar{e}_i^m\right). \tag{20}$$

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, π^m) implements $\bar{e}^m$, we must have

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

Using Theorem 2 in Barvinek et al. (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

$$\left|\left(\sum_{i=1}^{n} (c_i^m)^{-1}(u(x^m))\right) - x^m - \left(nc^{-1}(u(x^m)) - x^m\right)\right| < \epsilon. \tag{21}$$

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 (21) would imply that $\left(\sum_{i=1}^n (c_i^m)^{-1}(u(x^m))\right) - x^m < 0$, which in turn implies $\Pi_P(\bar{e}^m \mid (y^m, \pi^m)) < 0$. This is in contradiction to the assumption that $\Pi_P(\bar{e}^m \mid (y^m, \pi^m)) \geq \underline{\Pi}^m > 0$. Hence we know that $x^m \leq \tilde{B}$ for all $m \geq \underline{m}'$. Now simply let $B = \max\{x^1, \dots, x^{\underline{m}'-1}, \tilde{B}\}$. $\square$

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

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 $m \to \infty$.

[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 $e < 0$.

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

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

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

$$\lim_{m\to\infty}(c_i^m(\hat{e}^m) - c(\hat{e}^m)) = 0 \quad \text{and} \quad \lim_{m\to\infty}(c_i^m(\bar{e}_i^m) - c(\bar{e}_i^m)) = 0 \ \ \forall i \in I, \tag{22}$$

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 $c(\hat{e}^m) - \frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i^m) \leq 0$ for all $m \in \mathbb{N}$, and by condition (20) we have $\delta^m \geq 0$ for all $m \in \mathbb{N}$. Hence, we must also have

$$\lim_{m\to\infty}\left(c(\hat{e}^m) - \frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i^m)\right) = 0, \tag{23}$$

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

$$\lim_{m\to\infty}(c(\bar{e}_i^m) - c(\hat{e}^m)) = 0 \ \ \forall i \in I. \tag{24}$$

To establish (24), 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, π^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 $m \geq \underline{m}'$,

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

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 (23). Finally, (24) 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 $[0, \tilde{e}]$. $\square$

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 $\Pi_P(\bar{e}^m \mid (y^m, \pi^m)) \geq \underline{\Pi}^m$ 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 $m \geq \underline{m}'''$,

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

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. $\square$

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

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

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, π^m) and the effort profile $\bar{e}^m$ satisfy

$$\max_{k\in I} c_k^m(\bar{e}_k^m) \leq u\left(\frac{x^m}{n-1}\right). \tag{25}$$

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

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

Proof. We construct the allocation rule π 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, π) 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 (25) 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 $\sum_{i=1}^{n} p_i = 1$. $\square$

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

A.11. Proof of Proposition 8

Step 1. Take any contest $((y^m, \pi^m))_{m=1,\dots,n}$ that implements some $(\bar{z}, e^1, \dots, e^n)$. Fix any $m \in \{2, ..., n\}$. Let $\bar{y}^m = \frac{1}{m-1}\sum_{k=1}^{m-1} y_k^m$ and let $\hat{e}^m$ be such that

$$\frac{m-1}{m}u(\bar{y}^m) + \frac{1}{m}u(y_m^m) - c(\hat{e}^m) = \frac{1}{m}\sum_{k=1}^{m} u(y_k^m) - c(e^m).$$

It follows that $\hat{e}^m \geq e^m$, because by concavity we have

$$\frac{m-1}{m}u(\bar{y}^m) = \frac{m-1}{m}u\left(\frac{1}{m-1}\sum_{k=1}^{m-1} y_k^m\right) \geq \frac{1}{m}\sum_{k=1}^{m-1} u(y_k^m).$$

We now modify (if at all) the contest when there are m active agents. The modified prize profile $\hat{y}^m$ is given by $\hat{y}_1^m = \dots = \hat{y}_{m-1}^m = \bar{y}^m$ and $\hat{y}_m^m = y_m^m$. The (anonymous) allocation rule $\hat{\pi}^m$ is as follows. If all agents exert $\hat{e}^m$, then $\hat{y}_m^m$ is randomly and uniformly allocated. If an agent unilaterally deviates to some $e' \neq \hat{e}^m$, then he gets $\hat{y}_m^m$ with probability one. For all other effort profiles, the allocation rule can be chosen arbitrarily.

We claim that the modified contest implements $(\bar{z}, e^1, ..., \hat{e}^m, ...e^n)$. To see this, note first that

$$\begin{aligned}
\Pi^m\left(\hat{e}^m, \hat{e}_{-i}^m \mid (\hat{y}^m, \hat{\pi}^m)\right) &= \frac{m-1}{m}u(\bar{y}^m) + \frac{1}{m}u(y_m^m) - c(\hat{e}^m) \\
&= \Pi^m\left(e^m, e_{-i}^m \mid (y^m, \pi^m)\right) \\
&\geq u(y_m^m) \\
&\geq \Pi^m\left(e, \hat{e}_{-i}^m \mid (\hat{y}^m, \hat{\pi}^m)\right)
\end{aligned}$$

for all $e \neq \hat{e}^m$, where the second equality holds by construction and the first inequality follows because (y^m, π^m) implements $(e^m, ..., e^m)$. Hence, $(\hat{y}^m, \hat{\pi}^m)$ implements $(\hat{e}^m, ..., \hat{e}^m)$. By construction, the expected payoff of the agents remains unchanged (irrespective of the number of entrants). Therefore, the condition defining the cutoff $\bar{z}$ is also unaffected.

The principal gives away the same prize sum and collects weakly higher efforts from the agents in the modified contest with m entrants. Therefore, the expected payoff of the principal must be weakly higher. Repeating the argument for all m establishes property (i) of the proposition.

Step 2. Now consider the contest constructed in Step 1, where each prize profile satisfies $\hat{y}_1^m = \dots = \hat{y}_{m-1}^m = \bar{y}^m \geq \hat{y}_m^m$, which has the above defined allocation rule, and which implements $(\bar{z}, \hat{e}^1, \dots, \hat{e}^n)$. Suppose that for some $m \in \{2, ..., n\}$ we have

$$\Pi^m\left(\hat{e}^m, \hat{e}_{-i}^m \mid (\hat{y}^m, \hat{\pi}^m)\right) = \frac{m-1}{m}u(\bar{y}^m) + \frac{1}{m}u(\hat{y}_m^m) - c(\hat{e}^m) > u(\hat{y}_m^m),$$

which requires $\hat{y}_m^m < \bar{y}^m$. Let $y_L^m = \hat{y}_m^m + \varepsilon$ and $y_H^m = \bar{y}^m - \varepsilon/(m-1)$ and choose ε such that $u(y_L^m) = \Pi^m\left(\hat{e}^m, \hat{e}_{-i}^m \mid (\hat{y}^m, \hat{\pi}^m)\right)$. It is easy to see that such ε exists and $y_L^m \leq y_H^m$ still holds. Further, let $\tilde{e}^m$ be such that

$$\frac{m-1}{m}u(y_H^m)+\frac{1}{m}u(y_L^m)-c(\tilde{e}^m)=\Pi^m\left(\hat{e}^m,\hat{e}^m_{-i}\mid(\hat{y}^m,\hat{\pi}^m)\right).$$

Again by concavity, we have $\tilde{e}^m \geq \hat{e}^m$.

We now modify the contest when there are m active agents. The modified prize profile $\tilde{y}^m$ is given by $\tilde{y}_1^m=\ldots=\tilde{y}_{m-1}^m=y_H^m$ and $\tilde{y}_m^m=y_L^m$. The allocation rule $\tilde{\pi}^m$ is as follows. If all agents exert $\tilde{e}^m$, then $\tilde{y}_m^m$ is randomly and uniformly allocated. If an agent unilaterally deviates to some $e'\neq\tilde{e}^m$, then he gets $\tilde{y}_m^m$ with probability one. For all other effort profiles, the allocation rule can be chosen arbitrarily. Now observe that

$$\begin{aligned}\Pi^m\left(\tilde{e}^m,\tilde{e}^m_{-i}\mid(\tilde{y}^m,\tilde{\pi}^m)\right)&=\frac{m-1}{m}u(y_H^m)+\frac{1}{m}u(y_L^m)-c(\tilde{e}^m)\\&=\Pi^m\left(\hat{e}^m,\hat{e}^m_{-i}\mid(\hat{y}^m,\hat{\pi}^m)\right)\\&=u(y_L^m)\\&\geq\Pi^m\left(e,\tilde{e}^m_{-i}\mid(\tilde{y}^m,\tilde{\pi}^m)\right)\end{aligned}$$

for all $e\neq\tilde{e}^m$, which implies that $(\tilde{y}^m,\tilde{\pi}^m)$ implements $(\tilde{e}^m,...,\tilde{e}^m)$. The rest of the proof is analogous to Step 1. ■

A.12. *Proof of Proposition 9*

We first prove three intermediate results that apply to the case of perfect observability of efforts. Thereby, we do not make a distinction between the allocation rule π and the CSF p but use the notation π. We prove the proposition for an arbitrary observational structure afterwards.

Lemma 17. *Suppose that efforts are perfectly observable. For any contest (y,π) that implements a strategy profile σ such that there exists some $j\in I$ for which σ_j is not a Dirac measure, there exists a contest $(y,\hat{\pi})$ that implements a pure-strategy profile $\bar{e}$ and yields a strictly higher expected payoff to the principal.*

Proof. We will show that there exists an $\epsilon>0$ and an allocation rule $\hat{\pi}$ such that $(y,\hat{\pi})$ implements the pure-strategy profile $\bar{e}=(\bar{e}_1,\ldots,\bar{e}_n)$, where $\bar{e}_i=\mathbb{E}_\sigma[e_i]$ for all agents $i\neq j$ and $\bar{e}_j=\mathbb{E}_\sigma[e_j]+\epsilon$ for agent j. We construct the allocation rule $\hat{\pi}$ by letting

$$\hat{\pi}_i^k(\tilde{e})=\begin{cases}\mathbb{E}_\sigma\left[\pi_i^k(e)\right] & \text{if } \tilde{e}=\bar{e},\\ \mathbb{E}_\sigma\left[\pi_i^k(0,e_{-j})\right] & \text{if } \tilde{e}_j\neq\bar{e}_j \text{ and } \tilde{e}_\ell=\bar{e}_\ell\ \forall\ell\neq j,\\ \pi_i^k(\tilde{e}) & \text{otherwise,}\end{cases}$$

for all $i,k\in I$.

We now show that, in the contest $(y,\hat{\pi})$, there exists an $\epsilon>0$ such that for each agent $i\in I$ it is a best response to play $\bar{e}_i$ when the remaining agents are playing $\bar{e}_{-i}$, which implies that $(y,\hat{\pi})$ implements $\bar{e}$. This claim holds for each agent $i\neq j$ because, $\forall e_i'\neq\bar{e}_i$,

$$\begin{aligned}\Pi_i(\bar{e}\mid(y,\hat{\pi}))&=\sum_{k=1}^n\hat{\pi}_i^k(\bar{e})u(y_k)-c(\bar{e}_i)\\&=\sum_{k=1}^n\mathbb{E}_\sigma\left[\pi_i^k(e)u(y_k)\right]-c\left(\mathbb{E}_\sigma\left[e_i\right]\right)\end{aligned}$$

$$
\begin{aligned}
&\geq \sum_{k=1}^{n} \mathbb{E}_{\sigma}\left[\pi_i^k(e)u(y_k)\right] - \mathbb{E}_{\sigma}\left[c(e_i)\right] \\
&\geq \sum_{k=1}^{n} \mathbb{E}_{\sigma}\left[\pi_i^k(0,e_{-i})u(y_k)\right] \\
&\geq \sum_{k=1}^{n} \mathbb{E}_{\sigma}\left[\pi_i^k(0,e_{-i})u(y_k)\right] - c(e_i') \\
&= \sum_{k=1}^{n} \hat{\pi}_i^k(e_i',\bar{e}_{-i})u(y_k) - c(e_i') \\
&= \Pi_i((e_i',\bar{e}_{-i}) \mid (y,\hat{\pi})),
\end{aligned}
$$

where the first inequality follows the convexity of c and the second inequality follows from the fact that (y,π) implements σ. For agent j, $c\left(\mathbb{E}_{\sigma}[e_j]\right) < \mathbb{E}_{\sigma}[c(e_j)]$ since c is strictly convex and σ_j is not a Dirac measure. Then, there exists some $\epsilon > 0$ such that

$$
\sum_{k=1}^{n} \mathbb{E}_{\sigma}\left[\pi_j^k(e)u(y_k)\right] - c\left(\mathbb{E}_{\sigma}\left[e_j\right]+\epsilon\right) \geq \sum_{k=1}^{n} \mathbb{E}_{\sigma}\left[\pi_j^k(e)u(y_k)\right] - \mathbb{E}_{\sigma}[c(e_j)]
$$

from which, analogously to the argument above, it follows that for all $e_j' \neq \bar{e}_j$ we have $\Pi_j(\bar{e} \mid (y,\hat{\pi})) \geq \Pi_j(e_j',\bar{e}_{-j} \mid (y,\hat{\pi}))$. Thus, $(y,\hat{\pi})$ implements $\bar{e}$, resulting in a strictly higher payoff for the principal. □

Lemma 18. *Suppose that efforts are perfectly observable. For any contest (y,π) that implements a pure-strategy profile $\bar{e}$ and in which one of the following conditions is satisfied:*

(i) $y_2 > 0$, *or*
(ii) $\bar{e}_i \neq \bar{e}_j$ *for some* $i, j \in I$,

there exists a contest $(\tilde{y},\tilde{\pi})$ with $\tilde{y}_2 = 0$ that implements a symmetric pure-strategy profile $(\tilde{e},\ldots,\tilde{e})$ and yields a strictly higher expected payoff to the principal.

Proof. Starting from (y,π) that implements $\bar{e}$, we construct $(\tilde{y},\tilde{\pi})$ as follows. Let the prize profile be $\tilde{y} = (\tilde{y}_1,\ldots,\tilde{y}_n)$ with $\tilde{y}_1 = \sum_{i=1}^n y_i \equiv x$ and $\tilde{y}_2 = \ldots = \tilde{y}_n = 0$. Let e^x be the solution to $\frac{1}{n}u(x) - c(e^x) = 0$, as defined in the body of the text. Let $\tilde{\pi}$ be such that the prize $\tilde{y}_1 = x$ is allocated randomly and uniformly among the agents when $e = (e^x,\ldots,e^x)$. If some agent i unilaterally deviates, then agent i receives the prize $\tilde{y}_n = 0$ for sure, while the prize $\tilde{y}_1$ is allocated randomly among the non-deviating agents. For all other effort profiles, the allocation of the prizes can be chosen arbitrarily. It follows immediately from the definition of e^x that this contest implements $(e^x,\ldots,e^x)$.

We now claim that $\sum_{i=1}^n \bar{e}_i < ne^x$, so that the principal's payoff is strictly higher with $(\tilde{y},\tilde{\pi})$ than with (y,π). Using the fact that (y,π) implements $\bar{e}$, we have, $\forall i \in I$,

$$
\Pi_i(\bar{e} \mid (y,\pi)) = \sum_{k=1}^{n} \hat{\pi}_i^k(\bar{e})u(y_k) - c(\bar{e}_i) \geq 0.
$$

Summing over all agents, we obtain

$$\sum_{i=1}^{n}\sum_{k=1}^{n}\pi_i^k(\bar{e})u(y_k) - \sum_{i=1}^{n} c(\bar{e}_i) = \sum_{k=1}^{n} u(y_k) - \sum_{i=1}^{n} c(\bar{e}_i) \geq 0.$$

Now assume by contradiction that $\sum_{i=1}^{n} \bar{e}_i \geq ne^x$. Then, we have

$$c(e^x) \leq c\left(\frac{1}{n}\sum_{i=1}^{n}\bar{e}_i\right) \leq \frac{1}{n}\sum_{i=1}^{n} c(\bar{e}_i), \tag{26}$$

where the second inequality follows from strict convexity of c, and the inequality is strict whenever the original contest (y, π) satisfied condition (ii) in the lemma. In addition, strict convexity of u together with $u(0) = 0$ implies that

$$\sum_{k=1}^{n} u(y_k) = \sum_{k=1}^{n} u\left(\frac{y_k}{x}\cdot x\right) \leq \sum_{k=1}^{n}\frac{y_k}{x}\cdot u(x) = u(x), \tag{27}$$

and the inequality is strict whenever the original contest (y, π) satisfied condition (i) in the lemma. Taken together, we have

$$\sum_{k=1}^{n} u(y_k) - \sum_{i=1}^{n} c(\bar{e}_i) < u\,(x) - nc(e^x) = 0,$$

because either condition (i) or (ii) in the lemma is satisfied, which is a contradiction. □

Lemma 19. *Suppose that efforts are perfectly observable. For any fixed prize sum $x > 0$, a contest (y, π) maximizes the principal's payoff if and only if the prize profile is $y = (x, 0, ...0)$ and the allocation rule satisfies, for each $i \in I$,*

(i) $\pi_i^1(e^x, e_{-i}^x) = \frac{1}{n}$, *and*
(ii) $\pi_i^1(e_i, e_{-i}^x) \leq \frac{c(e_i)}{u(x)}$, $\forall e_i \neq e^x$.

Proof. It follows from Lemma 18 that we can constrain attention to contests with a prize profile given by $y_1 = x$ and $y_2 = \ldots = y_n = 0$ and which implement a symmetric pure-strategy effort profile (as any contest violating these conditions cannot be optimal). Any such contest must satisfy $u(x) - nc(\bar{e}) \geq 0$, where $\bar{e}$ is the implemented individual level of effort, as otherwise unilateral deviations to zero effort would be profitable. Hence, for a given prize sum $x > 0$, the maximization problem

$$\max_{\bar{e}}\ n\bar{e} - x \quad \text{s.t.} \quad u(x) - nc(\bar{e}) \geq 0$$

describes an upper bound for the principal's payoff. The unique solution to this problem is $\bar{e} = e^x$.

The contest described in the proposition implements $(e^x, ..., e^x)$ and hence achieves the upper bound and is optimal, which proves the if-statement. Any contest not satisfying conditions (i) and (ii) in the proposition does not implement $(e^x, ..., e^x)$ and thus does not achieve the upper bound, which proves the only-if-statement. □

We now prove the proposition. Fix an arbitrary observational structure (S, η) and a total budget $x > 0$ and consider a contest (y, π) as described in the proposition. It clearly implements the effort profile $(e^x, ..., e^x)$. Suppose by contradiction that (y, π) is not optimal, i.e., there exists a contest $(\tilde{y}, \tilde{\pi})$ with the same total prize budget x and which implements some strategy profile σ such that

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

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

$$\hat{\pi}_i^k(e) = \mathbb{E}_{\eta^e}\left[\tilde{\pi}_i^k(s)\right]$$

for all $i, k \in I$ and all $e \in E$. It follows that the induced CSF $\hat{p}$ of the contest $(\tilde{y}, \hat{\pi})$ with perfect observation is identical to the induced CSF $\tilde{p}$ of the contest $(\tilde{y}, \tilde{\pi})$ with the original observational structure (S, η). Since the prize profiles are also identical, it follows that $(\tilde{y}, \hat{\pi})$ implements σ under perfect observation and achieves a payoff for the principal strictly larger than $ne^x - x$. This is a contradiction to Lemma 19. ■

References

Akerlof, R.J., Holden, R.T., 2012. The nature of tournaments. Econ. Theory 51, 289–313.

Barvinek, E., Daler, I., Francu, J., 1991. Convergence of sequences of inverse functions. Arch. Math. 27, 201–204.

Benkert, J.-M., Letina, I., 2020. Designing dynamic research contests. Am. Econ. J. Microecon. 12, 270–289.

Chang, S.-H., Cosman, P.C., Milstein, L.B., 2011. Chernoff-type bounds for the Gaussian error function. IEEE Trans. Commun. 59, 2939–2944.

Che, Y.-K., Gale, I., 2000. Difference-form contests and the robustness of all-pay auctions. Games Econ. Behav. 30, 22–43.

Che, Y.-K., Gale, I., 2003. Optimal design of research contests. Am. Econ. Rev. 93, 646–671.

Clark, D.J., Riis, C., 1996. A multi-winner nested rent-seeking contest. Public Choice 87, 177–184.

Clark, D.J., Riis, C., 1998. Influence and the discretionary allocation of several prizes. Eur. J. Polit. Econ. 14, 605–625.

Cornes, R., Hartley, R., 2005. Asymmetric contests with general technologies. Econ. Theory 26, 923–946.

Dixit, A., 1987. Strategic behavior in contests. Am. Econ. Rev. 77, 891–898.

Drugov, M., Ryvkin, D., 2020a. How noise affects effort in tournaments. J. Econ. Theory 188.

Drugov, M., Ryvkin, D., 2020b. Tournament rewards and heavy tails. J. Econ. Theory 190.

Drugov, M., Ryvkin, D., 2021. Optimal Prizes in Tournaments under Nonseparable Preferences. Mimeo.

Erkal, N., Xiao, J., 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. Econ. Lett. 157, 167–170.

Fang, D., Noe, T., 2022. Less competition, more meritocracy? J. Labor Econ. 40, 669–701.

Fang, D., Noe, T., Strack, P., 2020. Turning up the heat: the discouraging effect of competition in contests. J. Polit. Econ. 128, 1940–1975.

Feng, X., Lu, J., 2018. How to split the pie: optimal rewards in dynamic multi-battle competitions. J. Public Econ. 160, 82–95.

Franke, J., Leininger, W., Wasser, C., 2018. Optimal favoritism in all-pay auctions and lottery contests. Eur. Econ. Rev. 104, 22–37.

Fu, Q., Jiao, Q., Lu, J., 2015. Contests with endogenous entry. Int. J. Game Theory 44, 387–424.

Fu, Q., Wang, X., Wu, Z., 2021a. Multi-prize contests with risk-averse players. Games Econ. Behav. 129, 513–535.

Fu, Q., Wu, Z., Zhu, Y., 2021b. On Equilibrium Uniqueness in Generalized Multi-Prize Nested Lottery Contests. Mimeo.

Fu, Q., Wu, Z., Zhu, Y., 2022. On equilibrium existence in generalized multi-prize nested lottery contests. J. Econ. Theory 200.

Glazer, A., Hassin, R., 1988. Optimal contests. Econ. Inq. 26, 133–143.

Green, J.R., Stokey, N.L., 1983. A comparison of tournaments and contracts. J. Polit. Econ. 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., Skaperdas, S., Vaidya, S., 2013. Contest functions: theoretical foundations and issues in estimation. Int. J. Ind. Organ. 31, 211–222.

Johnson, J.P., Myatt, D.P., 2006. On the simple economics of advertising, marketing, and product design. Am. Econ. Rev. 96, 756–784.
Konrad, K.A., 2009. Strategy and Dynamics in Contests. Oxford University Press.
Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. J. Polit. Econ. 89, 841–864.
Lemus, J., Marshall, G., 2021. Dynamic tournament design: evidence from prediction contests. J. Polit. Econ. 129, 383–420.
Lemus, J., Temnyalov, E., 2021. Diversification and Information in Contests. Mimeo.
Letina, I., Liu, S., Netzer, N., 2018. Delegating Performance Evaluation. University of Zurich, Department of Economics. Working Paper No. 266.
Letina, I., Liu, S., Netzer, N., 2020. Delegating performance evaluation. Theor. Econ. 15, 477–509.
Lewis, T.R., Sappington, D.E., 1994. Supplying information to facilitate price discrimination. Int. Econ. Rev. 35, 309–327.
Liu, B., Lu, J., 2019. The optimal allocation of prizes in contests with costly entry. Int. J. Ind. Organ. 66, 137–161.
Liu, B., Lu, J., 2021. Optimal Orchestration of Rewards and Punishments in Rank-Order Contests. Mimeo.
Liu, B., Lu, J., Wang, R., Zhang, J., 2018. Optimal prize allocation in contests: the role of negative prizes. J. Econ. Theory 175, 291–317.
McFadden, D.L., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.
Meyer, M.A., 1991. Learning from coarse information: biased contests and career profiles. Rev. Econ. Stud. 58, 15–41.
Moldovanu, B., Sela, A., 2001. The optimal allocation of prizes in contests. Am. Econ. Rev. 91, 542–558.
Moldovanu, B., Sela, A., Shi, X., 2012. Carrots and sticks: prizes and punishments in contests. Econ. Inq. 50, 453–462.
Morgan, J., Orzen, H., Sefton, M., 2012. Endogenous entry in contests. Econ. Theory 51, 435–463.
Morgan, J., Tumlinson, J., Várdy, F., 2022. The limits of meritocracy. J. Econ. Theory 201.
Nalebuff, B.J., Stiglitz, J.E., 1983. Prizes and incentives: towards a general theory of compensation and competition. Bell J. Econ. 14, 21–43.
Olszewski, W., Siegel, R., 2016. Large contests. Econometrica 84, 835–854.
Olszewski, W., Siegel, R., 2019. Pareto Improvements in the Contest for College Admissions. Mimeo.
Olszewski, W., Siegel, R., 2020. Performance-maximizing large contests. Theor. Econ. 15, 57–88.
Pinelis, I., 2002. L'Hospital type rules for monotonicity, with applications. J. Inequal. Pure Appl. Math. 3, 5 (electronic).
Sahm, M., 2022. Optimal accuracy of unbiased Tullock contests with two heterogeneous players. Games 13, 24.
Schweinzer, P., Segev, E., 2012. The optimal prize structure of symmetric Tullock contests. Public Choice 153, 69–82.
Sela, A., 2020. Optimal allocations of prizes and punishments in Tullock contests. Int. J. Game Theory 49, 749–771.
Shi, X., 2012. Optimal auctions with information acquisition. Games Econ. Behav. 74, 666–686.
Siegel, R., 2009. All-pay contests. Econometrica 77, 71–92.
Sisak, D., 2009. Multiple-prize contests – the optimal allocation of prizes. J. Econ. Surv. 23, 82–114.
Skaperdas, S., 1996. Contest success functions. Econ. Theory 7, 283–290.
Taylor, C.R., 1995. Digging for golden carrots: an analysis of research tournaments. Am. Econ. Rev. 85, 872–890.
Wang, Z., 2010. The optimal accuracy level in asymmetric contests. B. E. J. Theor. Econ. 10, 1–18.
Xiao, J., 2016. Asymmetric all-pay contests with heterogeneous prizes. J. Econ. Theory 163, 178–221.
Zhang, M., 2021. Optimal Contests with Incomplete Information and Convex Effort Costs. Mimeo.