Eliciting Performance: Deterministic versus Proportional Prizes

Pradeep Dubey* and Siddhartha Sahi†

12 December 2014

1 Abstract

Two standard schemes for awarding a prize have been examined in the literature. The prize is awarded

(π_D) deterministically: to the contestant with the highest output;

(π_P) probabilistically: to all contestants, with probabilities proportional to their outputs.

Our main result is that if there is sufficient diversity in contestants' skills, and not too much noise on output, then π_P will elicit more output on average than π_D. Indeed if contestants know each others' skills (the complete information case) then the expected output at *any* Nash equilibrium selection under π_P exceeds that at *any* individually rational selection under π_D. If there is incomplete information, the inequality continues to hold when we restrict to Nash selections for both schemes.

JEL Classification: C70, C72, C79, D44, D63, D82.

Key Words: principal, contestant, proportional and deterministic prizes, complete and incomplete information, Nash equilibrium, individually rational strategies

2 Introduction

We take the point of view, not uncommon to game theory, that the purpose of a prize is not so much to reward performance as to inspire it.

*Center for Game Theory, Department of Economics, Stony Brook University and Cowles Foundation for Research in Economics, Yale University

†Department of Mathematics, Rutgers University, New Brunswick, New Jersey

Our focus is on two different traditions for awarding a prize, both much discussed in the literature (see 2.1). The deterministic prize scheme π_D awards the entire prize, or "pot of gold[1]", to the best performer, and has been extensively examined in relation to tournaments. Sharply juxtaposed to π_D is the proportional prize scheme π_P, which awards the pot to all the contestants with probabilities that are proportional to their outputs. It has often been the center of attention in the context of studies on "lobbying", but not much besides, even though a case can be made for it on *a priori* natural grounds. Indeed let the gold be simply put "on market" to be exchanged against the output that the contestants have produced. How the gold gets allocated is then left to "market forces". Suppose that contestants $1, ..., n$ have put up $x_1, ..., x_n$ units of output; and that y units of gold is present on the other side of the market. The only price p, of the output in terms of gold, which will "clear" the market is[2] $p = y/(x_1 + \cdots + x_n)$, and this is tantamount to handing out the gold y to the contestants in proportion to the quantities they have put up[3].

Note that π_P also makes sense when the pot is indivisible. In this event, what is being marketed is the *probability* of winning the whole pot y; or, equivalently, lottery tickets for the prize. Our analysis is in fact couched in terms of the indivisible prize rather than the divisible pot of gold (the two are isomorphic). And, for this reason, when the entire pot goes to the highest output, we shall refer to it as the "deterministic" prize, though it is deterministic only in the outputs, and not necessarily in the effort undertaken by the contestants, since output may be a random function of effort.

In this paper, we shall delineate certain circumstances under which the proportional prize π_P elicits more performance than the deteministic prize π_D, which in turn is often better than multiple *a priori* fixed prizes[4] (see [22], and also subsection 8.3). These circumstances are roughly as follows. Suppose that the characteristics[5] of the contestants are relatively bounded (see Assumption 1 below); and that there is complete information amongst them regarding each others' characteristics. Then,

[1] This terminology is from [22].

[2] The total demand for gold is $px_1 + \cdots + px_n$ which must equal the supply y.

[3] Consider, as in [12], the set of *all* possible probabilistic prizes that are based on outputs. (This set includes both π_P and π_D as particular elements.) The proportional prize is the only one in the set which is *non-manipulable* in the following sense: if an agent pretends to be several agents by splitting his output to be sent out in different names, this can be of no benefit to him; nor can several agents benefit by merging their outputs and pretending to be one agent (see [16]).

[4] This is not to say that there will not be other circumstances where the same result holds, or yet others where it is reversed. Our analysis is far from being comprehensive, and it is our hope that this paper will stimulate further inquiry into the relative merits of π_P and π_D.

[5] i.e., productive skill, cost of effort, valuation of the prize

fixing the other parameters of the model, the (expected total) output at any Nash Equilibrium of the strategic game under π_P is of the order of $\min\{N, v\}$, where N is the number of contestants and v is their minimum valuation of the prize (see Theorem 5). Suppose furthermore that there is an elite coterie of the highly skilled, whose performance — via maximal effort — significantly outstrips that of the others (in the sense made precise in section 7, and subsection 8.5). Then, at any individually rational choice of strategies in the game under π_D, the output is of the order of the size K of the coterie (see Theorem 8). We conclude that whenever K is a small fraction of N and v is large enough, *the proportional prize π_P elicits more output from the contestants than the deterministic prize.π_D*. The hypothesis of a small elite, with *no other* constaints on the skill distribution of the rest of the contestants, thus lies at the heart of our analysis. We believe that it is pertinent to many real world situations, and well worth the analysis.

In particular, small K can arise as follows. Think of $\pi = \pi_P$ or π_D as a fixed *tradition* for awarding the prize, and *not* as a scheme designed by a "principal[6]", *i.e.*, π is to be applied across generations to come, or disjoint cohorts of contestants in different places. In this context, one might model contestants' skills as if they are drawn at random from a large domain, in a manner that does not correlate them to be similar (e.g., they are drawn i.i.d.). We show that, if there is not too much noise on output, then the *average* size of K will be small (see Theorem 11). It then follows, from our above result for fixed skills, that the superiority of π_P over π_D is maintained in terms of the average output on this domain (Theorem 12).

The intuition for our result is simple and best brought out with two contestants who have complete information about each other's characteristics. (We show, in section 9, that our results are not marred when there is incomplete information, i.e., each contestant is informed only of his own characteristics and has a probability distribution over those of his rivals.) Suppose the deterministic prize π_D is in use and that the two contestants' skills are sufficiently disparate so that the weak cannot produce more than the strong, with any significant probability, even if he works hard and the other slackens. Since effort is costly, the upshot is an equilibrium at which both contestants undertake low effort, so that total output is also low. In contrast, the proportional prize π_P generates better incentives to work. By increasing effort and producing more output, the weak contestant is able to achieve a decent increment in his probability of winning the prize, even when his output always lags behind his rival's. Therefore he is inspired to work and creates the competition which also spurs

[6]Such a principal may even be in a position to modify π away from π_P or π_D based on his knowledge of the characteristics of his "agents", a possibility that is examined in [12] in connection with "optimal" π on a given domain.

his rival to work, culminating in an equilibrium where effort and output are high. That an egalitarian scheme, which distributes rewards commensurate with output produced, will often generate better incentives to work than an elitist scheme in which the rewards are reserved for the top few — this, in our view, is a theme of wide-ranging application in the presence of disparate contestants, and it runs like a leitmotif in the design of mechanisms in different contexts (see, *e.g.,* [13],[11],[10]).

On the other hand, when skills are similar (think of athletic stars competing in the Olympics), π_D may elicit more effort than π_P. For if both work, they come out with nearly equal probabilities of winning the prize under either scheme. But if anyone slackens, his probability can drop sharply under π_D, and less so under π_P, especially when going down the effort ladder causes significant drops in output[7]. Thus there is more to lose by slackening when π_D is in use. (See especially the last example of section 9 in this regard, computed in the case of large noise.)

Now if contestants' skills are picked at random from a sufficiently large domain, and the noise on output is not so large as to overwhelm skills and make them count for little, then the probability that contestants are similar will tend to be low. Therefore the average output will go up when π_P replaces π_D. There is *a trade-off between the accuracy of reporting performance and the quantum of performance that is elicited.* Accuracy is best achieved by bestowing the indivisible prize according to π_D. In contrast, π_P gives the prize to the best performer only with some — albeit the highest — probability; and it rewards others with the remaining probability, compromising the accuracy of the report in the process but simultaneously boosting contestants' effort and output[8].

In section 9, we give examples to show that our theme remains intact when there is incomplete information: the NE-selection under π_P elicits more output compared to the NE-selection under π_D, as long as the noise on output is sufficiently small compared to the diversity of contestants' skills. (We write "*the* NE" because, in the more structured binary games that we examine in Section 8, NE's do turn out to be unique.)

What is clearly essential for our analysis is that contestants' performance be susceptible to quantification in terms of some tangible output produced or, more generally, a "score*"*. This often obtains in practice. For instance, a manager can

[7] which is not so unnatural an assumption when there are a few discrete levels of effort (rather than continuous effort)

[8] Going a bit beyond our model, if one were to grant that skill is not rigidly fixed for individuals, but is a dynamic thing which gets enhanced by their effort, then the superiority of π_P over π_D may be rephrased in more dramatic terms: *too much accuracy in reporting performance, i.e., in certifying the underlying skill of the agents, can inhibit the very production of that skill.*

consider total revenue earned as the criterion to award a badge of honor, or promotion to a higher echelon, to the best salesman of the year. In a race, the time taken for completion comes naturally to mind. Sometimes scores are of a more subtle structure: in a gymnastics contest each member of a jury gives subjective scores to different aspects of performance which are then aggregated to come up with final scores. (The reader can no doubt think of many other examples.) One upshot of assigning numerical scores, and perhaps the reason why they are so prevalent, is that they enable us to evaluate not only who beat whom, but by how much. Was the race keenly contested or one-sided? What was the margin of victory? These are questions that are often not without meaning, and amenable to plausible answers, which is reflected in the way scores get defined in practice.

The proportional prize π_P is our proxy for awarding the prize in a manner that is less drastic than the deterministic π_D, and more commensurate with performance. Any scheme close to π_P will inherit its properties (see subsection 8.2). So, for our purposes, the precision with which probabilities of winning the prize are defined does not really matter, so long as they do not stray too far from proportionality; and, in the same vein, minor differences in the measurement of the scores do not disturb our conclusions (see again subsection 8.2). Needless to say, if performances are incapable of being sensibly quantified by scores, and can only be ranked, then the proportional prize has no meaning and only ordinal prizes (*i.e.*, π_D and its variants with multiple deterministic prizes) make sense. (For an excellent treatment of the ordinal case, see [22].) In our model here, as in much of the literature, the purpose of the prize is to maximize the total score (output) of all the contestants, so *a fortiori* the individual scores that make up the total are taken to be observable It is not so much a matter of observability, but that the cost of observation is small enough to be ignored. This assumption underlies our analysis.

Let us also note that this paper is self-contained but, to round off the perspective, we shall often allude to its expanded version [12], which contains several variants and extensions of the results described here.

Finally, a word about the numbering system used in this paper: all assumptions, remarks, theorems, lemmas etc. are arranged in a *single* grand sequence: Assumption 1, Remark 2, Remark 3, Assumption 4, Theorem 5 etc. (Thus Theorem 5 is not the fifth theorem, in fact it is the first theorem, but it has fifth position in the grand sequence. This indeed makes it easier to locate the theorem, and should cause no confusion.)

2.1 Related Literature

There is a considerable literature on lobbying, where contestants put up bids of money and are awarded the prize either via π_P or π_D, called often "lottery" or "all-pay auctions", respectively.(see, *e.g.*, [28],[19],[14],[26],[27],[3],[4], [7],[8],[24],[15] and the references therein). In much of this literature contestants are assumed to have complete information about each other, and in all of it there is no issue of "moral hazard", *i.e.*, the bids submitted by the contestants are perfectly observable.

The literature on tournaments is vast and does often emphasize moral hazard, *i.e.*, the setting in which observable outputs depend stochastically on unobservable effort. However proportional prizes do not seem to have received attention there. For tournaments with a single deterministic prize, see [21],[18],[23],[25]; and for the general case of multiple deterministic prizes, see ([17],[5],[1],[9],[20],[6],[2],[22]).

In both strands of literature the focus is on analyzing Nash Equilibria (NE), which are often unique and susceptible of being described by explicit formulae, given the special structural assumptions of the models.

What is new in our approach is that we compare π_P and π_D in the presence of moral hazard. Our setting is sufficiently general so as to neither preclude multiple NE, nor guarantee pure-strategy NE. No assumptions are made on disutility or productivity other than the fact that they are monotonic in effort in the appropriate sense; in particular they are *not* required to be concave or convex. Nevertheless we are able to show that the worst NE selection under the proportional prize π_P elicits more output than the best NE under the deterministic prize π_D. In fact we show more, since our comparison is based on "*Weak Nash Strategies*" for π_P (see subsection **??**) and individually rational (IR) strategies for π_D, which are looser notions than NE (indeed IR is so mild a reqirement that *any* solution concept would be expected to satisfy it). To the extent that this constrains contestants' behavior less, our comparison is that much stronger (more credible?). Of course, the price we pay for our generality is that we stop at this comparison, and are unable to discern any finer structure in contestants' behavior, which would come to the fore were one to confine attention to NE, especially in simple scenarios where they are unique (as happens in some of the structured examples we study here in section 9, or in [12]).

3 The Model

There is a finite set N of contestants. Each **contestant** $n \in N$ has access to a fixed finite subset $E^n \subset [0, 1]$ of effort levels. We assume $0 \in E^n$ and $1 \in E^n$; these represent no effort and maximal effort respectively.

Contestant n may choose any effort $x \in E^n$. In doing so, he incurs disutility $\delta(x) \geq 0$ and produces stochastic output given by a non-negative random variable $\tau(x)$ with finite mean $\mu(x)$. (We allow for the possibility that the range of $\tau(x)$ is discrete, even finite.) Effort 0 incurs disutility $\delta(0) = 0$ and produces output $\tau(0) = 0$ with certainty: it is just a proxy for "not participating" in the game.

Contestants are driven to work by the lure of an indivisible prize. If a contestant places valuation $v > 0$ on the prize, and is awarded it with probability p, this yields him expected utility pv. (See, however, the subsection 8.2, where it is shown that the tenor of our results remains unchanged for a wider class of utilities.)

The triple (δ, τ, v) characterizes a contestant. We make *throughout* the following monotonicity and boundedness assumptions on the space[9] X of possible **characteristics** (δ, τ, v):

Assumption 1 *Both δ, τ are weakly monotonic in x and there exist universal positive constants c, C, d, D such that, for all $x \in E^n \setminus \{0\}$, and all $n \in N$*

$$cx < \delta(x) < Cx \tag{1}$$

and

$$dx < \tau(x) < Dx \tag{2}$$

Remark 2 *On account of* weak *monotonicity, there is no loss of generality in supposing that all contestants have the same set E of effort levels. The case of an arbitrary allocation of subsets of E across contestants can be embedded in this framework, with* 0 *and* 1 *representing non-paricipation and maximal effort for each contestant. So from now on we take $E^n = E$ for all $n \in N$.*

Suppose the population of contestants has characteristics $(\delta^n, \tau^n, v^n)_{n \in N}$. The prize is awarded on the basis of the realizations $t = (t^n)_{n \in N}$ of the random outputs $(\tau^n(e^n))_{n \in N}$.

The **deterministic prize** π_D is shared equally among the winners

$$W(t) := \{k \in N : t^k \geq t^n \text{ for all } n \in N\}.$$

[9] This space X is defined after fixing the domain and range of τ. It will shortly be taken to be measurable. A key scenario we have in mind is that X is a *finite* set, as spelled out in Remark 3 below. Or else one can confine attention to random variables τ which are characterized by finitely many parameters, so that (δ, τ, v) is a finite-dimensional vector; and then the Euclidean space generates the Borel sets. In this case the space X consists of all (δ, τ, v) that satisfy (1) and (2) of Axiom1 below. More generally, without such restrictions, the Levy-Prokhorov metric on the random variables τ is understood to define the Borel sets.

In other words

$$\pi_D^n(t) = \begin{cases} 1/|W(t)| & \text{if } n \in W(t) \text{ and } t \neq 0 \\ 0 & \text{else} \end{cases} .$$

Note that π_D is deterministic only in the outputs, not necessarily in the effort levels. (Also note that we have set $\pi^n(t) = 0$ for all $n \in N$ if $t = 0$, otherwise contestants would be rewarded for not participating in the game.)

The **proportional prize** π_P is awarded to each contestant in proportion to his output, *i.e.*,

$$\pi_P^n(t) = \begin{cases} t^n / \left(\sum_{k\in N} t^k\right) & \text{if } \ t \neq 0 \\ 0 & \text{else} \end{cases} .$$

4 The Strategic Game of Complete Information

We suppose that, in addition to knowing $\pi = \pi_D$ or π_P, the contestants also know each others' characteristics $(\delta^n, \tau^n, v^n)_{n\in N}$. This seems to be a tenable hypothesis if contestants compete in close proximity with one another. (In section 9, we consider the case when a contestant knows his own characteristics but is unsure about those of his rivals.)

Given $(\delta^n, \tau^n, v^n)_{n\in N}$, a strategic game is induced among the contestants by the choice of an allocation scheme π. The set of pure strategies of each contestant $n \in N$ is E. Any N-tuple of pure strategies $e = (e^n)_{n\in N}$ gives rise to a random vector

$$\tilde{t} = \tilde{t}\,(e) = (\tau^n(e^n))_{n\in N}$$

of outputs. The expected value p^k of $\pi^k(\tilde{t})$ represents the probability of k winning the prize and we define k's payoff to be

$$F^k(e) = p^k v^k - \delta^k(e^k).$$

Denote by Γ the mixed extension of this game; and by Σ^k the set of (mixed) strategies of k in Γ, *i.e.* Σ^k is just the set of probability distributions on E. (Without confusion, $F^k(\sigma)$ will continue to denote k's payoff, when the mixed strategy N-tuple $\sigma \equiv (\sigma^n)_{n\in N} \in \prod_{n\in N} \Sigma^n \equiv \Sigma$ is played in Γ.) For any $\sigma\in\Sigma$, denote

$$\sigma^{-n} \equiv \left(\sigma^k\right)_{k\in N\setminus\{n\}} \in \Sigma^{-n} \equiv \prod\nolimits_{k\in N\setminus\{n\}} \Sigma^k.$$

Recall that the choice $\sigma\in\Sigma$ is called **individually rational** (IR) in Γ if

$$F^n(\sigma) \geq \max_{u\in\Sigma^n} \min_{w\in\Sigma^{-n}} F^n(u, w)$$

for all $n \in N$; and is called a **Nash Equilibrium** (NE) of Γ if

$$F^n(\sigma) = \max_{u \in \Sigma^n} F^n(u, \sigma^{-n})$$

for all $n \in N$. Denote by $IR(\Gamma)$, $NE(\Gamma)$ the set of all strategies $\sigma \in \Sigma$ that are IR, NE in the game Γ, and note $NE(\Gamma) \subset IR(\Gamma)$.

5 Spaces of Games

Suppose characteristics $\chi \equiv (\delta^n, \tau^n, v^n)_{n \in N}$ are picked from $X \times \cdots \times X \equiv \mathbf{X}$ according to some probability distribution ξ on $\mathbf{X}$. (Recall that the underlying set X satisfies Assumption 1; and that X is a Borel space as explained in footnote 4, so that ξ is a measure on the Borel sets of $\mathbf{X}$,using the product topology from X.) Fix an allocation scheme $\pi = \pi_D$ or π_P. Then any $\chi \in \mathbf{X}$ induces a mixed-strategy game among the contestants (as discussed in section 4), which we shall denote $\Gamma_\pi(\chi)$. We wish to extend our solution concepts to the space of games specified by ξ. Our focus will be on what happens for *almost all* χ according to ξ, denoted $a.a.\chi(\xi)$, *i.e.*, for all χ except perhaps for those in a set of ξ-measure zero.

Let $f : \mathbf{X} \to \Sigma$ be a measurable function. Denoting $f(\chi) \equiv (\sigma^n)_{n \in N} \in \Sigma$, the total output at χ is

$$T(f, \chi) \equiv \sum_{n \in N} \sum_{x \in E} \sigma^n(x) \mu^n(x). \tag{3}$$

and integrating over $\mathbf{X}$ according to ξ, the **expected total output** is

$$T(f) \equiv \int_{\mathbf{X}} T(f, \chi)\, d\xi(\chi) \tag{4}$$

Given a prize scheme π we will say that $f : \mathbf{X} \to \Sigma$ is an ξ-**NE selection under** π if f is measurable and if $f(\chi)$ is a Nash Equilibrium of the game $\Gamma_\pi(\chi)$ for $a.a.\chi(\xi)$. The notion of a ξ-**IR selection under** π is defined similarly, substituting "IR" for "NE".

Remark 3 *We eschew the discussion as to when a measurable NE or IR selection exists, as it would be a technical digression from the main thrust of this paper. If they do not exist, our theorems are vacuously true. But in the key scenario where the space* $\mathbf{X}$ *of characteristics is* finite, *measurable selections are not an issue. This scenario is, in particular, realized when the relevant real intervals* $[c, C]$, $[d, D]$, $[v_{\min}, v_{\max}]$ *and* $[0, 1]$, *i.e., the range of* δ, τ, v *and of the probabilities occuring in* τ, *are all restricted* *to finite grids. (Alternatively, the finitely many parameters that characterize τ — see footnote 9 — could be taken to vary over a finite grid.) For some of our results, as will be self-evident, it is further needed to suppose that the mesh of the grids are sufficiently small — see, e.g., Remark 13. We invite the reader to focus on this key scenario.*

6 Proportional Prize: Expected Total Output from Nash Equilibria

It is clear *a priori* that, for any $\chi \in \mathbf{X}$ and any scheme π, the total expected output in $\Gamma_\pi(\chi)$, at any $\sigma \in \Sigma$, cannot exceed $|N|D$ since no contestant produces more than D when he chooses maximal effort 1 (see Assumption 1). Also[10], supposing $v^n = v$ for all $n \in N$, the total expected disutility incurred by the contestants at any individually rational strategy selection cannot exceed v, otherwise some contestant is incurring negative utility and would be better off not participating in the game. But then expected total output (see, again, Assumption 1) is at most Dv/c. Thus, the most this total can be is "of the order of" $\min(v, |N|)$, since D and c are constants of our model.

This is the flavor of our estimate in Theorem 5 below, showing that the proportional prize elicits a "decent quantum" of output from the contestants. However the theorem requires an additional assumption, which we now describe.

For $\chi = (\delta^n, \tau^n, v^n)_{n \in N}$ we denote

$$\underline{v}(\chi) = \min\{v^n : n \in N\}$$

and define $\underline{v}$ to be the essential infimum of $\underline{v}(\chi)$ with respect to ξ.

Assumption 4 *(Minimum valuation)* $\underline{v} > DC/d$.

This basically says that, for any two individuals picked from the population, if both work at maximal effort and are awarded the prize *proportionately,* then neither will have incentive to unilaterally quit the game — each values the prize sufficiently highly to want to stay in. Indeed, by Assumption 1 the most disadvantaged such individual produces d, incurs disutility C, and values the prize at $\underline{v}$ (while his rival

[10] Given $\chi = (\delta^n, \tau^n, v^n)_{n \in N}$, and a vector $\alpha \equiv (\alpha^n)_{n \in N} >> 0$ of positive scalars, let $\chi(\alpha) \equiv (\alpha^n \delta^n, \tau^n, \alpha^n v^n)$. Then the games $\Gamma_\pi(\chi)$ and $\Gamma_\pi(\chi(\alpha))$ are "strategically equivalent" and all our solution concepts remain the same for them. So w.l.o.g., scaling utilities appropriately, one could imagine $v^n = v$ for all $n \in N$.

produces D).Thus his reward is $\underline{v}d/(d+D)$ which must exceed C. Our Assumption 4 is somewhat milder.

The following result shows that Nash Equilibria (NE) elicit a decent quantum of output under the proportional prize.

Theorem 5 *Suppose Assumptions 1 and 4 hold; and denote* $e_{\min} = \min\{x \in E : x \neq 0\}$ *and*

$$a = |N| d e_{\min}, \quad b = (d\underline{v}/C) - D$$

Let $\chi \in \mathbf{X}$ *be arbitrary and let* Z *be the expected total output (see equation 4) at any NE at* χ *under* π_P. *Then*

$$Z \geq \min\{a, b\}.$$

The following Corollary is immediate.

Corollary 6 *Let* f *be a* ξ*-NE selection under* π_P. *Then* $T(f) \geq \min\{a, b\}$.

6.1 Some extensions of Theorem 5

The presence of "$e_{\min}$" is a dampener on our lower bound, but unavoidable given our extremely weak assumptions. Indeed there is nothing to preclude the scenario that every contestant incurs sharply rising disutility of effort as he advances above $e_{\min}$,while his output hardly goes up; and then the best one can hope for is to inspire everyone to work at $e_{\min}$. Were we to strengthen our assumption on productivity, requiring output to go up in significant chunks as we go up the effort ladder from $e_{\min}$ to 1, sharper estimates could be reached by the methods of this paper. (We leave this to the reader). Incidentally notice that, in the special case of binary effort levels, *i.e.*, $E = \{0, 1\}$, we automatically have $e_{min} = 1$ in Theorem 1 above, producing a sharp bound without further ado.

With this strengthened assumption, it can further be shown (see [12]) that under the proportional prize, there are increasing thresholds such that, as the valuation of the prize exceeds these thresholds, maximal effort successively becomes NE, unique NE, and "strictly dominant strategy upto error ϵ" (*i.e.*, maximal effort is the best reply of every contestant provided his rivals' aggregate output is at least ϵ — the threshold obviously needing to be raised as ϵ is lowered.) In this sense, the proportional scheme permits more certainty (predictability) about contestants' behavior at the cost of enhancing the prize This is *not* a feature of the deterministic prize.

Finally, we note that Theorem 1 remains valid — by the same proof — if we replace NE by WNS ("Weak Nash Strategies"). WNS are defined just like NE, but with unilateral deviations of a contestant restricted to shifting probabilities, albeit in whatever manner he desires, from his current strategy onto maximal effort. (Thus, in particular, the choice of maximal effort level 1 by each contestant constitutes a WNS.) Since NE are clearly a subset of WNS, this generalizes Theorem 1. (For details, see again [12].)

7 Deterministic Prize: Expected Output from Individually Rational Strategies

Theorem 8 below provides the crucial insight as to why the deterministic prize π_D elicits limited output. Indeed it shows that only the most productive contestant, along with those who stand a chance of beating him, set the bound on the output at any individually rational strategy-tuple.

Fix $\chi = (\delta^n, \tau^n, v^n)_{n \in N}$. Denote by h a contestant (**the "hero"**) who has maximal mean output under effort level 1, *i.e.*, for all $n \in N$, we have $\mu^h(1) \geq \mu^n(1)$ (where, recall, $\mu^n(x)$ is the mean of $\tau^n(x)$). Define $K(\chi)$ to be the **set of "elite contestants"** whose outputs at effort 1 have a positive probability of exceeding that of h, *i.e.*,

$$K(\chi) = \{n \in N : Pr[\tau^n(1) \geq \tau^h(1)] > 0\}.$$

We can show that the output under deterministic prize is commensurate with $|K(\chi)|$. First we need

Assumption 7

1. ***(Bounded relative valuations)*** *There exists a universal constant B such that for a.a.$\chi(\xi)$, if $\chi = (\delta^n, \tau^n, v^n)_{n \in N}$, then $v^n/v^k < B$ for all $n, k \in N$.*

2. ***(Stochastic dominance)*** *If $x > y$ in E then $\tau^n(x) \succeq \tau^n(y)$, where "$\succeq$" denotes first order stochastic dominance*[11].

Theorem 8 *Suppose Assumptions 1 and 7 hold. Let f be a ξ- IR-selection under π_D; then for a.a. $\chi(\xi)$*

$$T(f, \chi) \leq 2|K(\chi)|B^2CD/c.$$

[11] Recall that $X \succeq Y$ if $\Pr\{X \geq z\} \geq \Pr\{Y \geq z\}$ for all z

7.1 Average Size of the Elite Set $K(\chi)$

Let I_n denote the support of $\tau^n(1)$. *Any* population distribution of I_n, with a thin tail at the top, will have a small elite set. Indeed, as an extremal measure, just consider moving the highest I_n sufficiently to the right. This will have the upshot that the elite set has size one, consisting of the "superhero" alone. By way of a more robust example, suppose: (a) $I_n \subset (1.9, 2)$ for no more than 5% of the population (the highly skilled); (b) $I_n \subset (1, 1.7)$ for the remaining 95%; (c) every I_n is contained in an interval of size 0.1. (Here (c) ensures that the stochastic variation of 0.1 in the performance of any contestant is small compared to the domain $(1, 2)$ which describes the spread in performance across the entire population, *i.e.*, idiosyncratic stochasticity does not seriously dampen overall diversity.) Consider any noise on $\tau^n(1)$ whose size is 100% of the initial stochasticity of $\tau^n(1)$. Such a noise will cause each I_n to expand, but to at most thrice its initial size. Even with the onset of this noise, it is clear that the elite will not exceed 5% of the population. Of course with *very* large noise, every pair of I_n will overlap and the elite set will be all of N, rendering our analysis irrelevant.

Another natural scenario is that contestants' characteristics are not correlated to be similar but are sufficiently *"diverse"* (*e.g.*, drawn i.i.d. from a large set[12]). We shall, of course, require this diversity only on their productivities $(\tau^n(1))_{n\in N}$ under maximal effort. This is embodied in Assumption 10 below. First, a definition:

Definition 9 ***(Normalized Density)*** *Let Z be a random variable taking values in the n-cube $C_{|N|} = [d, D]^{|N|}$. Let λ denote the standard Lebesgue measure on $C_{|N|}$ scaled by $(D - d)^{-|N|}$. (so that $\lambda(C_{|N|}) = 1$). We say that Z has normalized density function ρ if ρ is Borel-measurable, nonnegative and*

$$\Pr(Z \in A) = \int_A \rho(x) d\lambda(x)$$

for all Borel sets $A \subset C_{|N|}$;and we define the upper bound of ρ to be the essential supremum of ρ on $C_{|N|}$.

We are ready to state

Assumption 10 ***(Diversity of Skills)***

[12] The i.i.d. assumption, though not always "realistic", is mathematically convenient and has come to constitute a benchmark.

1. *There exists $\epsilon > 0$ such that, for a.a. $\chi(\xi)$, if $\chi = (\delta^n, \tau^n, v^n)_{n \in N}$, then*

$$\text{support } \tau^n(1) \subset [\mu^n(1) - \epsilon, \mu^n(1) + \epsilon] \text{ for all } n \in N.$$

2. *As we vary χ on $\mathbf{X}$ according to ξ, the marginal distribution of the random variable*[13] *$(\mu^n(1))_{n \in N}$ has a normalized density function with finite upper bound β.*

Condition 2 of this assumption rules out the possibility that $(\mu^n(1))_{n \in N}$ is concentrated on the "diagonal" $\left\{(z, ..., z) \in C_{|N|} : d \leqq z \leqq D\right\}$ of the cube $C_{|N|}$. As the random variables $\mu^1(1), ..., \mu^N(1)$ go from being iid, with uniform density on $[d, D]$, to being concentrated on smaller and smaller neighbourhoods of the diagonal, β rises from 1 to ∞.In this scenario β is a measure of how likely it is that the contestants are similar. We should expect a threshold β^*such that π_P outperforms π_D if $\beta < \beta^*$,and π_D outperforms π_P if $\beta > \beta^*$. This is not to say that high β is *necessarily* bad for π_P. Indeed if β were high in regions of $C_{|N|}$ where contestants are *disparate* (*e.g.*,towards the northwest or southeast corners of the square, when $|N| = 2$), this would only accentuate the superiority of π_P over π_D. We do not follow this general line of inquiry here , wherein β would be allowed to become unbounded in selective regions of $C_{|N|}$, and remain bounded only where contestants are similar. Instead we restrict attention to the scenario where β is universally bounded on $C_{|N|}$, thereby only preventing contestants from being similar (or dissimilar!) with high probability.

Consider first (by way of motivation) the iid case. We can think of ϵ as the random noise on output, and then the "diversity" of contestants' productive skills is reflected for us in how *small* the term $\beta\epsilon = \epsilon/(D - d)^{|N|}$ is. (Diversity in skills is dampened by the noise ϵ. Indeed suppose noise ϵ is symmetric across the two contestants and let ϵ grow, keeping skills fixed. The two contestants will become increasingly similar since their output will depend essentially on the identical noise term and their skills will count for little when ϵ is sufficiently large).

Even in the non-i.i.d setting, the term $\beta\epsilon$ serves as a measure of diversity; and Lemma 1 below shows that the average size of the elite set, is no more than $1 + \beta|N|\epsilon$.

Lemma 11 *Suppose the distribution ξ satisfies Assumption 10. Then the expected size, under ξ, of the elite set $K(\chi)$ is at most $1 + \beta|N|\epsilon$.*

We are ready to state the main conclusion of this section.

[13] Recall that $(\mu^n(1))_{n \in N} \in C_{|N|}$ by (2).

Theorem 12 *Suppose Assumptions 1,7 and 10 hold. If f is any ξ-IR-selection on* $\mathbf{X}$ *under π_D then*

$$T(f) \leq \frac{2B^2CD}{c}(1+\beta|N|\epsilon).$$

Proof. Immediate from Theorems 8 and 11. ■

Remark 13 *With a sufficiently fine finite grid on $[d,D]$ the above estimate holds with any degree of accuracy desired, and our analysis goes through*

8 Deterministic versus Proportional Prizes

Theorems 5 and 8 enable an immediate comparison between the (expected total) outputs elicited by NE, IR strategy selections under π_P, π_D respectively. For any *fixed* $\chi \in \mathbf{X}$, the two theorems imply that π_P is better than π_D, without further ado, provided only that:

(a) highly-skilled "elite" contestants exist, whose best performance outstrips that of the rest by a significant margin;

(b) such elite contestants constitute a small fraction of the population

We believe that conditions (a) and (b) are pertinent in many real world situations.

Now we turn to *variable* χ. Suppose $\chi \in \mathbf{X}$ is chosen at random, and we are interested in the average output on $\mathbf{X}$. Fix, for example, all the parameters $c, C, d, D, b, B, \underline{v}$ of the model and suppose that Assumptions 1,4,7,10 hold..Then, for large enough N and v, there exists a threshold $\bar{\epsilon}(N)$ such that, if $\epsilon < \bar{\epsilon}(N)$, we have

$$T(f) > T(g)$$

for any ξ-NE-selection f under π_P, and any ξ-IR-selection g under π_D.This is so because the lower bound on output given by Theorem 5 (and its Corollary) is independent of the noise ϵ, and rises with N, v ; while the upper bound given by Theorem 12 goes to $2B^2CD/c$ as $\beta|N|\epsilon$ goes to 0.

To get a better feel, it might help to consider a numerical example. Let

$$B = C = c = d = 1, D = 2, |N| = 7, \underline{v} = 30, \epsilon = 0.05.$$

Further let the set of effort levels be $E = \{0,1\}$ so that $e_{\min} = 1$; and let the contestants' skills be picked iid with uniform probability in the interval $[d,D] = [1,2]$ so that $\beta = 1$.

By Theorem 5, the output is bounded *below* (noting $a = 7, b = 28$) by 5.6 at any NE-selection under the proportional prize. On the other hand, by Theorem 12, the output is bounded *above* by

$$(2B^2CD/c)(1 + \beta|N|\epsilon)) = 4(1 + 7(0.05)) = 5.4$$

at any IR- selection under the deterministic prize. Thus the proportional prize outperforms the deterministic.

8.1 Welfare

For simplicity we take $\beta = 1/(D - d)^{|N|}$ in this, and the next, subsection, *i.e.*, the random variables $\mu^n(1)$ are iid with uniform distribution on $[d, D]$. When the deterministic prize is used, only the contestants in the elite coterie $K(\chi)$ (whose average size is $1 + [|N|\epsilon/(D - d)^{|N|}]$) get the prize with significant probability under any IR strategy tuple. More precisely, the remaining contestants in $N \setminus K(\chi)$ get the prize with probablity at most $\underline{v}(\chi)B \sum_{k \in K(\chi)} \delta^k(1)$ (See the proof of Theorem 8 in the appendix for this estimate.)

If the proportional prize is used then, at any NE, not only does the expected total output go up as we just saw, but each contestant in $N \setminus K(\chi)$ wins the prize with much greater probability than before (at least $de_{min}/|N|D \equiv O(1/|N|)$, provided $de_{min}\underline{v}(\chi)/|N|D > Ce_{min}$, *i.e.*, provided $\underline{v}(\chi) > C|N|D/d$). Thus, *provided the minimum valuation* $\underline{v}(\chi)$ *of the prize is large enough, all the contestants in* $N \setminus K(\chi)$*, who constituted the impoverished majority under the deterministic scheme, suddenly find their prospects brighten when the proportional scheme is introduced and are able to become better off by working hard.* The elite coterie $K(\chi)$, of course, loses its status: the probabilities of winning the coveted prize drops from $O(1/|K(\chi)|)$ to $O(1/|N|)$ for each of its members, though they still must work so as to not lag behind the others. In short, the proportional prize inspires all contestants to work hard and considerably raises total output, as well as the payoffs of the impoverished majority.

8.2 Bounded Deviation

Suppose that, when a contestant produces a fraction x of total output, he wins the prize with probability $h(x)$, with $h(0) = 0$ and $h(1) = 1$; and that h is of bounded deviation from the linear function π_D, *i.e.*, there are positive constants m, M such that

$$m(x - y) < h(x) - h(y) < M(x - y), \text{ for all } y < x$$

Then a careful rereading of the proofs reveal that the estimates of Theorems 1 and 2 survive, though in slightly weakened form: lower bounds need to be diminished by a factor of m/M and upper bounds to be raised by a factor of M/m. In the same vein, a contestant's utility from winning the prize with probability p could be $f(p)$ instead of the standard expected value $pf(1)$.If f is of bounded deviation from the linear expectation $pf(1)$, we can accomodate f just like h. Finally the quantification of output can be altered without disrupting our results, so long as the alteration is of bounded deviation.

8.3 Multiple Prizes

One might wonder what happens when $l \leq |N|$ apriori fixed deterministic prizes are used instead of a single prize. When $|N| = 2$ it is evident that using two prizes is wasteful since the loser will always get the second prize for free. In general, if $l << |N|$, then again the proportional prize will perform better. The reason is as follows. Assume everyone works hard. Define l "heroes" by the top l mean outputs (as in section 7); and then define the coterie K to consist of those contestants whose outputs have a positive probability of overtaking the weakest hero. Arguing as in the proof of Theorem 8, the maximal effort in K will effectively bound the total IR output, regardless of the values of the l prizes. Also, as in the previous section, the expected size of K will be small. Thus the proportional prize will outperform l deterministic prizes when $l << |N|$. We leave the case of general l for future work.

8.4 Interdependent Production

The discerning reader will notice that our analysis remains valid even if the random output produced by a contestant is influenced by the effort (possibly factored through output) of the *others*.Various assumptions will need to be recast (somewhat cumbersomely) but the same method of proof applies. We skip the details.

8.5 More General Elite

We need not be so cut and dried as to require that non-elite contestants cannot overlap with the hero. This was done for ease of exposition. But, more generally, overlap with small probability does not disturb our conclusions. Say that $K(\chi)$ is an "$(1-\epsilon)-$ elite" set if the probability of any contestant in $N \setminus K(\chi)$ producing output equalling or exceeding the hero's, is at most ϵ. (This probability is to be of course considered under the scenario that the contestant and the hero are both at effort level 1; and, in the case of interdependent production, that everyone in $K(\chi)$ is also at effort level 1.) Then Theorem 8 holds, replacing c by $c/(1-\epsilon)$ in the upper bound; and so Theorem 12, and hence also the comparison being carried out in this section, holds with the same amendment.

9 The Strategic Game of Incomplete Information

Our main theme, namely that π_P elicits better performance than π_D when contestants' characteristics are sufficiently diverse and noise is small, has been established under the hypothesis that contestants know each others' characteristics. Now we present some evidence that the theme remains intact even when a contestant knows only his own characteristics with certainty and has a probability distribution over those of his rivals. This is the standard scenario of incomplete information. Our analysis will be in terms of illustrative examples, and not at the level of generality of the complete information case. (We hope it will spur others to carry out a more thoroughgoing.analysis.)

Let $E = \{0, 1\}$ and $N = \{1, 2\}$. Let $\delta^n(1) = 1$ and[14] $v^n = v > 1$ for $n = 1, 2$; *i.e.*, the incompleteness of information pertains only to the productivities τ^1, τ^2. Of course, $\tau^n_z(0) = 0$ as always, no matter what the "skill" z of contestant n may be. Suppose that $\tau^n_z(1)$ is uniformly distributed on the interval $[z, z+\epsilon]$, where ϵ is a measure of the noise on the output. Furthermore suppose that the skills of the contestants $n = 1, 2$ are drawn independently from the intervals $[a_1, b_1]$ and $[a_2, b_2]$, with uniform probability (and that all this is common knowledge to the contestants).

Since contestant n is informed of only his own skill, a strategy for him is given by a function

$$\sigma^n : [a_n, b_n] \to [0, 1]$$

where $\sigma^n(x)$ is the probability with which n chooses effort 1 when his skill is x.

For any prize allocation scheme π, the game of incomplete information Γ^*_π is then defined in the standard manner. It depends not only on $\pi = \pi_P$ or π_D but also on the parameters $v, a_1, b_1, a_2, b_2, \epsilon$ which we suppress because they are fixed.

First suppose ex-ante symmetry between the contestants and no noise: $[a_1, b_1] = [a_2, b_2] = [0, 1]$ (say) and $\epsilon = 0$.

Let $F^n_\pi((p, \sigma')|x)$ denote the payoff of n in the game Γ^*_π, when he chooses effort 1 with probability p and his skill level is x, while his rival chooses the strategy σ',*i.e.*,

[14] If $v \leq 1$ then the only NE in $\Gamma^*_{\pi_D}$ or $\Gamma^*_{\pi_P}$ is that both agents never work (since effort 1 costs 1 which cannot be compensated by any probability of winning the prize)

if n's strategy is σ, his payoff in Γ_π^* will be

$$F_\pi^n(\sigma, \sigma') = \int_0^1 F_\pi^n((\sigma(x), \sigma')|x)dx.$$

Notice that $F_\pi^n((1, \sigma')|x)$ increases[15] in x (for fixed n, π, σ'), since n's disutility of effort stays constant at 1 while his probability of winning the prize goes up[16]. Thus n's best reply to σ' is to switch from 0 to 1 at some "threshold" skill c, which solves $F_\pi^n((1, \sigma')|c) = 0$ *i.e.*, denoting by σ_c the strategy which assigs effort 1 if $x \geq c$ and effort 0 if $x < c$, we see that σ_c is a best reply to σ' in the game Γ_π^* if $F_\pi^n((1, \sigma')|c) = 0$. We conclude that (σ_c, σ_c) is a[17] (symmetric) NE in Γ_π^* if $F_\pi^n((1, \sigma_c)|c) = 0$. The unique $c(\pi)$ that solves this equation is computed rather easily for $\pi = \pi_P$ or π_D. Indeed we have $F_{\pi_D}^n((1, \sigma_c)|c) = cv - 1$ and

$$F_{\pi_P}^n((1, \sigma_c)|c) = cv + \int_c^1 (\frac{cv}{x+c})dx - 1 = cv[1 + ln\frac{1+c}{2c}] - 1$$

which gives (denoting $c(\pi_D) \equiv c_D$ and $c(\pi_P) \equiv c_P$)

$$c_D = \frac{1}{v} \tag{5}$$

and

$$v = \frac{1}{c_P[1 + ln(\frac{1+c_P}{2c_P})]} \tag{6}$$

When $c_P = 0$, the right hand side of (6) is infinity by L'Hospital's rule while at $c = 1$, it is 1. Since $v > 1$ the solution of (6) is $c_P < 1$, hence we have $ln(\frac{1+c_P}{2c_P}) > 0$. Thus, for any $v > 1$, we deduce that $c_P > c_D$. In short, more contestant-types are working at NE under π_P than under π_D and hence π_P elicits more expected output.

Next let us consider the effect of allowing for ex-ante asymmetry of the incomplete information. To this end, let $[a_2, b_2] = [\Delta, 1+\Delta]$ for $0 < \Delta < 1$[18] and $[a_1, b_1] = [0, 1]$, *i.e.*, contestant 2's skills are Δ-higher than 1's, so that Δ denotes the degree of asymmetry. As before, fix the noise $\epsilon = 0$. Arguing as in the ex-ante symmetric case,

---

[15] weakly in $\Gamma_{\pi_D}^*$ and strictly in $\Gamma_{\pi_P}^*$

[16] weakly in $\Gamma_{\pi_D}^*$ and strictly in $\Gamma_{\pi_P}^*$

[17] also "the", i.e., there is only one symmetric NE as the reader may easily verify.

[18] If $\Delta > 1$ then we have the trivial situation that the highest skill-type of 1 cannot beat the lowest skill type of 2 which renders the deterministic prize ineffective, while the proportional still continues to elicit effort.

there again exist thresholds $c^n_D(\Delta), c^n_P(\Delta)$ such that $(\sigma^1_{c_D(\Delta)}, \sigma^2_{c_D(\Delta)}), (\sigma^1_{c_P(\Delta)}, \sigma^2_{c_P(\Delta)})$ constitute the symmetric NE of the games $\Gamma^*_{\pi_D}, \Gamma^*_{\pi_P}$ respectively; and, moreover,

$$c^n_P(\Delta) < c^n_D(\Delta)$$

for $n = 1, 2$ and all Δ (unless v is so small that no contestant ever works in NE$-$ we implicitly eliminate such trivial NE by presuming v is high enough). Thus π_P always outperforms π_D and, as anticipated, *the superiority of π_P becomes more pronounced as the degree Δ of the asymmetry rises.* (For an example, similar in spirit, in the context of complete information, see [12]: there the dissimilarity of contestants is increased from 0 to ∞, and a threshold is shown to exist above which π_P supplants π_D in eliciting more output.)

The exact calculations for the asymmetric case emerge from the following lemma. Suppose a contestant is informed that his rival's output is uniformly distributed in some interval $[z, z + \eta] \subset R_+$ and that his own skill is x. Fix x and think of z, η as variable. We can compute two critical values $z_D \equiv z_D(x, \eta)$, $z_P \equiv z_P(x, \eta)$ such that the expected payoff of the contestant is zero in $\Gamma^*_{\pi_D}$, $\Gamma^*_{\pi_P}$ if he chooses effort 1 and if $z = z_D$, $z = z_P$ respectively. Since this payoff varies inversely in z, the contestant's best reponse to the rival is to choose effort 1 if $z < z_D$ and effort 0 if $z > z_D$ in the game Γ_D (or, effort 1 if $z < z_P$ and 0 if $z > z_P$, in the game Γ_P). The critical values z_D, z_P are as follows .

Lemma 14 *The critcal z-values are*

$$z_D = x - \eta/v \text{ and } z_P = \frac{\eta}{\exp(\eta/vx) - 1} - x.$$

Moreover we have

$$x(v-1) - \eta \leq z_P \leq x(v-1).$$

We leave it to the reader to see how our results for the asymmetric case can be straightforwardly derived from this proposition. In fact, this proposition suffices also for the analysis of games of "partial information" which lie between what we, following others, have called games of "complete" and "incomplete" information. To be concrete suppose $[a_n, b_n]$ is partitioned into k (for simplicity, equal) subintervals $[a_n + i\Delta, a_n + (i+1)\Delta]$ where $\Delta = (b_n - a_n)/k$ and $i = 0, 1, 2, ...k-1$. (When $k = 1$ we have "incomplete" information and as $k \to \infty$ we converge to "complete" information.) Each contestant is now informed of his own exact skill and of the subinterval of $[a_n, b_n]$ in which his rival's skill lies. This defines a game of partial information in the obvious way (from his initial probability distribution on $[a_n, b_n]$, the contestant can infer conditional probabilities of his rival's skill given the subinterval of $[a_n, b_n]$ in which it lies).

We have not done the exact calculations, but it seems reasonably clear that π_P outperforms π_D for every k, not just for the two extreme points $k = \infty$ and $k = 1$ that have already been checked.

9.1 The Effect of Noise

The purpose of our last example is to show that if the noise on output becomes so large as to make skills count for little, then π_D is more efficacious in eliciting output compared to π_P, whereas with small noise it is the other way around.

As before let $N = \{1, 2\}$. Now there are three levels of effort $E = \{0, 0.5, 1\} = \{0, S, W\}$ where S means "shirk" (exert little effort) and W means "work"; and 0 of course means non-participation in the game, as usual. The disutilities for these are $0, \delta, \Delta$ respectively, with δ small and Δ large. The skills are represented by n equally spaced points $a_1, \ldots, a_n$ in the interval $[2, 2 + \gamma]$ with $a_1 = 2$ and $a_n = 2 + \gamma$. Each contestant, independently of his rival, has probability $1/n$ of having any skill a_i,

Efforts $0, S$ lead to outputs $0, 1$ respectively, regardless of skill. Effort W leads to output uniformly distributed in the closed interval $[a_i, a_i + \varepsilon]$ if the skill is a_i,where ε represents the level of noise. To make noise large relative to skills, it will be simpler to restrict $0 \leq \varepsilon \leq 2$ and to fix γ small. (The smaller γ is fixed to be, the more it is the case that noise overwhelms skill when it approaches its upper bound.) Thus the game of incomplete information $\Gamma_\pi(\varepsilon, V)$ now depends only on the prize scheme $\pi = \pi_P$, π_D, on the noise ε and on the (common) valuation of the prize V.

Both for ease of calculation, and for better perspective, we devise a new measure for the efficacy of π in eliciting performance. Let $V_\pi(\varepsilon)$ be the smallest value[19] of the prize above which (W, W) is an NE in the game $\Gamma_\pi(\varepsilon, V)$, *i.e.*,

$$V_\pi(\varepsilon) = \inf \{V : (W, W) \text{ is an NE of } \Gamma_\pi(\varepsilon, V)\}$$

We claim that, by fixing n large enough, γ small enough, and δ sufficiently smaller than Δ, the following result obtains:

$$V_{\pi_P}(\varepsilon) < V_{\pi_D}(\varepsilon) \text{ for small enough } \varepsilon,$$

and

$$V_{\pi_P}(\varepsilon) > V_{\pi_D}(\varepsilon) \text{ for large enough } \varepsilon.$$

[19]The measure V_π is used in [12] to compare not just π_P and π_D, but to rank order a class of schemes π of which π_P and π_D are two instances.

The first (second) inequality says that prizes which induce (W, W) — as an NE — via π_D form a *strict* subset (superset) of prizes that induce (W, W) via π_P; and, in this sense, π_P is superior (inferior) to π_D in eliciting performance.

We shall check the first inequality for $\varepsilon = 0$ and the second for $\varepsilon = 2$, and the claim will then follow from continuity considerations.

Let $\varepsilon = 0$. Consider a contestant of the lowest skill level of 2, contemplating the choice of $0, S, W$ under the scenario that his rival has chosen W at all skill levels. It must be that W is his best reply in order to induce (W, W) an NE. of $\Gamma_\pi(0, V)$

When $\pi = \pi_D$, he cannot win the prize with effort S since he produces 1 while the rival produces at least 2. If he chooses effort W, then he can only win the prize (with probability 1/2) in the event that the rival has matching lowest skill 2, and this event occurs with probability $1/n$. Thus his situation is summed up in the following table (where "incentive to work" means the change in payoff by switching to W)

Effort	Prob of win	Payoff	Incentive to work
0	0	0	$(1/2n)V - \Delta$
S	0	$-\delta$	$(1/2n)V - \Delta + \delta$
W	$1/2n$	$(1/2n)V - \Delta$	

We conclude that, in order to make (W, W) an NE of $\Gamma_{\pi_D}(0, V)$, it is *necessary* to have

$$V \geq 2n\Delta$$

Now let $\pi = \pi_P$. Assume $\gamma \leq 1$. Noting that we get an upper (resp. lower) bound on his probanility of winning the prize by imagining his rival to always be endowed with the lowest (resp. highest) skill, and overestimating the output $2 + \gamma$ of the highest skill by 3, we get

Effort	Prob of win	Payoff	Incentive to work
0	0	0	$(2/5)V - \Delta$
S	$\leq 1/3$	$\leq (1/3)V - \delta$	$\geq (1/15)V - \Delta + \delta$
W	$\geq 2/5$	$\geq (2/5)V - \Delta$	

Thus, to make (W, W) an NE of $\Gamma_{\pi_P}(0, V)$, it *suffices* to have

$$V \geq 15\Delta$$

So $V_{\pi_P}(0) < V_{\pi_D}(0)$ provided $n \geq 8$ and $\gamma \leq 1$.

Next suppose $\varepsilon = 2$. Choose $\gamma \leq 0.5$ small enough to ensure that, if he chooses W, he wins the prize with probability at least 0.4. (a number close to half) under either $\pi = \pi_P$ or $\pi = \pi_D$. (This is clearly feasible since both probabilities converge to 0.5 as γ goes to 0.).

Under π_D we get the table

Effort	Prob of win	Payoff	Incentive to work
0	0	0	$(0.4)V - \Delta$
S	0	$-\delta$	$(0.4)V - \Delta + \delta$
W	≥ 0.4	$\geq (0.4)V - \Delta$	

Thus, to make (W, W) an NE of $\Gamma_{\pi_D}(2, V)$, it *suffices* to have

$$V \geq (0.4)^{-1}\Delta$$

Now consider π_P. Recall that the rival of the strongest skill produces uniformly in the interval $[2+\gamma, 4+\gamma]$, *i.e.*, produces at most $2+\gamma+1 \leq 3.5$ with probability 0.5.and at most $2+\gamma+2 \leq 4.5$ with the remaining probability 0.5. Hence the effort S ensures that the prize is won with probability at least $0.5(1/(4.5)) + 0.5(1/(5.5)) \geq 0.2$, and we get the table

Effort	Prob of win	Payoff	Incentive to work
0	0	0	
S	$\geq 1/5$	$\geq (0.2)V - \delta$	$\leq (0.3)V - \Delta + \delta$
W	$\leq 1/2$	$\leq (0.5)V - \Delta$	

Thus it is *necessary* to have, taking $\delta \leq (0.2)\Delta$,

$$V \geq (0.3)^{-1}(\Delta - \delta) \geq (0.3)^{-1}(0.8)\Delta > 2.6\Delta$$

Since $(0.4)^{-1} = 2.5 < 2.6$, we conclude that $V_{\pi_P}(2) > V_{\pi_D}(2)$ provided γ is small enough (*i.e.*, the noise is large enough).

Remark 15 *It is evident that if we introduce noise ε' on the output of S, with range in $[1, 1+\varepsilon']$, then so long as $\varepsilon' < 1$, our example above is not just unhampered, but in fact reinforced. It is also easy to see that increasing n and decreasing γ (which is our proxy for increasing noise) will amplify the conclusions of our example. Finally it suffices, for our example to work, that the probability of any of the n skill levels should go to zero as n goes to infinity (it need not precisely be $1/n$).*

10 Proofs

This section contains proofs that were postponed.

10.1 Theorem 5

Proof. *Let* $\sigma = \left(\sigma^1, \ldots, \sigma^N\right)$ be an NE under π_P at χ. Denote by Z the total expected output at σ. If $\sigma^n(0) = 0$ for all n then, by Assumption 1,

$$Z \geq |N| de_{min}$$

Now suppose $\sigma^n(0) > 0$ for some n. Let $\mathbf{W}$ denote the (stochastic) total output produced by all the contestants *other* than n at σ and let Z' denote the expectation of $\mathbf{W}$. Now if n is choosing effort 0 with positive probability, for a payoff of zero, it must be that effort 1 does not get him a positive payoff, *i.e.*, (with $\mathbb{E}$ for expectation)

$$0 \geq \mathbb{E}\left[\frac{d\underline{v}}{D + \mathbf{W}}\right] - C \geq \left[\frac{d\underline{v}}{D + Z'}\right] - C$$

where the first inequality comes again from Assumption 1 and the second is Jensen's inequality[20]. This yields

$$Z' \geq \frac{d\underline{v}}{C} - D$$

Since $Z \geq Z'$, we conclude that

$$Z \geq \min\left\{|N| de_{min}, \frac{d\underline{v}}{C} - D\right\}$$

■

10.2 Theorem 8

Proof. Since $\chi \equiv (\delta^n, \tau^n, v^n)_{n \in N}$ is fixed, we shall suppress it and write $K \equiv K(\chi)$. Imagine the scenario when every contestant in K chooses effort 1. In this scenario an $j \notin K$ has 0 probability of winning the prize at effort level 1 and hence, by the stochastic dominance condition of Assumption 7, at any effort level. This defines certain probabilities $\pi_*^k > 0$ for $k \in K$ to win the prize, and (recalling that by Assumption 1 each contestant produces at least $d > 0$ with probability 1) it is evident that $(i) \sum_{k \in K} \pi_*^k = 1$ and $(ii)\, \pi_*^k$ is independent of the mixed strategies chosen by the contestants in $N \setminus K$. Furthermore for $k \in K$, again by stochastic dominance, the probability that k wins can only increase if any contestants in $K \setminus \{k\}$ change to strategies other than 1. Hence we deduce that every contestant $k \in K$ can

[20] Jensen's inequality states that, if G is a convex function and X is a random variable, then $\mathbb{E}G(X) \geq \mathbb{E}G(\ X)$. We apply it here to the convex function $d/(X + D)$.

guarantee himself the payoff $\pi_*^k v^k - \delta^k(1)$ by playing 1. Thus, if $\sigma \in IR(\Gamma_{\pi_D}(\chi))$, the payoff $F^k(\sigma)$ of k at σ satisfies $F^k(\sigma) \geq \pi_*^k v^k - \delta^k(1)$ for all $k \in K$. But clearly $F^k(\sigma) \leq \bar{\pi}^k(\sigma)v^k$ (denoting $\bar{\pi}^k(\sigma) \equiv k$'s probability of winning the prize under σ), so we have

$$\bar{\pi}^k(\sigma) \geq \pi_*^k - (\delta^k(1)/v^k)$$

for all $k \in K$, which implies

$$\sum_{k\in K} \bar{\pi}^k(\sigma) \geq \sum_{k\in K} \pi_*^k - \sum_{k\in K} \frac{\delta^k(1)}{v^k} = 1 - \sum_{k\in K} \frac{\delta^k(1)}{v^k}$$

But then, putting $v \equiv v^1$ and observing $B^{-1}v \leq v^n \leq Bv$ for all $n \in N$ by part 1 of Assumption 7 , we have

$$\sum_{n\in N\setminus K} \bar{\pi}^n(\sigma) = 1 - \sum_{k\in K} \bar{\pi}^k(\sigma) \leq \sum_{k\in K} \frac{\delta^k(1)}{v^k} \leq \frac{B}{v} \sum_{k\in K} \delta_k(1)$$

So we obtain

$$\begin{aligned}
\sum_{n\in N\setminus K} F^n(\sigma) &= \sum_{n\in N\setminus K} \left[\bar{\pi}^n(\sigma)v^n - \sum_{e\in E} \sigma^n(e)\delta^n(e)\right] \\
&\leq Bv \sum_{n\in N\setminus K} \bar{\pi}^n(\sigma) - \sum_{n\in N\setminus K} \sum_{e\in E} \sigma^n(e)\delta^n(e) \\
&\leq B^2 \sum_{k\in K} \delta^k(1) - \sum_{n\in N\setminus K} \sum_{e\in E} \sigma^n(e)\delta^n(e)
\end{aligned}$$

But each $n \in N \setminus K$ can guarantee a payoff of at least 0 by choosing effort level 0, so each $F^n(\sigma)$ is non-negative since $\sigma \in IR(\Gamma_{\pi_D}(\chi))$, and thus $\sum_{n\in N\setminus K} F^n(\sigma) \geq 0$. Combining the above two inequalities, we have

$$\sum_{n\in N\setminus K} \sum_{e\in E} \sigma^n(e)\delta^n(e) \leq B^2 \sum_{k\in K} \delta^k(1)$$

Since $\delta^k(1) \leq C$ and $\delta^n(e) \geq ce$ by Assumption 1 , we get

$$\sum_{n\in N\setminus K} \sum_{e\in E} \sigma^n(e)e \leq B^2|K|\frac{C}{c}$$

Recalling also that $\mu^n(e) \leq De$ by Assumption 1, we obtain

$$\sum_{n\in N\setminus K} \sum_{e\in E} \sigma^n(e)\mu^n(e) \leq B^2|K|\frac{C}{c}D$$

Clearly, by our definition of h and Assumption 1,

$$\sum_{k\in K}\sum_{e\in E}\sigma^n(e)\mu^k(e) \leq B^2|K|\mu^h(1) \leq B^2|K|\frac{C}{c}D$$

(using the fact that $C > c$ in the last inequality). The above two inequalities prove the Key Theorem. ■

10.3 Lemma 11

For the proof of Lemma 11, it will be useful to first establish some auxiliary results. First, some notation. Let $\square = [0,1]^n$ be the unit cube in $\mathbb{R}^n$ and let $0 < \varepsilon < 1$ be fixed. For $x = (x_1, \ldots, x_n)$ in $\square$ we define

$$N_\varepsilon(x) = |\{i : x_i \in [M-\varepsilon, M)\}|, \text{ where } M = \max(x_i).$$

If X is a $\square$-valued random variable with density $\rho(x)$, we write N_ε^ρ for the random variable

$$N_\varepsilon^\rho = N_\varepsilon(X)$$

If $\rho(x) \equiv 1$ then the x_i are iid with uniform density on $[0,1]$. In this case we will show that N_ε^1 is closely related to the binomial random variable B_ε, which counts the number of successes in n independent trials with individual success probability ε:

$$\Pr(B_\varepsilon = k) = \binom{n}{k}\varepsilon^k(1-\varepsilon)^{n-k}.$$

Lemma 16 *If* $\rho(x) \equiv 1$ *then*

$$\Pr\left(N_\varepsilon^1 = k\right) = \begin{cases} \Pr(B_\varepsilon = k) & \text{if } k < n-1 \\ \Pr(B_\varepsilon = n-1) + \Pr(B_\varepsilon = n) & \text{if } k = n-1 \end{cases}$$

Moreover

$$E\left(N_\varepsilon^1\right) \leq n\varepsilon \tag{7}$$

Proof. It suffices to establish the first statement, since it implies that B_ε stochastically dominates N_ε^1, which in turn implies the second statement. For the proof of the first statement we note that the possible values of N_ε^1 are $0, 1, \ldots, n-1$, while those of B_ε are $0, 1, \ldots, n$. Therefore it suffices to prove that

$$\Pr\left(N_\varepsilon^1 = k\right) = \Pr(B_\varepsilon = k) \text{ for } k < n-1$$

Ignoring ties, which occur with probability 0, the event $N_{\varepsilon}^{1}=k$ is a disjoint union of $n\binom{n-1}{k}$ events, corresponding to the choice of the maximum index (in n ways) and the choice of the next k indices (in$\binom{n-1}{k}$ ways). By symmetry, each of these events has probability $\Pr(E_k)$, where E_k is the event

$$E_k=\{x_1 \text{ is largest}\}\&\{x_2,\ldots,x_{k+1}\in(x_1-\varepsilon,x_1)\}\&\{x_{k+2},\ldots,x_n\in[0,x_1-\varepsilon]\}$$

Thus its suffices to show that

$$\Pr(E_k)=\frac{\Pr(B_{\varepsilon}=k)}{n\binom{n-1}{k}}=\frac{\binom{n}{k}\varepsilon^k(1-\varepsilon)^{n-k}}{n\binom{n-1}{k}}=\frac{\varepsilon^k(1-\varepsilon)^{n-k}}{n-k}$$

Now writing $q(x)=\Pr(E_k|x_1=x)$ we have

$$\Pr(E_k)=\int_0^1 q(x)\,dx$$

Since $x_2,\ldots,x_n$ are independent and uniform on $[0,1]$ we get

$$q(x)=\begin{cases}\varepsilon^k(x-\varepsilon)^{n-k-1} & \text{if} \quad x>\varepsilon\\ 0 & \text{if} \quad x\leq\varepsilon\end{cases}$$

Integrating over x, making a change of variable $y=x-\varepsilon$, we get, as desired

$$\Pr(E_k)=\int_{\varepsilon}^{1}\varepsilon^k(x-\varepsilon)^{n-k-1}dx=\varepsilon^k\int_0^{1-\varepsilon}y^{n-k-1}dy=\frac{\varepsilon^k(1-\varepsilon)^{n-k}}{n-k}$$

■

Lemma 17 *Suppose $\rho(x)$ is bounded above by a constant β. Then we have*

$$E(N_{\varepsilon}^{\rho})\leq\beta n\varepsilon.$$

Proof. Using (7) we get

$$E(N_{\varepsilon}^{\rho})=\int_{\square}N_{\varepsilon}(x)\,\rho(x)\,dx\leq\beta\int_{\square}N_{\varepsilon}(x)\,dx=\beta E\left(N_{\varepsilon}^{1}\right)\leq\beta n\varepsilon$$

■

We can now prove Lemma 11

Proof. Transform Y, distributed uniformly on $[d,D]$, to $X=[Y-d]\,[D-d]^{-1}$ which is uniform on $[0,1]$. The average size of the elite set is unaffected by this transformation. Thus the result follows from Lemma 17 ■

10.4 Lemma 14

Proof. First consider π_D. Then $z = z_D$ implies $x = z + \eta/v$, and thus the player wins if the opponent's output lies in the interval $[z, z + \eta/v]$. This event has probability $(\eta/v)/\eta = 1/v$ and gives expected payoff $v(1/v) - 1 = 0$.

Now consider π_P. The expected payoff is

$$\frac{1}{\eta}\int_z^{z+\eta}\left(\frac{xv}{x+y}\right)dy - 1 = \frac{xv}{\eta}\ln\left(\frac{x+\eta+z}{x+z}\right) - 1$$

Setting this equal to zero and solving for z we get

$$z = \frac{\eta}{\exp(\eta/xv) - 1} - x = z_P$$

For the bounds on z_P we note that for an opponent of skill exactly $y^* = x(v-1)$ the payoff under π_P is $\frac{xv}{x+y^*} - 1 = 0$. Thus if $z + \eta < y^*$ the payoff at each y in $[z, z+\eta]$ is ≥ 0, which implies $z_P \geq y^* - \eta$. Similarly if $z > y^*$, the payoffs in $[z, z+\eta]$ is ≤ 0, which implies $z_P \leq y^*$. ■

References

[1] Anton, J., and Yao, D. (1992). Coordination in split award auctions. *Quarterly Journal of Economics* 107:681-701.

[2] Barut, Y. and Kovenock, D. (1998). The symmetric multiple prize all-pay auction with complete information. *European Journal of Political Economy.* 14:627-644.

[3] Baye, M., Kovenock, D. and De Vries, C.G. (1993). Rigging the lobbying process: An application of the all-pay auction. *American Economic Review* 83:289-294.

[4] Baye, M., Kovenock, D. and De Vries, C.G (1994). The solution to the Tullock rent-seeking game when R is greater than 2: Mixed strategy equilibria and mean dissipation rates. *Public Choice* 81:363-380.

[5] Broecker, T. (1990). Credit-worthiness tests and interbank competition. *Econometrica.* 58:429-452.

[6] Bulow, J., and Klemperer, P. (1999). The genenralized war of attrition. *American Economic Review.* 89:175-189.

[7] Che, Y.K. and Gale, I. (1997). Rent dissipation when rent seekers are budget constrained. *Public Choice* 92:109-126.

[8] Che, Y.K. and Gale, I. (1998). Caps on political lobbying. *American Economic Review* 88:643-651.

[9] Clark, D., and Riis, C. (1998). Competition over more than one prize. *American Economic Review.* 88:276-289.

[10] Dubey, P., and Geanakoplos, J. (2010). Grading exams: 100,99,98 .... or A,B,C ? *Games and Economic Behavior*, Vol 69, Issue 1,pp 72-94, Special Issue in Honor of Robert Aumann

[11] Dubey, P., and Haimanko, O. (2003). Optimal scrutiny in multi-period promotion tournaments. *Games and Economic Behavior.* 42(1):1-24

[12] Dubey, P., and Sahi, S. (2012). The Allocation of a Prize (Expanded), Working Paper, Department of Economics, Stony Brook University

[13] Dubey, P., and Wu, C. (2001). When less scrutiny induces more effort. *Journal of Mathematical Economics.* 36(4):311-336..

[14] Ellingsen, T. (1991). Strategic buyers and the social cost of monopoly. *American Economic Review* 81:648-657.

[15] Fang Hanming (2002) Lottery versus All-pay Auction Models of Lobbying. *Public Choice*, pp 351-371

[16] M.A.de Frutos (1999) Coalitional Manipulations in a Bankruptcy Problem, *Review of Economic Design*, Vol 4, No 3, pp 255-272.

[17] Glazer, A., and Hassin, R. (1988). Optimal contests. *Economic Inquiry.* 26:133-143.

[18] Green, J., and Stokey, N. (1983). A comparison of tournaments and contracts. *Journal of Political Economy.* 91(3):349-364.

[19] Hillman, A.L. and Riley, J.G. (1989) Politically contestable rents and transfers. *Economics and Politics* 1:17-39.

[20] Krishna, V., and Morgan, J. (1998). The winner-take-all principle in small tournaments. *Advances in Applied Microeconomics.* 7:61-74.

[21] Lazaer, E., and Rosen, S. (1981). Rank order tournaments as optimum labor contracts. *Journal of Political Economy.* 89:841-864.

[22] Moldovanu, B. and Sela, A. (2001). The optimal allocation of prizes in contests. *American Economic Review.* 91(3):542-558.

[23] Nalebuff, B., and Stiglitz, J. (1983). Prizes and incentives: Towards a general theory of compensation and competition. *Bell Journal of Economics.* 14:21-43.

[24] Nti, K.O. (1999) Rent-seeking with asymmetric valuations, *Public Choice, 98 (3-4), 415-430.*

[25] Rosen, S. (1986). Prizes and incentives in elimination tournaments. *American Economic Review.* 76:701-715.

[26] Rowley C.K. (1991) Gordon Tullock: Entrepeneur of public choice. *Public Choice* 71:149-169.

[27] Rowley C.K. (1993) *Public Choice Theory.* Edward Elgar Publishing.

[28] Tullock, G. (1975) On the efficient organization of trails. *Kyklos* 28:745-762.