Optimal Prizes

Pradeep Dubey[*] and Siddhartha Sahi[†]

9 June 2016

Abstract

Consider agents who undertake costly effort to produce stochastic outputs observable by a principal. The principal can award a prize to them with probabilities that depend on their outputs. His goal is to elicit maximal effort from the agents for the least prize.

We show that, given a general class Π of probabilistic schemes for awarding the prize and a domain of agents' characteristics on which the goal is to be achieved, there is a natural total order $\succeq$ on Π from the principal's point-of-view; as well as an "approximately $\succeq$-optimal" (if not optimal) scheme in Π, to any desired level of accuracy.

By way of illustration, we compute such optimal schemes for certain domains of binary games, i.e., games with two agents, each of whom has two effort levels (low, high). The optimal scheme is a monotonic step function, which is "in between" the well-known "deterministic" and "proportional" prizes. In the special scenario where the competition is over small fractional increments, as happens in the presence of strong contestants whose base levels of production are high even with low effort, it turns out that the optimal scheme awards the prize according to the "log of the odds", with odds based upon the proportional prize.

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

1 Introduction

Consider a principal who has to hire a given number N of agents to work for him. Each agent can undertake costly effort to produce a stochastic output that accrues

[*]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

to, and is observable by, the principal. The principal, in exchange, has a prize[1] to hand out that is valued by the agents. The question is: how should the principal award the prize in order to best elicit effort from his agents?

Two standard schemes have been studied in the literature (see section 1.1). The deterministic scheme π_D awards the prize to the agent with the highest output (breaking ties with equal probability); whereas the proportional scheme π_P awards it to all the agents, with probabilities proportional to their outputs. Both π_D and π_P are special instances of "probabilistic" schemes π which award the prize to agent[2] $n \in N = \{1, \ldots, N\}$ with probability $\pi^n(t)$ when outputs $t = (t^n)_{n \in N} \in \mathbf{R}_+^N$ are produced.

In an earlier paper [10], we examined the relative merits of π_D and π_P for a variable population N of agents. Here, in contrast, the set N will be *fixed* and we shall examine the class $\Pi = \Pi(N)$ of all probabilistic schemes on N.

The question then becomes: which scheme is "optimal" in Π from the standpoint of the principal?

A natural criterion suggests itself in order to compare schemes in Π. First notice that, given agents' characteristics[3] and a scheme for awarding the prize, a strategic game is induced among the agents in the obvious manner. Thus one could say that scheme π' is better for the principal than π'' if, at every point in the given domain $\mathbf{X}$ of agents' characteristics, the (expected, total) output at *any* Nash equilibrium (NE) under π' is at least as much as the output at *any* NE under π''. The trouble is that most pairs π', π'' are rendered incomparable in Π by this criterion. Typically there are multiple Nash Equilibria (NE) under π' and π'' at several points in the domain $\mathbf{X}$; and, at many such points, π' may be better than π'' or worse, depending on *which* pair of NE is selected for comparison. Even if — via suitable restrictions on $\mathbf{X}$ — it were possible to ensure uniqueness of NE, there would remain a second problem, which has nothing to do with the first: the output (at the unique NE) of π' may exceed that of π'', or vice versa, depending on the point chosen in $\mathbf{X}$, i.e, on the characteristics of the agents. In short, it may easily come to pass that π' is better than π'' on part of $\mathbf{X}$, and worse on the rest of it. The ambiguity of comparison here arises from the diversity of elements in $\mathbf{X}$, and this is not easy to curtail; for, precisely

[1] We take the prize to be indivisible, though perfectly divisible prizes can be easily accomodated in our framework (see footnote 9).

[2] The symbol N will be used for the set of agents, for the name of the last agent, and also for the number of agents; but its meaning will always be clear from the context. We assume throughout that $N \geq 2$.

[3] The characteristic of an agent consists of his skill and the value he places on the prize; and N-tuples of agents' characteristics are elements which constitute the domain $\mathbf{X}$ on which any scheme π is applied.

because characteristics are unobservable by the principal, $\mathbf{X}$ must be permitted to be sufficiently diverse.

We consider a "dual" problem where this ambiguity disappears. So far, the scheme π was thought of as fixed and the variable behavior (NE) induced by π was examined. In contrast, let us now fix behavior at maximal effort "$\mathbf{1}$", and focus on variable schemes which implement $\mathbf{1}$ as an NE. Scheme π' is said to be better than π'' (in symbols: $\pi' \succeq \pi''$) if, whenever π'' implements $\mathbf{1}$ as an NE on the given domain $\mathbf{X}$, so does π'.

Once there is concurrence amongst all agents as to when a prize is better than another[4] (even though they may hold widely divergent opinions as to *how* much better), the comparison $\succeq$ becomes a *total* order on Π. In this event, an optimal scheme is well-defined as a $\succeq$-maximal element of Π. We show that if Π contains at least one "responsive" scheme (such as π_P), then there exists an "approximately $\succeq$-optimal" scheme in Π, to any desired degree of accuracy. Responsive schemes are those in which increases in an agent's probability of winning the prize are "commensurate" with unilateral increases in his output (see section 4 for a precise definition); and it turns out that they are generic in Π in the standard topological sense of forming an open, dense subset of Π (see section 4.2). It further turns out, for any responsive π, that as the value of the prize is enhanced, maximal effort $\mathbf{1}$ progressively becomes an NE, the unique NE, and the unique NE with "almost dominant" strategies (see Section 5). In short, a responsive scheme enables *the principal to create more certainty that his agents will work hard, by the simple expedient of enhancing the value of the prize*[5].

Thus, unless one eliminates the generic set of responsive schemes from Π, or else becomes obsessive about saving every "ε-penny" as $\varepsilon \to 0$, there is no problem regarding the *existence* of an optimal scheme in Π (see Theorem 30 in section 6.3). The challenge is to uncover its *structure*. We do so for certain domains of binary games (with the hope that our analysis will spur the reader to consider more general domains.) The first domain consists of two agents and two effort levels (low, high), in which agents' skills can be ordered so as to exhibit "decreasing, or increasing,

[4] This would be the case if the prize is a pot of gold, or money, or indeed anything susceptible to quantification on a real scale, where all agents have strictly monotonic utility.

[5] Even when such enhancement is deemed infeasible, a different view could be taken of this result, in the spirit of "comparative statics" over a population of prizes. First *fix* agents' characteristics; and identify a prize with the point in $\mathbf{R}_+^N$ which represents its value to the different players in N. For any scheme π and $\varepsilon > 0$, denote by $V^+(\pi,\varepsilon) \subset V^*(\pi) \subset V(\pi) \subset \mathbf{R}_+^N$ the set of prizes that (respectively) implement $\mathbf{1}$ as the unique NE with "ε-dominant strategies" (see Section 5 for the precise definition), as the unique NE, and finally as just an NE. Then our result may be restated as follows: $\pi \succeq \pi'$ if, and only if, $V(\pi') \subset V(\pi)$; and, furthermore, $V^+(\pi,\varepsilon) \neq \emptyset$ if π is responsive.

returns" (see section 6.4 for a precise definition). The optimal scheme here turns out to be a monotonic step function, which is "intermediate" between the proportional and the deterministic schemes (see Remark 37).

Next we consider the binary model with the added proviso that agents' base skills are so strong (think of champions, stars, experts) that the *percentage* gain in output is small, when either agent switches from low to high effort, even though — on the absolute scale — these gains may be substantial. It turns out that the optimal scheme awards the prize according to the "log of the odds", with odds based upon the proportional scheme. Our analysis here is a little sketchy, and yet to be fully formalized, so we have placed it in a "Second Appendix" as work in progress.

It is worth noting that the optimal schemes in both these binary models are "robust", in that they are valid for a large class of skill distributions. In the first scenario, the optimal scheme does not depend upon the distribution in the interior of the domain, but only on its *boundary;* while in the second scenario, it does not depend upon the distribution *at all,* but only on the qualitative feature of "decreasing or increasing returns" in the skills (see sections 6.4 and 8.0.1).

Our model restricts attention to *finitely* many effort levels. This is meaningful when effort entails indivisible tasks. Think, for example, of a salesman-agent who makes trips to call upon his clients, with each trip taking up significant time. Here fractional trips have no meaning, since they incur costs with no results, and effort levels may be taken to be finite.

In such a setting, maximal effort is well-defined and its implementation is consistent with the maximization of expected total profit, provided that the principal values agents' outputs sufficiently highly compared to the prize he hands out. This too can often obtain in practice. To see the possibility in the context of our example, suppose that the items for sale are *highly*-priced (such as luxury cars or precious stones). The agents, coordinating sales on behalf of the owner-principal, are assigned disjoint sets of clients. The effort of an agent corresponds to the set of clients he calls upon. His skill is reflected in the probabilities with which he converts his client-calls into sales. Let us assume that these probabilities are all above some decent threshold, so that each agent achieves a significant jump in expected sales whenever he increases his effort and calls upon additional clients. The sales revenue generated by any agent is credited into the company account and hence it is perfectly observable by the principal and accrues to him in its entirety (i.e., the agent cannot market his output elsewhere). The prize that the agents are striving for is promotion to a coveted higher echelon in the company, which brings considerable status-utility and is accompanied by a wage increase. In this scenario, the principal will want to incentivize all his agents to exert maximal effort. An optimal scheme will minimize the wage increase that he needs to announce in order to achieve his goal. (The reader can no doubt think of other examples within the contours of our model.)

In conclusion, let us point out that, for the most part, we suppose that there is complete information among the agents about each others' characteristics (though the principal is ignorant of them). This is a tenable hypothesis when the agents compete in close proximity with one another. However the case of incomplete information, when an agent knows only his own characteristics and has limited information regarding those of the others, is also clearly important. We touch upon it in Section 8, and show via examples that our analysis remains intact; but these examples, while suggestive, do point to the need for further work.

1.1 Related Literature

There is a sizeable theoretical literature[6] on lobbying, where agents put up bids of money and are awarded the prize either via the proportional scheme or the deterministic scheme, called often "lottery" or "all-pay auctions", respectively (see, *e.g.*,[26],[16],[11],[23],[24],[3],[4], [8],[9],[21],[12] and the references therein). In much of this literature agents 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 agents 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 ("bids"). However proportional prizes do not seem to have received attention there. For tournaments with a single prize, see [18],[15],[20],[22]. Subsequent writers have considered multiple prizes whose number and sizes are fixed prior to the contest, and which are then awarded to the agents based upon the rank-order of their performance ([14], [5], [1], [7], [17], [6], [2], [19]).

In both these strands of literature, attention is mostly restricted to the scenario where the NE (of the game played for the prize) are unique.

A comparison between the deterministic and proportional prizes was carried out in [10], allowing for multiplicity of NE under either scheme. Circumstances were delineated in [10] in which the worst NE under π_P had higher output than the best NE under π_D. But the specific nature of π_P and π_D played an important role there. We have not been able to extend style of comparison to general probabilistic schemes.

[6] Quite apart from the theory, proportional schemes are ubiquitous in practice in the form of *lottery tickets* for prizes.

2 The Model

Recall that the set N of agents is fixed, and the principal can observe only the outputs $t = (t^n)_{n \in N}$ produced by them, based on which he must award the prize.

Definition 1 *A* ***probabilistic scheme*** π *for awarding the prize is given by a (measurable) map satisfying*

$$\mathbf{R}_+^N \xrightarrow{\pi} \mathbf{R}_+^N, \quad \sum_{n \in N} \pi^n(t) \leq 1, \quad \pi^n(t) = 0 \text{ if } t^n = 0$$

where the component $\pi^n(t)$*, of the vector* $\pi(t)$*, denotes the probability with which* $n \in N$ *is awarded the prize.*[7]

Notation 2 $\Pi = \Pi(N)$ *denotes the set of all probabilistic schemes on the set* N *of agents.*

There is a fixed finite set $E \subset [0, 1]$ of effort levels**.** We assume $0 \in E$ and $1 \in E$. These represent no effort and maximal effort respectively.

If an **agent** chooses effort $e \in E$, he incurs disutility $\delta(e) \geq 0$ and produces stochastic output given by a non-negative random variable $\tau(e)$ with finite mean $\mu(e) = \mathrm{Exp}\left(\tau\left(e\right)\right)$. (We allow for the possibility that the range of $\tau(e)$ is discrete, even finite.) Effort 0 is a proxy for "not participating" in the game; it incurs disutility $\delta(0) = 0$ and produces $\tau(0) = 0$ (*i.e.*, the output is 0 with certainty).

The pair (δ, τ) will be referred to as the **skill** of the agent[8].

There is an indivisible prize[9], which is handed out to the agents by a prinicpal. If an agent places **value** $v > 0$ on the prize, and is awarded it with probability p, this yields him expected utility pv.

The triple $\chi = (\delta, \tau, v)$ thus characterizes an agent.

[7] The first condition on π allows for the possibility that the prize is withheld from all agents; the second *requires* it to be withheld from any agent who produces zero output, otherwise he would be rewarded for not participating in the game.

[8] The skill of an agent increases if he can produce the same with less disutility of effort, or more with the same disutility.

[9] It is for simplicity that we take the prize to be indivisible. Perfectly divisible prizes can be easily accomodated in our framework. Indeed, denote by $u^n(q)$ the value (utility) to n of getting a fraction q of the total prize. We need only assume that u^n is strictly increasing and that the difference quotients $(u^n(q) - u^n(r))/(q - r)$ are uniformly bounded away from 0 and from ∞ on the domain of agents' characteristics. Then our entire analysis remains intact with the obvious modifications and at the cost of extra notation.

As was said in the introduction, the set N of agents is fixed, but their **characteristics** $(\chi^n)_{n\in N} = (\delta^n, \tau^n, v^n)_{n\in N}$ can vary, subject to various condititions that are about to be spelled out.

The **principal** cannot observe these characteristics, nor the effort levels $(e^n)_{n\in N}$ undertaken by the agents; as was said, all he can observe are the realizations $t = (t^n)_{n\in N}$ of the random outputs $(\tau^n(e^n))_{n\in N}$. .

We now introduce some assumptions on agents' characteristics.

Notation 3 *Let X denote a set of characteristics $\chi = (\delta, \tau, v)$; and X^N the N-fold product of X.*

Definition 4 *Given real-valued random variables Y and Z, we say that Y* stochastically dominates[10] *Z (and write $Y \succsim_{st} Z$)if, for every real number α,*

$$\Pr\{Y > \alpha\} \geq \Pr\{Z > \alpha\}.$$

Axiom 5 ***(Boundedness)*** *There exist universal positive constants c, C, d, D, β such that, for all $e \in E \setminus \{0\}$ and $(\delta, \tau, v) \in X$*

$$ce < \delta(e) < Ce, \quad de < \mu(e) < De, \quad \tau(e) < \beta; \tag{1}$$

and, moreover, δ and τ are both non-decreasing in e.

Axiom 6 ***(Productivity)*** *There exists a positive constant κ such that, for all $e \in E \setminus \{1\}$ and $(\delta, \tau, v) \in X$*

$$\mu(1) - \mu(e) > \kappa;$$

and, moreover,

$$\tau(1) \succsim_{st} \tau(e).$$

3 The Strategic Game of Complete Information

Given characteristics $\boldsymbol{\chi} = (\chi^n)_{n\in N} \in X^N$ and the principal's choice of a scheme π, a strategic game is induced among the agents as follows.

The set of pure strategies of each agent $n \in N$ is E. Any N-tuple of pure strategies $\mathbf{e} = (e^n)_{n\in N}$ gives rise to a random vector $\tilde{t} = \tilde{t}(\mathbf{e}) = (\tau^n(e^n))_{n\in N}$ of outputs. The expected value p^k of $\pi^k(\tilde{t})$ represents the probability of k winning the prize and agent k's payoff is then $F^k(\mathbf{e}) = p^k v^k - \delta^k(e^k)$.

[10] in the "first order" sense

Denote by $\Gamma = \Gamma(\boldsymbol{\chi},\pi)$ the mixed extension of this game. The set Σ of probability distributions on E then corresponds to the mixed strategies available to each agent. (Without confusion, $F^k(\sigma)$ will continue to denote k's payoff, when the mixed strategy N-tuple $\sigma = (\sigma^n)_{n\in N} \in \prod_{n\in N} \Sigma \equiv \Sigma^N$ is played in Γ.)

Notation 7 *Given a vector $u = \left(u^k\right)_{k\in N}$ we denote by u^{-n} the vector obtained from u by deleting the component u^n (i.e., $u^{-n} = \left(u^k\right)_{k\in N\setminus\{n\}}$); and by (u^{-n}, w) the vector obtained from u by replacing the component u^n with w.*

Recall that $\sigma \in \Sigma^N$ is called a **Nash Equilibrium** (NE) of Γ if, for all $n \in N$,

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

It will be useful also to develop the notion of a **Weak Nash Equilibrium** (WNE). This is defined exactly as an NE but with one amendment: the unilateral deviations ρ of n (from σ^n) are restricted to the set

$$\{\rho \in \Sigma : \rho(e) \leq \sigma^n(e) \text{ for all } e \neq 1\},$$

i.e., to shifting probabilities from σ^n onto n's maximal effort 1. Clearly every NE is a WNE; and $\mathbf{1} = (1, \ldots, 1)$ is a WNE.

4 Monotonic and Responsive Schemes

The very purpose of a prize is to encourage each agent to produce more output. Thus it is natural to restrict attention to schemes that do not penalize an agent when his output goes up, or those of his rivals goes down.

Definition 8 *The scheme π is called* monotonic *if, for all scalars $x \geq 0$ and vectors $v \geq 0$, we have*

$$\pi^n(t^{-n}, t^n + x) \geq \pi^n(t^{-n}, t^n) \geq \pi^n(t^{-n} + v, t^n); \tag{2}$$

and it is called weakly monotonic *if only the first inequality of (2) holds.*

A distinctive role is played in our model by schemes in which increases in output are "never in vain", *i.e.* which guarantee that rewards are commensurate with output in the following sense: if the output of any agent goes up unilaterally by a decent quantum, then so does his probability of winning the prize, provided the total output of others is bounded away from 0 and ∞.

Definition 9 *A scheme* π *is* responsive *if for any* $A, B, C > 0$ *and* $n \in N$*, there exists* $\alpha > 0$ *such that*

$$\pi^n(t^{-n}, t^n + x) - \pi^n(t^{-n}, t^n) > \alpha$$

for all (t^{-n}, t^n) *and* x *satisfying* $A < \sum_{k \in N \setminus \{n\}} t^k < B$ *and* $x > C$.

Notation 10 *Denote*

$$\Pi_m = \{\pi \in \Pi : \pi \textit{ is monotonic}\}$$

$$\Pi_w = \{\pi \in \Pi : \pi \textit{ is weakly monotonic}\}$$

$$\Pi_r = \{\pi \in \Pi : \pi \textit{ is responsive}\}$$

$$\Pi_{rm} = \Pi_r \cap \Pi_m$$

$$\Pi_{rw} = \Pi_r \cap \Pi_w$$

4.1 The Proportional Scheme

Recall that the **proportional scheme** π_P awards the prize to each agent in proportion to his output, *i.e.*,

$$\pi_P^n(t) = \begin{cases} \frac{t^n}{\sum_{k \in N} t^k} & \text{for} \quad t^n > 0 \\ 0 & \text{for} \quad t^n = 0 \end{cases}$$

Remark 11 *It is evident that* π_P *is monotonic and responsive. (The deterministic scheme* π_D*, in contrast, is monotonic but fails to be responsive.)*

4.2 Genericity of Responsive Schemes

For any two schemes π', π'' in the class Π of all probabilistic schemes (see Notation 2), and any scalar $0 < \varepsilon < 1$, we may consider the convex combination $\pi^* = (1 - \varepsilon)\pi' + \varepsilon\pi''$ which is the scheme obtained by using π' with probability $1 - \varepsilon$ and π'' with probability ε. It is clear that $\pi^* \in \Pi$, i.e., the set Π is convex.

Fix a scheme $\pi \in \Pi_{rm}$ that is both monotonic and responsive (*e.g.*, in the light of Remark 11, fix $\pi = \pi_P$). Given any monotonic scheme $\widetilde{\pi} \in \Pi_m$, consider $\widetilde{\pi}_\varepsilon = (1 - \varepsilon)\widetilde{\pi} + \varepsilon\pi$. It is obvious that $\widetilde{\pi}_\varepsilon$ is monotonic. One may also readily check that $\widetilde{\pi}_\varepsilon$ is responsive. Indeed, suppose A, B, C, α apply to π in accordance with Definition 9. Then, since a unilateral increase of an agent's output does not *reduce* his probability of winning in $\widetilde{\pi}$ (by the weak monotonicity of π, see Definition 8), we see that $A, B, C, \varepsilon\alpha$ meet the requirement of Definition 9 for $\widetilde{\pi}_\varepsilon$.

(The observations of the above paragraph are also valid if we replace "monotonic" by "weakly monotonic" throughout.)

Now consider the metric on Π derived from the sup norm. It is obvious that $\widetilde{\pi}_\varepsilon$ coverges to $\widetilde{\pi}$ as $\varepsilon \to 0$. Thus responsive schemes are dense. It is clear also that they form an open set. We have therefore established:

Remark 12 *The set Π_{rm} is open and dense in Π_m (and, likewise, Π_{rw} is open and dense in Π_w).*

In view of this remark, we will restrict attention to responsive schemes in the next section.

5 Implementability of Maximal Effort via Responsive Schemes

Our aim in this section is to show that, if π is monotonic and responsive (i.e., $\pi \in \Pi_{rm}$), then maximal effort $\mathbf{1} \equiv (1, ..., 1)$ can be implemented as an NE for sufficiently high valuations of the prize; and indeed as the unique NE for even higher valuations.

One can, in fact, go a step further and show that maximal effort, while not exactly a strictly dominant strategy, can come close to being such. To this end, we need some notation.

Notation 13 *For any N - tuple $\boldsymbol{\chi} = (\delta^n, \tau^n, v^n)_{n \in N} \in X^N$ of agents' characteristics, denote*

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

Notation 14 *Recall that $t^{-n} = \left(t^k\right)_{k \in N \setminus \{n\}}$ represents a vector of deterministic outputs by the rivals of n, and denote*

$$Pr\left(n, e; \boldsymbol{\chi}, \pi, t^{-n}\right) = Exp\left[\pi^n\left(t^{-n}, \tau^n(e)\right)\right];$$

and

$$\Delta\left(n, e; \boldsymbol{\chi}, \pi, t^{-n}\right) = Pr\left(n, 1; \boldsymbol{\chi}, \pi, t^{-n}\right) - Pr\left(n, e; \boldsymbol{\chi}, \pi, t^{-n}\right).$$

In other words, $\Pr(n, e; \boldsymbol{\chi}, \pi, t^{-n})$ is the probability that $n \in N$ wins the prize in $\Gamma(\boldsymbol{\chi}, \pi)$ when he chooses effort e and outputs of his rivals is *fixed* at t^{-n}; and $\Delta(n, e; \boldsymbol{\chi}, \pi, t^{-n})$ is the increment in this probability when n deviates unilaterally from e to 1.

Definition 15 *Let* $\varepsilon > 0$. *We say that* $\mathbf{1}$ *is* ε-dominant *in the game* $\Gamma(\boldsymbol{\chi}, \pi)$ *if, for all* $n \in N$ *and all* $e \in E \setminus \{1\}$ *and all* v, *we have*

$$\sum_{k \in N \setminus \{n\}} t^k > \varepsilon \Longrightarrow \left[\Delta\left(n, e; \boldsymbol{\chi}, \pi, t^{-n}\right)\right] v^n > \delta^n(1) - \delta^n(e).$$

In other words, if $\mathbf{1}$ is ε-*dominant*, this means that it is beneficial for each player to unilaterally deviate from every $e \in E \setminus \{1\}$ to 1, for any fixed output t^{-n} of his rivals, provided that the aggregate in t^{-n} is at least ε. (In other words, ε-dominant means: dominant upto inaccuracy ε.)

Theorem 16 below makes precise the sense in which maximal effort can be implemented. We emphasize that the first two parts of the theorem only require π to be responsive and weakly monotonic, while the third part requires the full force of the monotonicity assumption.

Theorem 16 *Suppose Axioms 5 and 6 hold.*

(i) There exists v_* *such that if* $v(\boldsymbol{\chi}) > v_*$ *and* $\pi \in \Pi_{rw}$, *then* $\mathbf{1}$ *is an NE of* $\Gamma(\boldsymbol{\chi}, \pi)$.

(ii) For all $\varepsilon > 0$ *there exists* $v_*(\varepsilon)$, *such that if* $v(\boldsymbol{\chi}) > v_*(\varepsilon)$ *and* $\pi \in \Pi_{rw}$, *then* $\mathbf{1}$ is ε-*dominant in* $\Gamma(\boldsymbol{\chi}, \pi)$.

(iii) There exists v^* *such that if* $v(\boldsymbol{\chi}) > v^*$ *and* $\pi \in \Pi_{rm}$, *then* $\mathbf{1}$ *is the unique NE of* $\Gamma(\boldsymbol{\chi}, \pi)$.

An NE is more likely to be played when it is unique, since it bypasses the problem that all agents have to "coordinate" upon the *same* NE (granting, of course, that each will play according to *some* NE to begin with). As for ε-dominant strategies (for small enough ε), they strain credulity *much* less than NE: the need no longer exists for any agent to correctly anticipate his rivals' behavior since his optimal strategy is independent of what they might be doing, at least so long as they are not producing very little (or, what comes to the same thing in view of Axiom 6, so long as at least one of them is playing 0 with low probability).

Thus the thrust of Theorem 16 is that *the principal can gain more certainty that his agents will work hard, by handing out prizes of increasingly higher value.* Indeed, denote $v^+(\varepsilon) = \max\{v^*, v_*(\varepsilon)\}$. Then clearly $v_* \le v^* \le v^+(\varepsilon)$ and, provided $\pi \in \Pi_{rm}$, Theorem 16 implies that, as the minimum value of the prize rises progressively from v_* to v^* to $v^+(\varepsilon)$, $\mathbf{1}$ is an NE, then a unique NE, and finally also ε-dominant. (Of course, $v_*(\varepsilon)$, and hence also $v^+(\varepsilon)$, rises as ε goes to zero; in fact, as one may readily verify, for π_P we have $v_*(\varepsilon) \uparrow \infty$ as $\varepsilon \downarrow 0$).

Remark 17 *It is worth noting that, as will become evident from its proof, claim (iii) of Theorem 16 can be strengthened with WNE in place of NE (recall $NE \subset WNE$).*

Remark 18 *Part (ii) of Theorem 16 cannot be sharpened to derive a finite threshold $v_{**} = \sup\{v_*(\varepsilon) : \varepsilon > 0\}$ which works for all $\varepsilon > 0$ (as would have to be the case if $\mathbf{1}$ were a strictly dominant N-tuple of strategies in the game $\Gamma(\boldsymbol{\chi}, \pi)$). Even the stronger hypothesis that $\pi \in \Pi_{rm}$ would not suffice for this purpose. To see this, consider $\pi_P \in \Pi_{rm}$. If $\sum_{k \in N \setminus \{n\}} t^k < \varepsilon$ and ε is sufficiently small, then with the smallest effort e^* above 0, agent n will already win the prize with probability very close to 1 (this follows from Axiom 6). Thus the increment in his probability of winning the prize will go up by an infinitesimal amount δ if he deviates from e^* to 1. To induce him to incur the disutility of this deviation, his valuation of the prize v^n will need to be very high. Indeed, as $\varepsilon \to 0$, it is clear that $\delta \to 0$ and so $v^n \to \infty$, establishing the impossibility of sharpening (ii) of Theorem 16.*

6 Optimal Prizes with Complete Information

6.1 The Preorder on Prizes

Up until now, the prize was fixed and each agent had a value $v \in \mathbf{R}_+$ placed on it. Now we shall extend our consideration to a *set* Ω of prizes and understand v to mean a *function*

$$v : \Omega \to \mathbf{R}_+$$

which will be referred to as a **valuation**. Also we shall, from now on, consider a general subset $\mathbf{X} \subset X^N$ as the *domain* of agents' characteristics on which the principal is seeking to implement maximal effort $\mathbf{1}$. And we shall denote by Π subset

Notation 19 *For $\boldsymbol{\chi} = (\delta^n, \tau^n, v^n)_{n \in N} \in \mathbf{X}$ and $\omega \in \Omega$, denote $\boldsymbol{\chi}(\omega) = (\delta^n, \tau^n, v^n(\omega))_{n \in N}$.*

Remark 20 *The game now requires $\boldsymbol{\chi} \in \mathbf{X}, \pi \in \Pi$* and *$\omega \in \Omega$ for its specification; and is denoted $\Gamma(\boldsymbol{\chi}(\omega), \pi)$.*

Notation 21 *Let $Pr\,(n, e; \boldsymbol{\chi}, \pi)$ be the probability that n wins the prize in the game $\Gamma(\boldsymbol{\chi}, \pi)$ when he chooses effort e and all his rivals choose 1. (Note: only the skills part of $\boldsymbol{\chi}$ are relevant here.)*

Definition 22 *Let $\Omega^n(\pi)$ denote the set of all $\omega \in \Omega$ such that, for all $e \in E \setminus \{1\}, n \in N$ and $\boldsymbol{\chi} = (\delta^n, \tau^n, v^n)_{n \in N} \in \mathbf{X}$, we have*

$$[Pr(n,1;\boldsymbol{\chi},\pi) - Pr(n,e;\boldsymbol{\chi},\pi)]\, v^n(\omega) \geq \delta^n(1) - \delta^n(e).$$

Thus $\Omega(\pi)$ is precisely the set of prizes at which 1 is a best response for each agent when all his rivals also choose 1, *no matter which point $\boldsymbol{\chi}$ in the domain $\mathbf{X}$ is under consideration.*

Definition 23 *We write $\widetilde{\pi} \succeq \pi$ if, for all $\boldsymbol{\chi} \in \mathbf{X}$ and $\omega \in \Omega$*

$$\mathbf{1} \text{ is an NE of } \Gamma(\boldsymbol{\chi}(\omega);\pi) \implies \mathbf{1} \text{ is an NE of } \Gamma(\boldsymbol{\chi}(\omega),\widetilde{\pi}).$$

Theorem 24 *The binary relation $\succeq$ is a preorder*[11] *on Ω.*

6.2 The Total Preorder on Prizes

We now pinpoint conditions under which the preorder $\succeq$ become a total preorder, *i.e.* complete. To this end, let us introduce:

Axiom 25 *(i) The set of prizes Ω is endowed with a total order*[12] *$\geq$; and for any $(\delta, \tau, v) \in X$, we have $v(\omega) > v(\omega')$ if, and only if, $\omega > \omega'$.*

In other words, there is *concurrence* among the agents' valuations as to whether a prize is better than another, reflected in the total order on Ω.

Theorem 26 *Suppose axioms 6 and 25 hold on X. Then $\succeq$ is a total preorder*[13] *on Π_w (for any $\mathbf{X} \subset X^N$).*

6.3 Approximately Optimal Prizes and their Existence

We specialize to the case: $(\Omega, \succsim) = (\mathbf{R}_+, \geq)$. What defines a prize is now its *size* $z \in \mathbf{R}_+$. Our Axiom 25 implies that all agents think size is "good", even though their views may be very different as to *how* good it is (think of the prize in terms of gold or money). In light of the metric structure of the reals, we strengthen Axiom 25 as follows (restating, for completeness, the strictly increasing property that is already implied by Axiom 25).

[11] i.e. a binary relation that is reflexive and transitive (not necessarily complete).

[12] i.e., $\geq$ is reflexive, antisymmetric, transitive, and complete. We write $\omega > \omega'$ if $\omega \geq \omega'$ but not $\omega' \geq \omega$.

[13] i.e., the preorder is complete

Definition 27 *(of the set $V^n(\boldsymbol{\zeta})$) Given skills $\boldsymbol{\zeta} = (\delta^n, \tau^n)_{n \in N}$ we say that the valuation $u \in V^n(\boldsymbol{\zeta})$ if there exists $\boldsymbol{\chi} = (\delta^n, \tau^n, v^n)_{n \in N} \in \mathbf{X}$ with $v^k = u$ for some $k \in N$.*

Axiom 28 *If $u \in V^n(\boldsymbol{\zeta})$, the function $u : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ is strictly increasing and continuous, with $u(0) = 0$ and $u(z) \rightarrow \infty$* uniformly[14] in u *as $z \rightarrow \infty$.*

Let $\Pi^* \subset \Pi$ be an arbitrary subset of probabilistic schemes which the principal is willing to utilize. We shall show that optimal schemes exist in Π^* up to arbitrarily small error $\varepsilon > 0$ so long as Π^* contains at least one scheme π that is weakly monotonic and responsive (i.e., $\pi \in \Pi^* \cap \Pi_{rw}$). To this end, first a

Definition 29 *We say that $\widetilde{\pi}$ is ε-optimal in Π^* if, for all $\boldsymbol{\chi} \in \mathbf{X}$ and all $z \in \mathbf{R}_+$ and all $\pi \in \Pi^*$,*

$$\mathbf{1} \textit{ is an NE of } \Gamma(\boldsymbol{\chi}(z); \pi) \implies \mathbf{1} \text{ is an NE of } \Gamma(\boldsymbol{\chi}(z + \varepsilon); \widetilde{\pi}).$$

This just says that $\widetilde{\pi}$ implements $\mathbf{1}$ as economically as any other scheme $\pi \in \Pi^*$, if we give $\widetilde{\pi}$ the minor leeway of boosting the size of the prize by ε for every agent. (Once again, ε-optimal means: optimal upto inaccuracy ε.)

Theorem 30 *Suppose Axioms 5, 6, and 28 hold on $\mathbf{X}$; and Π^* contains a weakly monotonic and responsive scheme. Then, for any $\varepsilon > 0$, there exists an ε-optimal scheme in Π^*. If, furthermore, Π^* is a finite set, then there exists an optimal scheme.*

By virtue of Theorem 30, the existence of ε-optimal (if not optimal) schemes — for arbitrarily small ε — is not in doubt. What calls for exploration is the *structure* of such a scheme. This will clearly depend upon the underlying pair Π^*, $\mathbf{X}$.

Our goal in the next two sections is to construct optimal schemes for two particular pairs Π^*, $\mathbf{X}$ (with $N = \{1, 2\}$ and $X^N = \mathbf{R}_+^{\{1,2\}}$ and all $u^n(x) = x$).

6.4 Optimal Prizes in Binary Contests

For the remainder of this paper, we restrict attention to the player set $\{1, 2\}$ and the class $\Pi^+ \subset \Pi^{\{1,2\}}$ consisting of all probabilistic prize schemes π which are represented

[14] i.e., for any $x > 0$.there exists $y > 0$ such that

$$z > y \text{ and } u \in V^n(\boldsymbol{\zeta}) \Longrightarrow u(z) > x$$

by some non-decreasing *prize function* $p : [0,1] \to [0,1]$ with $p(0) = 0$ and $p(1) = 1$, such that whenever $t^1 + t^2 > 0$, we have[15]

$$\pi^i(t^1, t^2) = \begin{cases} p\left(\frac{t^1}{t^1+t^2}\right) & \text{for} \quad i = 1 \\ 1 - p\left(\frac{t^1}{t^1+t^2}\right) & \text{for} \quad i = 2 \end{cases}$$

Notation 31 *Let* $\mathbf{Z}$ *denote the projection* $\{(\delta, \tau) : (\delta, \tau, v) \in X\}$ *of* X *onto its first two components, i.e.,* $\mathbf{Z}$ *is the set of skills that can occur in* X.

We shall first examine the binary case of two agents (*i.e.*, $N = \{1, 2\}$) with two effort levels and deterministic output. The effort levels are "shirk" ($e = 1/2$) and "work" ($e = 1$), in addition of course to effort level 0 for not participating in the game. So $E = \{0, 1/2, 1\}$. The disutility of effort is constant across $\zeta \in \mathbf{Z}$ (with[16] $\delta^n(1/2) = 0$ and $\delta^n(1) = \delta$ for $n = 1, 2$). What varies with $\zeta \in \mathbf{Z}$ is the skill (productivity) of an agent. Let $\tau(e, s)$ denote the deterministic output of each agent when he exerts effort $e \in \{1/2, 1\}$ and and is endowed with "skill" $s \in [k, K]$ (Thus $\mathbf{Z} \approx [k, K]^2$ here.)

For brevity, denote $\tau(1/2, s) \equiv \tau(s)$ and $\tau(1, s) \equiv \tau^*(s)$. We make some natural monotonicity assumptions on τ and τ^*, along with a form of "decreasing (or,later, increasing) returns to skill" :

Axiom 32 ***(Decreasing Returns to Skill)*** *Both* $\tau : [k, K] \to R_+$, $\tau^* : [k, K] \to R_+$ *are continuous and strictly monotonic; and* $\tau^*(s)/\tau(s) \leq \tau^*(s')/\tau(s')$ *if* $s' < s$. *Also* $\inf\{\tau^*(s) - \tau(s) : s \in [k, K]\} > 0$.

[15] For general N, one may consider weakly monotonic schemes which satisfy the following three conditions (where, for $w = (w^n)_{n \in N}$, we denote by θw the vector $(w^{\theta(n)})_{n \in N}$): (1) **Anonymity**: $\pi(\theta t) = \theta \pi(t)$ for any permutation $\theta : N \to N$; (2) **Scale Invariance**: $\pi(rt) = \pi(t)$ for all scalars $r > 0$; (3) **Disbursal**: $\sum_{n \in N} \pi^n(t) = 1$ if $t \neq 0$,and is 0 otherwise.

When $N = \{1, 2\}$, any scheme satisfying (1), (2), (3) is represented by some prize function p of the type we have described, and thus the three conditions characterize the class Π^+. They seem natural to us and hold for the familiar deterministic and proportional prizes. However prizes, that are awarded subject to meeting some *minimum threshold* on output, would violate both Scale Invariance and Disbursal.

For ease of computation, we shall work with Π^+ in our examples. But it is worth emphasizing that (approximately) optimal schemes exist in classes much larger than Π^+, by virtue of Theorem 30.

[16] We take $\delta^n(1/2) = 0$ for simplicity (recall that δ^n is permitted to be *weakly* increasing). But our analysis remains intact if $\delta^n(1)$ is sufficiently larger than $\delta^n(1/2) > 0$ (as can easily be checked.)

Axiom 32 says that the percentage gain in output, by switching from shirk to work, is a weakly decreasing function of the skill $s \in [k, K]$. (The case of increasing returns is entirely analogous; see Axiom 38 below.)

Our main result (see theorem 36 below) shows that, when Axiom 32 holds, *there exists an optimal scheme which takes the form of a monotonic step function.* The location of the jump points, and the sizes of the jumps, can be computed by an algorithm based on $r, R, \tilde{r}$, $\tilde{R}$,*i.e.*, on skill functions τ and τ^* restricted to the northeast boundary of the square $[k, K]^2$. And, graphically speaking, this optimal scheme lies " in between" the proportional scheme (whose graph is linear) and the deterministic scheme (whose graph has a single jump from 0 to 1 at 1/2).

To establish this result, first note that axiom 32 simplifies the analysis considerably, on account of:

Lemma 33 *Assume Axiom 32 holds. Let $s \in (k, K)$ and $t \in (k, K)$. Then there exist $s' \in [k, K]$ and $t' \in [k, K]$ such that*

$$\begin{aligned}
\frac{\tau^*(s')}{\tau^*(s') + \tau^*(t')} - \frac{\tau(s')}{\tau(s') + \tau^*(t')} &\leq \frac{\tau^*(s)}{\tau^*(s) + \tau^*(t)} - \frac{\tau(s)}{\tau(s) + \tau^*(t)} \\
\frac{\tau(s')}{\tau(s') + \tau^*(t')} &= \frac{\tau(s)}{\tau(s) + \tau^*(t)}
\end{aligned}$$

and either $s' = K$ or $t' = K$

(The proof is in the Appendix.)

Lemma 33 implies that our goal — of incentivizing an agent (of skill s) to switch from shirk to work, assuming his rival (of skill t) is working — will be achieved for every $(s, t) \in [k, K] \times [k, K]$ if it is achieved for (s, K) and (K, s) for all $s \in [k, K]$; in other words, we need only worry about incentivizing the agent in the following two extremal cases, corresponding to the north and east boundaries of the square $[k, K]^2$:

Case A His skill is $s \in [k, K]$ and his rival is working with skill K.

Case B His skill is K and his rival is working with skill $s \in [k, K]$

Denote

$$\begin{aligned}
R(s) &= \frac{\tau^*(s)}{\tau^*(s) + \tau^*(K)}, r(s) = \frac{\tau(s)}{\tau(s) + \tau^*(K)}, \\
\tilde{R}(s) &= \frac{\tau^*(K)}{\tau^*(K) + \tau^*(s)}, \widetilde{r}(s) = \frac{\tau(K)}{\tau(K) + \tau^*(s)}
\end{aligned}$$

When an agent switches from shirk to work, his fractional output goes up from $r(s)$ to $R(s)$ in Case A, $\tilde{r}(s)$ to $\tilde{R}(s)$ in Case B. Denote $q(s) = 1 - \tilde{r}(s)$. It is clear from our assumptions that $q > R > r$ and that $R(s) = 1 - \tilde{R}(s), R(K) = \tilde{R}(K) = 1/2$

It will be useful to introduce one more function, which captures the simple form of $\pi \in \Pi^{+}$ when there are only two agents.

Definition 34 ***(Prize function)*** *A prize function is a weakly increasing function $p : [0,1] \to [0,1]$ satisfying $p(1-x) = 1 - p(x)$ for all x.The function p is said to be* effective *at prize level v, if $\mathbf{1} = (1,1)$ is a Nash equilibrium for any pair $(s,t) \in [0,K] \times [0,K]$ of skills of the two agents in the associated game.*

(Note that our assumptions on Π^{+} imply that, if $|N| = 2$ and $\pi \in \Pi^{+}$, then there exists a prize function p such that $\pi^{n}(\tau^{1}, \tau^{2}) = p(\tau^{n}/(\tau^{1} + \tau^{2}))$, for $n \in N$, whenever $\tau^{1} + \tau^{2} \neq 0$,justifying our name for p). The lemma below will be handy:

Lemma 35 *The prize function p is effective at level v iff for all $s \in [0,K]$ we have*

$$p(q(s)) - \delta/v \geq p(R(s)) \geq p(r(s)) + \delta/v$$

Proof. As discussed earlier, $p(x)$ is effective iff $p\left(\tilde{R}(s)\right) \geq p(\tilde{r}(s)) + \delta/v$ and $p(R(s)) \geq p(r(s)) + \delta/v$ for all $s \in [0,K]$ Since $p\left(\tilde{R}(s)\right) = 1 - p(R(s))$, $p(\tilde{r}(s)) = 1 - p(q(s))$, the first inequality becomes $p(q(s)) - \delta/v \geq p(R(s))$ which proves the result. ■

Define a sequence of points $0 = x_0, x_1, \ldots, x_l$ in $[0, 1/2]$ by $x_i = R(0)$ for $i = 1$;and $x_i = \rho(x_{i-1})$ for $1 < i \leq l$.where $\rho(x) = \min\left(R\left(r^{-1}(x)\right), q\left(R^{-1}(x)\right)\right)$ and l is the smallest index i for which $r^{-1}(x_i)$ is undefined. Note that since q, R, r are all strictly increasing functions, so is ρ, and therefore $x_1, \ldots, x_l$ is an increasing sequence.

Now define $p^{*} : [0,1] \to [0,1]$ as follows (where $i = 0, 1, ..., l$):

$$p^{*}(x) = \begin{cases} i/2l & \text{for} \quad x_i \leq x < x_{i+1} \\ 1/2 & \text{for} \quad x_l \leq x \leq 1/2 \\ 1 - p^{*}(1-x) & \text{for} \quad 1/2 < x \leq 1 \end{cases}$$

We are now ready to state and prove

Theorem 36

1. (i) Any effective scheme has prize level $\geq 2l\delta$; (ii) $x \to p^{*}(x)\delta$ is an effective scheme with prize $2l\delta$.

Proof. Let p be effective with prize level v. By Lemma 35 with $s = 0$, we get $p(x_1) = p(R(0)) \geq p(r(0)) + \delta/v \geq \delta/v$. Next let $s = r^{-1}(x)$ or $s = R^{-1}(x)$ according as $\rho(x) = R(r^{-1}(x))$ or $q(R^{-1}(x))$. Then, again by Lemma 35, we get $p(\rho(x)) \geq p(x) + \delta/v$ whenever $x, \rho(x) \in [0,1]$. Applying this formula repeatedly we get

$$1/2 = p(x_l) \geq p(x_{l-1}) + \delta/v \geq \cdots \geq p(x_1) + (l-1)\,\delta/v \geq l\delta/v$$

which proves (i). For (ii) we first show that, for any s, each of the two intervals $[r(s), R(s)]$ and $[R(s), q(s)]$ contains some "jump" point x_i. Indeed if $x = r(s)$ is in $[x_{i-1}, x_i)$, then $R(s) = R(r^{-1}(x)) \geq \rho(x) > \rho(x_{i-1}) = x_i$, hence $x_i \in [r(s), R(s)]$. The argument is similar for $[R(s), q(s)]$. Now by the definition of p^* it follows that

$$p^*(q(s)) - 1/2l \geq p^*(R(s)) \geq p^*(r(s)) + 1/2l,$$

which is precisely the condition of Lemma 35 with $v = 2l\delta$. ■

Remark 37 *Theorem 36 shows that the graph of the optimal prize scheme is a monotonically increasing step function with l jump points. Note that the graph of π_D has one jump point at $1/2$ where the jump is from 0 to 1, while the graph of π_P may be said to have a continuum of jump points of equal jump sizes since it is a straight line with slope 1. In this graphical sense, the optimal schedule is "in between" π_D and π_P.*

One might define "increasing returns" as in Axiom 32 , substituting " $s' > s$ " in place of "$s' < s$

Axiom 38 ***(Increasing Returns to Skill)*** *Both $\tau : [k,K] \to R_+$, $\tau^* : [k,K] \to R_+$ are continuous and strictly monotonic; and $\tau^*(s)/\tau(s) \leq \tau^*(s')/\tau(s')$ if $s' > s$. Also* $\inf\{\tau^*(s) - \tau(s) : s \in [k,K]\} > 0$.

With Axiom 38 in place of Axiom 32, the natural variant of Lemma 33 holds, substituting k for K.

Lemma 39 *Assume Axiom 38 holds. Let $s \in (k,K)$ and $t \in (k,K)$, Then there exist $s' \in [k,K]$ and $t' \in [k,K]$ such that*

$$\frac{\tau^*(s')}{\tau^*(s') + \tau^*(t')} - \frac{\tau(s')}{\tau(s') + \tau^*(t')} \leq \frac{\tau^*(s)}{\tau^*(s) + \tau^*(t)} - \frac{\tau(s)}{\tau(s) + \tau^*(t)} \text{ and } \frac{\tau(s')}{\tau(s') + \tau^*(t')} = \frac{\tau(s)}{\tau(s) + \tau^*(t)}$$

and either $s' = k$ or $t' = k$.

(The proof of this is the same as the proof of Lemma 33 in the Appendix, with $s-\kappa, t-\kappa, k, s' < s$ in place of $s+\kappa, t+\kappa, K, s' > s$ respectively.)

Thus the whole analysis for optimal prizes can be replicated for this dual case, focusing on the southwest boundary of the square $[k, K]^2$, in place of the northeast boundary.

7 Optimal Prizes with Incomplete Information

In principle, optimal prizes can be defined in an analogous manner when there is incomplete information, *i.e.*, when each agent is informed precisely of his own characteristics but has limited information regarding those of his rivals. Indeed, in our context, he does not need to even know the probability distribution of others' characteristics but only the *support* of that distribution. Finer information would be simply irrelevant since we want to incentivize him to work hard for *every* realization of his rivals' characteristics.

By way of an illustrative example, recall the binary contest of section and suppose there is incomplete information in the above sense. It is evident that no agent will work if $v < 1$, no matter which prize function we select (recall that v is the value of the prize and that the disutility of working is 1).

Consider π^* which is defined as follows:

$$\pi^*(x) = \begin{cases} 0 & \text{for} \quad x = 0 \\ 1/2 & \text{for} \quad 0 < x < 1 \\ 1 & \text{for} \quad x = 1 \end{cases}$$

It is equally evident that, when $v \geq 2$, all agents will work[17] under π^*, except for the agent with skill 0; and that no other prize function can elicit more effort than π^*, so that π^* is optimal. In fact π^* is also optimal for $1 \leq v < 2$, as was pointed out by Mayank Goswami and Biligbaatar Tumendemberel at Stony Brook. We quote their argument. First observe that, for any $\pi \in \Pi^+$, it is obvious that exists a *smallest* threshold c_π such that all agents whose skill is *strictly* below c_π will shirk; and, from the fact that the agent with skill c_π is indifferent between shirking and working, we must have:

$$c_\pi + \int_{c_\pi}^{1} \pi(\frac{c_\pi}{c_\pi + x})dx = \frac{1}{v}$$

[17] We assume throughout that if an agent is indifferent between shirking and working, he will work (the principal paying him a "sliver" more in the background). This takes care of the case $v = 2$.

Notice that, for any $\pi \in \Pi^+$ and $x > c_\pi$, we have $c_\pi/(c_\pi + x) < 1/2$, and hence (recalling that π is weakly monotonic and $\pi(1/2) = 1/2$ for any $\pi \in \Pi^+$)

$$\pi\left(\frac{c_\pi}{c_\pi + x}\right) \leq \pi\left(\frac{1}{2}\right) = \frac{1}{2}$$

which implies

$$\frac{1}{v} = c_\pi + \int_{c_\pi}^{1} \pi(\frac{c_\pi}{c_\pi + x})dx \leq c_\pi + \int_{c_\pi}^{1} \frac{1}{2}dx = \frac{1}{2}(1 + c_\pi)$$

i.e.,

$$c_\pi \geq \frac{2}{v} - 1$$

But a simple calculation reveals that, when $1 < v < 2$, the unique NE under π^* requires that every agent whose skill is $(2/v) - 1$ or more must work (recall our convention that, if indifferent between shirking and working, an agent works) while the rest must shirk. This shows that π^* elicits as much effort as any $\pi \in \Pi^+$ even when $1 \leq v < 2$.

The discussion of optimal π became so simple for this binary example precisely because we could infer, from observing a positive output, that an agent must have worked. (Recall that the output of shirk was taken to be 0.) This enabled us to reward him *precisely* when he worked, via π^*.Were we to more generally postulate (as was done in section 9, for games of complete information) that the output of shirk, work is given by two functions τ, τ^* (that map skills to positive outputs) with $\tau < \tau^*$, the situation would become much more intriguing. For now it would no longer be possible to tell from an output whether it has been produced by a low-skill agent who is working or a high-skill agent who is shirking (when the intersection of the ranges of τ, τ^* is nonempty). Computation (hence comparison) of NE under π_D and π_P, in this scenario, as well as the computation of optimal π, are both interesting problems, which we leave for future research.

Appendix

This section contains proofs that were postponed.

7.1 Theorem 16

Recall, from Notation 14, that $\Delta(n, e; \boldsymbol{\chi},\pi, t^{-n})$ is the increment in n's probability of winning the prize in the game $\Gamma\left(\boldsymbol{\chi},\pi\right)$ when he deviates from e to 1, while his rivals' output is fixed at the (non-negative, $(N-1)$- dimensional) deterministic vector $t^{-n} = \left(t^k\right)_{k\in N\setminus\{n\}}$.

Suppose π is monotonic (see Definition 8), and $\tau^n(1) \succsim_{st} \tau^n(e)$. Then it is evident that

$$\Delta(n, e; \boldsymbol{\chi}, \pi, t^{-n}) \geq 0 \text{ for all } t^{-n}. \tag{3}$$

This inequality can be given a sharper, quantitative form if $t^{-n} \neq 0$.

Lemma 40 *Suppose Axioms 5 and 6 hold and $\pi \in \Pi_{rw}$. Let $\varepsilon > 0$. Then, for all $n \in N$, $e \in E \setminus \{1\}$, $\boldsymbol{\chi} \in X^N$ and t^{-n}, there exists $\alpha > 0$ such that*

$$\sum_{k\in N\setminus\{n\}} t^k \geq \varepsilon \Longrightarrow \Delta\left(n, e; \boldsymbol{\chi}, \pi, t^{-n}\right) > \alpha.$$

Proof. Let $X = \tau^n(1)$, $Y = \tau^n(e)$ denote the random output of n when he exerts effort $1, e$ respectively, where $e \in E \setminus \{1\}$. By Axiom 6, $X \succsim_{st} Y$. Then, by Theorem 1.A.2 of [25], there exists a random variable Z and functions ψ_2 and ψ_1 such that $\psi_2(z) \geq \psi_1(z)$ for all real numbers z , and

$$X \sim \psi_2(Z) \text{ and } Y \sim \psi_1(Z)$$

(where $\sim$ means equality in distribution). Denote

$$p = \Pr\left\{\psi_2(Z) - \psi_1(Z) > \kappa/2\right\}$$

Then, recalling from Axiom 6 that $\mu^n(e) = \operatorname{Exp}\left(\tau^n(e)\right)$ goes up by at least κ when n deviates from $e \in E \setminus \{1\}$ to 1 and from Axiom 5 that the non-negative random variables $\tau^n(1)$ and $\tau^n(e)$ are both bounded above by β, we have

$$\begin{aligned}
\kappa &\leq \operatorname{Exp}(X) - \operatorname{Exp}(Y) = \operatorname{Exp}(\psi_2(Z)) - \operatorname{Exp}(\psi_1(Z)) \\
&= \operatorname{Exp}(\psi_2(Z) - \psi_1(Z)) \leq (1-p)\frac{\kappa}{2} + p\beta \leq \frac{\kappa}{2} + p\beta
\end{aligned}$$

Hence

$$p \geq \frac{\kappa}{2\beta}$$

Now *fix* the outputs t^{-n} of the rivals of n as stipulated, i.e., with the proviso that the total produced by them is at least ε. Then, again invoking the bound β from Axiom 5, this total is in the interval $[\varepsilon, (N-1)\beta]$. Denote

$$f(t^n) = \pi^n(t^{-n}, t^n)$$

By the responsiveness of π, there exists $\alpha' > 0$ such that

$$f(\psi_2(Z)) - f(\psi_1(Z)) \geq \alpha' \text{ with probability } p;$$

and, moreover, α' does not depend on t^{-n} but only on κ, ε and $(N-1)\beta$ (see Definition 9). Furthermore, since (by the monotonicity of π, see Definition 8) f is non-decreasing and (recall) $\psi_2(Z) \geq \psi_1(Z)$ with probability 1, we have

$$f(\psi_2(Z)) - f(\psi_1(Z)) \geq 0 \text{ with probability } 1. \tag{4}$$

The four previous displays imply

$$\Delta(n, e; \boldsymbol{\chi}, \pi, t^{-n}) = \text{Exp}\left(f(\psi_2(Z)) - f(\psi_1(Z))\right) \geq (\frac{\kappa}{2B})\alpha'$$

for every $t^{-n} = \left(t^k\right)_{k \in N \setminus \{n\}}$ in which the total output is at least ε. Take

$$\alpha = (\frac{\kappa}{2\beta})\alpha'$$

to get the desired conclusion. ■

Corollary 41 *Suppose (i) the total output Z of the rivals of n is given by a non-negative random variable Z which has positive expectation $\eta > 0$ and which is bounded above by B; (ii) $\pi \in \Pi_{rw}$; (iii) Axiom 6 holds. Then the increment in n's probability of winning the prize, when he deviates from $e \in E \setminus \{1\}$ to 1 (holding fixed the random output of the others) goes up by a positive constant α^* that is independent of $\boldsymbol{\chi} \in X^N$ or $n \in N$, i.e.,*

$$\textit{Exp } \Delta(n, e; \boldsymbol{\chi}, \pi, Z) \geq \alpha^*$$

for all $n \in N, e \in E \setminus \{1\}, \boldsymbol{\chi} \in X^N$.

Proof. Denote

$$q = \Pr\left\{Z > \frac{\eta}{2}\right\}.$$

Then

$$\eta \leq (1-q)\frac{\eta}{2} + qB$$

which implies

$$q \geq \frac{\eta}{2B - \eta} > 0$$

By Lemma 40 (taking $\varepsilon = \eta/2$) there exists $\alpha > 0$ such that, for any $e \in E \setminus \{1\}$ and $\boldsymbol{\chi} \in X^N$,

$$\text{Exp } \Delta(n, e; \boldsymbol{\chi}, \pi, Z) \geq \left(\frac{\eta}{2B - \eta}\right)\alpha$$

establishing the corollary. ■

Notation 42 *We set*

$$\Delta^*(n, e; \boldsymbol{\chi}, \pi) = \Pr\ (n, 1; \boldsymbol{\chi}, \pi) - \Pr\ (n, e; \boldsymbol{\chi}, \pi).$$

In other words, recalling Notation 21, $\Delta^*(n, e; \boldsymbol{\chi}, \pi)$ represents the increment in n's probability of winning the prize in the game $\Gamma(\boldsymbol{\chi}, \pi)$, when he unilaterally deviates from effort e to 1, while the effort levels of his rivals are all fixed at 1.

Corollary 43 *Suppose Axioms 5 and 6 hold and $\pi \in \Pi_{rw}$. Then*

$$\inf\left\{\Delta^*(n, e; \boldsymbol{\chi}, \pi) : e \in E \setminus \{1\}, \boldsymbol{\chi} \in X^N\right\} > 0$$

Proof. Denote the total output of the rivals by Z and denote $\eta = \text{Exp } Z$. By Axiom 5,

$$\eta \geq (N-1)d > 0 \text{ and } Z \leq (N-1)\beta.$$

The conclusion immediately follows from Corollary 41. ■

Lemma 44 *Suppose Axioms 5 and 6 hold and $\pi \in \Pi_{rm}$. Then there exists $\alpha^* > 0$ such that, in any game $\Gamma(\boldsymbol{\chi}, \boldsymbol{\pi})$ for $\boldsymbol{\chi} \in X^N$ and for any agent $n \in N$, we have (for an arbitrary choice of strategies by the rivals of n):*

1. *if none of the rivals of n choose effort 0 with positive probability, then the increment in n's probability of winning the prize goes up by at least α^* when he switches from effort $e \in E \setminus \{1\}$ to 1;*

2. *n wins the prize with probability at least α^* by choosing effort 1 (regardless of the strategies of his rivals).*

Proof. First suppose that every rival $k \in N \setminus \{n\}$ chooses effort 0 with probability 0. Let $e^* = \min\{e : e \in E \setminus \{0\}\}$ be the smallest effort level above 0. Since every agent exerts effort e^* or more, Axiom 5 implies that the rivals of any agent n produce at least $(N-1)de^*$ by way of expected total output. Then the desired conclusion (i) follows at once from Corollary 41.

Next let the strategy σ^k be arbitrary for $k \in N \setminus \{n\}$. Imagine the scenario in which every rival $k \in N \setminus \{n\}$ of n shifts from playing 0 to playing 1, i.e., replaces his strategy σ^k by ρ^k where

$$\rho^k(e) = \begin{cases} 0 & \text{for} \quad e = 0 \\ \sigma^k(e) & \text{for} \quad e \in E \setminus \{1, 0\} \\ \sigma^k(1) + \sigma^k(0) & \text{for} \quad e = 1 \end{cases}$$

In this imaginary "ρ-scenario", agent n gets the prize with probability α^* when he chooses effort 1, as we just saw. But note that in the real "σ-scenario", when all his rivals $k \in N \setminus \{n\}$ are playing σ^k instead of ρ^k, the rivals' outputs are *reduced* compared to the imaginary scenario. By the monotonicity of π, it follows that agent n wins the prize with even higher probability $\alpha^{**} \geq \alpha^*$ via effort 1. ■

7.1.1 Part (i) (of Theorem 16)

Proof. By Axiom 5,

$$\delta^n(1) - \delta^n(e) < \delta^n(1) < C \text{ at all } \boldsymbol{\chi} \in X^N.$$

Let I^* denote the infimum in the conclusion of Corollary 43. Then it suffices to choose v_* so that

$$I^* v_* \geq C.$$

■

7.1.2 Part (ii)

Proof. Suppose a agent n unilaterally deviates from $e \in E \setminus \{1\}$ to 1 when the total output due to the others is at least $\varepsilon > 0$. By Lemma 40 and Axiom 5, there exists $\alpha' > 0$ such that his gain in payoff is at least

$$\alpha' v^n - C.$$

So if we take

$$v^{**}(\varepsilon) > \frac{C}{\alpha'}$$

it will establish (ii). ■

7.1.3 Part (iii)

Proof. The NE of $\Gamma(\boldsymbol{\chi}, \boldsymbol{\pi})$ can be of two types:

Type I: each agent chooses effort 0 with probability 0;

Type II: some agent chooses effort 0 with probability strictly more than 0.

We shall show that there exists v^* such that, if $v(\boldsymbol{\chi}) \geq v^*$, then NE of type II cannot occur, whereas all NE of type I are $\mathbf{1}$, thus establishing part (iii).

First consider an NE $\sigma = (\sigma^k)_{k \in N}$ of type I. If $\sigma^n(1) < 1$ for some agent n, then $\sigma^n(e) > 0$ for some $e \in E \setminus \{1\}$. Let n unilaterally deviate by shifting probability $\sigma^n(e)$ from e onto 1. By Lemma 44 the increment in n's probability of winning the prize is at least α^*. On the other hand, by Axiom 5, the extra disutility incurred in the deviation is at most C, so his net gain in payoff is at least $\sigma^n(e)\,(\alpha^* v^* - C)$. Thus if we take

$$v^* > \frac{C}{\alpha^*},$$

the deviation will be profitable, a contradiction. We conclude that all NE of type I are **1**.

Next suppose $\sigma = (\sigma^n)_{n\in N}$ is an NE of type II, with some agent n choosing 0 with probability $\sigma^n(0) > 0$. Let agent n shift probability $\sigma^n(0)$ from 0 onto 1. Again by 44, he wins the prize with probability at least α^* when he chooses 1. On the other hand, when he choose effort 0 he wins the prize with probability 0 (by definition). The cost of effort 1 is at most C. Thus the benefit to n from his deviation is again $\sigma^n(e)\,(\alpha^* v^* - C)$, which is positive on account of the previous display. This shows that NE of type II cannot occur. ■

7.2 Theorem 24

Proof. It is evident that

$$\widetilde{\pi} \succeq \pi \iff \Omega(\pi) \subset \Omega(\widetilde{\pi}).$$

Since $\subset$ is a partial order[18] on subsets of Ω, we conclude that $\succeq$ is a preorder on Ω. ■

7.3 Theorem 26

Proof. First note that, by Corollary 43, $\Delta^*(n, e; \boldsymbol{\chi},\pi) = \Pr(n, 1; \boldsymbol{\chi},\pi) - \Pr(n, e; \boldsymbol{\chi},\pi) > 0$ for $e \in E \setminus \{1\}$. Thus, by Axiom 25, the set

$$\Omega^n(\boldsymbol{\chi},\pi) = \{\omega \in \Omega : [\Delta^*(n, e; \boldsymbol{\chi},\pi)]\, v^n(\omega) \geq \delta^n(1) - \delta^n(e) \text{ for all } e \in E \setminus \{1\}\}$$

is "*upper-monotone*" in the following sense: if $y \in \Omega^n(\boldsymbol{\chi}, \pi)$ and $x \succ y$, then $x \in \Omega^n(\boldsymbol{\chi}, \pi)$. The property of being upper-monotone is preserved under intersections, hence $\Omega(\pi) = \cap\left\{\Omega^n(\boldsymbol{\chi},\pi) : \boldsymbol{\chi} \in X^N, n \in N\right\}$ is upper-monotone for all $\pi \in \Pi$. As already observed in the proof of 24: $\widetilde{\pi} \succeq \pi$ if, and only if, $\Omega(\pi) \subset \Omega(\widetilde{\pi})$. But the relation $\subset$ clearly constitutes a total order on the collection of upper-monotone subsets of the totally ordered set Ω, hence $\succeq$ is a total preorder. ■

[18] i.e. a binary relation that is reflexive, antisymmetric and transitive.

7.4 Theorem 30

Proof. Recall, from Notation 31, that $\mathbf{Z}$ is the projection of $\mathbf{X}$ onto the skill space. Further recall, from Notation 42, that $\Delta^*(n, e; \boldsymbol{\chi},\pi)$ depends only on the skill $\boldsymbol{\zeta} = \boldsymbol{\zeta}(\boldsymbol{\chi}) \in \mathbf{Z}$ that underlies $\boldsymbol{\chi}$, so we may (more properly) write

$$\Delta^*(n, e; \boldsymbol{\chi},\pi) = \Delta^*(n, e; \boldsymbol{\zeta},\pi).$$

For any $\pi \in \Pi^*$, $n \in N$, $\boldsymbol{\zeta} = (\delta^n, \tau^n)_{n\in N} \in \mathbf{Z}$ and $u \in V^n(\boldsymbol{\zeta})$, define

$$z^n(\boldsymbol{\zeta};\pi, u) = \inf \{z \in \mathbf{R}_+ : [\Delta^*(n, e; \boldsymbol{\zeta},\pi)]\, u\,(z) \geq \delta^n(1) - \delta^n(e) \text{ for all } e \in E \setminus \{1\}\}$$

and

$$z(\pi) = \sup \{z^n(\boldsymbol{\zeta};\pi, u) : u \in V^n(\boldsymbol{\zeta}), \boldsymbol{\zeta} \in \mathbf{Z}, n \in N\}.$$

By Axiom 28, 1 is a best-reply of n precisely when the size of the prize is at least $z^n(\boldsymbol{\zeta};\pi, u)$, assuming his valuation is u, his skill is in accordance with $\boldsymbol{\zeta}$, and — as usual — all his rivals have chosen 1.

A little reflection reveals

$$z(\widetilde{\pi}) \leq z(\pi) \iff \Omega(\widetilde{\pi}) \subset \Omega(\pi)$$

which implies

$$z(\widetilde{\pi}) \leq z(\pi) \iff \widetilde{\pi} \succeq \pi.$$

Now let $\pi^* \in \Pi^*$ be a weakly monotonic and responsive scheme (by assumption, one such exists in Π^*). Take any agent k and any $\boldsymbol{\chi} \in \mathbf{X}$. Consider the situation where all the rivals of k are exerting maximal effort 1. If k unilaterally deviates from $e \in E \setminus \{1\}$ to 1, Corollary 43 implies that his probability of winning the prize under π^* will go up by at least some constant $\alpha > 0$ via this deviation; and his consequent increase in payoff, if the prize is of size z, will go up by at least $\alpha v^k(z) \to \infty$ as $z \to \infty$ (on account of Axiom 28). In contrast (by Axiom 5) the extra disutility of effort incurred by k from his deviation is bounded above by C. It follows that $z(\pi^*) < \infty$. Hence

$$y = \inf \{z\,(\pi) : \pi \in \Pi^*\} < \infty.$$

Thus there exist π in Π^* for which $z(\pi)$ is arbitrarily close to y, and thus ε-optimal for arbitrary $\varepsilon > 0$. In the case that Π^* is finite, the inf in the above display is achieved, hence an exact optimal exists. ■

7.5 Lemma 33

Proof. Since τ^* and τ are strictly monotonic and continuous, there exist $\kappa > 0$ and $\kappa' > 0$ such that $s' \equiv s + \kappa \in [k, K]$,and $t' \equiv t + \kappa' \in [k, K]$ and

$$\frac{\tau(s')}{\tau(s') + \tau^*(t')} = \frac{\tau(s)}{\tau(s) + \tau^*(t)} \tag{5}$$

Hence there exists a *maximal* pair κ, κ' satisfying (5), and then either $s' = K$ or $t' = K$ (otherwise both κ and κ' could be increased slightly, still maintaining (5), and contradicting the maximality of κ, κ').

In view of (5), to prove (b) it suffices to show that

$$\frac{\tau^*(s')}{\tau^*(s') + \tau^*(t')} \leq \frac{\tau^*(s)}{\tau^*(s) + \tau^*(t)} \tag{6}$$

which is equivalent to

$$\frac{\tau^*(t')}{\tau^*(s')} \geq \frac{\tau^*(t)}{\tau^*(s)} \tag{7}$$

as can be seen by dividing the numerator and the denominator of the LHS and RHS of (6) by $\tau^*(s'^*)$ and $\tau^*(s)$ respectively.

But a similar maneuver shows that (5) is equivalent to

$$\frac{\tau^*(t')}{\tau(s')} = \frac{\tau^*(t)}{\tau(s)} \tag{8}$$

And, since $s' > s$, decreasing returns (Assumption AIV) imply

$$\frac{\tau^*(s')}{\tau^*(s)} \leq \frac{\tau(s')}{\tau(s)} \tag{9}$$

From (8) and (9), we get

$$\frac{\tau^*(s')}{\tau^*(s)} \leq \frac{\tau(s')}{\tau(s)} = \frac{\tau^*(t')}{\tau(t)} \tag{10}$$

establishing (7), and thereby (6) ■

7.6 Lemma 35

Proof. As discussed earlier, $p(x)$ is effective iff $p\left(\tilde{R}(s)\right) \geq p(\tilde{r}(s)) + \delta/v$ and $p(R(s)) \geq p(r(s)) + \delta/v$ for all $s \in [0, K]$ Since $p\left(\tilde{R}(s)\right) = 1 - p(R(s))$, $p(\tilde{r}(s)) = 1 - p(q(s))$, the first inequality becomes $p(q(s)) - \delta/v \geq p(R(s))$ which proves the result. ■

7.7 Theorem 36

Proof. Let p be effective with prize level v. By Lemma 35 with $s = 0$, we get $p(x_1) = p(R(0)) \geq p(r(0)) + \delta/v \geq \delta/v$. Next let $s = r^{-1}(x)$ or $s = R^{-1}(x)$ according as $\rho(x) = R(r^{-1}(x))$ or $q(R^{-1}(x))$. Then, again by Lemma 35, we get $p(\rho(x)) \geq p(x) + \delta/v$ whenever $x, \rho(x) \in [0,1]$. Applying this formula repeatedly we get

$$1/2 = p(x_l) \geq p(x_{l-1}) + \delta/v \geq \cdots \geq p(x_1) + (l-1)\delta/v \geq l\delta/v$$

which proves (i). For (ii) we first show that, for any s, each of the two intervals $[r(s), R(s)]$ and $[R(s), q(s)]$ contains some "jump" point x_i. Indeed if $x = r(s)$ is in $[x_{i-1}, x_i)$, then $R(s) = R(r^{-1}(x)) \geq \rho(x) > \rho(x_{i-1}) = x_i$, hence $x_i \in [r(s), R(s)]$. The argument is similar for $[R(s), q(s)]$. Now by the definition of p^* it follows that

$$p^*(q(s)) - 1/2l \geq p^*(R(s)) \geq p^*(r(s)) + 1/2l,$$

which is precisely the condition of Lemma 35 with $v = 2l\delta$. ■

8 Second Appendix (Work in Progress)

8.0.1 Contests between Experts: the Case of Small Fractional Increments

There are many contests where all the agents are so strong — think of experts, champions, stars — that their output levels are already very high at low effort. In this scenario, the exertion of extra effort causes a negligible *fractional* increase in output. However the increases can still be large on the absolute scale, enabling us to observe the enhanced performance due to extra effort (e.g., $0.1\% \times 10^6$ is ten times greater than $10\% \times 10^3$).

The ε-model We retain for simplicity the deterministic binary scenario of the previous section. We assume that an agent of skill t produces t units of output if he shirks; and $t + \psi(t)\varepsilon$ units if he works, where $\psi(t)$ and ε are both positive. As for skills, we suppose they lie in the compact interval $t \in [k, K]$ for $0 < k < K$. We also make the natural assumption that productivity increases with skill:

Axiom 45 *$\psi(t)$ is a continuous and strictly increasing function of t.*

Both ε and k/K are understood to be very small positive numbers. The closer these are to 0, the more accurately will the "continuum model" below portray the contest of the ε-model. We could think of k as fixed and suppose $K \to \infty$ and $\varepsilon \to 0$ in the asymptotic approximation of the continuum model. (Alternatively, for fixed K, we could suppose $k \to 0$ and $\varepsilon/k \to 0$.)

Notation 46 *Let $\alpha = \alpha(t,x)$ and $\beta = \beta(t,x)$ denote the fractions of total output produced by an agent of skill t when he works and shirks respectively, and his rival is of skill x and working; i.e.,*

$$\alpha = \frac{t + \psi(t)\varepsilon}{t + \psi(t)\varepsilon + x + \psi(x)\varepsilon}; \text{ and } \beta = \frac{t}{t + x + \psi(x)\varepsilon}$$

Recall the definition of Π^+ from the beginning of previous section 6.4.We shall, from now on, also denote by π the prize functon associated with any scheme in $\pi \in \Pi^+$.(There should be no confusion; the meaning will always be clear from the context.). Thus any prize function $\pi \in \Pi^+$ is weakly monotonic on $[0,1]$, with $\pi(0) = 0$, $\pi(1) = 1$ and $\pi(x) = 1 - \pi(1-x)$ for any $0 < x < 1$.

Our focus will be on the subclass of Π^+ consisting of differentiable prize functions.

Notation 47 *Let*

$$\Pi^* = \{\pi \in \Pi^* : \pi \text{ is differentiable}\}$$

Given a prize function $\pi \in \Pi^*$, define $I(\pi, t, x)$, the t-agent's **incentive to work** by

$$I(\pi, t, x) \equiv \pi(\alpha(t,x)) - \pi(\beta(t,x))$$

The minimum fraction is $b_* = k/(k + K + \psi(K)\varepsilon)$ while the maximum fraction is $b^* = (K + \psi(K)\varepsilon)/(k + K + \psi(K)\varepsilon)$. Thus, in our context, we assume $\pi : [b_*, b^*] \mapsto [0,1]$, with $\pi(b_*) = 0$, $\pi(b^*) = 1$ since π is continuous and since (on account of k/K being very small) $b_* \approx 0$ and $b^* \approx 1$.

For any $\pi \in \Pi^*$, the minimum prize that will incentivize agents to work at all realizations $(t,x) \in [k,K] \times [k,K] \equiv [k,K]^2$, is given by

$$V(\pi) = \frac{d}{m(\pi)}$$

where d is the disutility of switching from shirk to work and

$$m(\pi) = \min \left\{ I(\pi, t, x) : (t,x) \in [k,K]^2 \right\}$$

Thus to minimize $V(\pi)$ on Π^* we must maximize the minimum incentive $m(\pi)$ on Π^*.

Fix $\varepsilon > 0$. First observe that, for small enough ε,

$$\alpha - \beta \simeq \left(\frac{d}{du} \left(\frac{u}{u+x} \right) \Big|_{u=t} \right) (\Delta u) = \frac{x}{(t+x)^2} (\psi(t)\varepsilon) = \frac{x\psi(t)}{(t+x)^2}\varepsilon$$

So, if $\pi \in \Pi^*$,

$$I(\pi, t, x) \equiv \pi(\alpha) - \pi(\beta) \simeq \pi'(\beta) \frac{x\psi(t)}{(t+x)^2}\varepsilon \tag{11}$$

The Continuum Model

Definition 48 *π is **continuum-optimal** in Π^* if*

$$\min \left\{ \pi'(\beta(t,x)) \frac{x\psi(t)}{(t+x)^2} : (t,x) \in [k,K]^2 \right\} \geq \min \left\{ \widehat{\pi}'(\beta(t,x)) \frac{x\psi(t)}{(t+x)^2} : (t,x) \in [k,K]^2 \right\}$$

for all $\widehat{\pi} \in \Pi^*$.

Since ε is a constant in $I(\pi, t, x)$, the above definition implies that, given the approximation of (11), π achieves $\min \left\{ I(\widehat{\pi}, t, x) : (t,x) \in [k,K]^2 \right\}$ over $\widehat{\pi} \in \Pi^*$.

Although we have not formally verified this, intuition suggests that: if V^ε denotes the minimum prize required in Π^* to incentivize work (in the "ε-model" wherein the work output of the t-agent is given by $t + \psi(t)\varepsilon$), and if $V(\pi)$ denotes the corresponding quantity for a continuum-optimal π in Π^*, then $V^\varepsilon / V(\pi)$ converges to 1 as ε goes to 0. In this sense, a π that is (idealistically) continuum-optimal in Π^* is (realistically) nearly optimal in $\Pi^*$for small ε. This motivates Theorem 49 below. First recall

Strictly decreasing (increasing) returns to skill:

$\dfrac{t+\psi(t)}{t}$ is strictly decreasing (increasing) in t, *i.e.*, $\dfrac{\psi(t)}{t}$is strictly decreasing (increasing) in t

Theorem 49 *Assume that ψ has strictly decreasing returns to skill. There is a unique π (that does not depend on ψ) that is continuum-optimal in Π^*; and it is given by:*

$$\pi(x) = \frac{1}{2} + B \ln \frac{x}{1-x}$$

where $1/2 \leq x \leq K/(k+K)$ (the rest of π being determined by reflection around 1/2: $\pi(x) = \pi(1-x)$) and the constant B chosen to satisfy $\pi(K/(k+K)) = 1$. In the case of strictly increasing returns, an entirely analogous result holds with $1/2 \geq x \geq k/(k+K)$ in place of $1/2 \leq x \leq K/(k+K)$, and $\pi(k/(k+K)) = 0$ in place of $\pi(K/(k+K)) = 1$.

Proof. First we focus on decreasing returns. Then, by Lemma 33, we need only consider the two cases below.

Case A. Agent is at t and the rival at K. Then consider[19]

$$I(\pi, t, K) = \pi' \left(\frac{t}{t+K} \right) \frac{K\psi(t)}{(t+K)^2}.$$

Case B. Agent is at K and the rival at t. Then consider

$$I(\pi, K, t) = \pi' \left(\frac{K}{t+K} \right) \frac{t\psi(K)}{(t+K)^2}.$$

Since $\pi(x) = 1 - \pi(1-x)$ for all x, we get

$$\pi' \left(\frac{t}{t+K} \right) = \pi' \left(\frac{K}{t+K} \right)$$

which, in conjunction with $K\psi(t) > t\psi(K)$ (decreasing returns), implies $I(\pi, K, t) < I(\pi, t, K)$ for all $t \in [k, K]$. Thus it suffices to incentivize the t-agent to switch from shirk to work in Case B (for all $t \in [k, K]$). Since we want to maximize the minimum incentive, we must arrange for $I(K, t) = \sigma$, for some constant σ, and for all $t \in [k, K]$ .To see this, denote

$$G(t) = \frac{t\psi(K)}{(t+K)^2}$$

and let π be a solution to the differential equation :

$$\pi'(t/(t+K)) = \sigma/G(t) \text{ for all } t \in [k, K]$$

with (recalling $k/(k+K) \approx 0$) the boundary conditions

$$\pi \left(\frac{k}{k+K} \right) = 0 \text{ and } \pi \left(\frac{1}{2} \right) = \left(\frac{1}{2} \right)$$

Suppose there is a scheme $\widetilde{\pi}$ which does *not* satisfy the differential equation. If $\widetilde{\pi}'(t_1/(t_1+K)) > \pi(t_1/(t_1+K)) = \sigma/G(t_1)$ for some $t_1 \in [k, K]$, then since[20] $\int \pi'(y)dy = \int \widetilde{\pi}'(y)dy = 1/2$, we see at once that there exists $t_2 \in [k, K]$ such that

[19] We suppress ε in the term $I(\pi, t, K)$ since it is constant across $t \in [k, K]$.

[20] Here we have denoted $y = t/(t+K)$. The range of integration is from $y = k/(k+K) \approx 0$ to $y = 1/2$ or equivalently, recalling that $\pi(x) = \pi(1-x)$, from $y = 1/2$ to $y = K/(k+K) \approx 1$.

$\widetilde{\pi}'\left(t_2/\left(t_2+K\right)\right) < \pi'\left(t_2/\left(t_2+K\right)\right) = \sigma/G(t_2)$. But then the incentive (to work) at t_2 under $\widetilde{\pi}$, which is given by $\widetilde{\pi}'(t_2/(t_2+K))G(t_2)$, is strictly less than σ, which is the constant incentive under π at all $t \in [k, K]$. We conclude that the *minimum* incentive to work under $\widetilde{\pi}$ is less than that under π, establishing the superiority of π over $\widetilde{\pi}$. So an optimal scheme must satisfy the following differential equation (where $\widetilde{C}$ is another constant):

$$\pi'\left(\frac{K}{t+K}\right) = \widetilde{C}\frac{(t+K)^2}{t\psi(K)}, \textit{ i.e.}, \pi'\left(\frac{K}{t+K}\right) = \frac{\widetilde{C}}{\psi(K)}\left(\frac{t+K}{t}\right)^2 t.$$

For $x > 1/2$, let $x = K/(t+K)$, so $1-x = t/(t+K)$ and $t = K(1-x)/x$,enabling us to rewrite our differential equation:

$$\pi'(x) = \frac{\widetilde{C}}{\psi(K)}\left[\frac{1}{(1-x)^2}\right]\left[\frac{K(1-x)}{x}\right] = \frac{C}{x(1-x)}$$

where C is another constant and $1/2 \le x \le K/(k+K)$.The solution is

$$\pi(x) = A + B\ln\frac{x}{1-x}$$

where A, B are determined from the boundary conditions $\pi(1/2) = 1/2$ and $\pi(K/(k+K)) = 1$.(Thus $A = 1/2$.) Then, in the range $(k/(k+K)) \le x < 1/2$, the value of π is determined by reflection around $1/2$, *i.e.*, $\pi(x) = 1 - \pi(1-x)$.

The analysis for strictly increasing returns is entirely analogous. Indeed, by Lemma 39 for increasing returns, we need only consider two cases:

Case C. Agent is at t and the rival at k,where

$$I(\pi, t, k) = \pi'\left(\frac{t}{t+k}\right)\frac{k\psi(t)}{(t+k)^2}.$$

Case D. Agent is at k and the rival at t,where

$$I(\pi, k, t) = \pi'\left(\frac{k}{t+k}\right)\frac{t\psi(k)}{(t+k)^2}.$$

Strictly increasing returns imply $k\psi(t) > t\psi(k)$, hence $I(\pi, k, t) < I(\pi, t, k)$ for all $t \in [k, K]$, from which we derive as before that $\pi'(x) = C/(x(1-x))$ where C is another constant, $x = k/(t+k)$ and $1/2 \ge x \ge k/(k+K)$.The solution is

$$\pi(x) = A' + B'\ln\frac{x}{1-x}$$

for $1/2 \ge x \ge k/(k+K)$ and $1 - \pi(1-x)$ for $1/2 < K/(k+K)$, where A', B' are determined via the boundary conditions $\pi(k/(k+K)) = 0$ and $\pi(1/2) = 1/2$. ■

Remark 50 *In the above proof, suppose that* $\widetilde{\pi}$ *is a differentiable prize function with countably many jumps in the amounts* $\gamma_i > 0$, *for* $1 \leq i < \infty$, *in the interval* $I = [1/2, K/(k+K]$. *Then*

$$\frac{1}{2} = \int_I \pi'(y)dy = \sum_{1 \leq i < \infty} \gamma_i + \int_I \widetilde{\pi}'(y)dy$$

hence

$$\int_I \pi'(y)dy > \int_I \widetilde{\pi}'(y)dy$$

implying that there exists $y_2 \in I$, *and a corresponding* t_2 *with* $y_2 = t_2/\left(t_2 + K\right)$ *such that*

$$\widetilde{\pi}'\left(t_2/\left(t_2+K\right)\right) < \pi\left(t_2/\left(t_2+K\right)\right) = \sigma/G(t_2).$$

Then, arguing exactly as in the above proof we see that the incentive (to work) at t_2 *under* $\widetilde{\pi}$, *which is given by* $\widetilde{\pi}'(t_2/(t_2+K))G(t_2)$, *is strictly less than* σ, *which is the constant incentive under* π *at all* $t \in [k, K]$. *Thus* π *supersedes* $\widetilde{\pi}$. *(This argument can be replicated the case where* $\widetilde{\pi}$ *is an an arbitrary monotonic prize function, since such a function is almost everywhere differentiable except for countably many jump points.)*

Universality of the "Log Odds" Solution The term $x/(1-x)$ gives the "odds" of winning for the agent who produces the fraction x of the total output (while his rival produces the fraction $1-x$), assuming that lotteries are handed out in proportion to the outputs. Thus in the upper (lower) half of its domain, the optimal π awards the prize through "log of the odds" for strictly decreasing (increasing) returns to skill, completing π on the complementary half by the requirement $\pi(x) + \pi(1-x) = 1$. *What is noteworthy is that, apart from the type of returns (decreasing or increasing) exhibited by* ψ, *the solution is independent of the precise form of* ψ.The solution is first convex and then concave for strictly decreasing returns, and the other way round for strictly increasing returns , changing shape at the midpoint 1/2. In fact these two solutions are mirror images of each other if we reflect around the diagonal.

Also worthy of note is the fact (easily verified , and left to the reader) that, for constant returns to skill, we get the strictly increasing returns solution.

Interpretation of the Model with Small Fractional Increments The idea of an optimal scheme here is *not* that it maximizes expected total output. That would be much ado about nothing, since the output of each person goes up by only

$\varepsilon\%$ (at an extra disutility also of the order of $\varepsilon\%$) when he switches from shirk to work. The emphasis instead is on maximal effort *without* regard to the ensuing output. We have here an interpretation in mind that is, quite bluntly, *non-economic.* Output corresponds to a "score" that measures performance of "agents" in a "game" (think of the average score assigned by different judges to each person in a diving contest). The agents, who are all of star quality, are being incentivized to put in the final extra burst of effort to perform to the best of their ability.They value the prize enormously compared to the disutility incurred for the extra effort (the fame of being winner, perhaps also the money that fame might bring in the future). The interest is in finding a scheme π that is optimal in the sense that it most frugally (hence, frequently[21]) *creates competition and inspires maximal effort,* for its *own* sake (for the glory of the human spirit, for the sport). The minimum value $V(\pi)$ of the prize (which implements maximal effort under π) does entail *significant* savings, even though output rises very little: the ratio $V(\pi')/V(\pi) >> 1$ when we compare the optimal log-odds π with arbitrary $\pi' \in \Pi^+$.

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.

[21] i.e., π inspires maximal effort whenever any other scheme in Π does so (as we vary disutility of effort and valuation of the prize)

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

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

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

[10] Dubey, P., and Sahi, S. (2015). Eliciting Performance: Deterministic versus Proportional Prizes, forthcoming *International Journal of Game Theory,* Special Issue in Honor of Abraham Neyman.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[25] Shaked M. and J.G. Shanthikumar (1994). Stochastic Orders and their Applications. *Academic Press, Inc.*

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