

Contests for Experimentation*

Marina Halac[†] Navin Kartik[‡] Qingmin Liu[§]

August 27, 2014

Abstract

We study the design of contests for specific innovations when there is learning: contestants’ beliefs dynamically evolve about both the innovation’s feasibility and opponents’ success. Our model builds on exponential-bandit experimentation. We characterize contests that maximize the probability of innovation when the designer chooses how to allocate a prize and what information to disclose over time about contestants’ successes. A “public winner-takes-all contest” dominates public contests—those where any success is immediately disclosed—with any other prize-sharing scheme as well as winner-takes-all contests with any other disclosure policy. Yet, it is often optimal to use a “hidden equal-sharing contest”.

*We thank Martin Cripps, Alex Frankel, Qiang Fu, Kobi Glazer, Hugo Hopenhayn, Johannes Hörner, Ilan Kremer, Jingfeng Lu, Pauli Murto, Motty Perry, Sven Rady, Francesco Squintani, Juuso Välimäki, Tom Wiseman, and various seminar and conference audiences for helpful comments. Andrew Kosenko and Teck Yong Tan provided excellent research assistance. Kartik gratefully acknowledges financial support from the NSF.

[†]Columbia University and University of Warwick. Email: mhalac@columbia.edu.

[‡]Columbia University. Email: nkartik@columbia.edu.

[§]Columbia University. Email: qingmin.liu@columbia.edu.

Contests or prize awards are practical and proven mechanisms to procure innovations. In 1795, the French government offered 12,000 Francs for the development of a reliable food preservation method that would allow Napoleon Bonaparte to better feed his traveling army. After almost fifteen years of experimentation with various methods, the confectioner Nicolas Appert claimed the prize in 1810. His solution of heating food at elevated temperatures followed by airtight sealing remains the basis of modern-day canning.

Recent decades have seen a surge in the use of innovation contests ([McKinsey & Company, 2009](#)). The internet television company Netflix generated significant buzz in 2006 by announcing a \$1 million prize to induce a 10% improvement in the accuracy of its movie recommendation algorithm. In the philanthropic domain, there is an ongoing \$30 million Google Lunar X Prize for landing a private spacecraft on the surface of the Moon and sending “Mooncasts” back to Earth. In the public sector, President Barack Obama signed the America COMPETES Reauthorization Act in 2011 to grant U.S. government agencies the authority to conduct contests to spur innovation. There have also been renewed calls to reform the patent system by using prizes to avoid the ex-post deadweight losses of monopoly power (e.g. [Stiglitz, 2006](#)).¹

This paper studies the design of contests to procure specific innovations. Previous work on contest design has focused on settings in which there is no uncertainty about the environment. By contrast, we emphasize endogenous learning about the desired innovation. We are motivated by applications in which the viability or feasibility of the innovation is uncertain at the outset. Agents update their beliefs over time through their own *experimentation*—exerting costly effort and observing their outcomes—and also based on what they learn about other agents’ outcomes. In such contexts, how should one design a contest to maximize the probability of obtaining the innovation?

The model we develop in [Section 2](#) builds on the workhorse exponential-bandit framework ([Keller, Rady, and Cripps, 2005](#)). A principal and a set of ex-ante homogenous agents (or contestants) share a common prior over a binary state of the world. In the bad state, the innovation is infeasible. In the good state, an agent’s instantaneous probability of obtaining the innovation—hereafter, synonymous with

¹Debates about the merits of patents versus prizes (versus grants) to encourage innovation date back to at least the nineteenth century. In 2011, the U.S. Senator Bernie Sanders proposed two bills that together would create innovation prize funds of 0.57% of U.S. GDP—over \$80 billion at the time—for targeted medical research; the bills have not yet been put to vote in Congress.

“a success”—depends on the agent’s effort. At a linear cost, each agent chooses how much effort to covertly exert at each instant of time. Whether an agent succeeds or not is only directly observed by that agent and the principal, not by any other agent. All parties are risk neutral.

In our main specification, the principal has a fixed budget of prize money and her objective is to maximize the probability of a success. That is, the principal simply values obtaining the innovation; multiple successes provide her no additional benefit and she does not gain from saving any portion of the budget. While we also discuss how to endogenize the prize, it may be helpful to have the following interpretation in mind: innovation contests are now often funded by donors or philanthropic organizations (McKinsey & Company, 2009). In such cases, the firm designing the contest may not internalize the cost of prize money; rather, the firm garners reputation and related benefits by successfully procuring the innovation.²

There are two available design instruments. First, the principal chooses a *prize-sharing scheme* that specifies how the total prize will be divided amongst successful agents, possibly as a function of when each agent succeeds and with some specified deadline. For example, a “winner-takes-all” (WTA) contest awards the entire prize to the first agent who succeeds, whereas an “equal-sharing” (ES) contest gives every successful agent an equal share of the prize. Second, the principal also chooses an *information disclosure policy*, which specifies what information she discloses over time about agents’ outcomes. For example, a “public information” contest—or just public contest for short—reveals publicly, at every point of time, whether each agent has succeeded or not. A “hidden information” contest—or just hidden contest for short—does not reveal any information until the contest deadline.

In light of the agents’ risk neutrality and the principal valuing only one success, an intuitive solution to the design problem is to use a *public WTA* contest. After all, sharing the prize in any other fashion lowers a contestant’s expected reward from success, which should depress effort incentives. Not disclosing success immediately would lead contestants to fear that another contestant may have already succeeded, which should also lower incentives to exert effort. Consistent with this intuition, we show that a public WTA contest dominates a public contest with any other sharing

²Even if the same organization furnishes the prize and designs the contest, internal organizational structure and corresponding incentives may lead to similar considerations. There are also other interpretations. For example, contestants’ incentives often stem from factors such as the publicity received from recognition by the contest (cf. MacCormack, Murray, and Wagner, 2013, p. 27). Contest design can control how the publicity is allocated amongst successful contestants. In a different setting but with related motivation, Easley and Ghosh (2013) study a problem of “badge design”.

scheme, as well as a WTA contest with any other disclosure policy.

However, we find that a *hidden ES* contest can dominate a public WTA contest. The intuition turns on a tradeoff that arises in the current setting of experimentation. On the one hand, the principal wants to increase each agent’s expected reward from success; this familiar force pushes in favor of using a public WTA contest. On the other hand, the principal also wants to buttress agents’ beliefs about the innovation’s feasibility; this force, which owes entirely to learning, pushes in favor of hiding information—specifically, not disclosing the lack of success by other agents. Crucially, though, the gains from hiding information can only be harnessed by also sharing the prize.

We develop the above intuition in a simple two-period example in [Section 1](#). Our main results in [Section 3](#) characterize the optimal contest within a class of information disclosure policies. Among hidden and public information disclosure, the optimal contest among all prize-sharing rules is either public WTA or hidden ES.³ More generally, among *simple information disclosure policies*—policies that specify an arbitrary set of times at which the contest will publicly disclose whether each contestant has succeeded up until that time—we prove that it is optimal to use a “*mixture contest*”. A mixture contest takes the form of public WTA until some pre-specified date, at which point the contest switches to hidden ES (if no agent has yet succeeded).

[Section 4](#) discusses some formal extensions and [Section 5](#) concludes. The overarching message from our work is that when learning is important, the common presumption in favor of WTA contests for innovation and R&D must be qualified. We highlight that aside from learning, all other features of our model—e.g. seeking only one success—are geared in favor of public WTA contests.

Related literature

A subset of the prior work on contest design concerns research contests rather than innovation contests. The distinction is articulated well by [Taylor \(1995, p. 874\)](#): “in a research tournament, the terminal date is fixed, and the quality of innovations varies, while in an innovation race, the quality standard is fixed, and the date of discovery is variable.” The research-contest literature includes both static ([Fullerton](#)

³To be more precise: our analysis imposes anonymity on the sharing rule and focuses, when relevant, on symmetric equilibria. We view these as natural requirements. Dropping them would, if anything, only strengthen our main points.

and McAfee, 1999; Moldovanu and Sela, 2001; Che and Gale, 2003) and dynamic models (Taylor, 1995). Our approach of studying contest design when the principal has a fixed budget available has an antecedant in Krishna and Morgan (1998).^{4,5}

There is a sizable literature on different aspects of innovation or patent races, pioneered by Loury (1979) and Dasgupta and Stiglitz (1980). The focus in this literature is typically only on a winner-takes-all structure and much of it is without learning. Design questions are addressed, for example, by Bhattacharya, Glazer, and Sappington (1990), Moscarini and Smith (2011) and Judd, Schmedders, and Yeltekin (2012); see also the references therein.

Our work is more closely related to the following papers on innovation contests with learning or experimentation: Choi (1991), Malueg and Tsutsui (1997), Mason and Välimäki (2010), and Moscarini and Squintani (2010). All four papers share an exponential-bandit formulation, like us, but they focus solely on winner-takes-all contests rather than contest design.⁶ As Choi (1991) considers a multi-stage innovation process (whereas the other three papers do not), he notes that learning about the state through a competitor's success has a "positive effect" of making a firm more optimistic about the return to its own effort. Even though our setting is one of a single-stage innovation, the same logic is also essential to why a hidden equal-sharing contest can dominate a public winner-takes-all contest in our framework.

More broadly, the exponential-bandit framework has recently been used in many papers to study multi-agent strategic experimentation. Keller et al. (2005), Keller and Rady (2010), Klein and Rady (2011), and Murto and Välimäki (2011) feature only an informational interdependence across agents; Bonatti and Hörner (2011) and Cripps and Thomas (2014) also have a payoff interdependence, but both take a

⁴Some of the papers just cited (e.g. Taylor, 1995; Fullerton and McAfee, 1999) assume the principal only values the best innovation, while others (Krishna and Morgan, 1998; Moldovanu and Sela, 2001) assume the principal values the sum of all innovations or efforts. The latter assumption is standard in tournament theory pioneered by Lazear and Rosen (1981). In our framework, innovation is a binary variable—stochastically related to both effort and the underlying state—and the principal only values obtaining one innovation.

⁵We also note that there is a literature characterizing equilibria or payoffs in all-pay contests (or wars of attrition or all-pay auctions); see Siegel (2009) and the references therein. Kim and Lee (2014) study costly information acquisition in a war of attrition. Besides focusing on contest design, our framework differs in multiple ways from those typically considered in these papers; in particular, the total prize money paid to the agents in our setting is not a constant because no agent receives a reward when he does not obtain the innovation.

⁶Choi (1991) and Moscarini and Squintani (2010) effectively assume a linear cost of effort, like us, whereas Malueg and Tsutsui (1997) and Mason and Välimäki (2010) consider a convex cost. Malueg and Tsutsui (1997) assume that effort choices are publicly observed; Mason and Välimäki (2010) assume covert effort, as we do.

different form than the contests we study.⁷

Finally, questions about how much information a principal should disclose about agents' outcomes have been raised in various contexts. Regarding feedback in multi-stage contests (albeit where there is no learning about an exogenous state), see [Aoyagi \(2010\)](#), [Ederer \(2010\)](#), [Goltsman and Mukherjee \(2010\)](#), and [Wirtz \(2013\)](#).⁸ Outside of contests, [Che and Hörner \(2013\)](#) and [Kremer, Mansour, and Perry \(2013\)](#) study information disclosure when a designer wants to promote experimentation by a sequence of short-lived agents; [Bimpikis and Drakopoulos \(2014\)](#) study information disclosure to promote experimentation in a team setting.⁹

1 The Main Idea

This section explains the core intuition for our results in a simplified setting. A principal wants to obtain a specific innovation. The innovation's feasibility depends on a binary state—either good or bad—that is persistent and unobservable. The prior probability of the good state is $p_0 \in (0, 1)$. There are two periods, $t \in \{0, 1\}$, no discounting, and two risk-neutral agents. In each period each agent covertly chooses whether to work or shirk. If an agent works in a period and the state is good, the agent succeeds in that period with probability $\lambda \in (0, 1)$; if either the agent shirks or the state is bad, the agent does not succeed.¹⁰ Working in a period costs an agent $c > 0$. Successes are conditionally independent across agents given the state and observed only by the principal and the agent who succeeds. The principal wants to induce both agents to work until at least one succeeds; an additional success provides no extra benefit. The principal has a prize budget \bar{w} to pay the agents.

In this illustrative setting, we consider four contests. They vary by whether the entire prize is allocated to the first successful agent or divided equally among all

⁷Various authors have analyzed contracting in single-agent problems in the exponential-bandit framework, including [Bergemann and Hege \(1998, 2005\)](#), [Gomes, Gottlieb, and Maestri \(2013\)](#), [Halac, Kartik, and Liu \(2013\)](#), [Kwon \(2013\)](#), and [Hörner and Samuleson \(2014\)](#). [Strulovici \(2010\)](#) studies a game of experimentation with multiple agents and both payoff and informational interdependence.

⁸[Yildirim \(2005\)](#), [Gill \(2008\)](#), [Rieck \(2010\)](#), and [Akcigit and Liu \(2014\)](#) address the incentives for contestants to themselves disclose their outcomes to opponents. [Campbell, Ederer, and Spinnewijn \(2014\)](#) consider related issues in a moral-hazard-in-teams setting.

⁹[Lizzeri, Meyer, and Persico \(2002\)](#) and [Fuchs \(2007\)](#) study how much feedback an agent should get about his own performance in dynamic moral-hazard settings. Dynamic information disclosure about an exogenous state variable is addressed by [Ely, Frankel, and Kamenica \(2013\)](#) and [Ely \(2014\)](#).

¹⁰Without loss, an agent who succeeds in the first period cannot—more precisely, will not—exert effort in the second period. Whenever we refer to an agent working in the second period, it is implicitly in the event that he has not succeeded in the first period.

agents who succeed by the end of the second period, and by whether an agent's success in the first period is publicly disclosed or kept hidden.

Public winner-takes-all. Suppose the principal awards the entire prize \bar{w} to the first agent who obtains a success, and the principal publicly discloses all successes at the end of each period. If both agents succeed simultaneously, the prize is equally divided (or allocated to either agent with equal probability).¹¹ In this mechanism, neither agent will work in the second period if either succeeded in the first period. Thus, in any period, if there has been no earlier success and the opponent is exerting effort, an agent's expected reward for success is $\hat{w} := \lambda \frac{\bar{w}}{2} + (1 - \lambda)\bar{w}$. If $p_0 \lambda \hat{w} > c$, it is a dominant strategy for an agent to work in the first period; assume for the rest of this section that this condition holds. If neither agent succeeds in the first period, both agents work in the second period if and only if

$$p_1 \lambda \hat{w} \geq c, \tag{1}$$

where $p_1 := \frac{p_0(1-\lambda)^2}{p_0(1-\lambda)^2 + 1 - p_0}$ is the agents' belief in the second period that the state is good given that neither succeeded in the first period having exerted effort.

Hidden winner-takes-all. Suppose the principal still awards the entire prize \bar{w} to the first successful agent (and splits the prize in case of simultaneous success) but now she does not disclose any information about first-period successes until the end of the game. Plainly, an agent works in the first period if he is willing to work in the second period. However, because first-period successes are hidden, an agent's second-period decision must now take into account the possibility that the opponent may have already succeeded and secured the entire prize. When both agents work in the first period, an agent i 's incentive constraint for effort in the second period following his own lack of success in the first period is

$$\Pr[j \text{ failed} \mid i \text{ failed}] p_1 \lambda \hat{w} \geq c, \tag{2}$$

where j denotes i 's opponent. Clearly, constraint (2) is more demanding than (1). In other words, for any set of parameters, a public winner-takes-all (WTA) contest dominates a hidden WTA contest in the sense that the latter cannot induce effort by both agents in both periods when the former cannot; moreover, for some pa-

¹¹Simultaneous success will be a moot issue in our general model, which is cast in continuous time.

rameters, a public WTA contest induces effort by both agents in both periods while hidden WTA does not.

Public equal-sharing. Suppose the principal discloses all successes at the end of each period, but she now divides the prize \bar{w} equally between the two agents if both succeed by the end of the second period no matter their order of success (and continues to allocate the entire prize to an agent if he is the only one to succeed). If an agent succeeds in the first period, the opponent is certain in the second period that the state is good and, due to the shared-prize scheme, the opponent's reward for success is $\frac{\bar{w}}{2}$. It follows that when $\lambda\frac{\bar{w}}{2} < c$, an agent does not work in the second period if his opponent succeeds in the first period; in this case, the contest is equivalent to public WTA. On the other hand, if $\lambda\frac{\bar{w}}{2} > c$, an agent will work in the second period if the opponent succeeds in the first period. This "duplication effort" does not benefit the principal because she only cares about obtaining one success; moreover, as compared to public WTA, agents' incentives to work in the first period can now be lower due to two reasons: free-riding—an agent may want to wait for the other agent to experiment and reveal information about the state—and a lower expected reward for first-period success due to the opponent's duplication effort. In any case, observe that if both agents work in the first period and neither succeeds, the incentive constraint in the second period is still given by (1). Therefore, for any set of parameters, a public WTA contest dominates a public equal-sharing (ES) contest; it can also be shown that for some parameters, a public WTA contest induces effort by both agents in both periods while public ES does not.

Hidden equal-sharing. Suppose the principal uses the equal-sharing prize scheme as above, but does not disclose any information about first-period successes until the end of the game. Although the prize is shared, there is now no free-riding concern: an agent cannot learn from his opponent's success when that is not disclosed. In fact, because there is nothing to be learned about the opponent after the first period, it is without loss that an agent works in the first period if he works at all.¹² Suppose there is an equilibrium in which both agents work in the first period and consider an agent's incentive to work in the second period following his own lack of success in the first period. The agent does not know whether the opponent succeeded or failed

¹²"Without loss" in the sense that if there is an equilibrium in which an agent works in the second period (and possibly in the first period too), then there exists an outcome-equivalent equilibrium in which this agent works in the first period (and possibly in the second period too).

in the first period. In the event that the opponent succeeded, the agent’s posterior belief that the state is good is one while his reward for success becomes half the prize. On the other hand, in the event the other agent failed in the first period, the agent’s posterior belief is $p_1 < 1$ but the expected reward for success is $\hat{w} > \frac{\bar{w}}{2}$. Hence, an agent i ’s incentive constraint in the second period is

$$\Pr[j \text{ succeeded} \mid i \text{ failed}] \lambda \frac{\bar{w}}{2} + \Pr[j \text{ failed} \mid i \text{ failed}] p_1 \lambda \hat{w} \geq c. \quad (3)$$

It can be checked that there are parameters such that (3) holds while (1) does not. In other words, there are parameters under which both agents work in both periods in a hidden ES contest but not in a public WTA contest. For these parameters, a hidden ES contest dominates public WTA.

What is the intuition behind why hidden ES can dominate public WTA although neither hidden WTA nor public ES can? On the one hand, holding fixed an agent’s belief about the state, it is clear that a WTA prize scheme maximizes effort incentives. On the other hand, the nature of learning—specifically, failure is bad news—implies that the principal would like to hide information about the opponent’s first-period outcome to bolster an agent’s second-period belief in the only event that matters to the principal, viz. when the opponent fails in the first period. Hiding information but still using WTA is counter-productive, however, because when an agent conditions on obtaining a reward in the second period, he deduces that the opponent must have failed. Consequently, harnessing the benefits of hiding information requires some sharing of the prize. On the flip side, sharing the prize while maintaining public disclosure is not beneficial either because this change from public WTA only alters second-period incentives when the principal does not value additional effort, viz. when the innovation has obtained in the first period.

Regarding learning, it bears emphasis that public WTA would always be an optimal contest were it certain that the state is good—if there is no learning, there is no benefit to hiding information. Public WTA would also be always optimal if agents did learn but only from their own outcomes and not from others’, i.e. if their experimentation “arms” were uncorrelated.¹³

These intuitions in hand, we now turn to our main model and analysis.

¹³Regarding the gains from innovation: if the principal were to value obtaining a success by each agent, then naturally even a public ES contest may dominate public WTA for some parameters.

2 The Model

2.1 Setup

A principal wants to obtain a specific innovation. Whether the innovation is feasible depends on the state of nature, $\theta \in \{G, B\}$, where G represents “good” and B represents “bad”. This state is persistent and unobservable to all parties. There are $N \geq 1$ agents who can work on the principal’s project. Time is continuous and runs from 0 up to some end date $T \geq 0$, which is chosen by the principal. At every moment $t \in [0, T]$, each agent $i \in \mathcal{N} := \{1, 2, \dots, N\}$ covertly chooses effort $a_{i,t} \in [0, 1]$ at instantaneous cost $ca_{i,t}$, where $c > 0$. We let $A_t := a_{1,t} + \dots + a_{N,t}$. When an agent exerts effort $a_{i,t}$, he succeeds with instantaneous probability $\lambda_\theta a_{i,t}$, where $\lambda_G > \lambda_B \geq 0$ are constants commonly known to all parties. Successes are conditionally independent given the state. We assume that successes are observable only to the agent who succeeds and to the principal; [Section 4](#) discusses alternative scenarios. We assume $\lambda_B = 0$ and denote $\lambda := \lambda_G$.¹⁴

When a success is obtained, the principal receives a lump-sum payoff $v > 0$; the agents do not intrinsically care about project success. The principal values only one success: additional successes have no social value. The agents are risk neutral, have quasi-linear preferences, and are expected-utility maximizers. For tractability, we assume no discounting.¹⁵

Let $p_0 \in (0, 1)$ be the prior probability that the state is good, commonly known to all parties. Assume that the ex-ante expected marginal benefit of effort is larger than the marginal cost: $p_0 \lambda v > c$. This means that some experimentation is efficient, even though conditional on the bad state the marginal benefit of effort is zero.

Denote by p_t the posterior probability that the state is good when no agent has succeeded by time t given a (measurable) effort profile $\{a_{i,t}\}_{i,t}$. We refer to p_t as the *public belief*. By Bayes’ rule,

$$p_t = \frac{p_0 e^{-\int_0^t \lambda A_z dz}}{p_0 e^{-\int_0^t \lambda A_z dz} + 1 - p_0}. \quad (4)$$

¹⁴The key consequence of assuming $\lambda_B = 0$ is that a single success fully resolves the fundamental uncertainty about the innovation. This is a standard assumption in the exponential-bandit framework; see [Keller and Rady \(2010\)](#) for an exception.

¹⁵In this exponential-bandit model, agents become more pessimistic about the state over time as they experiment without success. This serves as an endogenous source of discounting, so payoffs are well defined even without an explicit discount factor. Our result that the optimal public WTA contest can be dominated by a hidden ES contest is robust to adding a small discounting friction.

The evolution of this belief is governed by the differential equation¹⁶

$$\dot{p}_t = -p_t(1 - p_t)\lambda A_t.$$

Let $a_i^t := \int_0^t a_{i,z} dz$ be the cumulative effort, or experimentation, by agent i up to time t conditional on him not having succeeded by t , and $A^t := \sum_i a_i^t = \int_0^t A_z dz$ the aggregate cumulative effort up to t given no success by t . The (aggregate) probability of success is

$$p_0 \left(1 - e^{-\lambda A^T}\right). \quad (5)$$

The principal does not internalize the cost of effort (nor payments to the agents, as explained in more detail below), so her objective is simply to maximize the probability of obtaining a success, i.e. to maximize the expectation of (5). By (4), (5) is equivalent to $p_0 - (1 - p_0)\frac{p_T}{1-p_T}$. Hence, as is intuitive:

Remark 1. For any set of parameters, the probability of success is increasing in aggregate cumulative effort, A^T . Moreover, for any prior belief p_0 , a lower public belief at the deadline, p_T , corresponds to a higher probability of success.

2.2 First best

Social welfare combines both the value of success and the total cost of effort. Since it is socially optimal to never exert effort after a success has been obtained, the social optimum is derived by maximizing

$$\int_0^\infty (p_t \lambda v - c) A_t e^{-\int_0^t p_z \lambda A_z dz} dt.$$

To interpret this expression, note that $e^{-\int_0^t p_z \lambda A_z dz}$ is the probability that no success is obtained by time t , and p_t is the probability that the state is good given no success by t . Conditional on the good state, a success then occurs at t with probability λA_t , yielding a value v . Since the public belief p_t is decreasing over time, a social-welfare maximizing effort profile is $a_{i,t} = 1$ for all $i \in \mathcal{N}$ if $p_t \lambda v \geq c$, and $a_{i,t} = 0$ for all $i \in \mathcal{N}$

¹⁶The following heuristic derivation provides intuition: since the probability of not obtaining a success conditional on the good state and effort level A_t in a small time interval $[t, t + dt)$ is $1 - \lambda A_t dt$, we have $p_{t+dt} = \frac{p_t(1 - \lambda A_t dt)}{p_t(1 - \lambda A_t dt) + 1 - p_t}$, which, after some simplification, yields $\frac{p_{t+dt} - p_t}{dt} = \frac{-p_t(1 - p_t)\lambda A_t}{1 - p_t \lambda A_t dt}$. Taking $dt \rightarrow 0$ yields the stated formula.

otherwise. The first-best *stopping (posterior) belief* is given by

$$p^{FB} := \frac{c}{\lambda v}. \quad (6)$$

2.3 Contests and strategies

The principal has a budget $\bar{w} > 0$ that can be used to incentivize the agents. We assume limited liability, i.e. each agent must receive a non-negative payment, and to avoid trivialities, $p_0 \lambda \bar{w} > c$. In general, a mechanism specifies a deadline $T \geq 0$ and a vector of payments $(w_1, \dots, w_N) \in \mathbb{R}_+^N$ that, without loss, are made at T as a function of the principal's information at T (when each agent succeeded, if ever) and subject to the ex-post constraint $\sum_{i \in \mathcal{N}} w_i \leq \bar{w}$. In addition, the principal chooses an information disclosure policy, which in its most general form specifies for each agent i and at each time t , a signal of the entire history of successes.

We are interested in a sub-class of mechanisms, which we call contests. With regards to payments, we allow for any (deterministic) payment scheme subject to an anonymity requirement. Formally, let s_i denote the time at which agent i succeeds; by convention, $s_i = \emptyset$ if i does not succeed. We consider payment schemes defined by a function $w(s_i, \mathbf{s}_{-i})$, whose value specifies the payment to agent i as a function of all agents' success times.¹⁷ For any \mathbf{s} , let $W(\mathbf{s})$ denote the total payment made by the principal. We impose the following constraints on $w(\cdot)$: (i) $w(s_i, \mathbf{s}_{-i}) = w(s_i, \sigma(\mathbf{s}_{-i}))$ for any permutation σ ; (ii) $w(\emptyset, \cdot) = 0$; and (iii) $\mathbf{s} \neq (\emptyset, \dots, \emptyset) \implies W(\mathbf{s}) = \bar{w}$.¹⁸

With regards to information disclosure, the principal specifies a set of times at which the entire history of successes to date is publicly disclosed; at other times, there is no information disclosed. Formally, let $o_{i,t} = 1$ if agent i succeeds at time t and $o_{i,t} = 0$ otherwise. An information disclosure policy is an arbitrary set $\mathcal{T} \subseteq [0, T]$ such that: (i) at any time $t \in \mathcal{T}$, the principal sends a public message $m_t = (o_{1,z}, \dots, o_{N,z})_{z < t}$; (ii) at any time $t \notin \mathcal{T}$, the principal sends a public message $m_t = \emptyset$. We refer to this class of policies as *simple information disclosure policies*, postponing

¹⁷As usual, bold symbols denote vectors. Note that an agent is paid the same regardless of whether he succeeds once or more than once; recall that the principal only values obtaining one success.

¹⁸Restrictions (ii) and (iii) are without loss of generality: (ii) because of limited liability and (iii) because the principal must satisfy the budget constraint and does not value unspent prize money. Given a fixed set of agents, restriction (i) may entail some loss of generality, but we conjecture that it would not if the principal can also specify the number of agents, which is a variation we discuss in [Subsection 3.3](#). In any case, relaxing (i) would, if anything, only strengthen the result that a public winner-takes-all contest can be dominated by a hidden equal-sharing contest.

a discussion of more general policies to [Section 4](#). A contest with $\mathcal{T} = [0, T]$ is a *public contest*—the principal immediately discloses any success—whereas a contest with $\mathcal{T} = \emptyset$ is a *hidden contest*—the principal discloses no information at all about any success until the deadline.

Denote by $h_i^t := (m_z, o_{i,z}, a_{i,z})_{z < t}$ the private history of agent i at time t ; note that it includes the public message. An agent i 's (pure) strategy is a measurable function that specifies, for each history h_i^t , a choice of effort at time t , $a_{i,t}$. Without loss, we interpret $a_{i,t}$ as agent i 's effort at t conditional on him not having succeeded by t , as an agent never exerts effort after succeeding. Our solution concept is Nash equilibrium.¹⁹ We restrict attention to symmetric equilibria (viz., equilibria in which all agents use the same strategy); we will indicate subsequently where this restriction is used (see, in particular, [fn. 21](#)). We say that an agent i uses a *stopping strategy* with *stopping time* z if the agent exerts full effort until time z (so long as he has not learned that any agent, including himself, has succeeded) followed by no effort.

3 Optimal Contests

We solve for the contest that maximizes the probability of success. [Subsection 3.1](#) shows that a public winner-takes-all (WTA) contest is optimal among public contests and [Subsection 3.2](#) shows that a hidden equal-sharing (ES) contest is optimal among hidden contests. We compare public WTA and hidden ES contests in [Subsection 3.3](#). In [Subsection 3.4](#), we show that an optimal contest (within the class of simple information disclosure policies) is a mixture of public WTA and hidden ES; we provide intuitive sufficient conditions for when the optimum is a degenerate mixture, i.e. either public WTA or hidden ES.

3.1 Public contests

In a public contest, an agent's success is immediately disclosed to all other agents. Agents therefore update their beliefs based on their outcomes as well as their opponents' outcomes, given the equilibrium strategies.

Consider a public contest with an arbitrary prize scheme $w(s_i, s_{-i})$. Let $A_{-i,z}$ denote (i 's conjecture of) the aggregate effort exerted by i 's opponents at time z so

¹⁹Our analysis would not be affected by imposing standard refinements such as (appropriately defined versions of) subgame perfection or perfect Bayesian equilibrium.

long as no agent has obtained a success by z . We denote by $w_{i,t}$ the expected reward agent i receives if he is the first one to succeed at t , which depends on $w(s_i, s_{-i})$ and the continuation strategies of the opponents who may continue to exert effort and share the prize. If some agent besides i is the first agent to succeed at t , we denote agent i 's expected continuation payoff by $u_{i,t}$. We suppress the dependence of the relevant variables on the continuation strategy profile to save on notation. Agent i 's problem can then be written as

$$\max_{(a_{i,t})_{t \in [0, T]}} \int_0^T [(p_{i,t} \lambda w_{i,t} - c) a_{i,t} + p_{i,t} \lambda A_{-i,t} u_{i,t}] e^{-\int_0^t p_{i,z} \lambda (a_{i,z} + A_{-i,z}) dz} dt, \quad (7)$$

where $p_{i,t}$ is i 's belief that the state is good at time t (so long as success has not been obtained), given by the following analog of (4):

$$p_{i,t} = \frac{p_0 e^{-\int_0^t \lambda (a_{i,z} + A_{-i,z}) dz}}{p_0 e^{-\int_0^t \lambda (a_{i,z} + A_{-i,z}) dz} + 1 - p_0}.$$

To interpret the objective (7), note that $e^{-\int_0^t p_{i,z} \lambda (a_{i,z} + A_{-i,z}) dz}$ is the agent's belief that no success will obtain by time t . Conditional on the good state and no success by t , the probability that the agent is the first to succeed at t is $\lambda a_{i,t}$, and the probability that an agent besides i is the first to succeed at t is $\lambda A_{-i,t}$.

Solving for equilibria in an arbitrary public contest is not straightforward; instead, we take an indirect but also more insightful approach. First, observe that agent i can ensure $u_{i,t} \geq 0$ (by shirking after t). It follows that the agent chooses $a_{i,t} = 0$ if $p_{i,t} \lambda w_{i,t} < c$, and thus $a_{i,t} > 0$ requires

$$p_{i,t} \geq \frac{c}{\lambda w_{i,t}} \geq \frac{c}{\lambda \bar{w}}, \quad (8)$$

where the second inequality is because $w_{i,t} \leq \bar{w}$ by the principal's budget constraint. Hence, the lowest belief at which agent i is willing to exert positive effort in a public contest is $p_{i,t} = \frac{c}{\lambda \bar{w}}$.

Consider now a public WTA contest, where the full prize is awarded to the first agent that succeeds: $w_{i,t} = \bar{w}$ and $u_{i,t} = 0$ for all $t \in [0, T]$. Since the agent's belief $p_{i,t}$ is decreasing over time, the unique solution to (7) in this case is $a_{i,t} = 1$ if $p_{i,t} \geq p^{PW}$ and $a_{i,t} = 0$ otherwise,²⁰ where

$$p^{PW} := \frac{c}{\lambda \bar{w}}. \quad (9)$$

²⁰Throughout, we break indifference in favor of exerting effort whenever this is innocuous.

It follows that in a public WTA contest with deadline T , there is a unique equilibrium in which all agents exert full effort until either a success is obtained, or the public belief, viz. expression (4) with $A_t = N$, reaches p^{PW} (or the deadline T binds), and they exert zero effort thereafter. To maximize experimentation, the deadline T is optimal if and only if $T \geq T^{PW}$, where T^{PW} is the time at which the public belief reaches p^{PW} given that all agents exert full effort, i.e.

$$\frac{p_0 e^{-N\lambda T^{PW}}}{p_0 e^{-N\lambda T^{PW}} + 1 - p_0} = \frac{c}{\lambda \bar{w}}. \quad (10)$$

Comparing with condition (8) above, we see that an optimal public WTA contest induces effort by all agents until their belief reaches the lowest belief at which any agent is willing to exert positive effort in a public contest. It follows that this contest yields the maximum possible aggregate cumulative effort as a solution to program (7), and thus maximizes the probability of success (see Remark 1). Hence, a public WTA contest is optimal within the class of public contests.

Proposition 1. *An optimal public winner-takes-all contest is optimal among public contests. In an optimal public winner-takes-all contest, each agent uses a stopping strategy with stopping time T^{PW} defined by (10). T^{PW} is increasing in p_0 and \bar{w} , decreasing in c and N , and non-monotonic in λ . The probability of success is increasing in p_0 , \bar{w} and λ , decreasing in c , and independent of N .*

(All proofs are in the [Appendix](#).)

The non-monotonicity of T^{PW} with respect to λ is due to two countervailing effects: on the one hand, for any given belief $p_{i,t}$, the marginal benefit of effort is larger if λ is higher; on the other hand, the larger is λ , the faster each agent updates his belief down following a history of effort and no success (cf. [Bobtcheff and Levy, 2010](#); [Halac et al., 2013](#)). Nevertheless, the stopping belief, p^{PW} , is decreasing in λ , as seen immediately from (9), and as a result the probability of obtaining a success is increasing in λ . It is also intuitive why p^{PW} is independent of the number of agents: the likelihood that multiple agents succeed at the same instant is second order, hence the only effect of higher N on an agent's incentives at time t (so long as no one has succeeded yet) is to lower the public belief at t .

Remark 2. Comparing (6) and (9), it is clear that a public WTA contest can implement the first-best solution if and only if $\bar{w} \geq v$.

3.2 Hidden contests

In a hidden contest, an agent's success is not disclosed until the deadline. Agents therefore update their beliefs based on their own outcomes only.

Consider a hidden contest with an arbitrary prize scheme $w(s_i, s_{-i})$. Denote by $w_{i,t}$ the expected reward agent i receives if he succeeds at time t , which depends on $w(s_i, s_{-i})$ and the strategies of the opponents. Then agent i 's problem reduces to

$$\max_{(a_{i,t})_{t \in [0, T]}} \int_0^T \left(p_{i,t}^{(1)} \lambda w_{i,t} - c \right) a_{i,t} e^{-\int_0^t p_{i,z}^{(1)} \lambda a_{i,z} dz} dt, \quad (11)$$

where $p_{i,t}^{(1)}$ is i 's belief that the state is good at time t given that he has not succeeded by t , which is given by the following analog of (4):

$$p_{i,t}^{(1)} = \frac{p_0 e^{-\int_0^t \lambda a_{i,z} dz}}{p_0 e^{-\int_0^t \lambda a_{i,z} dz} + 1 - p_0}.$$

To interpret the objective (11), note that $e^{-\int_0^t p_{i,z}^{(1)} \lambda a_{i,z} dz}$ is the agent's belief that he will not succeed by time t . Conditional on the good state and the agent not succeeding by t , the agent succeeds at t and receives $w_{i,t}$ with probability $\lambda a_{i,t}$.

Consider now a hidden ES contest, where the principal divides the prize equally among all the agents who succeed by some deadline. That is, given a deadline $T > 0$, agent i receives $\frac{\bar{w}}{n}$ if he succeeds at any time $t \in [0, T]$ and a total of $n \in \{1, \dots, N\}$ agents succeed by T . Since i 's expected reward for success, which we denote w_i^{HS} , is independent of when he succeeds, it is immediate from (11) that an optimal strategy for i in this case is a stopping strategy where $a_{i,t} = 1$ if $p_{i,t}^{(1)} \lambda w_i^{HS} \geq c$ and $a_{i,t} = 0$ otherwise. Under a stopping strategy, $p_{i,t}^{(1)} = \frac{p_0 e^{-\lambda t}}{p_0 e^{-\lambda t} + 1 - p_0}$.

Let p_i^{HS} and T_i^{HS} be the stopping belief and time respectively. Deadline T is optimal if and only if $T \geq \max_{i \in \mathcal{N}} T_i^{HS}$, in which case an agent i 's stopping belief satisfies

$$p_i^{HS} \lambda w_i^{HS} = c, \quad (12)$$

where

$$p_i^{HS} = \frac{p_0 e^{-\lambda T_i^{HS}}}{p_0 e^{-\lambda T_i^{HS}} + 1 - p_0}. \quad (13)$$

Consider symmetric equilibria. T_i^{HS} and w_i^{HS} are then independent of i and can be simply denoted T^{HS} and w^{HS} . Should agent i succeed at any time, he learns that

the state is good, in which event (i) the probability he ascribes to any opponent succeeding (resp., not succeeding) by T^{HS} is $1 - e^{-\lambda T^{HS}}$ (resp., $e^{-\lambda T^{HS}}$); and (ii) he views opponents' successes as independent. Thus, when all opponents use a stopping time T^{HS} , an agent's expected reward for success in a hidden ES contest is

$$w^{HS} = \bar{w} \left[\sum_{n=0}^{N-1} \frac{1}{n+1} \binom{N-1}{n} \left(1 - e^{-\lambda T^{HS}}\right)^n e^{-(N-1-n)\lambda T^{HS}} \right]. \quad (14)$$

The term in square brackets is the expected share of the prize that an agent receives for success. It can be shown (see the proof of [Proposition 2](#) in [Appendix A](#)) that this expected-share expression simplifies to $\frac{1 - e^{-N\lambda T^{HS}}}{(1 - e^{-\lambda T^{HS}})^N}$. Substituting this expression and (13) into (12) implicitly defines the equilibrium stopping time, T^{HS} , by

$$\frac{1 - e^{-N\lambda T^{HS}}}{(1 - e^{-\lambda T^{HS}})^N} \frac{p_0 e^{-\lambda T^{HS}}}{p_0 e^{-\lambda T^{HS}} + 1 - p_0} = \frac{c}{\lambda \bar{w}}. \quad (15)$$

Each of the two terms on the left-hand side above is strictly decreasing in T^{HS} ; hence, among symmetric equilibria in stopping strategies, there is a unique equilibrium. Moreover, it can be shown that any symmetric equilibrium must be outcome equivalent to this one in the sense that the probability of success by each agent—equivalently, the private belief reached by the deadline T by each agent, and hence the public belief at T (see [Remark 1](#))—is the same.²¹

Proposition 2. *In any hidden equal-sharing contest, all symmetric equilibria are outcome equivalent. In an optimal hidden equal-sharing contest, there is a symmetric equilibrium in which each agent uses a stopping strategy with stopping time T^{HS} defined by (15). T^{HS} is increasing in p_0 and \bar{w} , decreasing in c and N , and non-monotonic with respect to λ . The probability of success is increasing in p_0 , \bar{w} and λ and decreasing in c . An increase in N can increase or decrease the probability of success.*

²¹For $T \leq T^{HS}$, the unique symmetric equilibrium is in stopping strategies with stopping time T . Consider $T > T^{HS}$. Even though best responses always exist in stopping strategies, there will be symmetric equilibria in which agents do not play stopping strategies. The reason is that there are multiple strategies by which an agent can arrive at T with the private belief indicated in (15). However, in any symmetric equilibrium, the cumulative effort by each agent must be T^{HS} ; this is because an agent's expected share of the prize from success is strictly decreasing in each opponent's cumulative effort, and his own private belief at T is strictly decreasing in his own cumulative effort. While we do not study asymmetric equilibria, we conjecture that they do generally exist here because cumulative efforts are strategic substitutes. However, note that allowing for asymmetric equilibria—combined with letting the principal select her preferred equilibrium—cannot decrease experimentation in a hidden ES contest, and thus will only strengthen our message that hidden ES can dominate public WTA (as under public WTA, the unique equilibrium is symmetric).

The comparative statics in [Proposition 2](#) are largely intuitive, so we only note two points. First, the non-monotonicity of T^{HS} in λ owes to the same countervailing forces that were noted after [Proposition 1](#). Second, the probability of obtaining a success may increase or decrease when N increases because there are two competing effects. On the one hand, holding fixed T^{HS} , conditional on at least one opponent having succeeded (which reveals the state to be good), a larger number of opponents having succeeded only lowers the expected benefit of effort for an agent at T^{HS} . However, an increase in N also lowers T^{HS} , which by itself decreases the probability that any opponent has succeeded by T^{HS} conditional on the good state. [Figure 1](#) provides an example in which the probability of a success is non-monotonic in N .

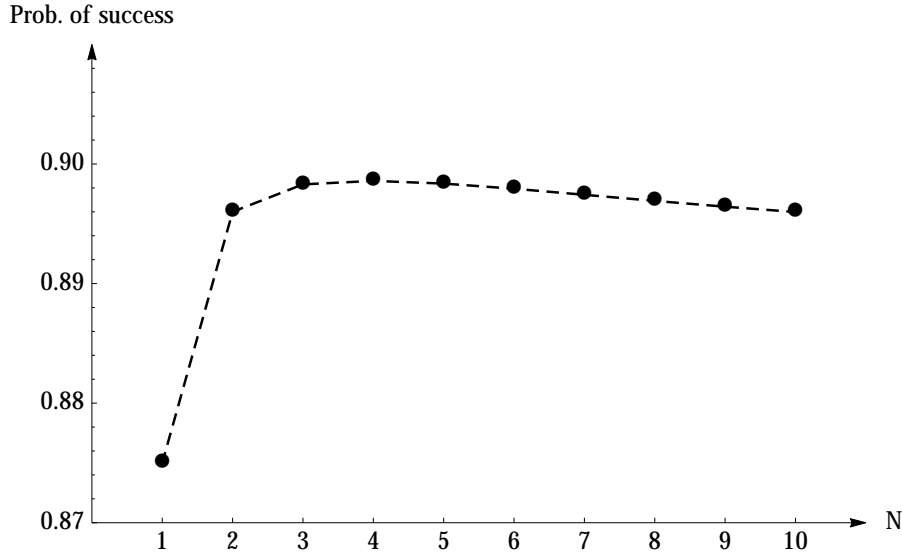


Figure 1 – Probability of success in an optimal hidden ES contest. Parameters are $p_0 = 0.9$, $c = 0.4$, $\bar{w} = 1$, and $\lambda = 2$; the maximum is obtained at $N = 4$.

[Proposition 2](#) speaks to hidden equal-sharing contests. The principal may also consider dividing the prize asymmetrically; for example, rewarding agents who succeed earlier with larger shares of the prize. However:

Proposition 3. *An optimal hidden equal-sharing contest is optimal among hidden contests.*

Let us sketch the argument. There is an optimal hidden ES contest in which each agent uses a stopping time T^{HS} ; each agent’s incentive constraint for effort binds at T^{HS} (see (15)). Since an agent’s expected reward for success is independent of when he succeeds in a hidden ES contest, the proof of [Proposition 3](#) further shows that each agent’s incentive constraint binds at each $t \in [0, T^{HS}]$; that is, at each moment

before the stopping time, an agent is indifferent over how much effort to exert given that he has exerted full effort in the past. Intuitively, shirking at time t precludes success at t but increases the continuation payoff after t (as the private belief does not decrease); these effects cancel when the reward for success is constant over time (cf. Halac et al., 2013). It then follows that a hidden contest in which an agent's expected reward for success is not constant over $[0, T]$ cannot induce more experimentation than the optimal hidden ES contest. To see why, suppose to the contrary that some hidden contest admits a symmetric equilibrium where each agent's cumulative effort is $T > T^{HS}$; for simplicity, suppose agents use a stopping strategy. The principal's budget constraint implies that the ex-ante expected reward for obtaining success cannot be greater than w^{HS} , the expected reward in the optimal hidden ES contest. Thus, given a non-constant expected reward sequence, there is a time $t \leq T$ such that an agent's expected reward for success is strictly less than w^{HS} at t but no less than w^{HS} at each $t' \in (t, T]$. But since the agent's incentive constraint binds given a constant reward w^{HS} , the constraint must be violated at t under the non-constant reward sequence, a contradiction.

3.3 Public winner-takes-all versus hidden equal-sharing

Proposition 1 shows that an optimal public WTA contest is optimal among all public contests and **Proposition 3** shows that an optimal hidden ES contest is optimal among all hidden contests. To compare these two contests, it suffices to compare their stopping times: the principal prefers hidden ES to public WTA if and only if $T^{HS} > T^{PW}$, where T^{PW} is given by (10) and T^{HS} is given by (15). Since the left-hand side of (10) and that of (15) are each decreasing as functions of T^{PW} and T^{HS} respectively, $T^{HS} > T^{PW}$ if and only if the left-hand side of (15) would be strictly larger than its right-hand side were $T^{HS} = T^{PW}$.

Proposition 4. *The principal strictly prefers an optimal hidden equal-sharing contest to an optimal public winner-takes-all contest if and only if*

$$\frac{1 - e^{-\lambda N T^{PW}}}{(1 - e^{-\lambda T^{PW}})N} \frac{p_0 e^{-\lambda T^{PW}}}{p_0 e^{-\lambda T^{PW}} + 1 - p_0} \lambda \bar{w} > c, \quad (16)$$

where T^{PW} is defined by equation (10).

To see the intuition behind (16), let

$$L := \sum_{n=1}^{N-1} \frac{\Pr[n \text{ opponents succeeded by } T^{PW} \mid G]}{\Pr[\text{at least one opponent succeeded by } T^{PW} \mid G]} \left(\frac{1}{n+1} \right). \quad (17)$$

In words, L is an agent's expected share of the prize for success at T^{PW} in a hidden ES contest given that all opponents use a stopping time T^{PW} and at least one opponent has succeeded. Inequality (16) is equivalent to, for any $i \in \mathcal{N}$,

$$\begin{aligned} c &< \Pr[\text{some } j \neq i \text{ succeeded by } T^{PW} \mid i \text{ did not}] L \lambda \bar{w} \\ &+ \Pr[\text{no } j \neq i \text{ succeeded by } T^{PW} \mid i \text{ did not}] \Pr[G \mid \text{no success by } T^{PW}] \lambda \bar{w}. \end{aligned}$$

The definition of T^{PW} in (10) implies that $\Pr[G \mid \text{no success by } T^{PW}] \lambda \bar{w} = c$. Thus, (16) is equivalent to just $L \lambda \bar{w} > c$, or using (17), to

$$\lambda \bar{w} \sum_{n=1}^{N-1} \frac{\Pr[n \text{ opponents succeeded by } T^{PW} \mid G]}{\Pr[\text{at least one opponent succeeded by } T^{PW} \mid G]} \left(\frac{1}{n+1} \right) > c. \quad (18)$$

The intuition for (18) is as follows. Assume all opponents are using a stopping time T^{PW} in a hidden ES contest. At T^{PW} , conditional on all opponents having failed, agent i is indifferent over his effort (by definition of T^{PW}). So, he strictly prefers to continue if and only if he strictly prefers to continue conditional on at least one opponent having succeeded. The left-hand side of (18) is i 's expected benefit from effort at T^{PW} conditional on some opponent having succeeded (which implies the state is good); the right-hand side is the cost.

When $N = 2$, condition (18) (and hence condition (16)) simplifies to

$$\lambda \frac{\bar{w}}{2} > c. \quad (19)$$

This condition is transparent: it says that an agent would continue experimenting if he knew his only opponent had already succeeded, in which case he infers the state is good but success will only earn him half the prize.

For $N > 2$, (19) is a necessary condition for the optimal hidden ES contest to dominate the optimal public WTA contest, but it is not sufficient. Inspecting (18), a simple sufficient condition is

$$\lambda \frac{\bar{w}}{N} \geq c. \quad (20)$$

This condition says that an agent would continue experimenting if he knew that all his opponents have succeeded. The example in [Figure 1](#) shows that condition (20) is not necessary, as there, hidden ES dominates public WTA for all $N \in \{2, \dots, 10\}$ even though (20) fails when $N > 5$. (Recall that, as shown in [Proposition 1](#), the probability of success in an optimal public WTA contest is independent of N ; hence, it is equal to that under an optimal hidden ES contest when $N = 1$.)

How do changes in parameters alter the principal’s choice between hidden ES and public WTA? An increase in p_0 decreases the left-hand side of (16), making the dominance of hidden ES less likely (in the sense of weakly shrinking the set of other parameters for which the inequality holds).²² This reinforces the intuition that the gains from using hidden ES stem from bolstering agents’ beliefs that the innovation may be feasible despite their own failures, which is more important to the principal when the prior is lower; were $p_0 = 1$, in which case there would be no learning, public WTA would be an optimal contest. Regarding parameters \bar{w} , λ , and c , the necessary and sufficient conditions (19) and (20) reveal that hidden ES will dominate (resp., be dominated by) public WTA when $\frac{c}{\lambda \bar{w}}$ is sufficiently small (resp., large).

The discussion above assumes a fixed number of agents. If the principal can instead choose the number of agents, then an optimal hidden ES contest always does at least as well as any public WTA contest. This is because the principal can replicate the public WTA outcome by setting $N = 1$ and using hidden ES. An implication of [Proposition 4](#) is that combining hidden ES with an optimally chosen $N > 1$ can be strictly better than using public WTA with any N ; this is the case if and only if condition (16) holds. Furthermore, as seen in [Figure 1](#), it can be optimal to set $N > 2$; this contrasts with a result of [Che and Gale \(2003\)](#) in a different contest environment.

3.4 Simple information disclosure policies

Moving beyond public and hidden disclosure, we next consider the class of contests with simple information disclosure policies: the principal specifies $\mathcal{T} \subseteq [0, T]$ such that the full history of outcomes is disclosed at each $t \in \mathcal{T}$ and nothing is disclosed at

²²To verify this, use (18): an increase in p_0 increases T^{PW} , which decreases the left-hand side of (18) because of a first-order stochastically dominant shift of the relevant probability distribution. Note that our analysis maintains the assumption that $p_0 < 1$. However, the current comparative static points to continuity with no learning: were $p_0 = 1$, a public WTA contest would always be optimal, as agents would work until a success is obtained; if condition (20) holds, then hidden ES would also be optimal.

$t \notin \mathcal{T}$.²³ The following result shows that an optimal contest in this class is a (possibly degenerate) *mixture* of the optimal contests under public and hidden disclosure.²⁴

Proposition 5. *Among contests with simple information disclosure policies, an optimal contest is a mixture contest that implements public winner-takes-all from time 0 until some time t_S and hidden equal-sharing from t_S until some time T , with $t_S \in [0, T]$. Moreover,*

1. *If $\frac{\lambda\bar{w}}{N} > c$, then $t_S = 0$ and $T = T^{HS}$ (so the contest is hidden equal-sharing),*
2. *If $\frac{\lambda\bar{w}}{2} < c$, then $t_S = T = T^{PW}$ (so the contest is public winner-takes-all).*

An intuition for the form of the optimal mixture contest stems from the earlier discussion about how changes in the prior alter the principal’s choice between public WTA and hidden ES: the latter is more beneficial when the agents’ beliefs are lower. The formal proof of [Proposition 5](#) is constructive, as follows. Take any contest $C = \{w(\cdot), \mathcal{T}, T\}$, and let t_C be the last time at which the principal discloses information in C , i.e. $t_C := \sup\{t : t \in \mathcal{T}\}$. As an agent’s belief at t_C , call it p_{t_C} , corresponds to the public belief, we show that there is a public WTA contest that induces experimentation until the public belief drops below p_{t_C} . Furthermore, because C has hidden disclosure from t_C on, [Proposition 3](#) implies that a hidden ES contest starting with prior p_{t_C} induces more experimentation than contest C does over $[t_C, T]$. Building on these two points, we construct a mixture contest that implements public WTA until the public belief reaches p_{t_C} and hidden ES from then on, and we show that this mixture contest dominates the original contest C .

[Proposition 5](#) provides simple sufficient conditions for either public WTA or hidden ES to be optimal among contests with simple disclosure policies; these conditions are intuitive given our discussion following [Proposition 4](#) in [Subsection 3.3](#). Plainly, for $N = 2$ it is always optimal to use either public WTA or hidden ES. When $\lambda\bar{w}/c \in (2, N)$, the optimal mixture contest can have a deadline T and a strictly interior switching time, $t_S \in (0, T)$. The intuition turns on the tradeoff between increasing an agent’s expected reward from success versus increasing his belief that he can succeed: as t_S increases, the agent’s belief about the innovation’s feasibility

²³The results of this section also apply if the principal randomizes over disclosure times, so long as randomization is independent of the history.

²⁴A virtue of the present analysis is that the principal does not need commitment power to implement an optimal contest. In particular, if an optimal mixture contest ends without success, the principal cannot benefit from then running another contest. On the other hand, if restricted to only use public or hidden contests, such “renegotiation” can be an issue.

from t_S on decreases, but his expected reward for success after t_S increases because in expectation the prize is shared with a smaller number of agents.

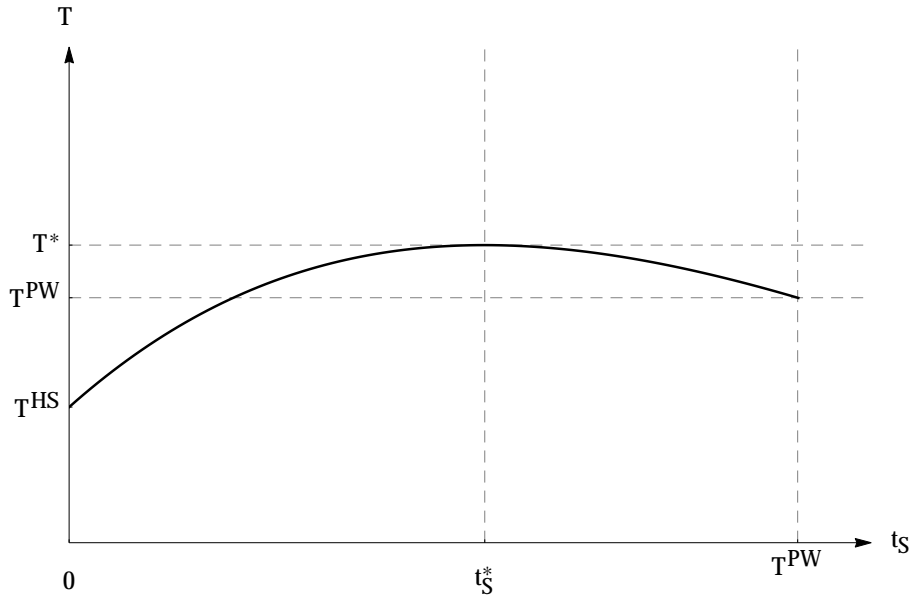


Figure 2 – Stopping time T as a function of switching time $t_S \in [0, T]$ in a mixture contest. Parameters are $p_0 = 0.9$, $c = 0.18$, $\bar{w} = 1$, $\lambda = 0.4$, and $N = 3$.

Figure 2 presents an example in which neither a public WTA nor a hidden ES contest is the optimal mixture contest. Given any switching time, t_S , the graph shows the (minimum) optimal deadline, T —equivalently, the agents’ stopping time—in a mixture contest. When $t_S = T$, the contest is public WTA (so $T = T^{PW}$), while $t_S = 0$ corresponds to hidden ES (so $T = T^{HS}$). It is evident that for the given parameters, the optimal mixture contest has a stopping time $T^* > \max\{T^{PW}, T^{HS}\}$ and a strictly interior switching time $t_S^* \in (0, T^*)$.

4 Extensions

4.1 Other information disclosure policies

As a general matter, the principal can improve upon simple information disclosure policies. This is true even if attention is restricted to disclosure that is non-stochastic and common in the sense that all agents are provided with the same information. Specifically, we demonstrate in the [Supplementary Appendix](#) that the principal can sometimes induce more experimentation by using a “cutoff disclosure” policy: for

some $1 \leq m < N$, the principal discloses at each time $t \in [0, T]$ only whether more than m agents have succeeded by t .

The complexity of agents' strategic interaction combined with learning precludes us from providing results about optimal information disclosure in full generality.²⁵ A plausible conjecture is that even allowing for an arbitrary disclosure policy, the sufficient conditions of Proposition 5 generalize: if $\frac{\lambda\bar{w}}{N} > c$, then hidden ES is an optimal contest; if $\frac{\lambda\bar{w}}{2} < c$, public WTA is an optimal contest.

Given the salience and widespread use of WTA contests, it is worth noting that:

Proposition 6. *An optimal public winner-takes-all contest is optimal among winner-takes-all contests with any disclosure policy.*

The logic is as discussed earlier: the principal only cares about agents' effort incentives following a history of no successes; hiding any information about this history in a WTA contest reduces effort incentives because when an agent conditions on some opponent having succeeded, his expected reward for success is zero.

4.2 Observability of success

Our model has posited that a success is observable to both the agent who succeeds and to the principal. We now consider what happens if only the principal or only the agent directly observes a success and can choose whether and when to verifiably reveal it to the other party. In both cases, we will see that a hidden ES contest dominates a public WTA contest under the same conditions as in our baseline model. We also discuss agents' ability to verifiably reveal a success to their opponents and the conditions under which our results continue to hold in this case.

²⁵It may be tempting to suggest that stochastic disclosure can always be used to improve on hidden ES, as follows. Suppose the sharing rule is ES and information is hidden until T^{HS} . At T^{HS} and beyond, the principal randomizes such that a signal (common to all agents) is only sent with some (non-stationary) probability when no-one has succeeded; if someone has succeeded, no signal is sent. Then, observing no signal will increase agents' beliefs and they would be willing to experiment beyond T^{HS} . (Cf. Ely (2014).)

The problem, however, is that this reasoning overlooks the dynamic strategic interaction. Specifically, in a hidden ES contest, cumulative efforts are strategic substitutes; hence, anticipating that his opponents may work longer than T^{HS} , an agent will not want to work early on because his expected share of the prize from success before T^{HS} is now lower. Furthermore, the possibility of learning at T^{HS} and beyond from his opponents' outcomes will also reduce an agent's incentive to exert effort before T^{HS} , just as in a public contest with a shared prize scheme.

Only principal observes success. Suppose that the principal observes an agent’s success but the agent does not. The principal can choose when to reveal a success to the agent. This scenario is relevant, for example, to the Netflix contest: there, contestants had to submit an algorithm whose performance was evaluated by Netflix on a proprietary “qualifying dataset” to determine whether the 10% improvement target had been achieved.²⁶ We suppose that if a success is obtained, the principal must reveal it by the end of the contest; she cannot harness the innovation’s benefits without paying out the prize. We now interpret the arguments of the prize-sharing scheme $w(s_i, s_{-i})$ as the times at which the principal reveals agents’ successes.

It is readily seen that our results extend to this setting. As the principal only values one success and does not benefit from saving prize money, she has no incentive to not reveal a success immediately to an agent who succeeds. At each time t , an agent conditions on not having obtained a success by t unless the principal has revealed otherwise. Consequently, the analysis of [Section 3](#) applies without change. Indeed, verifiable revelation by the principal is not essential: the same outcomes can be supported even if the principal is only able to make cheap-talk or unverifiable statements about an agent’s success.

Only agent observes success. Suppose next that the principal does not observe success directly; rather, any agent who succeeds can choose voluntarily when to verifiably reveal his success to the principal. This assumption is obviously relevant for many applications. The payments $w(s_i, s_{-i})$ are now interpreted as a function of the times at which agents reveal their success.

It is weakly dominant for an agent to immediately reveal his success in a WTA contest. Immediate revelation to the principal is also optimal in a hidden ES contest, as an agent’s expected reward for success is independent of when he reveals success. Given the analysis in [Section 3](#), it follows that in both public WTA and hidden ES contests, there exist symmetric equilibria where agents follow stopping strategies and reveal their successes immediately, inducing the same outcome as when both the principal and the agent directly observe success. Naturally, verifiability is important here; the same outcome cannot be obtained with cheap talk by the agents.

²⁶Netflix made available a “training dataset” for contestants’ use; we set aside considerations such as a contestant updating about the likelihood of his algorithm succeeding on the qualifying dataset based on its performance on the training dataset.

Agents can reveal success to opponents. Lastly, suppose agents can verifiably reveal their success to other agents. Would they have the incentive to do so? While the issue is moot in public contests, it is paramount in hidden contests, because such an incentive would unravel the principal’s desire to keep successes hidden.

In a hidden ES contest, a successful agent wants to deter opponents from continuing experimenting so that he can secure a larger share of the prize. Revealing a success has two opposing effects: it makes opponents more optimistic about the innovation’s feasibility but decreases their expected prize shares from their own success. An agent’s incentive to reveal that he has succeeded (so long as no other agent has already done so) will thus generally involve a tradeoff,²⁷ and the tradeoff’s resolution could potentially harm the principal. However, if condition (20) holds, the resolution is unambiguous: revealing a success always increases experimentation by other agents. Therefore, that sufficient condition for hidden ES to be optimal also ensures that agents have no incentives to reveal their success to opponents, and hence the principal can indeed implement hidden ES when (20) holds.

The foregoing discussion presumes that an agent can verifiably reveal a success to his opponents without actually making the innovation available to them. In many contexts, this would be difficult, however; for example, a contestant in the Netflix contest could probably not prove that he has succeeded without sharing his algorithm. Clearly, if verifiable revelation implies sharing the innovation, then an agent would never reveal a success to his opponents in a hidden ES contest. Moreover, revelation cannot be credibly obtained if messages are cheap talk.

4.3 Endogenizing the prize

We have studied contest design for a principal who has a fixed budget of prize money and whose objective is to maximize the probability of a success: the principal does not gain from saving any of the budget, or more broadly, does not value money. Trivially, if the value of the innovation is larger than the principal’s budget, i.e. $v \geq \bar{w}$, our problem is equivalent to one where the principal does value money (quasi-linearly, as usual) but is restricted to spend the whole budget \bar{w} upon a success.

Suppose now the principal values money and can choose the size of the prize.

²⁷Once an agent reveals success, all other successful agents will have strict incentives to reveal too: with the uncertainty about the innovation’s feasibility resolved, the only effect of revelation is to lower opponents’ expected prize shares.

That is, \bar{w} is now optimally chosen by the principal, but we maintain the assumption that the full amount \bar{w} must be paid to the agents if there is at least one success. Formally, the principal chooses \bar{w} and a contest to maximize (the expectation of)

$$(v - \bar{w})p_0 \left(1 - e^{-\lambda A^T}\right). \quad (21)$$

Plainly, this problem can be solved in two stages: first, for any given \bar{w} , solve for the optimal contest; second, given the optimal contest as a function of \bar{w} , solve for the optimal prize \bar{w} . The analysis of [Section 3](#) corresponds to the first stage and shows that, for any prize \bar{w} , the optimal contest (within the class of simple disclosure policies) is a mixture contest. Analyzing the second stage yields:

Proposition 7. *Fix any set of parameters $\{p_0, \lambda, c, N\}$. When the value of the innovation, v , is large (resp., small) enough, the principal maximizes (21) by choosing some prize \bar{w} and a hidden equal-sharing (resp., public winner-takes-all) contest.*

The logic is simple: the larger is v , the larger the gains to the principal from inducing more experimentation, and hence the larger is the prize the principal will choose. For v large enough, the principal optimally chooses a prize \bar{w} large enough that $\frac{\lambda \bar{w}}{N} > c$; for v small enough, the optimal prize is small enough that $\frac{\lambda \bar{w}}{2} < c$. The result then follows from [Proposition 5](#).

5 Concluding Remarks

This paper has studied contest design for specific innovations in an environment with learning, using the exponential-bandit framework of experimentation. In a nutshell, our main result is that a winner-takes-all contest in which any successful innovation is disclosed immediately (“public WTA”) is often dominated by a contest in which no information is disclosed until a deadline at which all successful agents equally share the prize (“hidden ES”). Within a class of salient disclosure policies, an appropriately-crafted mixture of these two contests is always optimal; simple sufficient conditions guarantee optimality of either public WTA or hidden ES.

Although our formal analysis is within the confines of a particular model, we believe the underlying intuition—a tradeoff between the reward an agent expects to receive should he succeed and his belief about the likelihood of success—is valid more generally. Our work suggests that the common default assumption (for theory) or prescription (for policy) of using WTA schemes deserves further scrutiny

when the feasibility of the innovation is uncertain and successful innovations are not automatically public information. In particular, alternative patent schemes that implement some version of “sharing the prize” may be warranted for certain kinds of R&D. On the other hand, in contexts where innovations are publicly observable, our analysis implies that a WTA contest is optimal; note though that the principal would sometimes be willing to pay a cost to alter the observability structure.

We conclude with some additional observations.

Socially efficient experimentation. We have focused on contest design for a principal who does not internalize agents’ effort costs. But our analysis also implies that public WTA contests can be dominated by a hidden ES contest even for a social planner who does internalize these costs. As noted in [Remark 2](#), a public WTA contest can implement the first-best solution if and only if the social value of a success, v , is no larger than the total prize available, \bar{w} . Thus, when condition (16) holds—so that an optimal hidden ES contest induces a later stopping time than any public WTA contest—and when v is sufficiently larger than \bar{w} , hidden ES will be preferred to public WTA net of effort costs: even though hidden ES induces wasteful effort after the innovation is first obtained, it increases agents’ incentives to experiment. It is plausible—even likely—that in various circumstances, the social value of innovation is substantially larger than the prize available to a contest designer, e.g. for medical innovations or scientific discoveries.

Discounting. For tractability, our analysis has ignored discounting. From a robustness perspective, small discounting frictions would not qualitatively alter our main points. More broadly, however, one advantage of public WTA contests is that terminating the contest when the first success is obtained does not affect agents’ incentives; by contrast, hidden ES contests must be run until their deadline. This suggests that if the principal cannot profit from an agent’s innovation until the end of a contest, then discounting would introduce a force in favor of using public WTA. A related point is that when there is discounting—or flow costs to prolonging contests—hidden ES contests do rely on commitment power.

Multistage contests. Innovations sometimes require a sequence of successes. For simplicity, suppose the principal only gains the profit v when an agent successfully completes two tasks or stages: first I , for intermediate, and then F , for final. Success in each stage $k \in \{I, F\}$ is uncertain and determined as in our baseline model, but

with a stage-specific state θ^k . The states are independently distributed (for simplicity, but this can be relaxed) with respective prior probabilities $p_0^k := \Pr(\theta^k = G)$.

Our results can be applied to contest design in this two-stage setting. Specifically, one can show that if p_0^F is sufficiently large—large enough that it would be optimal to use a public WTA contest were F the only stage (given any number of agents no larger than N)—then there is an optimal contest that has the following form. The principal initially runs a mixture contest during which agents can work on stage I until some pre-specified time T ; at time T , at most one agent is selected to advance and can work on stage F ; the entire prize \bar{w} is given to this agent if he succeeds on F . The advancing agent is either the first agent to succeed on I during the mixture contest’s public disclosure phase, or, if no-one succeeded in that phase, the advancing agent is determined by a uniform randomization among the successful agents in the hidden disclosure phase.²⁸ Intuitively, the “continuation value” from being selected for the second stage acts like the total prize for the first stage. As in our baseline model, the optimal mixture contest for stage I may be degenerate: public WTA or hidden ES. For further analysis of multistage settings, we refer readers to the contemporaneous work of [Bimpikis, Ehsani, and Mostagir \(2014\)](#).

Endogenous entry. We have taken the number of contestants, N , to be fixed (or to be specified by the principal). An alternative that may be more appropriate for some contexts would be to consider a fixed entry cost—any contestant must incur some cost to either register for the contest or to get started with experimentation—and endogenously determine N through free entry. We suspect our main themes would extend to such a specification, and hope this will be addressed by future research. Naturally, the entry cost itself may also be set endogenously by the principal, even if her objective is only to maximize the probability of a success.

Heterogeneity. It would also be interesting to incorporate heterogeneity among agents into the current framework. For example, agents may be privately informed about their “ability” (λ) or their cost of effort (c). While introducing the latter is likely to have intuitive effects—agents with higher c stop experimenting sooner—the former would be more subtle because of the countervailing effects on stopping times discussed following [Proposition 1](#).

²⁸If no agent succeeded on stage I , then no-one advances.

A Appendix: Proofs

A.1 Proof of Proposition 1

The first part of the result was proven in the text. The comparative statics of T^{PW} in the parameters are obtained through straightforward manipulation of equation (10); see the [Supplementary Appendix](#) for derivations.

The probability of obtaining a success in an optimal public WTA contest is given by expression (5) with $A^{TPW} = NT^{PW}$. By [Remark 1](#), the comparative statics for the probability of success follow immediately from the left-hand side (LHS, hereafter) of (9) being increasing in c , decreasing in λ and \bar{w} , and independent of p_0 and N .

A.2 Proof of Proposition 2

The first part of the proposition is proven in the text (see the discussion in [fn. 21](#) for the first sentence of the proposition). We prove the comparative statics in four steps.

Step 1. For any $\kappa \in [0, 1]$ and $N \geq 2$, we claim $\sum_{n=0}^{N-1} \left(\frac{1}{n+1}\right) \binom{N-1}{n} (1-\kappa)^n \kappa^{N-1-n} = \frac{1-\kappa^N}{(1-\kappa)N}$, which immediately implies that the expression in square brackets in (14) is equivalent to $\frac{1-e^{-N\lambda T^{HS}}}{(1-e^{-\lambda T^{HS}})^N}$ for any $N \geq 1$. The claim is verified as follows:

$$\begin{aligned} \sum_{n=0}^{N-1} \left(\frac{1}{n+1}\right) \binom{N-1}{n} (1-\kappa)^n \kappa^{N-1-n} &= \sum_{n=0}^{N-1} \frac{(N-1)! (1-\kappa)^n \kappa^{N-1-n}}{(N-1-n)! n! (n+1)} \\ &= \frac{1}{(1-\kappa)N} \sum_{n=0}^{N-1} \binom{N}{n+1} (1-\kappa)^{n+1} \kappa^{N-(n+1)} \\ &= \frac{1}{(1-\kappa)N} \left(1 - \binom{N}{0} (1-\kappa)^0 \kappa^N\right) = \frac{1-\kappa^N}{(1-\kappa)N}. \end{aligned}$$

Step 2. Letting $\kappa := e^{-\lambda T^{HS}} \in (0, 1)$ and $q := \frac{1-p_0}{p_0} > 0$, rewrite (15) as

$$\frac{1-\kappa^N}{(1-\kappa)N} \frac{\kappa}{\kappa+q} = \frac{c}{\lambda \bar{w}}. \quad (22)$$

We show that the LHS of (22) is increasing in κ . We compute $\frac{\partial \left(\frac{1-\kappa^N}{(1-\kappa)N}\right)}{\partial \kappa} = \frac{1-\kappa^N - N\kappa^{N-1}(1-\kappa)}{(1-\kappa)^2 N}$, which is positive if and only if

$$1 \geq N\kappa^{N-1} - (N-1)\kappa^N. \quad (23)$$

Differentiation shows that the right-hand side (RHS, hereafter) of (23) is increasing in κ since $\kappa \in (0, 1)$. As (23) holds with $\kappa = 1$, it follows that $\frac{\partial\left(\frac{1-\kappa^N}{(1-\kappa)^N}\right)}{\partial\kappa} > 0$. Moreover, $\frac{\kappa}{\kappa+q}$ is also increasing in κ because $q > 0$; hence the LHS of (22) is increasing in κ .

A similar argument establishes that the LHS of (22) is decreasing in N .

Step 3. Using Step 2, we now derive the comparative statics of T^{HS} . As the LHS of (22) is increasing in κ and κ is decreasing in T^{HS} , the LHS of (22) is decreasing in T^{HS} . Moreover, (i) the RHS of (22) is increasing in c and decreasing in \bar{w} while the LHS is independent of these parameters, and (ii) q is decreasing in p_0 and thus the LHS of (22) is increasing in p_0 , while the RHS is independent of this parameter. We thus obtain that T^{HS} is increasing in p_0 and \bar{w} and decreasing in c . Similarly, because the LHS of (22) is decreasing in both N and T^{HS} and the RHS is independent of both parameters, we obtain that T^{HS} is decreasing in N .

Lastly, we show that T^{HS} is non-monotonic with respect to λ by providing an example. Let $p_0 = \frac{1}{2}$ and $N = 2$. Then (15) becomes

$$\frac{1 - e^{-2\lambda T^{HS}}}{2(1 - e^{-\lambda T^{HS}})} \frac{e^{-\lambda T^{HS}}}{e^{-\lambda T^{HS}} + 1} = \frac{c}{\lambda \bar{w}},$$

which simplifies to $T^{HS} = \frac{1}{\lambda} \log\left(\frac{\lambda \bar{w}}{2c}\right)$. Differentiating, $\frac{\partial T^{HS}}{\partial \lambda} = \frac{1}{\lambda^2} (1 - \log\left(\frac{\lambda \bar{w}}{2c}\right))$, which is positive for small λ (i.e., $p_0 \lambda \bar{w} \approx c$) and negative for large λ .

Step 4. We can now show the comparative statics for the probability of obtaining a success. The probability of success in an optimal hidden equal-sharing contest is given by expression (5) with $A^{T^{HS}} = NT^{HS}$. The comparative static with respect to c is immediate: as shown in Step 3, if c increases, T^{HS} decreases, which implies that NT^{HS} and thus the probability of success decreases. An analogous argument shows that the probability of success is increasing in \bar{w} and p_0 .

We next show that the probability of success increases with λ . From Step 3, T^{HS} may increase or decrease when λ increases. However, note that λT^{HS} must increase when λ increases: if λT^{HS} decreases, the LHS of (15) increases, while the RHS decreases when λ increases, leading to a contradiction. Therefore, λT^{HS} increases with λ , implying that the probability of obtaining a success increases with λ .

Finally, the ambiguous effect of an increase in N on the probability of success is seen through the example reported in Figure 1.

A.3 Proof of Proposition 3

We analyze symmetric equilibria of hidden contests. We will use the following:

Lemma 1. *Take any symmetric equilibrium of a hidden contest with prize scheme $w(s_i, \mathbf{s}_{-i})$ and optimal deadline T . If this equilibrium does not have full effort by all agents from 0 to T , there is another scheme $w'(\cdot)$ with an optimal deadline T' that has a symmetric equilibrium in which each agent exerts full effort (so long as he has not succeeded) from 0 until T' , and where the aggregate cumulative effort is the same as under scheme $w(\cdot)$ and deadline T .*

Proof. See the [Supplementary Appendix](#). ■

We proceed in three steps.

Step 1. Consider a hidden contest with prize scheme $w(s_i, \mathbf{s}_{-i})$ and associated optimal deadline T as defined in Lemma 1. Given full effort from 0 to T , the contest induces a sequence of expected rewards for success at each time $t \in [0, T]$ as shown in the proof of Lemma 1. We show that any hidden contest that induces a constant expected reward sequence is weakly dominated by an optimal hidden ES contest. Note that for a constant expected reward sequence, the principal's budget constraint implies that for all $t \in [0, T]$,

$$w_{i,t} = \mathbb{E}_{s_{-i}} w(t, \mathbf{s}_{-i}) = \mathbb{E}_n \left[\frac{\bar{w}}{n} | n \geq 1, T \right]. \quad (24)$$

Consider an optimal hidden ES contest. The induced sequence of expected rewards for success is constant with $w_{i,t} = w^{HS} = \mathbb{E}_n \left[\frac{\bar{w}}{n} | n \geq 1, T^{HS} \right]$ for all $t \in [0, T^{HS}]$, where

$$\frac{p_0 e^{-\lambda T^{HS}}}{p_0 e^{-\lambda T^{HS}} + 1 - p_0} \lambda w^{HS} = c.$$

Note that T^{HS} and w^{HS} are unique.

Suppose there exists a hidden contest with associated optimal deadline T (as defined in Lemma 1) that induces a constant sequence of expected rewards and has $T > T^{HS}$. By (24), the induced sequence must have $w_{i,t} = \mathbb{E}_n \left[\frac{\bar{w}}{n} | n \geq 1, T \right]$ for all $t \in [0, T]$. But then if $T > T^{HS}$, the expected number of successful agents given full effort from 0 until T is larger than that under deadline T^{HS} , and hence the expected reward for success is $w_{i,t} < w^{HS}$. It follows that $p_{i,T} \lambda w_{i,T} < c$, but then agents do not exert effort continuously from 0 until T , a contradiction.

Step 2. We show that in an optimal hidden ES contest, the agent's incentive constraint for effort is binding at each time $t \in [0, T^{HS}]$.

An agent i 's continuation payoff at any time $t' \in [0, T]$ is

$$U_{t'} := p_{i,t'} \int_{t'}^T (w_{i,t} \lambda - c) a_{i,t} e^{-\int_{t'}^t \lambda a_{i,z} dz} dt - (1 - p_{i,t'}) c \int_{t'}^T a_{i,t} dt.$$

Consider the continuation payoff at t' from a strategy $\tilde{a}_{i,t} = 1 \forall t \in [0, T] \setminus (t', t' + \varepsilon)$ and 0 otherwise:

$$U_{t'}(\varepsilon) := p_{i,t'} \int_{t'+\varepsilon}^T (w_{i,t} \lambda - c) e^{-(t-t'-\varepsilon)\lambda} dt - (1 - p_{i,t'}) c (T - t' - \varepsilon).$$

We compute

$$\begin{aligned} U'_{t'}(\varepsilon) &= p_{i,t'} \left[- (w_{i,t'+\varepsilon} \lambda - c) + \int_{t'+\varepsilon}^T (w_{i,t} \lambda - c) e^{-(t-t'-\varepsilon)\lambda} \lambda dt \right] + (1 - p_{i,t'}) c, \\ U'_{t'}(0) &= - (p_{i,t'} w_{i,t'} \lambda - c) + p_{i,t'} \int_{t'}^T (w_{i,t} \lambda - c) e^{-(t-t')\lambda} \lambda dt. \end{aligned}$$

Note that $U_{t'}(0) \geq U_{t'}(\varepsilon)$ for any $\varepsilon \geq 0$ and thus $U'_{t'}(0) \leq 0$ because of Nash equilibrium. If $w_{i,t} = w_i$ for all t , the expression above simplifies to

$$\begin{aligned} U'_{t'}(0) &= - (p_{i,t'} w_i \lambda - c) + p_{i,t'} \int_{t'}^T (w_i \lambda - c) e^{-(t-t')\lambda} \lambda dt \\ &= - (p_{i,t'} w_i \lambda - c) + p_{i,t'} (w_i \lambda - c) \left(1 - e^{-(T-t')\lambda} \right) \\ &= c(1 - p_{i,t'}) - p_{i,t'} (w_i \lambda - c) e^{-(T-t')\lambda}. \end{aligned}$$

In an optimal hidden ES contest, $T = T^{HS}$, $w_i = w^{HS}$, and $p_{i,T^{HS}} w^{HS} \lambda = c$; hence

$$p_{i,t'} e^{-(T^{HS}-t')\lambda} (w^{HS} \lambda - c) = c(1 - p_{i,t'}).$$

Therefore, we obtain $U'_{t'}(0) = 0$, showing that each agent's incentive constraint is binding at each time $t' \leq T^{HS}$ in an optimal hidden ES contest.

Step 3. We show that any hidden contest that induces a non-constant sequence of expected rewards is weakly dominated by hidden ES. Suppose $\{w_{i,t}\}_{t \in [0, T]}$ is non-

constant. The incentive constraint for agent i at any $z \in [0, T]$ is given by

$$U'_z(0) = -(p_{i,z}w_{i,z}\lambda - c) + p_{i,z}\lambda \int_z^T (w_{i,t}\lambda - c) e^{-(t-z)\lambda} dt \leq 0. \quad (25)$$

Suppose, to contradiction, $T > T^{HS}$. For any $z \leq T^{HS} < T$, rewrite (25) as

$$U'_z(0) = -(p_{i,z}w_{i,z}\lambda - c) + p_{i,z}\lambda \left[\int_z^{T^{HS}} (w_{i,t}\lambda - c) e^{-(t-z)\lambda} dt + \int_{T^{HS}}^T (w_{i,t}\lambda - c) e^{-(t-z)\lambda} dt \right] \leq 0. \quad (26)$$

By Step 2, it holds that in an optimal hidden ES contest, $U'_z(0) = 0$ at any time $z \leq T^{HS}$. Subtracting this binding constraint from (26), we obtain for any $z \leq T^{HS}$:

$$-p_{i,z} (w_{i,z} - w^{HS}) \lambda + p_{i,z}\lambda \left[\int_z^{T^{HS}} (w_{i,t} - w^{HS}) \lambda e^{-(t-z)\lambda} dt + \int_{T^{HS}}^T (w_{i,t}\lambda - c) e^{-(t-z)\lambda} dt \right] \leq 0.$$

Define

$$t' = \sup \left\{ t : \left\{ z \in \left(t - \frac{1}{m}, t \right) : w_{i,z} \leq w^{HS} \right\} \text{ has positive Lebesgue measure for any } m > 0 \right\}.$$

Such a t' is well-defined because otherwise $w_{i,t} > w^{HS}$ almost everywhere, which contradicts the budget constraint.

Consider first the case where $t' \leq T^{HS}$. Then for any m , consider any $z \in \left(t' - \frac{1}{m}, t' \right) : w_{i,z} \leq w^{HS}$. Define

$$\begin{aligned} \Delta_z &:= -p_{i,z} (w_{i,z} - w^{HS}) \lambda + p_{i,z}\lambda \int_z^{t'} (w_{i,t} - w^{HS}) \lambda e^{-(t-z)\lambda} dt \\ &\quad + p_{i,z}\lambda \int_{t'}^{T^{HS}} (w_{i,t} - w^{HS}) \lambda e^{-(t-z)\lambda} dt + p_{i,z}\lambda \int_{T^{HS}}^T (w_{i,t}\lambda - c) e^{-(t-z)\lambda} dt. \end{aligned}$$

Note that $-p_{i,z} (w_{i,z} - w^{HS}) \lambda \geq 0$ and $\int_{t'}^{T^{HS}} (w_{i,t} - w^{HS}) \lambda e^{-(t-z)\lambda} dt \geq 0$ by the definition of t' . Moreover,

$$\left| \int_z^{t'} (w_{i,t} - w^{HS}) \lambda e^{-(t-z)\lambda} dt \right| \leq 2\lambda\bar{w} (t' - z)$$

because the budget constraint implies $|w_{i,t} - w^{HS}| \leq 2\bar{w}$. Finally, note that $\gamma := \int_{T^{HS}}^T (w_{i,t}\lambda - c) e^{-(t-z)\lambda} dt > 0$, because the incentive constraint implies that for any t , $p_{i,t}w_{i,t}\lambda - c \geq 0$, and $p_{i,t} < 1$. Now take $\frac{1}{m} < \frac{\gamma}{2\lambda\bar{w}}$. Then $\Delta_z > 0$, a contradiction.

Consider next the case where $t' > T^{HS}$. Then for m sufficiently large and $z \in \{z \in (t' - \frac{1}{m}, t') : w_{i,z} \leq w^{HS}\}$,

$$U'_z(0) = -(p_{i,z}w_{i,z}\lambda - c) + p_{i,z}\lambda \int_z^T (w_{i,t}\lambda - c) e^{-(t-z)\lambda} dt > 0.$$

The first term is strictly positive because $p_{i,T^{HS}}w^{HS}\lambda - c = 0$, $p_{i,z} < p_{i,T^{HS}}$, and $w_{i,z} \leq w^{HS}$. The second term is strictly positive for the same reasons as before. Thus, again, we reach a contradiction.

A.4 Proof of Proposition 4

Proposition 1 shows that in an optimal public WTA contest, agents follow a stopping strategy with stopping time T^{PW} given by (10). Proposition 2 shows that in an optimal hidden ES contest, agents follow a stopping strategy with stopping time T^{HS} given by (15). Hence, the principal strictly prefers hidden ES to public WTA if and only if $T^{HS} > T^{PW}$. Using (10) and (15), this condition is equivalent to (16) as shown in the text.

A.5 Proof of Proposition 5

We proceed in two steps, first showing that a mixture contest is optimal and then verifying the sufficient conditions for the optimality of public WTA and hidden ES.

Step 1. Consider an arbitrary contest C with prize scheme $w(s_i, s_{-i})$, simple information disclosure policy \mathcal{T} , and deadline T . We focus on symmetric equilibria and, without loss of generality, define the deadline T so that agents exert positive effort at T given no success by T . The aggregate cumulative effort up to T induced by contest C (given no success by T) is A^T . We want to show that there exist $t_S^* \geq 0$, $T^* \geq t_S^*$, and a mixture contest that implements public WTA from 0 until t_S^* and hidden ES from t_S^* until T^* such that the aggregate cumulative effort by T^* (given no success by T^*) induced by this contest is $A^* \geq A^T$. This mixture contest will therefore dominate C .

Suppose for the purpose of contradiction that $A^* < A^T$ for all $t_S^* \geq 0$ and $T^* \geq t_S^*$. This implies $A^{PW} < A^T$ and $A^{HS} < A^T$, where A^{PW} is the aggregate cumulative effort induced in an optimal public WTA contest and A^{HS} is the aggregate cumulative effort induced in an optimal hidden ES contest.

Let t_C be the last time at which information is disclosed in contest C , i.e. $t_C = \max\{t : t \in \mathcal{T}\}$. Consider a history of no success. The agents' belief at t_C corresponds

to the public belief; denote this belief by p_{t_C} . Since agents are willing to exert positive effort at some point $t \in [t_C, T]$ (recall that, without loss, T is defined so that agents exert positive effort at T given no success by T), and an agent's reward for success at any such point cannot be strictly larger than \bar{w} (by the principal's budget constraint), we must have $p_{t_C} \geq \frac{c}{\lambda \bar{w}}$. It follows that there exists a public WTA contest that induces full effort by each agent until the public belief reaches $p^{PW} \leq p_{t_C}$. Let t^* be the time at which the belief reaches p_{t_C} in a public WTA contest and denote by \tilde{A}^{PW} the aggregate cumulative effort induced by this contest over $[0, t^*]$.

Next, note that given the public belief p_{t_C} at t_C , the continuation equilibrium in contest C has hidden disclosure. Consider an optimal hidden ES contest starting with a prior belief $p_0 = p_{t_C}$ and denote by \tilde{A}^{HS} the total cumulative effort induced by such a contest. By [Proposition 3](#), \tilde{A}^{HS} is weakly larger than the aggregate cumulative effort induced by contest C over $[t_C, T]$.

We show that a mixture contest with switching time t^* induces aggregate cumulative effort $A^* = \tilde{A}^{PW} + \tilde{A}^{HS}$, and hence it dominates contest C . Let $p_{i,t}$ be agent i 's belief at time t , where this belief is updated given public disclosure from time 0 until t^* and hidden disclosure from t^* on. Denote by w the agent's expected reward for success at any time $t \geq t^*$ given no success by time t^* , and let $A_{-i,z}$ denote (i 's conjecture of) the aggregate effort exerted by i 's opponents at time z so long as they have not succeeded by z . The agent's problem is:

$$\begin{aligned} \max_{(a_{i,t})_{t \in [0, T^*]}} & \int_0^{t^*} (\bar{w} p_{i,t} \lambda - c) a_{i,t} e^{-\int_0^t p_{i,z} \lambda (a_{i,z} + A_{-i,z}) dz} dt \\ & + e^{-\int_0^{t^*} p_{i,z} \lambda (a_{i,z} + A_{-i,z}) dz} \int_{t^*}^{T^*} (w p_{i,t} \lambda - c) a_{i,t} e^{-\int_{t^*}^t p_{i,z} \lambda a_{i,z} dz} dt. \end{aligned}$$

The belief $p_{i,t}$ is decreasing and the expected reward for success is also (weakly) decreasing because $w \leq \bar{w}$. Hence, an optimal strategy for agent i is a stopping strategy where $a_{i,t} = 1$ if $\bar{w} p_{i,t} \lambda \geq c$ and $a_{i,t} = 0$ otherwise for $t \leq t^*$, and $a_{i,t} = 1$ if $w p_{i,t} \lambda \geq c$ and $a_{i,t} = 0$ otherwise for $t > t^*$ (so long as the agent has not succeeded). It follows that if a public WTA contest induces \tilde{A}^{PW} until time t^* given no continuation game, it also induces \tilde{A}^{PW} until time t^* when the continuation game is a hidden ES contest. Finally, starting from t^* , the continuation game is the same as that under a hidden ES contest starting at time 0 with prior belief $p_0 = p_{t_C}$. Therefore, this mixture contest induces $A^* = \tilde{A}^{PW} + \tilde{A}^{HS}$ and dominates contest C .

Step 2. We show the sufficient conditions for the optimality of hidden ES and public

WTA given in the proposition. Consider a mixture contest with switching time t_S . The stopping time T for any agent i is given by:

$$\begin{aligned} \frac{c}{\lambda \bar{w}} &= \underbrace{\Pr[\text{some } j \neq i \text{ succ. in } [t_S, T] \mid \text{no one did by } t_S, i \text{ didn't by } T]}_{\alpha(t_S)} L_{[t_S, T]} \\ &+ \underbrace{\Pr[\text{no } j \neq i \text{ succ. in } [t_S, T] \mid \text{no one did by } t_S, i \text{ didn't by } T]}_{1-\alpha(t_S)} \Pr[G \mid \text{no succ. by } T], \end{aligned}$$

where

$$L_{[t_S, T]} := \sum_{n=1}^{N-1} \Pr[n \text{ opponents succ. btw } t_S \text{ and } T \mid \text{at least one did, no succ. by } t_S] \left(\frac{1}{n+1} \right).$$

We can rewrite this condition as

$$\begin{aligned} \frac{c}{\lambda \bar{w}} &= \underbrace{\frac{p_0 e^{-N\lambda t_S} e^{-\lambda(T-t_S)} (1 - e^{-(N-1)\lambda(T-t_S)})}{p_0 e^{-N\lambda t_S} e^{-\lambda(T-t_S)} + (1 - p_0)}}_{\alpha(t_S)} \underbrace{\frac{\frac{1 - e^{-\lambda(T-t_S)N}}{(1 - e^{-\lambda(T-t_S)})^N} - e^{-\lambda(N-1)(T-t_S)}}{1 - e^{-\lambda(T-t_S)(N-1)}}}_{L_{[t_S, T]}} \\ &+ \underbrace{\left[1 - \frac{p_0 e^{-N\lambda t_S} e^{-\lambda(T-t_S)} (1 - e^{-(N-1)\lambda(T-t_S)})}{p_0 e^{-N\lambda t_S} e^{-\lambda(T-t_S)} + (1 - p_0)} \right]}_{1-\alpha(t_S)} \underbrace{\frac{p_0 e^{-N\lambda T}}{p_0 e^{-N\lambda T} + (1 - p_0)}}_{\Pr[G \mid \text{no success by } T]}. \quad (27) \end{aligned}$$

Note that $\alpha(t_S)$ is decreasing in t_S ; $L_{[t_S, T]}$ is decreasing in T and increasing in t_S ; $\Pr[G \mid \text{no success by } T]$ is decreasing in T ; and the RHS of (27) is decreasing in T .

Suppose first that $\frac{\lambda \bar{w}}{2} < c$. Then for any t_S and T , $L_{[t_S, T]} < \frac{c}{\lambda \bar{w}}$. Given T , it follows that $\Pr[G \mid \text{no success by } T] > \frac{c}{\lambda \bar{w}}$, as otherwise (27) would not hold with equality. Consequently, if t_S increases, $(1 - \alpha(t_S))$ and $L_{[t_S, T]}$ increase and thus the RHS of (27) increases. This implies that T must increase when t_S increases (so that the RHS decreases and remains equal to the LHS), and therefore setting $t_S = T$ is optimal.

Suppose next that $\frac{\lambda \bar{w}}{N} > c$. Then for any t_S and T , $L_{[t_S, T]} > \frac{c}{\lambda \bar{w}}$. Given T , it follows that $\Pr[G \mid \text{no success by } T] < \frac{c}{\lambda \bar{w}}$, as otherwise (27) would not hold with equality. Note that a change in t_S now causes two opposing effects on the RHS of (27): on the one hand, reducing t_S increases $\alpha(t_S)$, which increases the RHS of (27), but on the other hand it reduces $L_{[t_S, T]}$, which reduces the RHS of (27). We show that if $\frac{\lambda \bar{w}}{N} > c$, the net effect of reducing t_S to zero on the RHS of (27) is positive, which implies that T must increase and therefore setting $t_S = 0$ is optimal.

To show this, note that by (27), $(1 - \alpha(t_S)) \Pr[G \mid \text{no success by } T] = \frac{c}{\lambda \bar{w}} - \alpha(t_S) L_{[t_S, T]}$,

and hence $\alpha(0)L_{[0,T]} + (1 - \alpha(0)) \Pr[G|\text{no success by } T]$ is equal to

$$\alpha(0)L_{[0,T]} + \frac{1 - \alpha(0)}{1 - \alpha(t_S)} \left(\frac{c}{\lambda\bar{w}} - \alpha(t_S)L_{[t_S,T]} \right). \quad (28)$$

We need to show that (28) is greater than $\frac{c}{\lambda\bar{w}}$. (28) can be rewritten as

$$\begin{aligned} & \frac{p_0 e^{-\lambda T}}{p_0 e^{-\lambda T} + 1 - p_0} \left(\frac{1 - e^{-\lambda TN}}{(1 - e^{-\lambda T})N} - e^{-\lambda T(N-1)} \right) + \frac{p_0 e^{-\lambda T - \lambda t(N-1)} + (1 - p_0) \frac{c}{\lambda\bar{w}}}{p_0 e^{-\lambda T} + 1 - p_0} \\ & - \frac{p_0 e^{-\lambda T - \lambda t(N-1)} + (1 - p_0)}{p_0 e^{-\lambda T} + 1 - p_0} \frac{p_0 e^{-\lambda T - \lambda t(N-1)}}{p_0 e^{-\lambda T - \lambda t(N-1)} + (1 - p_0)} \left(\frac{1 - e^{-\lambda N(T-t)}}{(1 - e^{-\lambda(T-t)})N} - e^{-\lambda(T-t)(N-1)} \right). \end{aligned}$$

Some algebra then shows that (28) is greater than $\frac{c}{\lambda\bar{w}}$ if and only if

$$\frac{1 - e^{-\lambda TN}}{(1 - e^{-\lambda T})N} - \frac{e^{-\lambda t(N-1)}(1 - e^{-\lambda N(T-t)})}{(1 - e^{-\lambda(T-t)})N} - \frac{c}{\lambda\bar{w}} (1 - e^{-\lambda t(N-1)}) \geq 0.$$

By assumption, $\frac{c}{\lambda\bar{w}} < \frac{1}{N}$; thus, again doing some algebra, it suffices to show that

$$\frac{1 - e^{-\lambda TN}}{(1 - e^{-\lambda T})} - \frac{e^{-\lambda t(N-1)} - e^{-\lambda NT + \lambda t}}{(1 - e^{-\lambda(T-t)})} - 1 + e^{-\lambda t(N-1)} \geq 0.$$

This inequality holds with equality when $t = 0$, and a routine computation verifies that the derivative of the LHS with respect to t is non-negative for all $N \geq 2$.

A.6 Proofs of Proposition 6 and Proposition 7

See the [Supplementary Appendix](#).

References

- AKCIGIT, U. AND Q. LIU (2014): "The Role of Information in Innovation and Competition," Unpublished.
- AOYAGI, M. (2010): "Information Feedback in a Dynamic Tournament," *Games and Economic Behavior*, 70, 242–260.
- BERGEMANN, D. AND U. HEGE (1998): "Venture Capital Financing, Moral Hazard, and Learning," *Journal of Banking & Finance*, 22, 703–735.
- (2005): "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 36, 719–752.
- BHATTACHARYA, S., J. GLAZER, AND D. SAPPINGTON (1990): "Sharing Productive Knowledge in Internally Financed R&D Contests," *Journal of Industrial Economics*, 39, 187–208.
- BIMPIKIS, K. AND K. DRAKOPOULOS (2014): "Disclosing Information in Strategic Experimentation," Unpublished.
- BIMPIKIS, K., S. EHSANI, AND M. MOSTAGIR (2014): "Information Provision in Dynamic Innovation Tournaments," Personal communication.
- BOBTCHEFF, C. AND R. LEVY (2010): "More Haste, Less Speed: Signalling through Investment Timing," Unpublished.
- BONATTI, A. AND J. HÖRNER (2011): "Collaborating," *American Economic Review*, 101, 632–663.
- CAMPBELL, A., F. EDERER, AND J. SPINNEWIJN (2014): "Delay and Deadlines: Freeriding and Information Revelation in Partnerships," Forthcoming in *American Economic Journal: Microeconomics*.
- CHE, Y.-K. AND I. GALE (2003): "Optimal Design of Research Contests," *American Economic Review*, 93, 646–671.
- CHE, Y.-K. AND J. HÖRNER (2013): "Optimal Design for Social Learning," Unpublished.
- CHOI, J. P. (1991): "Dynamic R&D Competition under 'Hazard Rate' Uncertainty," *RAND Journal of Economics*, 22, 596–610.
- CRIPPS, M. AND C. THOMAS (2014): "Strategic Experimentation in Queues," Unpublished.
- DASGUPTA, P. AND J. STIGLITZ (1980): "Uncertainty, Industrial Structure, and the Speed of R&D," *The Bell Journal of Economics*, 1–28.

- EASLEY, D. AND A. GHOSH (2013): "Incentives, Gamification, and Game Theory: An Economic Approach to Badge Design," in *Proceedings of the fourteenth ACM conference on Electronic commerce*, ACM, 359–376.
- EDERER, F. (2010): "Feedback and Motivation in Dynamic Tournaments," *Journal of Economics & Management Strategy*, 19, 733–769.
- ELY, J. (2014): "Beeps," Unpublished.
- ELY, J., A. FRANKEL, AND E. KAMENICA (2013): "Suspense and Surprise," Forthcoming in *Journal of Political Economy*.
- FUCHS, W. (2007): "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97, 1432–1448.
- FULLERTON, R. AND R. P. MCAFEE (1999): "Auctioning Entry into Tournaments," *Journal of Political Economy*, 107, 573–605.
- GILL, D. (2008): "Strategic Disclosure of Intermediate Research Results," *Journal of Economics & Management Strategy*, 17, 733–758.
- GOLTSMAN, M. AND A. MUKHERJEE (2010): "Interim Performance Feedback in Multistage Tournaments: The Optimality of Partial Disclosure," Unpublished.
- GOMES, R., D. GOTTLIEB, AND L. MAESTRI (2013): "Experimentation and Project Selection: Screening and Learning," Unpublished.
- HALAC, M., N. KARTIK, AND Q. LIU (2013): "Optimal Contracts for Experimentation," Unpublished.
- HÖRNER, J. AND L. SAMULESON (2014): "Incentives for Experimenting Agents," *RAND Journal of Economics*, 44, 632–663.
- JUDD, K., K. SCHMEDDERS, AND Ş. YELTEKIN (2012): "Optimal Rules for Patent Races," *International Economic Review*, 53, 23–52.
- KELLER, G. AND S. RADY (2010): "Strategic Experimentation with Poisson Bandits," *Theoretical Economics*, 5, 275–311.
- KELLER, G., S. RADY, AND M. CRIPPS (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73, 39–68.
- KIM, K. AND F. Z. X. LEE (2014): "Information Acquisition in a War of Attrition," *American Economic Journal: Microeconomics*, 6, 37–78.
- KLEIN, N. AND S. RADY (2011): "Negatively Correlated Bandits," *Review of Economic Studies*, 78, 693–732.
- KREMER, I., Y. MANSOUR, AND M. PERRY (2013): "Implementing the 'Wisdom of the Crowd'," Forthcoming in *Journal of Political Economy*.

- KRISHNA, V. AND J. MORGAN (1998): "The Winner-Take-All Principle in Small Tournaments," in *Advances in Applied Microeconomics: Contests*, ed. by M. R. Baye, Jai Pr, vol. 7, 61–74.
- KWON, S. (2013): "Dynamic Moral Hazard with Persistent States," Unpublished.
- LAZEAR, E. AND S. ROSEN (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *The Journal of Political Economy*, 89, 841–864.
- LIZZERI, A., M. MEYER, AND N. PERSICO (2002): "The Incentive Effects of Interim Performance Evaluations," Unpublished.
- LOURY, G. (1979): "Market Structure and Innovation," *Quarterly Journal of Economics*, 93, 395–410.
- MACCORMACK, A., F. MURRAY, AND E. WAGNER (2013): "Spurring Innovation through Competition," *MIT Sloan Management Review*, 55, 25–32.
- MALUEG, D. AND S. TSUTSUI (1997): "Dynamic R&D Competition with Learning," *RAND Journal of Economics*, 28, 751–772.
- MASON, R. AND J. VÄLIMÄKI (2010): "Learning in Contests," Unpublished.
- MCKINSEY & COMPANY (2009): "And the Winner Is...: Capturing the Promise of Philanthropic Prizes," Sydney, Australia: McKinsey & Company, 2009.
- MOLDOVANU, B. AND A. SELA (2001): "The Optimal Allocation of Prizes in Contests," *American Economic Review*, 542–558.
- MOSCARINI, G. AND L. SMITH (2011): "Optimal Dynamic Contests," Unpublished.
- MOSCARINI, G. AND F. SQUINTANI (2010): "Competitive Experimentation with Private Information: The Survivor's Curse," *Journal of Economic Theory*, 145, 639–660.
- MURTO, P. AND J. VÄLIMÄKI (2011): "Learning and Information Aggregation in an Exit Game," *The Review of Economic Studies*, 78, 1426–1461.
- RIECK, T. (2010): "Information Disclosure in Innovation Contests," Unpublished.
- SIEGEL, R. (2009): "All-Pay Contests," *Econometrica*, 77, 71–92.
- STIGLITZ, J. (2006): "Give Prizes not Patents," *New Scientist*, 2569, 20.
- STRULOVICI, B. (2010): "Learning While Voting: Determinants of Collective Experimentation," *Econometrica*, 78, 933–971.
- TAYLOR, C. (1995): "Digging for Golden Carrots: An Analysis of Research Tournaments," *American Economic Review*, 85, 872–890.
- WIRTZ, J. (2013): "Feedback and Learning in Tournaments," Unpublished.
- YILDIRIM, H. (2005): "Contests with Multiple Rounds," *Games and Economic Behavior*, 51, 213–227.

B Supplementary Material for Online Publication Only

B.1 Omitted proofs

B.1.1 Details for the proof of [Proposition 1](#)

The comparative statics of T^{PW} with respect to c and \bar{w} follow from the fact that the RHS of (10) is increasing in c and decreasing in \bar{w} , while the LHS is independent of these parameters and decreasing in T^{PW} . Similarly, the comparative statics of T^{PW} with respect to p_0 and N follow from the fact that the LHS of (10) is increasing in p_0 , decreasing in N and decreasing in T^{PW} , while the RHS is independent of these parameters. Finally, to compute the comparative static with respect to λ , note that (10) gives

$$T^{PW} = \frac{1}{\lambda N} \log \left(\frac{p_0}{1-p_0} \left(\frac{\lambda \bar{w}}{c} - 1 \right) \right),$$

and thus,

$$\frac{\partial T^{PW}}{\partial \lambda} = \frac{\lambda \bar{w} - (\lambda \bar{w} - c) \log \left(\frac{p_0(\lambda \bar{w} - c)}{c(1-p_0)} \right)}{\lambda^2 N (\lambda \bar{w} - c)},$$

where the logarithm in the numerator is non-negative because $p_0 \lambda \bar{w} \geq c$. Hence, $\frac{\partial T^{PW}}{\partial \lambda}$ is positive for λ small enough (i.e., $p_0 \lambda \bar{w} \approx c$) and negative if λ is large.

B.1.2 Details for the proof of [Proposition 3](#)

We provide here a proof for [Lemma 1](#).

First, without loss, take $T = \sup \{t : a_{i,t} > 0\}$. Next, suppose that each agent's effort $a_{i,t}$ is not constantly 1 over $[0, T]$. Let each agent's private belief at T be $p_{i,T}$. We choose a sub-interval $[0, T']$ such that each agent's private belief at T' conditional on no success before T' and all agents exerting full effort for the whole sub-interval is $p'_{i,T'} = p_{i,T}$. We find a prize scheme $w'(\cdot)$ such that exerting full effort for the whole sub-interval $[0, T']$ is a Nash equilibrium.

To this end, define a function $\tau : [0, T] \rightarrow \mathbb{R}_+$ by

$$\tau(z) = \int_0^z a_{i,t} dt. \tag{29}$$

Note that τ is weakly increasing. Take $T' = \tau(T)$. By convention, for any $t \in [0, T']$, we let $\tau^{-1}(t) = \inf \{z : \tau(z) = t\}$. Denote by $p'_{i,t}$ the private belief at time $t \in [0, T']$ under full effort. It is straightforward that for $t \in [0, T']$, $p'_{i,t} = p_{i,\tau^{-1}(t)}$.

We find $w'(\cdot)$ such that for any t' , agent i 's payoff by following the new equilibrium, $a'_{i,t} = 1$, over $[t', T']$ is the same as his payoff from following the old equilibrium over $[\tau^{-1}(t'), T]$ under $w(\cdot)$. The latter payoff is:

$$p_{i,\tau^{-1}(t')} \int_{\tau^{-1}(t')}^T (w_{i,t}\lambda - c) a_{i,t} e^{-\int_{\tau^{-1}(t')}^t \lambda a_{i,z} dz} dt - (1 - p_{i,\tau^{-1}(t')}) c \int_{\tau^{-1}(t')}^T a_{i,t} dt,$$

where $w_{i,t}$ is the expected reward if agent i succeeds at t given scheme $w(\cdot)$, i.e. suppressing the dependence on equilibrium strategies,

$$\begin{aligned} w_{i,t} &= \mathbb{E}_{s_{-i}} w(t, \mathbf{s}_{-i}) \\ &= \int_{[0,T]^{N-1}} w(t, s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_N) \left(\prod_{j \neq i} a_{j,s_j} \right) \lambda^{N-1} e^{-\lambda \sum_{j \neq i} \int_0^{s_j} a_{j,z} dz} d\mathbf{s}_{-i}. \end{aligned}$$

The former payoff is:

$$p_{i,t'} \int_{t'}^{T'} (w'_{i,t}\lambda - c) a'_{i,t} e^{-\int_{t'}^t \lambda a'_{i,z} dz} dt - (1 - p_{i,t'}) c \int_{t'}^{T'} a'_{i,t} dt,$$

where we find scheme $w'(\cdot)$ such that $w'_{i,t} = w_{i,\tau^{-1}(t)}$:

$$\begin{aligned} w'_{i,t'} &= \mathbb{E}_{s_{-i}} w'(t', \mathbf{s}'_{-i}) \\ &= \int_{[0,T']^{N-1}} w'(t', \mathbf{s}'_{-i}) \lambda^{N-1} e^{-\lambda \sum_{j \neq i} s'_j} d\mathbf{s}'_{-i} \\ &= \int_{[0,T]^{N-1}} w'(\tau(t), \tau(\mathbf{s}'_{-i})) \left(\prod_{j \neq i} \tau'(s_j) \right) \lambda^{N-1} e^{-\lambda \sum_{j \neq i} \tau(s_j)} d\mathbf{s}_{-i} \\ &= \int_{[0,T]^{N-1}} w'(\tau(t), \tau(\mathbf{s}'_{-i})) \left(\prod_{j \neq i} a_{j,s_j} \right) \lambda^{N-1} e^{-\lambda \sum_{j \neq i} \tau(s_j)} d\mathbf{s}_{-i}, \end{aligned}$$

where $\tau(\mathbf{s}'_{-i}) := (\tau(s_1), \dots, \tau(s_{i-1}), \tau(s_{i+1}), \dots, \tau(s_N))$. We want to find $w'(\cdot)$ such that for any i and t ,

$$\begin{aligned} &\int_{[0,T]^{N-1}} w'(\tau(t), \tau(\mathbf{s}_{-i})) \left(\prod_{j \neq i} a_{j,s_j} \right) \lambda^{N-1} e^{-\lambda \sum_{j \neq i} \tau(s_j)} d\mathbf{s}_{-i} \\ &= \int_{[0,T]^{N-1}} w(t, \mathbf{s}_{-i}) \left(\prod_{j \neq i} a_{j,s_j} \right) \lambda^{N-1} e^{-\lambda \sum_{j \neq i} \int_0^{s_j} a_{j,z} dz} d\mathbf{s}_{-i}. \end{aligned} \quad (30)$$

Note that $\tau(s_j) = \int_0^{s_j} a_{j,z} dz$. Hence, (30) is equivalent to

$$\int_{[0,T]^{N-1}} [w'(\tau(t), \tau(\mathbf{s}_{-i})) - w(t, \mathbf{s}_{-i})] \left(\prod_{j \neq i} a_{j,s_j} \right) e^{-\lambda \sum_{j \neq i} \int_0^{s_j} a_{j,z} dz} d\mathbf{s}_{-i} = 0.$$

It thus suffices that for any t and \mathbf{s}_{-i} ,

$$w'(\tau(t), \tau(s_1), \dots, \tau(s_{i-1}), \tau(s_{i+1}), \dots, \tau(s_N)) = w(t, s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_N).$$

This defines the contest $w'(\cdot)$.

We now argue that the full-effort profile, $a'_{i,z}$ is optimal (i.e. a best response). Suppose, per contra, that there is a strictly better strategy $\hat{a}_{i,z}$. Let us show that there is a profitable deviation from the original equilibrium. For any $z \in [0, T]$, define

$$\tilde{a}_i^z := \int_0^{\tau(z)} \hat{a}_{i,t} dt \quad (31)$$

as the total amount of effort that agent i should exert by time z .

Note that

$$\tilde{a}_i^z = \int_0^z \tilde{a}_{i,t} dt, \quad (32)$$

where (29) and (31) imply that for any $z \in [0, T]$,

$$\tilde{a}_{i,z} := \frac{d\tilde{a}_i^z}{dz} = \frac{d\tau(z)}{dz} \hat{a}_{i,\tau(z)} = a_{i,z} \hat{a}_{i,\tau(z)} \in [0, 1].$$

It follows from (31) and (32) that

$$\text{for any } t', t \in [0, T]: \int_{t'}^t \hat{a}_{i,z} dz = \int_{\tau^{-1}(t')}^{\tau^{-1}(t)} \tilde{a}_{i,z} dz. \quad (33)$$

We claim that (33) implies that the payoff at any $t \in [0, T]$ under $\hat{a}_{i,z}$ and at $\tau^{-1}(t) \in [0, T]$ under $\tilde{a}_{i,z}$ are the same. This follows from the same argument as before, because we have used the same $\tau(\cdot)$ as before and hence, at any z , the total amount of effort by agents $-i$ at z is also the same as that at $\tau^{-1}(z)$, and, by construction, agent i 's private belief and total effort at z is the same as that at $\tau^{-1}(z)$.

Since $\hat{a}_{i,z}$ is a profitable deviation from $a'_{i,z}$, we conclude that $\tilde{a}_{i,z}$ is a profitable deviation from $a_{i,z}$, a contradiction.

B.1.3 Proof of Proposition 6

Consider a WTA contest $C = \{w(\cdot), D, T\}$, where $w(\cdot)$ is the WTA prize scheme, D is an arbitrary information disclosure policy, and T is the deadline. Denote by $I_{i,t}$ the information that the principal has disclosed to agent i by time t ; if D is stochastic, let $I_{i,t}$ correspond to any given realization of D . Denote agent i 's belief at time t by $p_{i,t}$ and his expected reward for success at time t by $w_{i,t}$. We let (agent i 's conjecture of) the aggregate cumulative effort up to t given no success by t be A^t .

At any time t , agent i can ensure a positive continuation payoff by shirking. The agent thus chooses $a_{i,t} > 0$ only if

$$c \leq p_{i,t} \lambda w_{i,t}. \quad (34)$$

Let

$$\alpha_{i,t} := \Pr[\text{some } j \neq i \text{ succeeded by } t \mid i \text{ did not, } I_{i,t}] \in [0, 1].$$

In a WTA contest, (34) is equivalent to

$$c \leq (1 - \alpha_{i,t}) \Pr[G \mid \text{no success, } A^t] \lambda \bar{w}.$$

Consider now an optimal public WTA contest. As shown in Section 3, agent i chooses $a_{i,t} = 1$ if

$$c \leq \Pr[G \mid \text{no success, } A^t] \lambda \bar{w}.$$

It follows that, given aggregate cumulative effort A^t , if agent i exerts positive effort in contest C , he exerts full effort in an optimal public WTA contest. Therefore, if the aggregate cumulative effort induced by contest C is A^T and that induced by an optimal public WTA contest is A^{PW} , then $A^{PW} \geq A^T$.

B.1.4 Proof of Proposition 7

The principal chooses \bar{w} and a contest, $C = \{w(\cdot), \mathcal{T}, T\}$, to maximize the expectation of (21). Let $C(\bar{w})$ be the optimal contest and $A^T(\bar{w})$ the aggregate cumulative effort induced by this contest as a function of the prize \bar{w} . By Proposition 5, $C(\bar{w})$ is a mixture contest that implements public WTA from time 0 until a switching time t_S and hidden ES from t_S on. Hence, each agent follows a stopping strategy with stopping condition given by (27) and, using this stopping condition, a routine computation shows that aggregate cumulative effort is strictly increasing in \bar{w} : $\frac{dA^T(\cdot)}{d\bar{w}} > 0$.

Consider the principal's payoff given an optimal prize \bar{w} and induced aggregate cumulative effort in an optimal contest $A^T(\bar{w})$, which we denote $U^P(\bar{w}, v, \lambda, p_0) := (v - \bar{w})p_0 \left(1 - e^{-\lambda A^T(\bar{w})}\right)$. This payoff has strictly increasing differences in (v, \bar{w}) , as $U_{v\bar{w}}^P = p_0 \lambda e^{-\lambda A^T(\bar{w})} \frac{dA^T(\bar{w})}{d\bar{w}} > 0$. Therefore, the optimal prize \bar{w} is increasing in v . For v large enough, the optimal prize \bar{w} is large enough that $\frac{\lambda \bar{w}}{N} > c$, so [Proposition 5](#) implies that a hidden ES contest is optimal. Similarly, for v small enough, the optimal prize \bar{w} is small enough that $\frac{\lambda \bar{w}}{2} < c$, so [Proposition 5](#) implies that a public WTA contest is optimal.

B.2 Beyond simple information disclosure

[Proposition 5](#) showed that a mixture contest is optimal within the class of contests with simple information disclosure policies. If $\frac{\lambda \bar{w}}{2} < c$, the optimal mixture contest is public WTA, while if $\frac{\lambda \bar{w}}{N} > c$, the optimal mixture contest is hidden ES. As mentioned in [Subsection 4.1](#), we conjecture that these conditions are sufficient for the optimality of public WTA and hidden ES respectively when allowing for any information disclosure policy.

Here we provide an example showing that if $\frac{\lambda \bar{w}}{N} < c < \frac{\lambda \bar{w}}{2}$ and any disclosure policy can be used, then a non-mixture contest can be optimal for the principal. In particular, we show that the principal can improve upon mixture contests by using “cutoff disclosure policies”.

Consider $N = 3$ and parameters such that (i) $\lambda \frac{\bar{w}}{3} < c$ and (ii) $\lambda \frac{\bar{w}}{2} > c$. We construct a contest C that dominates both public WTA and hidden ES. We show that any mixture contest (i.e., any contest that implements public WTA until some point and hidden ES from then on) is dominated by a mixture of public WTA and C .

Let the contest $C = \{w(\cdot), D, T\}$ be defined as follows. The prize scheme $w(\cdot)$ is equal-sharing: if n agents succeed by the deadline T , each successful agent receives a reward $\frac{\bar{w}}{n}$. The disclosure policy D is “two-agent cutoff disclosure”: at any time t , the principal sends public message m_2 if two agents have succeeded by t and she sends public message m_{01} if either no agent or only one agent has succeeded by t . Note that assumption (i) implies that if message m_2 is sent at t , no agent exerts effort at $z > t$. Furthermore, this implies that each agent i will follow a stopping strategy, and without loss we let T coincide with the stopping time in this strategy (as in our analysis of [Section 3](#), we focus on symmetric equilibria). Thus, each agent exerts full effort until either he succeeds, or the principal sends message m_2 , or time T is

reached. The stopping time T is then given by

$$\begin{aligned} c &= p_T^{(C)} \lambda \bar{w} \left[\frac{2e^{-\lambda T} (1 - e^{-\lambda T})}{1 - (1 - e^{-\lambda T})^2} \frac{1}{2} + \frac{e^{-2\lambda T}}{1 - (1 - e^{-\lambda T})^2} \right] \\ &= \frac{p_0 e^{-\lambda T} (1 - (1 - e^{-\lambda T})^2)}{p_0 e^{-\lambda T} (1 - (1 - e^{-\lambda T})^2) + 1 - p_0} \lambda \bar{w} \left[\frac{2e^{-\lambda T} (1 - e^{-\lambda T})}{1 - (1 - e^{-\lambda T})^2} \frac{1}{2} + \frac{e^{-2\lambda T}}{1 - (1 - e^{-\lambda T})^2} \right]. \end{aligned} \quad (35)$$

We first show that C dominates any hidden ES contest. In an optimal hidden ES contest, each agent follows a stopping strategy with stopping time T^{HS} given by

$$\begin{aligned} c &= p_{T^{HS}}^{(1)} \lambda \bar{w} \left[(1 - e^{-\lambda T^{HS}})^2 \frac{1}{3} + 2e^{-\lambda T^{HS}} (1 - e^{-\lambda T^{HS}}) \frac{1}{2} + e^{-2\lambda T^{HS}} \right] \\ &= \frac{p_0 e^{-\lambda T^{HS}}}{p_0 e^{-\lambda T^{HS}} + 1 - p_0} \lambda \bar{w} \left[(1 - e^{-\lambda T^{HS}})^2 \frac{1}{3} + 2e^{-\lambda T^{HS}} (1 - e^{-\lambda T^{HS}}) \frac{1}{2} + e^{-2\lambda T^{HS}} \right]. \end{aligned}$$

Note that by assumption (i), $\lambda \frac{\bar{w}}{3} < c$; hence,

$$c < \frac{p_0 e^{-\lambda T^{HS}} (1 - e^{-\lambda T^{HS}})^2}{p_0 e^{-\lambda T^{HS}} + 1 - p_0} c + \frac{p_0 e^{-\lambda T^{HS}}}{p_0 e^{-\lambda T^{HS}} + 1 - p_0} \lambda \bar{w} \left[2e^{-\lambda T^{HS}} (1 - e^{-\lambda T^{HS}}) \frac{1}{2} + e^{-2\lambda T^{HS}} \right],$$

which we can rewrite as

$$c < \frac{p_0 e^{-\lambda T^{HS}} (1 - (1 - e^{-\lambda T^{HS}})^2)}{p_0 e^{-\lambda T^{HS}} (1 - (1 - e^{-\lambda T^{HS}})^2) + 1 - p_0} \lambda \bar{w} \left[\frac{2e^{-\lambda T^{HS}} (1 - e^{-\lambda T^{HS}})}{(1 - (1 - e^{-\lambda T^{HS}})^2)} \frac{1}{2} + \frac{e^{-2\lambda T^{HS}}}{(1 - (1 - e^{-\lambda T^{HS}})^2)} \right]. \quad (36)$$

Since the RHS of (35) and that of (36) are each decreasing in T and T^{HS} respectively, it follows from these conditions that $T > T^{HS}$. Hence, since agents use stopping strategies in both contests and they stop at a later time in C than in an optimal hidden ES contest, we conclude that C dominates any hidden ES contest.

We next show that C dominates any public WTA contest. In an optimal public WTA contest, each agent also follows a stopping strategy with stopping time T^{PW} given by

$$c = p_{T^{PW}} \lambda \bar{w}. \quad (37)$$

Since the RHS of (35) and that of (37) are each decreasing in T and T^{PW} respectively, we have that $T > T^{PW}$ if and only if

$$c < \frac{p_0 e^{-\lambda T^{PW}} (1 - (1 - e^{-\lambda T^{PW}})^2)}{p_0 e^{-\lambda T^{PW}} (1 - (1 - e^{-\lambda T^{PW}})^2) + 1 - p_0} \lambda \bar{w} \left[\frac{2e^{-\lambda T} (1 - e^{-\lambda T^{PW}})}{1 - (1 - e^{-\lambda T^{PW}})^2} \frac{1}{2} + \frac{e^{-2\lambda T^{PW}}}{1 - (1 - e^{-\lambda T^{PW}})^2} \right].$$

We can rewrite this as, for any agent i ,

$$c < \Pr[\text{some } j \neq i \text{ succ. by } T^{PW} \mid i \text{ did not, not all } j \neq i \text{ did}] \lambda \frac{\bar{w}}{2} \\ + \Pr[\text{no } j \neq i \text{ succ. by } T^{PW} \mid i \text{ did not, not all } j \neq i \text{ did}] \Pr[G \mid \text{no success by } T^{PW}] \lambda \bar{w}.$$

Since $\Pr[G \mid \text{no success by } T^{PW}] \lambda \bar{w} = c$, this condition simplifies to:

$$c < \lambda \frac{\bar{w}}{2},$$

which is true by assumption (ii) above. Therefore, since agents use stopping strategies in both contests and they stop at a later time in C than in an optimal public WTA contest, we conclude that C dominates any public WTA contest.

We note that the results above are consistent with public WTA dominating hidden ES as well as with hidden ES dominating public WTA. Hidden ES dominates public WTA if and only if

$$c < \lambda \bar{w} \left[\frac{2(1 - e^{-\lambda T^{PW}}) e^{-\lambda T^{PW}}}{1 - e^{-2\lambda T^{PW}}} \frac{1}{2} + \frac{(1 - e^{-\lambda T^{PW}})^2}{1 - e^{-2\lambda T^{PW}}} \frac{1}{3} \right],$$

which, under assumptions (i) and (ii), may or may not be satisfied depending on parameters.

Finally, we show that any mixture contest that implements public WTA until a time t^* and hidden ES from t^* on is dominated by a mixture contest that implements public WTA until t^* and contest C from t^* on. Let T^* be the stopping time in the public WTA-hidden ES mixture. If $t^* = 0$ or $t^* = T^*$, the claim follows from the analysis above. Suppose then that $t^* \in (0, T^*)$. Note that in both contests agents use a stopping strategy; moreover, it follows that if all agents exert full effort until t^* in the public WTA-hidden ES mixture, they will do so as well in the public WTA- C mixture. Thus, all we need to do is compare the continuation from t^* on. Let p_{t^*} be the public belief at t^* . Clearly, if C dominates hidden ES starting with a prior belief $p_0 = p_{t^*}$, then the public WTA- C mixture will dominate the public WTA-hidden ES mixture. But note that the condition for C to dominate hidden ES is independent of p_0 , and is simply given by (i), i.e. $\lambda \frac{\bar{w}}{3} < c$ (this is the only condition that we used to obtain (36) above). Therefore, under this condition, C dominates hidden ES starting at any prior, and consequently the public WTA- C mixture dominates the public WTA-hidden ES mixture.