Races or Tournaments? Theory and Evidence from the Field[1]

[PRELIMINARY AND INCOMPLETE]

Andrea Blasco   Kevin J. Boudreau   Karim R. Lakhani   Michael Menietti

Last updated: 29 June, 2017

[1]Blasco: Harvard Institute for Quantitative Social Science, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138 (email: ablasco@fas.harvard.edu).

Abstract

We examine the performance of two different choices of contest design: the race (where the winner is the first to achieve a minimum quality) and the tournament (where the winner is the one with the highest quality in a given period). After characterizing the optimal design, we report results of a field experiment conducted to compare the performance of three alternatives motivated by theory: the race, the tournament, and the tournament with a minimum quality requirement. Outcomes in a race are of comparable quality, supplied faster, and with lower participation rates. Based on these findings, we show the optimal design under several counterfactual situations.

JEL Classification: M15; M52; O31.

Keywords: races; tournaments; contest theory; crowdsourcing; innovation.

1 Introduction

Contests have tremendous impact on economic growth. Historically, awards offered by government agencies have accelerated agricoltural innovation;[1] improved methods for navigation and cartography;[2] and kick-started the modern aviation sector.[3] Today, in addition to government agencies, also philanthropic organizations and private firms recur to contests for numerous purposes that include incentivizing employees and crowdsourcing a long list of internal activities (marketing campaigns, data analysis, software development) to large online communities of freelancers.

While the economic relevance of contests is well recognized (xxxx), contest designer xxx have this problem xxx. How to do xxx? How to xxxx? In this paper, we compare two different choices of design. One is the "race" competition, where prizes are offered to the competitor that *first* reaches a fixed performance target. The other is the "tournament" competition, where prizes are offered to the competitor that does *best* relative to others by a given deadline. Familiar examples of races are the Longitude rewards offered by the British government in 1714, the Orteig Prize in 1919, and, more recently, the Netflix prize in 2009; while most architectural competitions, the Golden Carrot Contest in 1992, the Defense Advanced Research Project Agency (DARPA) series of Grand Challenges between 2004 and 2013, and the Progressive Insurance Automotive X-Prize in 2010 are all examples of tournaments.[4]

Races and tournaments are very general forms of competition. And an extensive literature exists that has focused on races, in the context of patent races, and tournaments separetely. With only a few papers in economics comparing race and tournament in a single framework. A underxxxx that is perhaps becayse in many situations the choice seem imposed by the environment (as in the case of the legal environment defining a patent race) rather than a straightforward choice of contest design. In addition, and from a contest design perspective, the choice between race and tournament competitions is difficult to examine: xxxx. A natural way to examine the issue is via modeling choices as a function of the contest sponsors's preferences towards two desirable, but often incompatible, goals: i) maximizing revenues through raising the efforts of competitors while ii) lowering the time it takes to complete a given job. It is unclear, however, how either two

[1](**???**) shows how agricultural innovation in England went up due to the awards offered by the Royal Agricultural Society at the beginning of 20th century.

[2]A well known example is the Longitude rewards offered by the British government in 1714 that contributed to solve for the first time the problem of the precise determination of a ship's longitude at sea.

[3]Historically, the Orteig Prize xxxx. More recently, X-Prize inspired by xxxx.

[4]See Wikipedia article on "Inducement prize contest" (`https://en.wikipedia.org/wiki/Inducement_prize_contest`) for a list of these contests with links to their descriptions. The Golden Carrot Contest has been described by Taylor (1995).

competition formats will solve this trade-off. On the one hand, a fixed deadline may accelerate production in tournaments. On the other, a too short deadline may deter entry, thereby lowering revenues from competitors effort. Likewise, a time competition may encourage competitors to reach a given target faster, but it may also limit competition if the target can be reached only by a few competitors. Alternatively, the choice between tournaments and races can be seen as the response to "efficiency" concerns of the contest sponsors. Under this view, races and tournaments may lead to the same expected outcomes in terms of time or effort but lead to different duplication costs by regulating "entry" into the contest [as discussed by Fullerton Mcafee, xxxx]. Hence, the "time preferences" story and the "efficiency" are two possible explanation for using a race or a contest.

In this article, we investigate the choice between races and tournaments both theoretically, and empirically in the field. We proceed in two ways. First, we develop a contest model that encompasses both the race and the tournament in a single framework. Exploring the duality of the model, we compare equilibrium behaviors under both competitive formats and characterize the optimal choice for the contest designer. Then, we design and execute an experiment to test the implications of the theory in the field, and xxxx providing policy recommendations.

Our theoritical approach extends the contest model introduced by Moldovanu and Sela (2001) to a situation in which xxx decide both time and quality. Thus, contests have an all-pay structure by which participants pay an immediate cost for an uncertain future reward. The decision of timing and quality is made under the uncertainty of the costs of the rivals. The contest designer wants to maximize reveues and has preferences for both time and quality. Following the analysis of the model, we show that the optimal design depends on the number of participants and the concavity of their cost function. We also show that XXX, YYY, and ZZZZ.

To fix ideas, imagine a government willing to design an innovation contest aimed at finding solutions to a problem of public health, such as antibiotic resistance.[5] To minimize the risk that the threat of xxxx will materialize before a solution is found, one may choose a tournament competition format with a tight deadline for participants to provide their solutions. The problem is to find the right duration. When the duration of the competition is too short, incentives maybe insufficents for competitors to exert enough effort resulting in inadequate solutions. Alternatively, the government can set up a race competition with a prize being awarded to the first competitor who achieves, or goes beyond, a minimum quality threshold. Here

[5] This example is taken. . .

the problem of accelarating the timing of innovation should not be a big issue but competitors may work inefficiently, as they have no incentives to exceed the minimum threshold. Fixed the prize structure, both approaches have specific advantages and limitations. However, xxxx.

The context of the field experiment was an online programming competition run on Topcoder at the end of 2016. In a typical programming competition, participants compete writing source code that solves a given problem for winning a monetary prize. We worked together with researchers from the United States National Health Institute (NIH) and the Scripps Research Institute (SCRIPPS) to select a challenging problem for the contest. The selected problem was based on an algorithm called BANNER built by NIH (Leaman, Gonzalez, and others 2008) that uses expert labeling to annotate abstracts from a prominent life sciences and biomedical search engine, PubMed, so disease characteristics can be more easily identified. The goal of the programming competition was to improve upon the current NIH's system by using a combination of expert and non-expert labeling, as described by Good et al. (2014). The competition was hosted online on the platform Topcoder.com (about 1M registered users in 2016). Top submissions were awarded monetary prizes ranging between $100 to $5000 for a total prize pool of more than $40,000.

Our intervention consisted in sorting at random participants into independent virtual rooms of 10 or 15 people. These virtual rooms were then randomly assigned to one of three different competitive settings: a race, a tournament, and a tournament with a reserve score, which is the lowest acceptable score by the platform for a submission to be awarded a prize.

We find that xxxxx [participation in the tournament is xxx compared to the race the reserve.]

We also find that xxxx [submission are quicker in a race, whereas are equally distributed at the end of the competition in the the tournament and in the tournament with quality requirement.]

Another interesting finding is that xxxxx [No evidence trade-off between a race and a tournament in terms of higher scores vs faster submissions. We do find that scores are higher in the tournament but we do not find a strong trade-off in the sense that race had comparable good quality solutions than the tournament.]

2 Literature

This paper is related to the contest theory literature Dixit (1987) Baye and Hoppe (2003), Parreiras and Rubinchik (2010), Moldovanu and Sela (2001), Moldovanu and Sela (2006), Siegel (2009), Siegel (2014). It also relates to the literature on innovation contests Taylor (1995), Che and Gale (2003). And the personnel economics approach to contests Lazear and Rosen (1981), Green and Stokey (1983), Mary, Viscusi, and Zeckhauser (1984).

Empirically, Dechenaux, Kovenock, and Sheremeta (2014) provide a comprehensive summary of the experimental literature on contests and tourments. Large body of empirical works have focused on sports contests Szymanski (2003). More recently, inside firms (xxx) and online contest (xxxx).

This paper is also related to the econometrics of auctions Paarsch (1992), Laffont, Ossard, and Vuong (1995), Donald and Paarsch (1996) and more recently Athey, Levin, and Seira (2011), Athey and Haile (2002), and Athey and Haile (2007).

An extensive literature has discussed the reasons why contests are sometimes preferred to other forms of incentives (e.g., individual contracts). Typically, contests reduce monitoring costs [xxx], incentivize production with common risks [xxx], and deal with indivisible rewards [xxxx], among others. While there is not much debate on why contests should be used, the issue of how to effectively design and deploy a contest still attracts much research.

Several aspects of contest design have been investigated, including the optimal prize structure [XXX, xxxx, xxxx], number of competitors [XXX, XXX], and imposing restrictions to competition such as minimum effort requirements [XXX, XXX]. Also, a great deal of theoretical models of races and tournaments have been developed and applied to a wide range of economic situations including patent races [xxx], arms races [xxx], sports [xxx], the mechanism of promotions inside firms [xxxx], sales tournaments [xxxx], etc.

Harris and Vickers (1987), Grossman and Shapiro (1987) investigate the dynamics issues patent races where the interest is how firms compete for a patent. Bimpikis, Ehsani, and Mostagir (2014) looks at the problem of how to design an information structure that is optimal when the contest is a race and innovation is uncertain (encouragement and competition effect). In the laboratory, Zizzo (2002) finds poor support to predictions of dynamic xxxx. In general we do not know much about the dynamic aspect of contests.

The duality. As pointed out by Baye and Hoppe (2003), many of these models of tournament and race competitions are specific cases of a more general "contest games." And sometimes it is possible to design one or the other in a way to exploit a "duality." In other words, in theory, a competition can be designed as a tournament to do xxx or as a race to do xxx. While theoretically very useful, how to exploit this duality in practice remains largely unknown. Lack of data. As before, xxxx. The main challenge is self-selection. The answer to this optimal design question relates to the cost function of agents with respect to "time" and to "effort." It is hard to say which solution is better. However, it is easier to tell whether you should have one prize or multiple prizes.

3 The model

We now generalize the contest game introduced by Moldovanu and Sela (2001) to a situation where players simultaneously decide $i)$ the quality and $ii)$ how fast to produce a given output. Then we explore the problem of revenue maximization faced by a contest designer with preferences for both quality and time.

3.1 Basic setup

There are $i = 1, ..., n$ players competing for $k = 1, ..., q$ prizes of values $v_1 \geq v_2 \geq ... \geq v_q \geq 0$ and total value normalized to one, $\sum_{k=1}^{q} v_k = 1$. Prizes are awarded on the basis of a nonnegative variable y_i measuring performance and how much time t_i it takes each player to carry out a given task (e.g., solve a problem). Players simultaneously decide on their performance and time, and pay at once any cost of effort incurred. Players differ in an individual ability a_i, reflecting skills, time constraints, and other elements affecting their cost of effort. A function $C(\cdot)$ determines the cost each player incurs:

$$C(a, y, t) = a^{\alpha} y^{\beta} t^{\gamma} \qquad \text{with } \alpha, \gamma < 0, \beta > 1. \tag{1}$$

Thus, costs are decreasing in a, increasing in y, and decreasing in t, with constant elasticities α, β, and γ, respectively. (Alternatively, one can replace time t_i with a new variable "speed" s_i defined as the ratio of quality over time.)

Abilities are drawn at random from a common distribution function $F(\cdot)$ with density $f(\cdot)$ on a bounded interval. Information is asymmetric, as players observe privately the realization of their own ability variable before making their choices.

Players are risk neutral and maximize their expected payoffs

$$\pi_i = \sum_{k=1}^{q} p_i^k(y_i, t_i) v_k - C(a_i, y_i, t_i) \tag{2}$$

where $p_i^k(\cdot)$ denotes player i's conditional probability of winning the k-th prize given a performance y_i and a time t_i to complete the task.

For a given distribution of the actions of player i's opponents, the probability of winning a prize $p_i^k(\cdot)$ is fully specified by the "competition style" of the contest. We consider three competition styles: the race; the tournament; and the tournament with reserve.

To define the probability under each competition style we need first to define the distribution of the actions of the opponents for each player. As information is asymmetric and players move simultaneously, opponents' performances and timings can be treated as random variables $Y_1, ..., Y_{n-1}$ and $T_1, ..., T_{n-1}$ drawn from common distributions $F_Y(\cdot)$ and $F_T(\cdot)$, respectively. In a contest, realizations of these variables need to be ordered to award any prize. Let denote the j'th smallest of the order statistics of the Y's by $Y_{(j)}$, with $Y_{(1)}$ being the smallest, $Y_{(2)}$ being the second smallest, and so on; and the j'th smallest of the T's by $T_{(j)}$, with $T_{(1)}$ being the smallest, $T_{(2)}$ being the second smallest, and so on.

In a race competition, the first player to achieve a given minimum performance target $\underline{y}$ gets the first prize, the second to achieve the same target gets the second prize, and so on. Hence, the conditional probability of winning the k-th prize in a race is

$$p_i^{k,\text{race}}(y_i, t_i) = \begin{cases} \Pr\{T_{(k-1)} < t_i \leq T_{(k)}\} & \text{if } y_i \geq \underline{y} \\ 0 & \text{otherwise;} \end{cases}$$

where $T_{(0)} = 0$ and we use the convention that an opponent j's timing is infinite if the performance is (strictly) below the target ($y_j < \underline{y}$).

In a tournament competition, instead, the player having achieved the highest performance by a given deadline $\bar{t}$ gets the first prize, the player having achieved the second highest performance gets the second prize, and so on. Hence, the conditional probability of winning the k-th prize in a tournament is

$$p_i^{k,\text{tournament}}(y_i, t_i) = \begin{cases} \Pr\{Y_{(n-k)} > y_i \geq Y_{(n-k-1)}\} & \text{if } t_i \leq \bar{t} \\ 0 & \text{otherwise;} \end{cases} \tag{3}$$

where we use the convention that an opponent j's performance is zero if the time is (strictly) above the deadline ($t_j > \bar{t}$).

Finally, in a tournament with reserve competition, the conditional probability is similar to that in a tournament competition as in 3 with the only difference that it is zero when the performance is below a given "reserve" performance target $\underline{y}$.

The contest designer can influence player's choices by choosing the competition style of the contest and, therefore, one the above conditional probabilities. The contest designer is risk neutral agent and aims to maximize its revenues by increasing the winner's performance, while keeping low the time spent (alternatively, maximizing the speed of production). Let denote the actions of the contest winner by (Y^w, T^w). The contest designer's expected payoff (net of payments to competitors) is:

$$\pi_{cd} = E[Y^w + \tau T^w \mid Y^w \geq \underline{y}, T^w \leq \bar{t}]$$

where $\tau \leq 0$ denotes the contest designer's (negative) preferences towards time.

3.2 Equilibrium

In this section, we solve the model for the unique symmetric Bayesian Nash equilibrium of players. To simplify notation, we focus on a contest with two prizes $q = 2$.

3.2.1 Race

At equilibrium, each player i chooses y_i and t_i by maximizing π_i given beliefs about the equilibrium actions of the other players.

In a race competition, the key observation is that any performance below the target gives a zero probability of winning and any performance above the target gives a constant probability of winning. Thus, player i's optimal choice y_i^* is either zero (i.e., the lowest possible) or $y_i^* = \underline{y}$.

Assuming y_i^*, the optimal timing t_i^* is given by first order conditions

$$\partial xxxx = \sum$$

Here, the key observation is that, for a given level of quality, any time that is strictly below the deadline does not affect the probability of winning but is costly in terms of effort (working faster is costlier) and any time that is strictly above the deadline gives a negative payoff. Thus, choosing $t_i = \bar{t}$ is a (weakly) dominant strategy for each player. Then the first order condition with respect to quality is:

$$\sum_{k=1}^{q} \hat{p}_k'(y_i) v_k = c_a(a_i) c_y'(y_i) c_\tau(\bar{t}).$$

where $\hat{p} = p(\cdot, \bar{t})$. Then it can be show that xxxx.

$$
\begin{aligned}
0 = &\alpha f_{(1:N-1)}(\phi)\phi' + (1-\alpha)\phi'\{[1 - F_{(1:N-1)}(\phi)]f_{(1:N-2)}(\phi) + \\
&+ f_{(1:N-1)}(\phi)F_{(1:N-2)}(\phi)\} - c_a(a)c_y(\underline{y})c'_\tau(t_i)
\end{aligned} \tag{4}
$$

subject to the boundary condition $\phi(0) = \underline{a}$ (i.e., the lowest-ability competitor's optimal output quality is zero).

As shown by Moldovanu and Sela (2001), the solution is

$$
y^*(a_i) = c_y^{-1}\left[c_y(\underline{y}) + \frac{1}{c_\tau(\bar{t})}\left(\alpha \int_{a_i}^{\bar{a}} A(z)dz + (1-\alpha)\int_{a_i}^{\bar{a}} B(z)dz\right)\right] \tag{5}
$$

where

$$
A(x) = \frac{1}{c_a(x)} f_{(n-1:n-1)}(x) \tag{6}
$$

and

$$
B(x) = \frac{1}{c_a(x)}\left\{\left[1 - F_{(n-1:n-1)}(x)\right] f_{(n-1:n-2)}(x) + f_{(n-1:n-1)}(x)F_{(n-1:n-2)}(x)\right\}. \tag{7}
$$

Monotonicity of the equilibrium output quality implies that, for every $i = 1, ..., n$, the equilibrium expected payoff from the contest π_i^* depends on the rank of the player's ability relative to the others. As a result, the equilibrium expected payoff net of costs is

$$
R(a_i) = \alpha F_{n:n}(a_i) + (1-\alpha)[1 - F_{n:n}(a_i)]F_{n-1:n-1}(a_i). \tag{8}
$$

% payoffs

3.2.2 Equilibrium in a race

In a similar way, one can derive the equilibrium strategy in a race. Again the key observation is that any quality below the target gives a zero probability of winning and any quality above the target gives a constant probability of winning. Thus, player i's choice of optimal quality y^* is either zero (with $t_i = \bar{t}$ by convention) or $y^* = \underline{y}$.

Then, the equilibrium xxx for player i is

$$t^*(a_i) = c_\tau^{-1}\left[c_\tau(\bar{t}) + \frac{1}{c_y(\underline{y})}\left(\alpha\int_{a_i}^{\bar{a}} A'(z)dz + (1-\alpha)\int_{a_i}^{\bar{a}} B'(z)dz\right)\right] \tag{9}$$

where

$$A(x) = \frac{1}{c_a(x)} f_{(n-1:n-1)}(x) \tag{10}$$

and

$$B(x) = \frac{1}{c_a(x)}\left\{\left[1 - F_{(n-1:n-1)}(x)\right] f_{(n-1:n-2)}(x) + f_{(n-1:n-1)}(x) F_{(n-1:n-2)}(x)\right\}. \tag{11}$$

An important property of XX is that $y^*(a_i)$ has its upper bound in XX and lower bound in XX. Again payoffs are xxxx. Hence, by setting to zero and solving for the ability, gives the marginal ability $\underline{a}$ as

$$\underline{a} = h(n, V, F_A, C, d). \tag{12}$$

3.2.3 Tournament vs races

By comparing equilibrium xxx and xxx, we find that the race and the tournament do not (ex-post) dominate one another with respect to output quality. Whereas the race always dominates the tournament with respect to completion time. [This is only when the deadline is the same. Otherwise, there's always xxxx.] This result is stated below.

Proposition 1. *There always exist an interval of abilities where the output quality is higher in the race than in the tournament. By contrast, every player takes less completion time in the race than in the tournament.*

Proof. Marginal type has utility zero in a race but the same type has a strictly positive utility in the tournament. Since probability of winning is not different in the race or the tournament (the bid is a monotonic transformation of the individual ability or, in other words, rankings are virtually the same), expected payoffs in equilibrium differ only in the cost functions. Hence, to be an equilibrium, the player in the tournament should bid less than the player in the race to earn a strictly positive expected payoff. □

Let's make an example.

```
p <- plnorm   # pdf individual abilities
r <- rlnorm   # Simulate individual abilities
cy <- function(x) x^2 # Cost function performance
ct <- function(x) 2*exp(1-x)  # Cost function timing
```

FIGURE 1. Equilibrium bids in a race and a tournament.

Implications. The above proposition applies only if the target is higher in a race than in a tournament. But what if the two competitions had the same target ? In that case, tournaments and races have the same marginal type. Therefore, the performance of players in the tournament with reserve are always non-lower than those in the race. This does not imply that it is optimal to set the target. On the contrary, we will show that it is optimal to set an optimal target in a tournament that is below the optimal target in a race. Next section.

3.3 The contest designer's problem

Let us now focus on the contest designer's problem. Imagine the contest designer can choose the competition format to be either the race or the tournament. Imagine all other aspects of design are given. The prize structure α has been already chosen. There is a deadline $\bar{t}$, which is the same in both competition formats. [The quality requirement $\underline{y}_c$ in the tournament is smalle than that in the race $\underline{y}_{\text{race}} > \underline{y}_{\text{tour}}$)] We will relax this assumption later to consider a more general setting where these variables are also part of the contest designer's problem.

The contest designer has an objective function that is increasing in the expected quality of the winning solution and decreasing in the corresponding time to completion. Here, to do not complicate exposition, we assume that the contest designer cares about the winning submission only: second placed efforts are not considered. [If the principal values the diversity of the solutions ... but we assume it does not.]

XXX EQUATION XXXX

The optimal choice involves a comparison of the expected profits between the race and the tournament. Given xxxx, we can show that there will be a threshold on the cost of completion time $\hat{\tau}$ above which the race is a better choice than the tournament, and vice versa.

Proposition 2. *There's a tau above which ...*

Proof. In a tournament, the objective function is

$$
\begin{aligned}
R_{\text{tour}} &= \Pr(t_{(1:n)} \leq \bar{t}) \left\{ \int y^*(x \mid t_{(1:n)} \leq \bar{t}) dF_{n:n}(x) - \tau\bar{t} - 1 \right\} \\
&= \int_{\hat{a}}^{\bar{a}} y^*(x) dF_{n:n}(x) - \tau\bar{t} - 1. \qquad (13)
\end{aligned}
$$

That is, the contest designer's objective function is the sum of the expected output quality for a given deadline, minus the cost τ of having the winner working on the task until completion (i.e., until the deadline), and the cost of the prize pool (recall the prize pool is normalized to one).

[Implicitly, you're assuming that the prize is always large enough to ensure positive effort.] [Second prize too is stochastic!!!!]

In a race, the objective function is

$$
\begin{aligned}
R_{\text{race}} &= \Pr(a_{(N)} \geq \hat{a}) \left\{ \underline{y} - \alpha - \Pr(a_{(N-1)} \geq \hat{a})(1-\alpha) \right\} - \tau \int_{\hat{a}}^{\infty} t^*(x) dF_{N:N}(x) \\
&= [1 - F_{N:N}(\hat{a})] \left\{ \underline{y} - \alpha - [1 - F_{N-1:N}(\hat{a})](1-\alpha) \right\} - \tau \int_{\hat{a}}^{\infty} t^*(x) dF_{N:N}(x). \qquad (14)
\end{aligned}
$$

Note. $t^*(x) \leq \bar{t}$ for all x's. Thus, a lower bound for the above objective function can be computed:

$$
\underline{R_{\text{race}}} = [1 - F_{N:N}(\hat{a})] \left\{ \underline{y} - \alpha - [1 - F_{N-1:N}(\hat{a})](1-\alpha) - \tau\bar{t} \right\} \qquad (15)
$$

An even simpler lower bound is rewriting the above expression as if $\alpha = 1$ (note if the real alpha was set 1 then also mtype would change and therefore setting alpha hits a lower bound only when mtype does xxxx when alpha is 1).

Note. $y^*(x)$ is lower than $\underline{y}$ for all $a < \hat{a}$. Thus, a lower bound of the tournament's expression is

$$
\overline{R_{\text{tour}}} = [1 - F_{N:N}(\hat{a})]\underline{y} + \int_{\hat{a}}^{\infty} y^*(x) dF_{N:N}(x) - \tau\bar{t} - 1. \qquad (16)
$$

$$
\begin{aligned}
\underline{R_{\text{race}}} \geq& \overline{R_{\text{tour}}} \\
[1-F_{N:N}(\hat{a})](\underline{y}-1-\tau\bar{t}) \geq& [1-F_{N:N}(\hat{a})]\underline{y}+\int_{\hat{a}}^{\infty} y^*(x)dF_{N:N}(x)-\tau\bar{t}-1 \\
-[1-F_{N:N}(\hat{a})](\tau\bar{t}+1) \geq& \int_{\hat{a}}^{\infty} y^*(x)dF_{N:N}(x)-(\tau\bar{t}+1) \\
F_{N:N}(\hat{a})(\tau\bar{t}+1) \geq& \int_{\hat{a}}^{\infty} y^*(x)dF_{N:N}(x) \\
\tau \geq& \left[\frac{\int_{\hat{a}}^{\infty} y^*(x)dF_{N:N}(x)}{F_{N:N}(\hat{a})}-1\right]\frac{1}{\bar{t}} \qquad (17)
\end{aligned}
$$

End proof.

When the cost of time τ is sufficiently high, the race is preferred. Interestingly, the threshold is a function of the deadline to complete the job, as xxx. It also depends on the shape of xxxx.

3.3.1 Optimal minimum-entry

Now we turn to discuss the contest designer's choice of an optimal minimum requirement $\underline{y}$. So far, we have assumed that $\underline{y}_{\text{race}} > \underline{y}_{\text{tour}}$. Now, we show that the assumption that xxxx is indeed an optimal choice of the contest designer. This is summarized in the next proposition.

Proposition 3. *Suppose the contest designer can choose the target that max profits in both the race and the tournament. Then, the optimal $\underline{y}$ in tournament is generally lower than that in a race.*

To prove that it is indeed the case. We proceed in two steps. First, we assume that the contest designer does not care about minimizing the timing of the innovation by imposing $\tau = 0$. For simplicity, assume that $\alpha = 1$ (winner-takes-all). In a race, this means that the optimal target will be a value that makes equal the costs in terms of less participation versus the gains in terms of higher values of the winning solutions. Formally, the contest designer's problem in a race is

$$
\text{maximize } R^{\text{race}} = [1-F_{N:N}(\hat{a})](\underline{y}_{\text{race}}-1). \qquad (18)
$$

Note that $\hat{a}$ depends on the target. This is clearly concave in $\underline{y}_{\text{race}}$. Thus, the first order condition is also sufficient.

$$\text{FOC } \Rightarrow -F'_{N:N}(\hat{a})\hat{a}'(\underline{y}_{\text{race}} - 1) + [1 - F_{N:N}(\hat{a})] = 0. \tag{19}$$

In a tournament, ...

$$\text{maximize } R^{\text{race}} = \int_{\hat{a}}^{\infty} y^*(x, \underline{y}) dF_{N:N}(x) - [1 - F_{N:N}(\hat{a})]. \tag{20}$$

Convexity is not sure. If not, then the optimal target is zero. Which is lower than the optimal target in a race.

Instead. If the objective function is (strictly) concave then there's an internal solution.

$$\begin{aligned}
\text{FOC } \Rightarrow & \frac{d \int_{\hat{a}}^{\infty} y^*(x, \underline{y}) dF_{N:N}(x))}{d\underline{y}} + F'_{N:N}(\hat{a})\hat{a}' = 0 \\
& \text{(by using Leibniz rule)} \\
\Rightarrow & - y^*(\hat{a}, \underline{y})\hat{a}' F'_{N:N}(\hat{a}) + \int_{\hat{a}}^{\infty} \frac{\partial y^*(x, \underline{y})}{\partial \underline{y}} dF_{N:N}(x) - F'_{N:N}(\hat{a})\hat{a}' = 0 \\
\Rightarrow & - \underline{y}\hat{a}' F'_{N:N}(\hat{a}) + \int_{\hat{a}}^{\infty} \frac{\partial y^*(x, \underline{y})}{\partial \underline{y}} dF_{N:N}(x) - F'_{N:N}(\hat{a})\hat{a}' = 0. \qquad (21)
\end{aligned}$$

Using (19) with (21), the optimal target is the same in the race and the tournament only if

$$\int_{\hat{a}}^{\infty} \frac{\partial y^*(x, \underline{y})}{\partial \underline{y}} dF_{N:N}(x) = [1 - F_{N:N}(\hat{a})]. \tag{22}$$

$$\frac{\partial y^*(x, \underline{y})}{\partial \underline{y}} = \frac{c'_y(\underline{y})}{c'_y(y^*(x, \underline{y}))}.$$

Then.

- If $c_y(\cdot)$ is linear, we have that the ratio is one for all x.
- If $c_y(\cdot)$ is convex, then we have that it is less than one. If

- If $c_y(\cdot)$ is concave, then we have that it is higher than one.

As a result, if linear or convex the first order condition is lower than that in the race. Since the obj. function is concave (second order is decreasing), the target should be lower in a tournament than in a race to satisfy the first order condition. (a lower target increases the focs.).

Conjecture. If $\tau > 0$, the $\underline{y}$ in the race is higher.

3.4 Structural econometric model

Readings:

- The winner's curse, reserve prices, and endogenous entry: Empirical insights from eBay auctions
- Entry and competition effects in first-price auctions: theory and evidence from procurement auctions
- Auctions with entry

General two-step strategy:

- First step. Identify the marginal type from the data and the distribution of types.
- Second step. Using the estimated distribution of types.

Basic idea. Equilibrium condition gives:

$$y_i^* = y^*(a_i; F_{\mathcal{A}}). \tag{23}$$

with $y^*(\cdot)$ being an invertible function with ϕ denoting the inverse.

Hence the distribution of bids is

$$F_Y(y) = \Pr(y_i^* \leq y) = \Pr(y^*(a_i) \leq y) = \Pr(a_i \leq \phi(y)) = F_{\mathcal{A}}(\phi(y)). \tag{24}$$

Identification of the model. suggest

4 Experimental design

Over the past few years, an extensive literature has focused on naturally-occurring data to provide new insights into the relevance of contest theory in numerous settings. The use of naturally-occurring data for studying differences between races and tournaments, however, is problematic. As our simple theoretical model indicates, both competitors and contest designers will expect sensible differences in payoffs between a race and a tournament competition. This creates a problem of selection that may bias an analysis based on natural occurring data.

Instead of using naturally occurring data, we test our theory by designing and executing a field experiment. In doing so, one need an environment in which the same contest can be "replicated" under different competition styles, while maintaining constant all the other characteristics of the contest. It is also crucial to have competitors registering for the contest before learning about the competition style to avoid the selection problem. And there must be a way to observe the decisions made during the contest, including the choice of entering the contest, as well as the timing and quality of the submissions made by the entrants.

Such an environment was provided by the online platform "Topcoder.com" who agreed to provide i) access to its large member base of competitors (over 1 million registered users in 2016) and ii) access to its platform tools for managing online contests (web forums, leaderboards, payment methods) that we used for the execution of our experimental design.

A few key factors made this platform ideal and unique for our experiment.

- First, platform members are "sophisticated" competitors. They typically join the platform to participate in periodic programming competition, called "Marathon matches," where they can win prize money for solving a computational problem (e.g., xxxx). Hence, they are expected to understand well the competition process and to be familiar with the idea of competing for winning prizes. In addition, some of them appear fairly strategic in their behavior, as discussed by xxxx and xxxx.

- Second, the platform provides us with different measures of competitors' individual ability based on their past performance, including a "skill rating" that provides a metric of their ability as contestants [xxx]. The fact that these measures are publicly available for each member on the platform is particularly important because it is consistent with our theoretical approach that presumes competitors have common beliefs about the overall distribution of skills in the contest (such a common belief can be based on observing the distribution of ratings of competitors in a contest).

- Third, the platform provides us with a rich data analytics on participation, giving us the capability of measuring accurately timing and extent of participation of each competitor. It also provides an objective metric of submission quality in the form of automatic scores based on xxxx.

The timing of the experiment was as follows.

- As a first step, a four day registration period for a new competition was announced on XXXX. The announcement was inviting platform members to register and participate in a contest for solving a hard information extraction problem (the automatic extraction of structured information from biomedical research papers). As an incentive for participation, a total prize pool of $41,000 was offered to top submissions in the form of cash prizes. The announcement was sent via email to all newsletter subscribers and was publicized through posts on the platform's blog. Contest participation was limited to members who had some minimal experience in programming competitions on the platform (the requirement was to have at least one registration to a prior competition).

- The online registration involved signing an informed consent for the research, as well as responding to a short initial survey. After registration, we sorted all registered members into 24 "virtual rooms" (note that 24 was the largest number of concomitant virtual rooms allowed by the platform at that time). Each room can be viewed as an independent contest where we offered cash prizes of $1,000 and $100 to the first and second placed competitor, respectively. In each of these rooms, competitors had access to a leaderboard (that was updated about every 48 hours), a web forum to ask questions about the computational problem, and a submission system through which they could submit their codes. Submitted codes were then scored offline, generating preliminary scores that were then used to update the room leaderboards. Competitors could submit their codes during a 8 day submission period.

- Each room was then randomly assigned to a competition style and a room size. We examine three competitive styles: i) tournament, ii) race, and iii) tournament with minimum-quality requirement (hereinafter "reserve"); and two room sizes: i) 15 competitors and ii) 10 competitors. Hence, our experimental design can be xxxx by a 3x2 matrix, as shown in Table XXXX.

In the tournament, the winner was the submission with the highest score computed on the last submission made within the 8 day submission period. The second placed competitor was the second highest score. In the race, the winner was the first xxxx. in a period at the end of the competition was awarded a grand prize of $6,000. In the race condition, the first to achieve a score of xxxx was awarded the same grand prize of $6,000. Finally, in the tournament with a minimum-quality requirement, we awarded a grand prize of $6000 to the top submission achieving a score greater or equal than a given threshold (xxxxx).

- The threshold in the "race" and in the "reserve" was chosen following two main criteria. First, we run a pre-trail experiment that involved 4 coders solving the same problem in isolation for 5 days. This helped us forming basic predictions about xxxx. Second, we surveyed the NIH researchers who developed Banner asking for three percentage improvements they considered "useful," "desirable," and "unlikely."

- The "treatment" was announced at the start of the submission phase via email and in one section of the description of the computational problem as well.

- Finally, payments were administered by the platform. Additional rewards (limited-edition T-shirts) were offered to those responding to a final survey.

6

7

8

[6]Typical problems include a wide range of data science problems such as classification and regression, image processing, and natural language processing. Solutions are submitted in the form of computer programs to be run and scored by the platform. the problem demands good programming skills as well as a strong background in machine learning and statistics.

[7]The typical competition format on Topcoder is the tournament. At the end of a given submission period, the last submissions of players are scored and ranked by performance on a holdout dataset. Based on the final ranking, top submissions are awarded cash prizes. The extent of prizes depends on the nature and complexity of the problem but is generally between $5,000 and $20,000. In addition to monetary incentives, all active competitors attain a skill rating that provides a metric of their ability as contestants and sometimes play a role in signaling skills to potential employers [need ref.]. Topcoder also host non-rated events. On occasions, Topcoder also hosts race competitions, called first-to-finish. Compared to the marathon matches, these other competitions tend to be employed for less challenging problems and with smaller cash prizes. This practice, which makes problematic comparisons based on existing data, seems more motivated by tradition (with first-to-finish formats being introduced at a recent time and at a small scale) than by a proper calculation of the potentially different benefits associated with one or the other competition format.

[8]To select a challenging data science problem for our competition, we worked together with researchers from the United States National Health Institute (NIH) and the Scripps Research Institute (SCRIPPS). The selected problem was based on an algorithm called BANNER that was built by researchers at the NIH (Leaman, Gonzalez, and others 2008). The algorithm uses domain-expert manual labeling to train a natural language entity recognition model that performs automatic annotation of abstracts from a large corpus of biomedical research papers. Automatic annotations help disease characteristics to be more easily identified. The specific goal of the programming competition was to improve upon the current NIH's system by using a combination of domain-expert and non-expert manual labeling (e.g., Good et al. 2014).

9

4.1 Data

We collected basic platform data including membership registration and participation in past programming competitions. The online registration survey provided additional demographic information, including gender, age, geographic origins, education, and most preferred programming language. In addition, we asked registrants their willingness to take risks "in general," as a measure of risk aversion (Dohmen et al. 2011), and a forecast of how many hours they expected to be able to work on the problem during the submission phase of the challenge.[10]

Table 1: Descriptive statistics

	Mean	Median	St.Dev.	Min	Max	Obs.	P-value
year	2009.9	2010	4	2001	2015	299	0.596
rating	1322.4	1278	425	593	3071	205	0.989
registrations	17.6	9	23	1	161	299	0.626
submissions	7.2	2	12	0	91	299	0.867
lpaid	8.4	8	3	3	14	139	0.791
nwins	0.3	0	2	0	27	299	0.370
ntop10	1.6	0	5	0	64	299	0.273
risk	6.4	7	2	1	10	279	0.958
hours	31.3	24	25	0	192	277	0.995
male	1.0	1	0	0	1	276	0.404
timezone	2.1	2	5	-8	10	277	0.389
grad	0.5	0	1	0	1	278	0.208
below30	0.7	1	0	0	1	278	0.503

*Notes:*Variables definition: 'year' is the year of membership registration; 'rating' is the skill rating; 'registrations' is the count of registered competitions; 'submissions' is the count of registered competitions with a submission; 'lpaid' is the logarithm of the total prize money won; 'nwins' the count of wins; 'ntop10' the count of top 10 placements; 'risk' a measure of risk aversion; 'hours' forecast of total hours of work; 'male' indicates the gender; 'timezone' is of residence at the time of the competition; 'grad' is an indicator for graduate or post-graduate educational degree; and 'below30' indicates age below 30 years old.

Table 1 shows descriptive statistics. It also provides p-values from a series of Kruskal-Wallis and Fisher's exact non-parametric tests, showing no evidence of systematic differences across treatment groups (the lowest p-value was 0.208). Hence, our randomization appears successful.

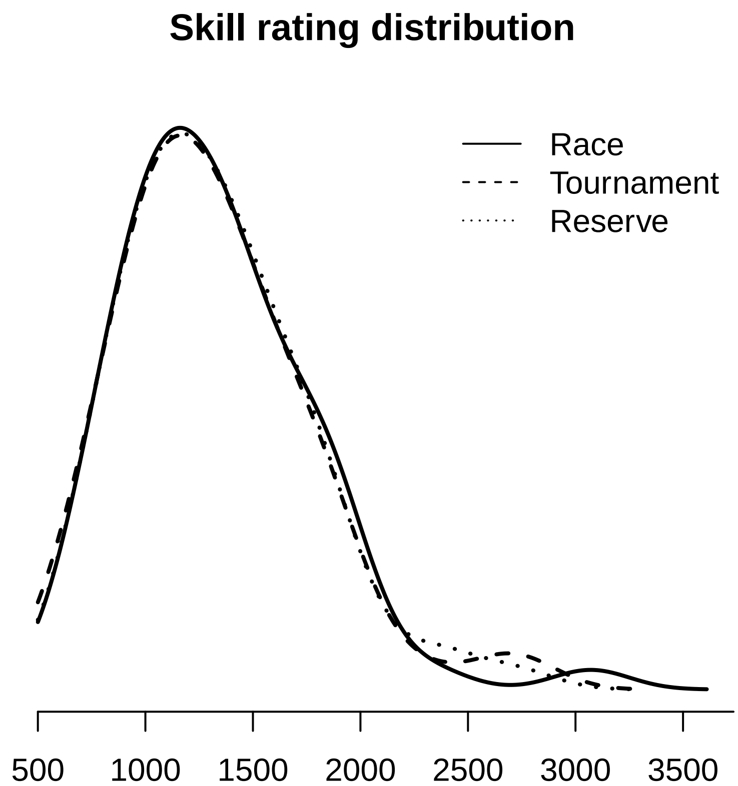
[9] A problem statement containing a full description of the algorithmic challenge, the rules of the game, and payoffs was published at the beginning of the submission phase. The submission phase was of 8 days in which participants could submit their computer programs. Each submission was automatically scored and feedback in the form of preliminary scores was published regularly on the website via the leaderboard.

[10] The exact question was: "The submission phase begins March 08. Looking ahead a week, how many hours do you forecast to be able to work on the solution of the problem?"

Overall, our sample constituted a group of expert members of the platform. The median registrant had been signed-up for more than 5 years and had registered to 9 competitions (about 2 competitions per year) of which 2 with submissions. Thus, each subject had in median 32 percent probability of making submissions after registration. Male (95 percent) and below-30-years-old (66 percent) registrants were predominant, reflecting the gender and age distribution of the platform.[11] Registrants also reported forecasting a median of 24 total hours of work over the eight day submission period, showing a strong commitment to the challenge.

A key pre-treatment variable was the individual ability of registrants. The platform provided us with several proxies. A sensible measure was the skill rating. The skill rating is an elo-type measure of a competitor's relative ability compared to other platform members. This measure is represented by a number that increases or decreases depending on the difference between an hypothetical expected rank — based on the skill rating of the opponents — and the actual rank achieved by a competitor at the end of a competition. If the actual rank is higher than the expected rank, the skill rating increases. If, instead, the actual rank is lower than the expected rank, the skill rating decreases.

[11] While such a large male predominance can be surprising, it is not specific to the platform under study but, as researchers have found [XXX], it is a common attribute of many online platforms, including very popular platforms such as Wikipedia and StackOverflow.



As shown in Figure **??**, the distribution of the skill rating was right-skewed reflecting the presence of a few individuals with very high ratings compared to the average. This suggests the presence of large skill differences in our sample. The value of this proxy, however, was missing for those who had no history of submissions (about 31 percent). Other proxies that we considered include the count of past top ten positions, the count of wins, and the total prize money won while being a member of the platform. These other proxies are highly correlated with but exhibit less variation than the skill rating and, therefore, appear less apt to differentiate between registrants with different abilities.

```
## Error in detach(races): invalid 'name' argument
```

5 Results

5.1 Room differences

Competitors in 24 rooms were randomly assigned to either a race, a tournament, or a tournament with reserve competition. After an eight day submission period, the count of competitors with submissions, scores, and timing of submissions were recored for each room. The distribution of room entrants, average scores, and average timing is shown in Figure 1.

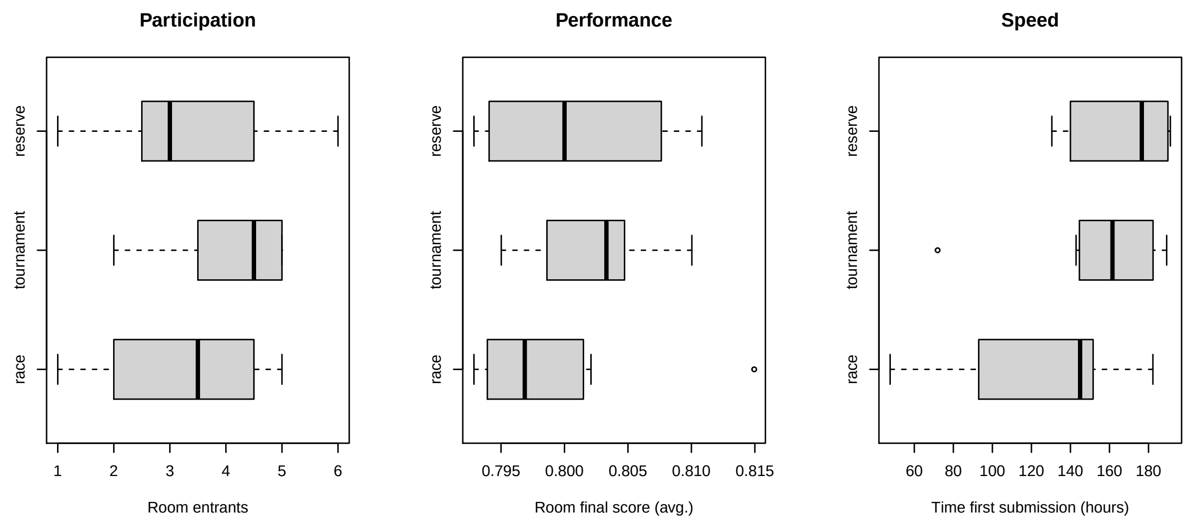


Figure 1: Distribution of room outcomes by competition style

Our theoretical model predicts that competitors will enter more a Tournament than a Race or a Tournament with Reserve competition, given the absence of any performance target. Consistent with this prediction, the left panel of Figure 1 shows a higher median count of room entrants in the Tournament than in the other groups. To test to see if the difference in means between room entrants in the Tournament and the other groups is greater than zero, we use a Welch Two Sample t-test which gives a (one-sided) p-value of 0.077. Bootstrap resampling gives very similar results, indicating robustness to problems due to our small sample size. Hence, there is evidence to suggest that the average count of room entrants in the Tournament was greater than in the other groups.

```
##  [1] 0.450 0.340 0.591 0.552 0.338 0.742 0.375 0.143 0.453 0.132 0.941
## [12] 0.248 1.000 0.759
```

According to our model, the higher proportion of entrants in the Tournament will be driven by low-

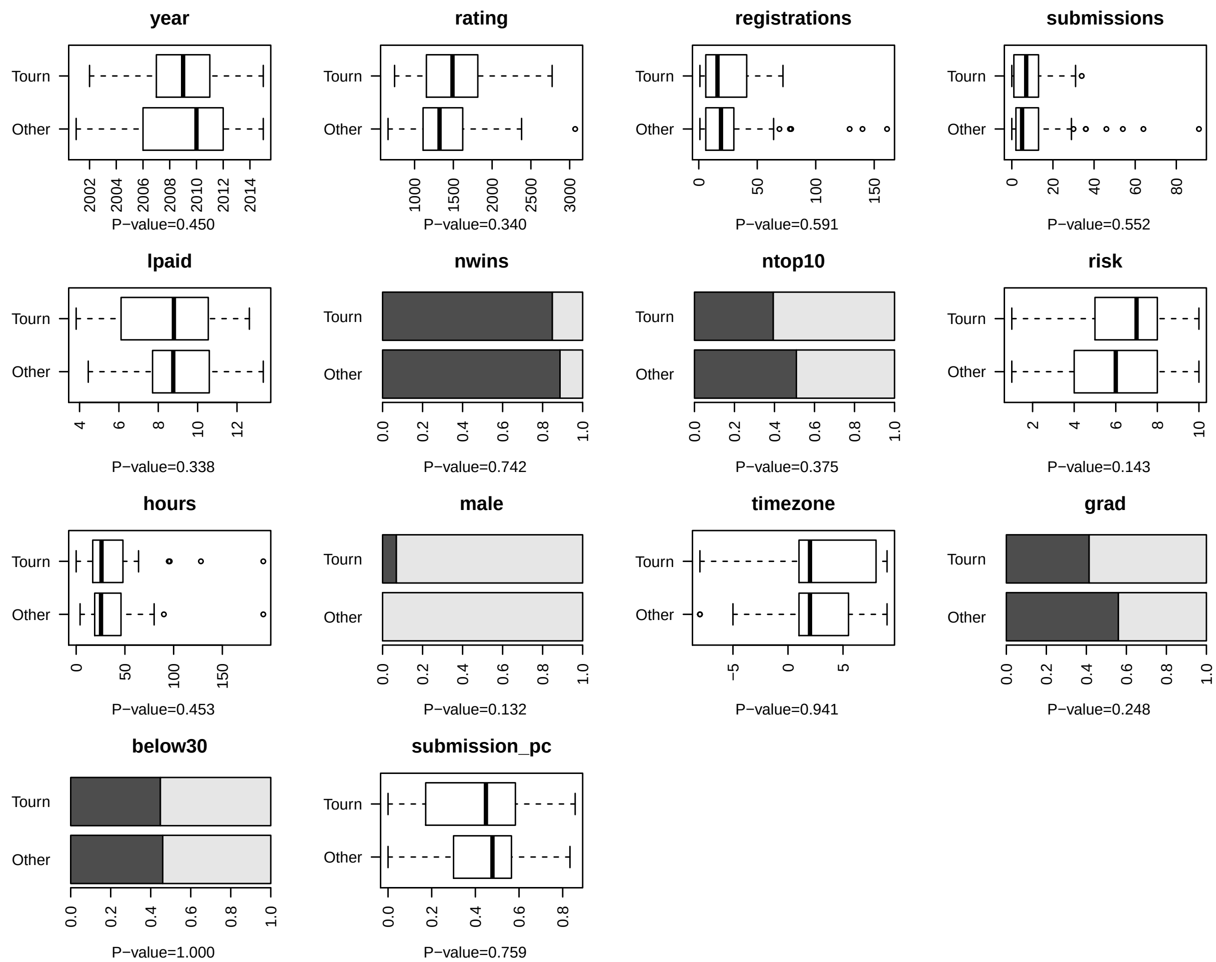


Figure 2: Conditional distribution of competitors' characteristics given entry

skilled competitors. We test this prediction on three main measures of individual skills — the skill rating (`rating`) and whether the competitor had ever won (`nwins`) or been in the top 10 rank (`ntop10`) of a competition before. While entrants had on average higher skills than non-entrants, we find no evidence supporting any skill-based selection across treatments using these measures (see Figure 2). Another possibility was that differences in time availability like being in a different timezone than the United States could have had a negative effect on competitors' participation in a Race. We find no evidence supporting this view by examining differences across treatments in the competitors' time zone automatically identified at registration (`timezone`) and self-reported, forecasted hours of work (`hours`) that we collected prior to the submission phase. Finally, we also find no differences in other demographics like age, gender, and platform registration year and experience. Thus, somewhat contrary to our model's predictions, the higher proportion of entrants in the Tournament did not seem to be associated with measures of skills, time availability, and experience. This result could be in part due to our imperfect measures of skills that do not consent fine-grained analysis. On the other hand, one may speculate that the extra participation was driven by an individual "taste" for the Tournament over the other competition styles.

```
##
##  Kruskal-Wallis rank sum test
##
## data:  with(final, split(final.cap, treatment))
## Kruskal-Wallis chi-squared = 2, df = 2, p-value = 0.4
```

```
##
## Call:
## lm(formula = final.cap ~ treatment, data = final)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.008018 -0.004903 -0.000712  0.002628  0.015901
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)          0.79905    0.00229  348.76   <2e-16 ***
## treatmenttournament  0.00325    0.00324    1.00     0.33
## treatmentreserve     0.00184    0.00324    0.57     0.58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00648 on 21 degrees of freedom
## Multiple R-squared:  0.0459, Adjusted R-squared:  -0.0449
## F-statistic: 0.506 on 2 and 21 DF,  p-value: 0.61
```

```
##
##  F test to compare two variances
##
## data:  final.cap by treatment == "reserve"
## F = 0.7, num df = 20, denom df = 7, p-value = 0.6
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.164 2.464
## sample estimates:
## ratio of variances
##              0.748
```

We next examine differences in performance across competition styles (see middle panel of Figure 1). Our theoretical model indicates the Tournament with Reserve will provide higher scores on average compared to the other groups. We test this prediction using the mean of the competitors' last scored submission in each room. One problem in computing the room means was the need to cope with extreme values in the distribution of scores, as any small bug in the code could generate very low scores (e.g., zero), as shown in Figure 3. To deal with extreme values we used two methods (1) we trimmed last scores before computing the room means and (2) we capped scores at the value of the baseline score (i.e., the score one would obtain by submitting the BANNER's algorithm without making any useful change). Using either measure, we do not find evidence supporting our hypothesis that mean room scores were higher in the Tournament with

Reserve compared to the other treatments. The reader may find this finding not entirely surprising given the lack skill-based selection across groups that we documented earlier.

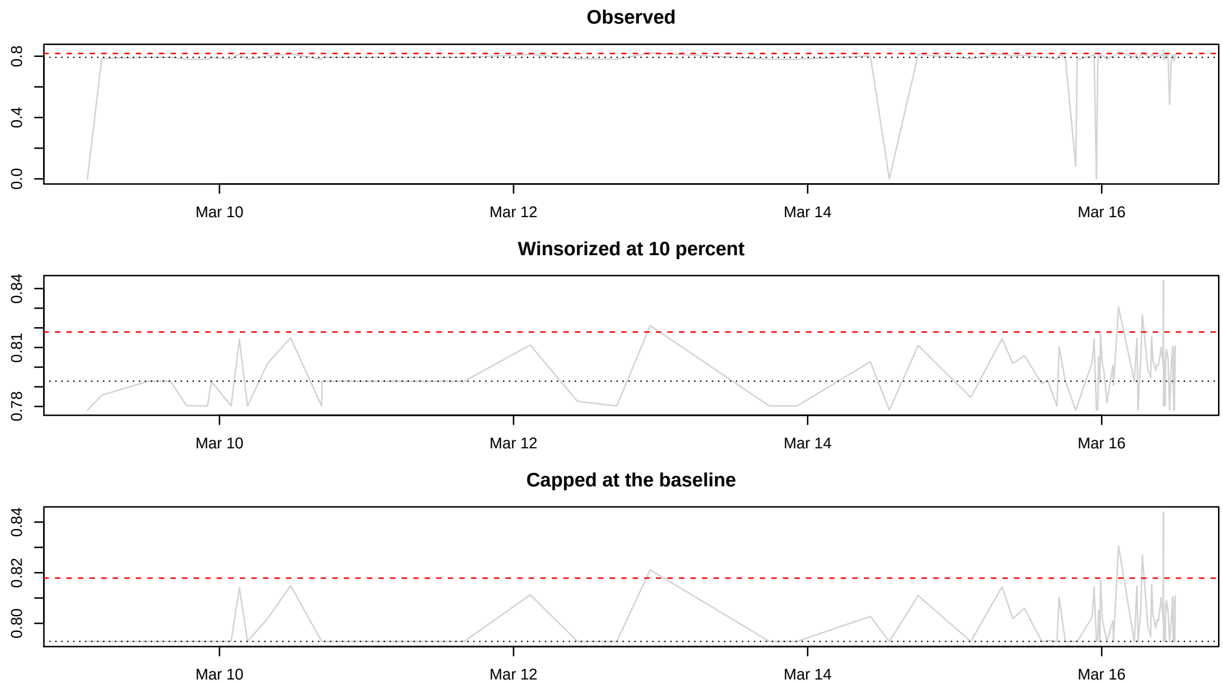


Figure 3: Scores over time

Finally, we examined differences in submission speed measured by the mean time of the first submission in each room. As shown in the right panel of Figure 1, the median room time-to-submit in the Race was about 20 and 40 hours shorter than the Tournament and the Tournament with Reserve, respectively. The variance, however, was also larger with values ranging from below 60 to 190 hours, whereas by comparison the other groups' distributions were both above the 120 hours. To test to see if the difference in means between speed in the Race and the other groups was greater than zero, we used a Welch Two Sample t-test which gives a (one-sided) p-value of 0.038. [Bootstrap] Thus, our data support the hypothesis that competitors' speed in a Race was higher than in the other groups. Taken together with the absence of skill-based selection and score differences, this result implies that competitors in the Race competition have exerted greater efforts (i.e., by lowering execution time while keeping performance at a comparable level) relative to the other groups.

To summarize the results obtained so far, we have found evidence supporting a higher participation in the Tournament. This higher participation, however, does not seem to be driven by low-skilled competitors. We have also found that competitors in the race made submissions faster without sacrificing performance, which suggests that they have paid higher costs from effort associated with the accelerated speed compared to other competitors. Finally, we reject the hypothesis that the Tournament with Reserve yields higher performance levels. [One explanation is that competitors were fishing to hit the threshold and then stopped exerting effort].

5.2 Self-reported measures of hours worked

To partially corroborate [final survey]

5.3 Panel data

5.4 Structural model

To understand individual propensities to enter the contest, we now specify a logistic regression model for the conditional probability of entry ($Y_i = 1$, entry; $Y_i = 0$, exit) given the assigned competition style (Z_i), and a matrix of control variables (X_i):

$$\Pr(Y_i = 1 \mid X_i = x_i, Z_i = z_i) = \text{probit}^{-1}(z_i + x_i\beta) \tag{25}$$

The estimates show that:

- the difference in individual propensity between tournaments and other competition style is positive but not significant (low power)
- individual skill ratings are good predictor of participation (an increase of 100 points corresponds to about a 6 percent increase in the odds of submitting).
- The stated hours are also important predictors (an increase of 1 hour corresponds to a 1 percent increase in probability).
- This model has a nice structural interpretation. Players have a latent "ability" variable that is based upon their own skill rating plus some error noise. They submit if and only if ability is higher than a threshold, which is determined by the competition styles and room size.

Interesting test. Is the skill rating distribution conditional on entry different across competition. A t.test rejects this hypothesis.

```
## Error in density.default(X[[i]], ...): 'x' contains missing values
```

Table 2:

	Dependent variable:					
	submit					
	(1)	(2)	(3)	(4)	(5)	(6)
treatmenttournament	0.195	0.191	0.166	0.161	0.163	0.167
	(0.188)	(0.223)	(0.226)	(0.227)	(0.228)	(0.226)
treatmentreserve	0.022	0.128	0.140	0.149	0.177	0.143
	(0.191)	(0.225)	(0.226)	(0.227)	(0.230)	(0.226)
rating.100		0.064^{***}	0.068^{***}	0.069^{***}	0.064^{***}	0.067^{***}
		(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
hours.imp			0.009^{**}	0.009^{**}	0.010^{**}	0.009^{**}
			(0.004)	(0.004)	(0.004)	(0.004)
timezone.imp				−0.009		
				(0.018)		
male.imp					0.088	
					(0.697)	
below30.imp					-0.320^{*}	
					(0.191)	
risk.imp						−0.008
						(0.042)
Constant	-0.635^{***}	-1.320^{***}	-1.650^{***}	-1.660^{***}	-1.530^{**}	-1.590^{***}
	(0.136)	(0.335)	(0.365)	(0.366)	(0.780)	(0.468)
Observations	299	205	205	205	205	205
Log Likelihood	−179.000	−129.000	−126.000	−126.000	−125.000	−126.000
Akaike Inf. Crit.	363.000	266.000	262.000	264.000	263.000	264.000

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

```
##
##  Kruskal-Wallis rank sum test
##
## data:  rating_l
## Kruskal-Wallis chi-squared = 2, df = 2, p-value = 0.4
```

```
##
##  Welch Two Sample t-test
##
## data:  rating_l$race and rating_l$tournament
## t = -0.4, df = 50, p-value = 0.7
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -331  218
## sample estimates:
## mean of x mean of y
##      1452      1509
```

Our model also predicts lower participation for individuals in large rooms (15 competitors) compared to those in small rooms (10 competitors). Under the model, the "marginal type" is increasing in group size and so the individual probability of entry is lower. However, our data proved only negligible differences in participation between large (28.9 percent) and small rooms (28.6 percent). Additionally, we found no significant treatment differences conditional on the room size being large or small (xxxxx). Hence, one may conclude that a group-size difference of 5 people is probably not large enough to be impactful.

```
##
##  Welch Two Sample t-test
##
## data:  nsub_l$race and nsub_l$tournament
## t = 0.6, df = 50, p-value = 0.5
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
##  -0.594  1.128
## sample estimates:
## mean of x mean of y
##      2.20      1.93
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  nsub_l$race and nsub_l$tournament
## W = 500, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
```

The median submission per participant was of 0 submissions, with a minimum of 0 and a maximum of 126 submissions.

```
##  setting  value
##  version  R version 3.3.2 (2016-10-31)
##  system   x86_64, darwin13.4.0
##  ui       X11
##  language (EN)
##  collate  en_US.UTF-8
##  tz       America/New_York
##  date     2017-06-29
##
##  package   * version date       source
##  backports   1.0.5   2017-01-18 CRAN (R 3.3.2)
##  boot      * 1.3-18  2016-02-23 CRAN (R 3.3.0)
##  codetools   0.2-15  2016-10-05 CRAN (R 3.3.2)
##  devtools    1.12.0  2016-06-24 CRAN (R 3.3.0)
##  digest      0.6.12  2017-01-27 CRAN (R 3.3.2)
```

```
##   evaluate    0.10    2016-10-11 CRAN (R 3.3.0)
##   highr       0.6     2016-05-09 CRAN (R 3.3.0)
##   htmltools   0.3.5   2016-03-21 CRAN (R 3.3.0)
##   knitr       1.15.1  2016-11-22 CRAN (R 3.3.2)
##   lattice     0.20-34 2016-09-06 CRAN (R 3.3.2)
##   magrittr  * 1.5     2014-11-22 CRAN (R 3.3.0)
##   Matrix      1.2-7.1 2016-09-01 CRAN (R 3.3.2)
##   memoise     1.0.0   2016-01-29 CRAN (R 3.3.0)
##   races     * 0.2     2017-06-08 local (@0.2)
##   Rcpp        0.12.9  2017-01-14 CRAN (R 3.3.2)
##   rmarkdown   1.3     2016-12-21 CRAN (R 3.3.2)
##   rprojroot   1.2     2017-01-16 CRAN (R 3.3.2)
##   stargazer * 5.2     2015-07-14 CRAN (R 3.3.0)
##   stringi     1.1.2   2016-10-01 CRAN (R 3.3.0)
##   stringr     1.2.0   2017-02-18 CRAN (R 3.3.2)
##   survival  * 2.40-1  2016-10-30 CRAN (R 3.3.0)
##   withr       1.0.2   2016-06-20 CRAN (R 3.3.0)
##   xtable    * 1.8-2   2016-02-05 CRAN (R 3.3.0)
##   yaml        2.1.14  2016-11-12 CRAN (R 3.3.2)
```

References

Athey, Susan, and Philip A Haile. 2002. "Identification of Standard Auction Models." *Econometrica* 70 (6). Wiley Online Library: 2107–40.

———. 2007. "Nonparametric Approaches to Auctions." *Handbook of Econometrics* 6. Elsevier: 3847–3965.

Athey, Susan, Jonathan Levin, and Enrique Seira. 2011. "Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions*." *The Quarterly Journal of Economics* 126 (1). Oxford University Press: 207–57.

Baye, Michael R, and Heidrun C Hoppe. 2003. "The Strategic Equivalence of Rent-Seeking, Innovation, and Patent-Race Games." *Games and Economic Behavior* 44 (2). Elsevier: 217–26.

Bimpikis, Kostas, Shayan Ehsani, and Mohamed Mostagir. 2014. "Designing Dynamic Contests." Working paper, Stanford University.

Che, Yeon-Koo, and Ian Gale. 2003. "Optimal Design of Research Contests." *The American Economic Review* 93 (3). American Economic Association: 646–71.

Dechenaux, Emmanuel, Dan Kovenock, and Roman M Sheremeta. 2014. "A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments." *Experimental Economics*. Springer, 1–61.

Dixit, Avinash Kamalakar. 1987. "Strategic Behavior in Contests." *The American Economic Review* 77 (5). American Economic Association: 891–98.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association* 9 (3). Wiley Online Library: 522–50.

Donald, Stephen G, and Harry J Paarsch. 1996. "Identification, Estimation, and Testing in Parametric Empirical Models of Auctions Within the Independent Private Values Paradigm." *Econometric Theory* 12 (03). Cambridge Univ Press: 517–67.

Good, Benjamin M, Max Nanis, CHUNLEI Wu, and ANDREW I Su. 2014. "Microtask Crowdsourcing for Disease Mention Annotation in Pubmed Abstracts." In *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 282–93. NIH Public Access.

Green, Jerry R, and Nancy L Stokey. 1983. "A Comparison of Tournaments and Contracts." *The Journal of Political Economy* 91 (3): 349–64.

Grossman, Gene M, and Carl Shapiro. 1987. "Dynamic R&D Competition." *Economic Journal* 97 (386). Royal Economic Society: 372–87.

Harris, Christopher, and John Vickers. 1987. "Racing with Uncertainty." *The Review of Economic Studies* 54 (1). Oxford University Press: 1–21.

Laffont, Jean-Jacques, Herve Ossard, and Quang Vuong. 1995. "Econometrics of First-Price Auctions." *Econometrica* 63 (4). Econometric Society: 953–80.

Lazear, Edward P, and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *The Journal of Political Economy* 89 (5): 841–64.

Leaman, Robert, Graciela Gonzalez, and others. 2008. "BANNER: An Executable Survey of Advances in Biomedical Named Entity Recognition." In *Pacific Symposium on Biocomputing*, 13:652–63.

Mary, O'Keeffe, W Kip Viscusi, and Richard J Zeckhauser. 1984. "Economic Contests: Comparative Reward Schemes." *Journal of Labor Economics* 2 (1). University of Chicago Press: 27–56.

Moldovanu, Benny, and Aner Sela. 2001. "The Optimal Allocation of Prizes in Contests." *The American Economic Review*. JSTOR, 542–58.

———. 2006. "Contest Architecture." *Journal of Economic Theory* 126 (1). Elsevier: 70–96.

Paarsch, Harry J. 1992. "Deciding Between the Common and Private Value Paradigms in Empirical Models of Auctions." *Journal of Econometrics* 51 (1-2). Elsevier: 191–215.

Parreiras, Sérgio O, and Anna Rubinchik. 2010. "Contests with Three or More Heterogeneous Agents." *Games and Economic Behavior* 68 (2). Elsevier: 703–15.

Siegel, Ron. 2009. "All-Pay Contests." *Econometrica* 77 (1). Wiley Online Library: 71–92.

———. 2014. "Contests with Productive Effort." *International Journal of Game Theory* 43 (3). Springer: 515–23.

Szymanski, Stefan. 2003. "The Economic Design of Sporting Contests." *Journal of Economic Literature* 41 (4). American Economic Association: 1137–87.

Taylor, Curtis R. 1995. "Digging for Golden Carrots: An Analysis of Research Tournaments." *The American Economic Review* 85 (4). American Economic Association: 872–90.

Zizzo, Daniel John. 2002. "Racing with Uncertainty: A Patent Race Experiment." *International Journal of Industrial Organization* 20 (6). Elsevier: 877–902.