Races or Tournaments?[1]

[PRELIMINARY AND INCOMPLETE]

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

Last updated: 05 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

Government agencies, research institutions, and commercial companies sponsor regularly prize-based competitions ("contests") to engage workers and other participants in solving hard problems, generating business ideas, thinking up new products. Competitions are typically used to maximize competitors efforts while minimizing the time it takes to complete the task. Balancing these two desirable, though often incompatible, aims is, however, a recurring issue in the design of these competitions. First, contest designers may lack adequate knowledge of costs and skills of participants. Hence, design choices are made under considerable uncertainty. Second, competitors do XXXX, YYY, ZZZ.

In this article, we investigate the trade-off betwen time and output quality byb comparing two different competition formats: the "race," where the first to finish an innovation project wins, and the "tournament," where the best finished project wins. While most of past literatured focused on xx and yy, there is no comparison of xxx and xxx.

Tough races and tournaments are widespread, the basic trade-off between xxx and yyy is not well studied. 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.[1] 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 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.

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

[1] This example is taken...

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.

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

Consider a contest game in which there are $i = 1, ..., n$ players willing to compete for $k = 1, ..., q$ prizes of decreasing value $v_1 \geq v_2 \geq ... \geq v_q \geq 0$. Players simultaneously decide how much quality y_i and how much time t_i they spend in a given production task. This decision dictates the cost they incur and the probability of winning a prize.

The cost is determined by a function $C(\cdot)$ that is increasing in q_i, decreasing in t_i, and varies based upon an individual ability a_i which is meant to reflect differences in skills, time constraints, and other elements affecting quality and time in production. The cost function is assumed to have the following (Cobb-Douglas) form:

$$C(a, y, t) = a^{\alpha} y^{\beta} t^{\gamma} \tag{1}$$

with $\alpha, \gamma < 0$ and $\beta > 1$. Thus, the higher the quality over time ratio (the "speed") or the lower the level of individual ability, the higher will be the production costs incurred by players.

There is asymmetric information about the cost parameters. While players know their own ability, they are not aware of the ability of the other players. It is, however, common knowledge that abilities are drawn at random from a common cumulative distribution function $F(\cdot)$ with density $f(\cdot)$ on a bounded interval $[\underline{a}, \bar{a}]$ with $\underline{a} > 0$.

Based on this information, players maximize the following expected payoff:

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

where $p_k(\cdot)$ denotes the conditional probability of winning a prize given player i's quality y_i and time t_i.

We further consider two general limitations to competition regarding time ("deadlines") and quality (minimum-quality "targets"). In a contest where players have to meet a fixed deadline $\bar{t}$ to be elegible for prizes, the conditional probability of winning a prize is zero when $t_i > \bar{t}$. In a contest where players have to achieve a minimum-quality target $\underline{y}$ to be elegible for prizes, the conditional probability of winning a prize is zero when $y_i < \underline{y}$.

And we examine two types of competition: the tournament and the race.

A tournament competition is a contest with a deadline (and possibly a minimum-quality target) where the player having achieved the highest quality before the deadline gets the first prize, the player having achieved the second highest output quality before the deadline gets the second prize, and so on. Let denote the k'th smallest of the y_i's by $y_{k,n}$ ($y_{1,n}$ being the smallest, $y_{2,n}$ being the second smallest, and so on) with the convention that, when a player passed the deadline, the corresponding quality is zero. Then, player i's conditional probability of winning the first prize is:

$$p_1^T(y_i, t_i) = \Pr(y_i \geq y_{n-1:n-1}) \tag{3}$$

when $t_i \leq \bar{t}$, and is zero otherwise; the conditional probability of winning the second prize is:[2]

$$p_2^T(y_i, t_i) = [1 - \Pr(y_i \geq y_{n-1:n-1})] \Pr(y_i \geq y_{n-2:n-2}) \tag{4}$$

when $t_i \leq \bar{t}$, and is zero otherwise; and so on.

A race competition is a contest with minimum-quality target $\underline{y}$ where the first player to achieve a given minimum quality target $\underline{y}$ gets the first prize, the player being the second to achieve the target gets the second prize, and so on. Let denote the k'th smallest of the t_i's by $t_{k,n}$ ($t_{1,n}$ being the smallest, $t_{2,n}$ being the second smallest, and so on). Player i's conditional probability of winning a first prize in a race is:

$$p_1^R(y_i, t_i) = \Pr(t_i \leq t_{n-1:n-1}) \tag{5}$$

when $t_i \leq \bar{t}$ and $y_i \geq \underline{y}$, and is zero otherwise.

In other words, races and tournaments are a special case of a general contest game but with different probabilities of winning.

Let denote the actions of the winner of the contest by the vector (y^w, t^w). From the point of view of the contest designer, the expected payoff (net of payments) is:

$$R = E[y^w] - \tau E[t^w]$$

where τ denotes the contest designer's preference for expected time of the output (e.g., in a tournament is the deadline).

3.2 Equilibrium

In this section, we solve the model for the unique symmetric Bayesian Nash equilibrium of players.

3.2.1 Tournament

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

[2]Here we use the fact that individual choices are simultaneous and, therefore, independent.

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.

$$\begin{aligned} 0 =& \alpha f_{(1:N-1)}(\phi)\phi' + (1-\alpha)\phi'\{[1 - F_{(1:N-1)}(\phi)]f_{(1:N-2)}(\phi) + \\ &+ f_{(1:N-1)}(\phi)F_{(1:N-2)}(\phi)\} - c_a(a)c_y(\underline{y})c'_\tau(t_i) \end{aligned} \tag{6}$$

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

$$y^*(a_i) = c_y^{-1}\left[c_y(\underline{y}) + \frac{1}{c_\tau(\bar{t})}\left(\alpha \int_{a_i}^{\bar{a}} A(z)dz + (1-\alpha)\int_{a_i}^{\bar{a}} B(z)dz\right)\right] \tag{7}$$

where

$$A(x) = \frac{1}{c_a(x)} f_{(n-1:n-1)}(x) \tag{8}$$

and

$$B(x) = \frac{1}{c_a(x)}\left\{\left[1 - F_{(n-1:n-1)}(x)\right] f_{(n-1:n-2)}(x) + f_{(n-1:n-1)}(x)F_{(n-1:n-2)}(x)\right\}. \tag{9}$$

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

$$R(a_i) = \alpha F_{n:n}(a_i) + (1-\alpha)[1 - F_{n:n}(a_i)]F_{n-1:n-1}(a_i). \tag{10}$$

% 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

$$t^*(a_i) = c_\tau^{-1}\left[c_\tau(\bar{t}) + \frac{1}{c_y(\underline{y})}\left(\alpha\int_{a_i}^{\bar{a}} A'(z)dz + (1-\alpha)\int_{a_i}^{\bar{a}} B'(z)dz\right)\right] \tag{11}$$

where

$$A(x) = \frac{1}{c_a(x)} f_{(n-1:n-1)}(x) \tag{12}$$

and

$$B(x) = \frac{1}{c_a(x)}\left\{\left[1 - F_{(n-1:n-1)}(x)\right] f_{(n-1:n-2)}(x) + f_{(n-1:n-1)}(x) F_{(n-1:n-2)}(x)\right\}. \tag{13}$$

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

$$\underline{a} = h(n, V, F_A, C, d). \tag{14}$$

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

$$
\begin{aligned}
R_{\text{tour}} &= \Pr(t_{(1:n)} \le \bar{t}) \left\{ \int y^*(x \mid t_{(1:n)} \le \bar{t}) dF_{n:n}(x) - \tau \bar{t} - 1 \right\} \\
&= \int_{\hat{a}}^{\bar{a}} y^*(x) dF_{n:n}(x) - \tau \bar{t} - 1. \qquad (15)
\end{aligned}
$$

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

$$
\begin{aligned}
R_{\text{race}} &= \Pr(a_{(N)} \ge \hat{a}) \left\{ \underline{y} - \alpha - \Pr(a_{(N-1)} \ge \hat{a})(1-\alpha) \right\} - \tau \int_{\hat{a}}^{\infty} t^*(x) dF_{N:N}(x) \\
&= [1 - F_{N:N}(\hat{a})] \left\{ \underline{y} - \alpha - [1 - F_{N-1:N}(\hat{a})](1-\alpha) \right\} - \tau \int_{\hat{a}}^{\infty} t^*(x) dF_{N:N}(x). \qquad (16)
\end{aligned}
$$

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

$$
\underline{R_{\text{race}}} = [1 - F_{N:N}(\hat{a})] \left\{ \underline{y} - \alpha - [1 - F_{N-1:N}(\hat{a})](1-\alpha) - \tau \bar{t} \right\} \qquad (17)
$$

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

$$
\overline{R_{\text{tour}}} = [1 - F_{N:N}(\hat{a})] \underline{y} + \int_{\hat{a}}^{\infty} y^*(x) dF_{N:N}(x) - \tau \bar{t} - 1. \qquad (18)
$$

$$
\begin{aligned}
\underline{R}_{\text{race}} \geq& \overline{R_{\text{tour}}} \\
[1 - F_{N:N}(\hat{a})](\underline{y} - 1 - \tau\bar{t}) \geq& [1 - F_{N:N}(\hat{a})]\underline{y} + \int_{\hat{a}}^{\infty} y^*(x) dF_{N:N}(x) - \tau\bar{t} - 1 \\
-[1 - F_{N:N}(\hat{a})](\tau\bar{t} + 1) \geq& \int_{\hat{a}}^{\infty} y^*(x) dF_{N:N}(x) - (\tau\bar{t} + 1) \\
F_{N:N}(\hat{a})(\tau\bar{t} + 1) \geq& \int_{\hat{a}}^{\infty} y^*(x) dF_{N:N}(x) \\
\tau \geq& \left[\frac{\int_{\hat{a}}^{\infty} y^*(x) dF_{N:N}(x)}{F_{N:N}(\hat{a})} - 1 \right] \frac{1}{\bar{t}} \qquad (19)
\end{aligned}
$$

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

$$
\text{maximize } R^{\text{race}} = [1 - F_{N:N}(\hat{a})](\underline{y}_{\text{race}} - 1). \qquad (20)
$$

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.

$$
\text{FOC } \Rightarrow -F'_{N:N}(\hat{a})\hat{a}'(\underline{y}_{\text{race}} - 1) + [1 - F_{N:N}(\hat{a})] = 0. \qquad (21)
$$

In a tournament, …

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

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.

$$\begin{aligned}
\text{FOC } \Rightarrow & \frac{d \int_{\hat{a}}^{\infty} y^*(x, \underline{y}) dF_{N:N}(x))}{d\underline{y}} + F'_{N:N}(\hat{a})\hat{a}' = 0 \\
& \text{(by using Leibniz rule)} \\
\Rightarrow & - y^*(\hat{a}, \underline{y})\hat{a}' F'_{N:N}(\hat{a}) + \int_{\hat{a}}^{\infty} \frac{\partial y^*(x, \underline{y})}{\partial \underline{y}} dF_{N:N}(x) - F'_{N:N}(\hat{a})\hat{a}' = 0 \\
\Rightarrow & - \underline{y}\hat{a}' F'_{N:N}(\hat{a}) + \int_{\hat{a}}^{\infty} \frac{\partial y^*(x, \underline{y})}{\partial \underline{y}} dF_{N:N}(x) - F'_{N:N}(\hat{a})\hat{a}' = 0.
\end{aligned} \tag{23}$$

Using (21) with (23), the optimal target is the same in the race and the tournament only if

$$\int_{\hat{a}}^{\infty} \frac{\partial y^*(x, \underline{y})}{\partial \underline{y}} dF_{N:N}(x) = [1 - F_{N:N}(\hat{a})]. \tag{24}$$

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

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:

$$y_i^* = y^*(a_i; F_{\mathcal{A}}). \tag{25}$$

with $y^*(\cdot)$ being an invertible function with ϕ denoting the inverse.
Hence the distribution of bids is

$$F_Y(y) = \Pr(y_i^* \leq y) = \Pr(y^*(a_i) \leq y) = \Pr(a_i \leq \phi(y)) = F_{\mathcal{A}}(\phi(y)). \tag{26}$$

Identification of the model. suggest

4 Experimental design

4.1 Context

The experiment was based on an eight-day programming competition hosted on the online platform Topcoder.com in 2016. The competition was inviting platform members to submit solutions to 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.

Programming competition with similar characteristics, called "Marathon matches," are hosted regularly on Topcoder. Most of these competitions are sponsored by organizations (governamental agencies, research institutions, corporations) seeking a solution to a hard computational problem and willing to connect with a large online community of potential solvers (Topcoder's registered members were over 1M in 2016). 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. Thus, the problem demands good programming skills as well as a strong background in machine learning and statistics.

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.

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, as described by Good et al. (2014).

4.2 Design

We study three competitive conditions: i) tournament, ii) race, and iii) tournament with minimum-quality requirement. All conditions were inviting competitors to solve the same problem under the same rules. The only difference was the structure of incentives. In the tournament condition, the top submission 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).

To focus on monetary from winning, we dropped ratings.

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."

A xxx-day preliminary registration phase resulted in 299 pre-registered members.[3] Registrants were then split into 24 separate rooms of either 10 or 15 people. Each room was then randomly assigned to a competitive condition.

Pre-registration To raise participation, additional . Note that in each condition competitors were sorted at random into 8 separate groups (four with 10 people and other four with 15 people). Hence, as shown in Table XXXX, our design generated a total of 3x8 = 24 groups.

Grand prizes of xxxx were awarded to the top xxx in every conditions.

The competition was announced on the platform and to all community members via email.

Table 1: Experimental design

	1	2	3	4	5	6	7	8
Race	9	10	10	10	15	15	15	15
Tournament	10	10	10	10	15	15	15	15
Reserve	10	10	10	10	15	15	15	15

The competition was announced on the platform and to all community members via email. A preliminary online registration was required to participate, which resulted in 340 pre-registered members. Among these pre-registered members, we selected the 299 with had registered to a programming contest at least once before the present contest. This choice was to ensure that participants were sufficiently experienced and understood the basic rules governing programming competitions.

In each of these groups, contestants were given access to a "virtual room" that is a private web page listing handles of the other participants in the group, a leaderboard to be updated regularly during the competition, and a common chat that they can use to ask clarifying questions with respect to the problem at hand.

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

[3]Here we excluded xxxx who xxxx requirements.

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.

4.3 Data

For each registered competitor, we collected basic data including membership registration to the platform and statistics about participation in past programming competitions. A key variable to measure was expected ability of competitors. We examined several proxies. For rated competitors, a sensible measure was the skill rating earned for the performance in past competitions. The value of this proxy, however, was missing for all unrated competitors who had no history of submissions. Other proxies that we considered include the count of past wins, top ten positions, and the total prize money won while being a member of the platform. Though complete for both rated and unrated individuals, these other proxies contain less information about expected ability because most competitors have never won a competition or earned a prize.

Additional demographic information was collected via a pre-registration survey where competitors were asked their gender, age, geographic origins, education, and the most preferred programming language.[4] We also asked their willingness to take risks "in general," as a measure of risk aversion (Dohmen et al. 2011); and a forecast on how many hours they expected to be able to work on the problem in the next few days of the challenge.[5]

Overall, our sample included an uneven mix of rated (69 percent) and unrated (31 percent) competitors. Since one requirement for study enrollment was the minimum of one past registration in a marathon match, all competitors had some experience. However, experience varied a lot between rated and unrated competitors. The rated competitors were very experienced being members of the platform for a median of 6 years and having made a median of 16 registrations to past competitions (about 4.4 competitions per year) of which 6 with submissions. The unrated ones had less years as members on the platform (a median of 2), less registrations (a median of 2), and no submissions.

As shown in Figure 1, the distribution of our main proxy of individual ability for the rated competitors (the skill rating) was right-skewed reflecting the presence of a few individuals with very high ratings compared to the average.[6] This asymmetry was mirrored by the other proxies as well (with 5 coders accounting for 62 percent of past wins). Hence, xxxx. The figure also shows that the distribution of skill ratings was the same in all three competitive conditions.

[4]This question seemed relevant because the system to be improved (BANNER) was written in Java.

[5]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?"

[6]This right-skewed property is not specific of our sample but seems to hold more generally for the distribution of

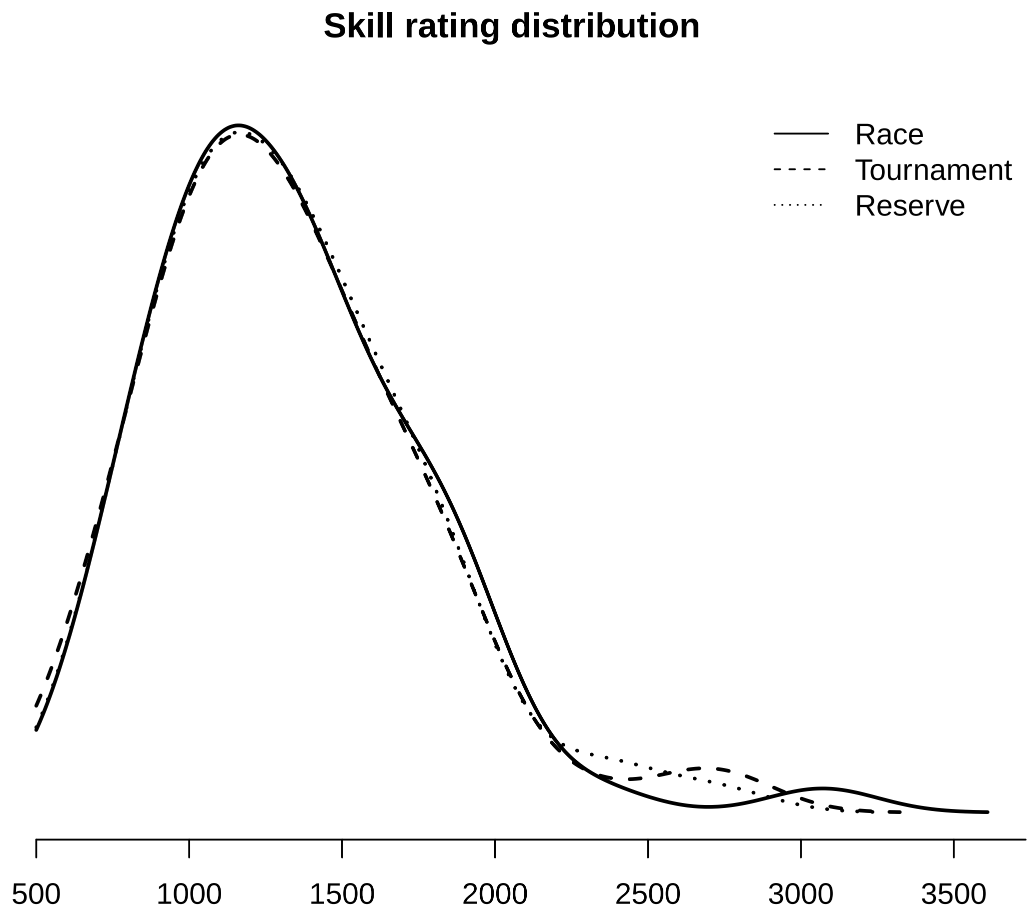

Figure 1: This picture shows kernel density estimates of the distribution of the skill ratings for each competitive condition. For testing whether samples originate from the same distribution we use a Kruskal-Wallis rank sum test that gives a pvalue of 0.989. Thus, we do not reject the null hypothesis of the data being drawn from the same distribution in each condition.

Table XXX reports summary statistics for our control variables. As shown in Figure ??, individuals reported being more willing to take risks than not (the median response was 7 out of 10) and prepared to work on the problem a median of 24 hours in total over the eight day submission period.

All experimental groups did not differ significantly in terms of the distribution of pre-treatment covariates. Using a Kruskal-Wallis rank sum test we find no difference in the distribution of participation measures (registrations, submissions) and ability proxies (skill rating, wins, top ten positions, earnings) across treatments (the lowest p-value was 0.273). Likewise, using a Pearson's Chi-squared test we find no association between each categorical variable (age, gender, education, programming language, risk attitudes,[7] and expected total hours of work) and treatments (the lowest p-value was 0.375). Hence, the randomization was successful in keeping pre-treatment variables balanced across competitive conditions.

```
##  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     2017-06-05
##
##  package    * version date       source
##  backports    1.0.5   2017-01-18 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)
##  magrittr   * 1.5     2014-11-22 CRAN (R 3.3.0)
##  memoise      1.0.0   2016-01-29 CRAN (R 3.3.0)
##  races      * 0.1     2017-05-28 local (@0.1)
##  Rcpp         0.12.9  2017-01-14 CRAN (R 3.3.2)
```

skill ratings of the entire platform.

[7]Risk attitudes can be also modeled as a continuous variable and tested using the Kruskal-Wallis rank sum test. Results of this test are reported in Figure ??.

```
##   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)
##   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.