Contingent Prizes in Dynamic Contests*

Jorge Lemus† Guillermo Marshall‡

May 5, 2022

Abstract

Firms and government agencies are increasingly procuring solutions to problems using online contests, which typically provide a real-time leaderboard and award prizes at their end date. Becoming the leader of the contest at any given time depends on how hard it is to replace the current leader ("current competition"). Remaining the leader of the competition depends on the level of competition during the rest of the contest ("future competition"). Current and future competition dynamically affect players' incentives and may discourage effort. This motivates us to investigate, empirically, the impact of contingent prizes (e.g., time- or score-dependent prizes) on contest outcomes. Why? Contingent prizes allow the contest designer to manage how these different competition effects impact incentives throughout the contest. To this end, we estimate a structural model using observational data from large online competitions to study the performance of contingent prize structures. We complement our model-based evidence with an experiment that randomizes prize structures across competitions. Evidence from both methodologies shows that contingent prizes can significantly improve contest outcomes.

Keywords: Contests, Tournament Design, Contingent Prizes, Dynamic Games

JEL codes: C51, C57, C72, O31.

*We thank workshop participants at DePaul, UBC (Sauder), and WUSTL (Olin) for helpful comments and suggestions. Guillermo Marshall is supported in part by funding from the Social Sciences and Humanities Research Council (Canada). All errors are our own.

†Department of Economics, University of Illinois at Urbana-Champaign, jalemus@illinois.edu

‡Sauder School of Business, University of British Columbia, guillermo.marshall@sauder.ubc.ca.

1 Introduction

Firms and government agencies are increasingly procuring solutions to problems using online contests. In particular, data-science competitions have become popular, and most of them share the following features: they provide well-defined evaluation rules, a real-time leaderboard, enable players to make multiple submissions throughout the competition, and, notably, award all the prizes at the end.[1] By contrast, the XPRIZE Carbon Removal competition, the largest incentive prize in history, splits a $100 million prize over time: After one year of competition, the competition will award up to 15 prizes of $1 million each. At the end of the competition, the winner will get $50 million, and three runner-ups will share $30 million.[2]

In a dynamic contest, *when* should prizes be allocated? Our contribution is to empirically investigate the impact of *contingent* prizes on contest outcomes. Contingent prizes are prize-allocation rules based on the history of the contest, e.g., to the leader at certain point in time, or to the first player to achieve a benchmark performance. Specifically, we focus on contests of a fixed length, well-defined rules, where players publicly observe a real-time leaderboard and can make multiple submissions throughout the contest, and where there is uncertainty between effort and progress. In such contests, we investigate the performance of time- and score-contingent prizes, or a combination of both schemes, for a fixed budget.

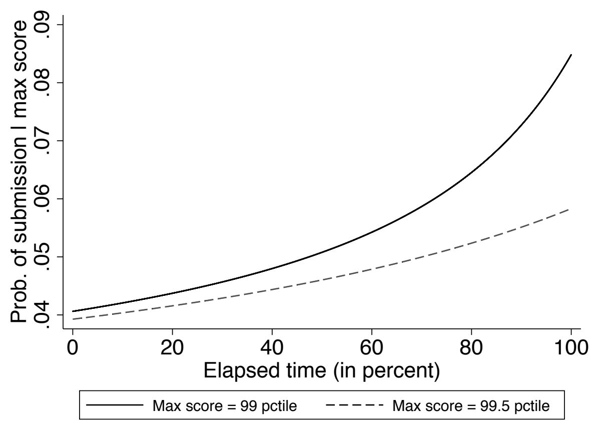
Contingent prizes matter because they can dynamically motivate players who would be discouraged to play had the prize been allocated at the end of the competition. In that case, two salient economic forces deter players from making a costly submission during the competition. First, at any point during the competition, a player's chance of rising to the top of the leaderboard could be small ("current-competition" effect), discouraging her from playing. Second, even if a player believes that she can become the competition leader at the current time, she anticipates that more submissions by rival players will come, which can displace her from the top of the leaderboard ("future-competition" effect). The current-competition effect is strongest near the end of the competition, when players have made many submissions, possibly exhausting most of the alternatives to improve scores. In contrast, the future-competition effect is strongest at the beginning of the competition, when many rival submissions are yet to come. By awarding contingent prizes, the contest designer can change the relative importance of these effects, potentially enhancing players' incentives to play.

[1] Kaggle.com is the most popular platform that permits a sponsor to host a data-science competition.

[2] https://www.xprize.org/prizes/elonmusk. This competition, sponsored by the Musk Foundation, seeks technologies for removing carbon dioxide from the atmosphere.

Figure 1 illustrates these forces for a contest that allocates the full prize pool at the end. The figure shows the (estimated) probability of playing, conditional on a fixed score, at different moments in time. As the maximum score increases (solid versus dashed line), the submission probability curve shifts down, capturing the discouragement effect of the maximum score, i.e., the current-competition effect. Fixing the maximum score, which corresponds to choosing one of the curves, the probability of making a submission increases over time (both curves slope upwards), showing that future-competition effect discourages players from making submissions earlier in the competition.

Figure 1: Probability of making a submission



Notes: The figure plots the probability of making a submission for given parameter values. Each curve plots this probability over time given a maximum score (fixed over time).

A contest designer can use contingent prizes to manipulate the current- and future-competition effects. While a prize structure that only rewards the leader at the end of the competition does little to incentivize participation early on, a prize structure that rewards competition leaders early on softens the future-competition effect. However, with a fixed budget, early rewards leave a smaller reward for later in the competition, when the current-competition effect is more pressing, and they also incentivize players to play earlier, which increases the maximum score faster, making the current-competition effect even more pressing towards the end of the competition.

The insight that allocating prizes throughout the competition can improve performance by motivating players in dynamic settings is in line with some findings in the literature (see, e.g., Harris and Vickers, 1987; Konrad and Kovenock, 2009). In multi-battle contests, for instance, the optimal inter-temporal allocation of prizes depends on how effort impacts the probability of winning a battle (see, e.g., Feng and Lu, 2018). Time-contingent prizes induce a particular multi-state contest, where the prizes partition the contest into sub-contests of predetermined length allocating prizes at the end.

Other settings focus on score-contingent prizes. For instance, in Ely et al. (2021) the principal (contest designer) observe players' progress, and he can choose prizes, feedback, and when to end the contest. The optimal design features a sequence of contests of fixed length. The competition ends when (at least) one player surpasses a benchmark score (equivalently, when a breakthrough occurs). In Benkert and Letina (2020) players privately observe their progress and choose when to report it to the principal. The optimal prize structure involves making interim transfers to all players while the competition has not ended and a prize to the first player who reveals a success to the principal, who then ends the contest. Unlike these papers, we focus on contests that resemble those used in practice, i.e., with a fixed length and where a leaderboard publicly discloses, in real-time, the players' performance. Similar to these papers, we study both time- and score-contingent prizes; with score-contingent prizes, the contest designer chooses both the size of the prize and the threshold score to award the prize.

We first present an illustrative model showing that allocating the prize only at the end of the competition might be optimal, but often it is not. The model also shows that time-contingent prizes do not dominate nor are dominated by score-contingent prizes (or by a hybrid design that combines both time- and score-contingent prizes). Thus, which prize structure leads to better outcomes is an empirical question.

We combine two empirical methodologies to investigate the impact of contingent prizes on performance. First, we estimate a structural model with observational data from Kaggle.com, the largest platform for hosting data-science competitions.[3] In these competitions, players can submit multiple solutions that are scored based on an objective criterion (e.g., prediction accuracy). A public leaderboard displays these scores in real time, and the participant with the highest score is awarded a prize at the end of the competition. Thus, these data allow us to estimate structural parameters of a dynamic-contest model where the prize is allocated at the end of the competition. Using these estimates, we simulate the equilibrium of each contest under counterfactual prize structures, where prizes are contingent on scores, times, or both, and compare outcomes across equilibria.

Second, we recruited students from the University of British Columbia and the University of Illinois and ran a randomized control experiment to assess the impact of different prize structures. This methodology provides an answer that is independent of our modelling choices. In the experiment, students competed in groups of up to five students in a prediction competition. Each group was assigned to one of three conditions: (1) The player with the highest

[3]Expedia, Google, Two Sigma, the NFL, among others, have sponsored competitions in Kaggle.

score at the end of the competition received the full prize pool (baseline). (2) The player with the highest score two days before the end of the competition received 30% of the prize pool, and the player with the highest score at the end of the competition received the remaining 70% of the prize pool, (time-contingent prizes). (3) The first player to surpass a benchmark score received 30% of the prize pool, and the player with the highest score at the end of the competition received the remaining 70% of the prize pool (score/time contingent prizes).

Our model-based and experimental evidence align in showing that contingent prizes can significantly improve contest outcomes relative to rewarding leaders only at the end of a competition. Our model estimates show that a hybrid design—one that provides one prize to the first player to surpass a benchmark score and another prize to the leader at the end of the competition—produces the best outcomes among six counterfactual prize structures because it lessens the discouragement caused by the future competition that comes after achieving the benchmark. Our experimental results show that the hybrid/benchmark prize structure causes similar gains in contest outcomes, but with larger magnitudes than what the empirical model predicts.

More specifically, using our model estimates, for each contest we consider six counterfactual designs, each one representing a different prize structure. The first three designs award $k \in \{2, 3, 6\}$ prizes, equally spaced over time, to the interim leader of the competition. That is, the player leading the competition at time $t_k = T/k$ receives a prize π_k. The fourth design ("2 timed prizes") allocates two prizes, one at the end of the competition and the other at an optimally-chosen time, with the magnitude of both prizes chosen optimally. The fifth design ("benchmark") awards the full prize pool to the first player surpassing an optimally-chosen benchmark score B, provided that this submission arrives before the end of the contest. The last design ("hybrid/benchmark") provides a prize to the first player surpassing a benchmark score and then a prize to the leader at the end of the competition, where the benchmark and the magnitude of both prizes are optimally chosen. In each case, we compute the equilibrium of the game under these alternative prize structures and compare it to the baseline equilibrium (i.e., when all the prize money is awarded at the end of the contest).

For each contest, we find the parameters for each one of these six prize structure that maximize the expected maximum score, given the contest's primitives. For instance, within the class of time-contingent prizes with 2 equally-spaced prizes (i.e., one at the middle of the competition and one at the end), we find the optimal prize split for each contest. These results are useful for a designer who has a good understanding of the contest's primitives, such as the contest's difficulty, player's submission cost, and the distribution of scores for a given submission.

While in some instances the contest designer can form reasonable estimates of these primitives, in others, the contest designer may be unable to do it. To provide an evaluation of different prize structures when a contest designer does not know the contest's primitives, for each one of the six different prize structure we find parameters that a designer can use "blindly." To this end, for each prize structure, we find parameters that maximize the average maximum score across all the contests in our data, i.e., for each prize structure the parameters are constrained to be identical for all contests. For instance, with two time-contingent prizes equally spaced over time, the optimal split is to allocate 30% of the prize at the middle of the contest and 70% at the end of the contest. However, when two prizes are optimally allocated over time, the optimal uniform prize structure is to allocate 25% of the prize when 68% of the competition time has elapsed, and the remaining 75% of the prize at the end date.

Our estimates show that the contest designer can achieve a large fraction of the gains of each contest's optimal prize structure by simply using a uniform prize structure. That is, even when contests are heterogeneous in their primitives, a simple uniform policy can lead to significant gains in contest outcomes. We also find that an optimal prize structure can achieve the same performance of a contest that allocates prizes at the end using roughly 52 percent of the budget. Thus, our results suggest that firms sponsoring online contests, which have paid million of dollars in prizes, may be leaving money on the table by using suboptimal prize structures.

1.1 Related Literature

A long-standing economic question is how to optimally design contests. Some of questions that have been investigated in the literature include: How many participants should enter the contest? (Taylor, 1995; Fu and Lu, 2010; Aycinena and Rentschler, 2019); How many participants should earn a prize? (Ehrenberg and Bognanno, 1990; Moldovanu and Sela, 2001; Olszewski and Siegel, 2020; Kireyev, 2020); Should contests be divided into multiple stages? (Moldovanu and Sela, 2006; Sheremeta, 2011); Should participants receive performance feedback during the contest? (Gross, 2017; Mihm and Schlapp, 2019; Lemus and Marshall, 2021). The answer to these questions depend on the nature of the contest (e.g., static or dynamic) and on the characteristics of the players (e.g., homogenous or heterogeneous).

One insight from multi-battle contests is that intermediate prizes can counteract the discouragement effect. For instance, Konrad and Kovenock (2009) examines this insight in multi-battle, all-pay contests with complete information. Iqbal and Krumer (2019) empirically explore it using team matches in tennis. In Feng and Lu (2018), prizes are contingent on the number of victories in a three-battle, Tullock contest. They find that the importance of players' efforts relative to noise in the winning probabilities determines whether it is optimal to allocate intermediate prizes, which moderate the discouragement effect, which is stronger when the discriminatory power is high. Alshech and Sela (2021) characterize optimal split of the budget in a two-stage Tullock contest where the designer chooses a "bonus" prize for the player who wins both stages, in addition to prizes for each battle. They show that the prize for winning in both stages should be allocated in the two-stage Tullock contest with two players, but not when there are more than two players. This result is related to the effect of the future competition effect in the allocation of prizes.

Other design choices for multi-battle contests include the number of stages and prizes. Clark and Nilssen (2020), for example, use this design choices when agents are ex-ante heterogeneous and find a design that achieves full rent dissipation. Fu and Lu (2009) explore a similar question with homogeneous agents and show that a grand contest induces more effort than any split version of the contest. This result resembles a time-contingent design that allocates all the prize at the end. Chowdhury and Kim (2017) show that this result reverses if the winner-selection mechanism is simultaneous rather than sequential. In a two-stage contest, Klein and Schmutzler (2021) show that it is optimal to allocate all the prize at the end rather than giving an intermediate prize. In a laboratory setting, they compare effort in a two-stage contest with a single prize versus a sequence of two independent contest (i.e., two one-stage contests). They find that a single prize increases effort by 41%.

Fu and Lu (2012) study the optimal design of multi-stage, sequential-elimination contests. Also Stracke et al. (2014), who compare one versus two prizes and show that two prizes may dominate one when players are risk-averse. They present laboratory evidence supporting this prediction. Cason et al. (2020) study the optimal allocation of prizes when performance is a noise measure of effort. Experimentally, they find that a noisy performance measure increases effort, which is contrary to the theory. Güth et al. (2016) study "output-dependent prizes," which relate to our setting with a benchmark prize. They show, both theoretically and experimentally, that output-dependent prizes outperform fixed-prize tournaments and piece rates. Liu et al. (2018) show that negative prizes for players with effort below a threshold can increase the expected total effort.

2 Dynamic Prize Allocation: A Simple Model

This section serves the purpose of illustrating tradeoffs for different prize allocations and showing contingent prizes may improve competition outcomes relative to prizes allocated at the terminal date of the contest. However, none of the schemes is dominant for all parameters.

Two players, A and B, compete in two stages, $t \in \{1, 2\}$. The competition starts with player A as the leader with a score of $\bar{s} > 0$ and player B as the follower with a score of 0. At $t = 1$, only player B can play. If she chooses to play, depending on the outcome of the play, she can become the leader or continue as the follower. At $t = 2$, only one of the players can play: player A can play with probability α and B with probability $1 - \alpha$.

When the score of the current leader is s, the player who plays increases the score by ε with probability $q(s)$. Thus, the score from playing is a random variable, s', with

$$s' = \begin{cases} s + \varepsilon & \text{with probability } q(s), \\ 0 & \text{with probability } 1 - q(s). \end{cases}$$

The function $q(\cdot)$ is *decreasing*: it is harder to replace the leader when her score is higher. Whenever a player can play, she draws a playing cost, c, where $c \sim K(\cdot)$ is i.i.d among players and stages. After observing this cost, the player decides whether to play.

Player B's incentive to play at time $t = 1$ is influenced by the current competition effect, captured by the parameter $\bar{s}$, and the future competition effect, captured by α. Different time-contingent prizes impact the incentives to play through their impact on the current and future competition effects and, therefore, the expected maximum score.

The contest designer chooses the prize allocation that maximizes the expected maximum score. We can find it by solving the game by backward induction.

2.1 Time-contingent prizes

The player who leads the competition at the end of stage t receives the monetary prize π_t. The contest designer has a fixed budget normalized to 1, and decides the size of the prize for each stage, i.e., $\pi_1 + \pi_2 = 1$. The identity and score of the leader of the competition at each stage is public information.

Lemma 1. *At stage 2, if the leader can play, she chooses not play and wins the prize π_2 for* *sure. If the follower can play, when her playing cost is c and the leader's score is s, she plays if and only if $q(s)\pi_2 > c$.*

After stage 1, the score is either $\bar{s}$ or $\bar{s} + \Delta$. Let $q_H = q(\bar{s})$ and $q_L = q(\bar{s} + \varepsilon)$. The leader is replaced with higher probability when the score is lower: $q_H > q_L$. Let $Q_H = K(q_H\pi_2)$ and $Q_L = K(q_L\pi_2)$. The probability that a follower who can play at stage 2 actually plays is higher when the score is lower: $Q_H > Q_L$.

Lemma 2. *At stage 1, Player B (the follower) plays if and only if*

$$q_H(\underbrace{\pi_1 + \pi_2(1 - \alpha q_L Q_L)}_{Leader} - \underbrace{(1-\alpha)Q_H(q_H\pi_2 - E[c|c \le q_H\pi_2])}_{Follower}) \ge c, \tag{1}$$

Player B's benefit of playing at stage 1 is the probability the play is effective (q_H) times the payoff difference of becoming the leader and continuing as the follower. Becoming the leader gives an immediate payoff of π_1 and a future payoff of π_2 whenever player A does not replace her in stage 2, which happens with probability $1 - \alpha q_L Q_L$. Continuing as the follower gives an expected payoff of $q_H\pi_2 - E[c|c \le q_H\pi_2]$ whenever player B can and does play at stage 2, which happens with probability $(1-\alpha)Q_H$.

Fiercer future competition (larger α) has an ambiguous effect on the incentive to play at $t = 1$ because it reduces both the payoff the becoming the leader (who will be replaced more often) and the payoff of continuing as the follower (who is less likely to play again in the future). By increasing the prize in stage 1, π_1, the contest designer can directly counteract the discouragement effect caused by the future competition. However, with a fixed budget, it also directly lowers π_2, reducing the follower's incentive to play at stage 2.

Lemma 3. *More future competition (larger α) decreases the incentive to play at $t = 1$ whenever $\pi_2 q_L Q_L < Q_H(q_H\pi_2 - E[c|c \le q_H\pi_2])$. A larger prize in stage 2 decreases Player B's incentive to play at stage 1 but increases the follower's incentive to play at stage 2.*

The contest designer trades off enhancing the incentives to play at stages 1 and 2. The expected maximum score is a non-linear function of π_2 that depends on the shape of the cost distribution, K, the intensity of future competition, α, and the difficulty of replacing the current leader, $q(\cdot)$. Proposition 1, in the Appendix, characterizes the optimal prize structure.

2.2 Simple Benchmark

Consider a prize structure in which the first player to surpass a benchmark score B wins a prize of 1 and ends the contest. In this simple framework, the contest designer can choose two possible benchmark scores : $B = \bar{s} + \varepsilon$ and $B = \bar{s} + 2\varepsilon$.

Lemma 4. *With a benchmark score of $B = \bar{s} + \varepsilon$,*

(i) Conditional on reaching stage 2 there is a play with probability $Q_H = K(q_H)$.

(ii) Player B plays at stage 1 if and only if

$$q_H(1 - (1 - \alpha)Q_H(q_H - E[c|c < q_H])) \geq c \tag{2}$$

To play at stage 1, the expected marginal value of playing must be larger than the cost. If player B plays and ends the contest at stage 1, she receives the full prize of 1. If the contest does not end, player B can play at stage 2 with probability $(1 - \alpha)$.

The comparison of the impact of time- and score-contingent prizes on the incentive to play at stage 1 is ambiguous. Comparing (1) and (2), note that both the value of being a leader and continue as a follower are larger in the benchmark contests. Fixing the time-contingent prizes, however, the benchmark contest incentivizes *early* plays when there is more future competition (larger α).

Lemma 5. *With a benchmark score of $B = \bar{s} + 2\varepsilon$,*

(i) Conditional on reaching stage 2 there is a play with probability $Q_L = K(q_L)$.

(ii) Player B plays at stage 1 if and only if

$$q_H(1 - \alpha)Q_L(q_L - E[c|c < q_L]) \geq c \tag{3}$$

When the benchmark score is $B = \bar{s} + 2\varepsilon$, the prize is allocated only when there are two successes in a row. At stage 2, either there is no play (when player B failed to increase the score at stage 1), or there is a play with probability $K(q_L)$. Compared with time-contingent prizes, the benchmark encourages later plays when scores are high, but discourages later plays when scores are low (when the benchmark is unattainable). In other words, the current competition effect is more prominent at later stages with time-contingent prizes.

Comparing (3) and (2), note that as α increases, player B's incentive to play at stage 1 decreases with a high benchmark but it increases with a low benchmark. Intuitively, when players expect high competition in the future, whoever has the opportunity to play has incentive to do so as soon as possible if the benchmark is low. However, high future competition hinders the incentive to play in early stages because a player might not benefit from a successful play at the current stage.

Proposition 2, in the Appendix, finds the optimal benchmark for any given set of parameters by comparing the expected maximum score in the two scenarios.

2.3 Hybrid: Benchmark and Prize for Leader at the End

We also consider a prize structure in which the first player to surpass a benchmark score B wins a prize π_B and the player with the highest score at the end of the competition wins $1-\pi_B$. In this simple framework, the designer chooses π_B and one of two possible benchmark scores: $B = \bar{s} + \varepsilon$ and $B = \bar{s} + 2\varepsilon$. The analysis of this case is analogous to the previews cases, but also more involved, so we relegate the analysis to Appendix A.

2.4 Comparing Contingent-Prize Structures

Figure 2 shows a numerical simulation comparing the expected maximum score under optimal time-contingent prizes, a simple benchmark design, and a hybrid design. The left panel plots the expected maximum score (up to a constant) as a function of α. The right panel shows the optimal policies for each design as a function of α: prize at stage 2, π_2, for stage contingent prizes; benchmark that, if surpassed, ends the contest; benchmark score and reward for reaching the benchmark, π_B, for hybrid contests.

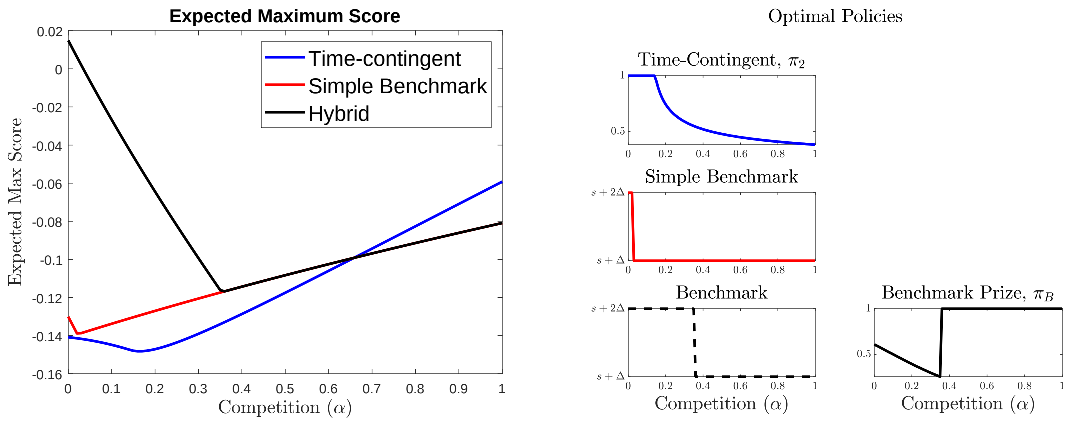


Figure 2: Optimal design for different prize systems. The parameters used in the numerical simulation are: $q_H = q_L = 0.8$, $K(c) = \sqrt{c}$.

A hybrid design, of course, weakly dominates the simple benchmark design (the simple benchmark is a particular case of the hybrid design when $\pi_B = 1$). Notably, time-contingent prizes are not dominated nor they dominate a hybrid design. Specifically, the top-right panel of Figure 2 shows that while allocating the prize at the end is optimal within the class of time-contingent prizes when α is small, a hybrid design dominates. On the other hand, with intense future competition (large α), time-contingent prizes dominate a hybrid design.

More generally, which design dominates depends on the level of future competition, how hard it is to replace the leader, and on the distribution of costs. Therefore, the optimal allocation of contingent prizes is an empirical question, which we answer in the following sections.

3 Data and Background Information

We use publicly available data on 57 featured competitions hosted by Kaggle.[4] These competitions received thousands of submissions, coming from an average of 894 players per contest, and offered average prize of $30,489. A partial list of competition characteristics are summarized in Table 1 (see Table A.1 in the Online Appendix for the full list).[5]

In the competitions, participants have access to a training and a test dataset. The test dataset includes both an outcome variable and covariates, while the test dataset only includes covariates. The goal of the contest is to generate the most accurate predictions of the outcome

[4]https://www.kaggle.com/kaggle/meta-kaggle

[5]Lemus and Marshall (2021) provide a detailed overview of the dataset as well as descriptive evidence.

variables for the covariates in the test dataset. A submission in a contest must include an outcome variable prediction for each observation in the test dataset. Kaggle scores each submission, according to an objective evaluation rule, in a public leaderboard and allocates prizes at the end of the competition to players with the highest scores.[6]

We have information at the contest-level on all submissions in a contest, including the time of the submission, who made them (team identity), and their score (public and private scores). We are able to reconstruct both the public and private leaderboard at every instant of time for every contest. Using the same approach as in Lemus and Marshall (2021), we standardize the distribution to have a mean of zero a standard deviation of one.

4 Empirical Model

N players compete in a contest of length T. We divide the length of the contest into time intervals of length δ, so time is discrete $t = 0, \delta, 2\delta, ..., T$. Payoffs are not discounted. Players enter at an exogenously given time, so that at time t, N_t players have already entered. They have perfect foresight and stay until the end of the contest. Making a submission when the current maximum score is s increases the score to $s' = s + \varepsilon$ with probability q_s. We assume that q_s decreases in s, i.e., it becomes increasingly difficult to replace the leader as the maximum score increases.

Players publicly observe and keep track of three state variables: the time period (t), the current maximum score (s), and the identity of the current leader (ℓ). The state space is

$$\mathcal{S} = \{(t, s, \ell) \,:\, t = 0, \delta, ..., T,\ s = 0, \varepsilon, 2\varepsilon, ... \text{ and } \ell = 1, ..., N\}$$

At period $t = 0, ..., T - 1$, at most one player can play: with probability $\lambda \in (0, 1)$ one (randomly selected) player can play, and with probability $1 - \lambda$ no one can. We denote $t' = t + \delta$ and $s' = s + \varepsilon$.

A play corresponds to an instantaneous submission at cost c, which is a random variable

[6] Kaggle partitions the test dataset into two subsets and does not inform participants which observations correspond to each subset. The first subset is used to generate the *public score*; the second subset is used to generate the *private score*. The public score is posted in real-time on a public leaderboard, whereas the private score is never made public during the contest. The private score is used to determine the winner of the competition and to allocate the prize at the end of the competition, making the public score a noisy signal of the private score (i.e., performance). The coefficient of correlation between public and private scores in our sample is 0.99, which motivates us to abstract away from the noise in the public signals in our empirical model for tractability.

distributed according to $K(\cdot)$ and i.i.d. across players and time. Players observe their cost realization before choosing whether to play.

4.1 Time-contingent Prizes

Consider a contest that awards a prize π_t to the competition's leader at time t (i.e., the player with the highest score at that time). At time t, a player is either the leader or one of the $N_t - 1$ followers. Let $L_{t,s}$ be the value of being the leader at time t when the maximum score is s, and let $F_{t,s}$ be the value of being one of the followers. When the competition ends, the leader gets π_T and the followers 0, so the terminal values are $L_{T,s} = \pi_T$ and $F_{T,s} = 0$.

The leader's expected value at $t = 0, 1, ..., T-1$ is

$$L_{t,s} = \pi_t + \left(1 - \lambda \frac{N_t}{N}\right) L_{t',s} + \frac{\lambda}{N} L_{t,s}^{\text{own play}} + \frac{\lambda(N_t - 1)}{N} L_{t,s}^{\text{rival play}} \tag{4}$$

In (4), the leader at time t receives the prize π_t, and then there are three cases. First, with probability $1 - \lambda \frac{N_t}{N}$ none of the players who have entered the contest can play, so the current leader remains the leader and receives the continuation payoff $L_{t',s}$. Second, with probability $\frac{\lambda}{N}$ the current leader can play, in which case she receives the continuation payoff $L_{t,s}^{\text{own play}}$, defined in (5). Third, with probability $\frac{\lambda(N_t-1)}{N}$ one of the $N_t - 1$ followers can play, in which case the current leader receives the continuation payoff $L_{t,s}^{\text{rival play}}$, defined in (8).[7]

The value $L_{t,s}^{\text{own play}}$ is given by

$$L_{t,s}^{\text{own play}} = E_c \left[\max\{q_s L_{t',s'} + (1 - q_s) L_{t',s} - c, L_{t',s}\} \right], \tag{5}$$

The leader at state (t, s) chooses whether to play after observing the cost realization, c, playing if only if the expected marginal value of increasing the score is larger than the cost of making the submission, i.e.,

$$q_s(L_{t',s'} - L_{t',s}) \geq c. \tag{6}$$

From this condition, the probability that the leader plays is

$$p_{t,s}^L = K(q_s(L_{t',s'} - L_{t',s})). \tag{7}$$

[7]Note that the probability that any given player gets to play does not depend on the number of players that have actually entered (N_t), as any one player gets the opportunity to play with probability $1/N$, where N is the time-independent number of players that will participate in the competition.

The last term in (4), $L_{t,s}^{\text{rival play}}$, is the leader's continuation payoff when one of the current $N_t - 1$ followers can play, which happens with probability $\frac{\lambda(N_t-1)}{N}$. We have

$$L_{t,s}^{\text{rival play}} = p_{t,s}^F \left(q_s F_{t',s'} + (1-q_s) L_{t',s}\right) + (1-p_{t,s}^F) L_{t',s}. \tag{8}$$

A follower who can play makes a submission and replaces the current leader with probability $p_{t,s}^F q_s$, in which case the maximum score increases to s', and the current leader becomes a follower, obtaining $F_{t',s'}$. With probability $p_{t,s}^F(1-q_s)$, the follower makes a submission but fails to replace the leader, so the current leader remains the leader, obtaining $L_{t',s}$. With probability $1-p_{t,s}^F$ the follower chooses not to play, so the current leader remains the leader, obtaining $L_{t',s}$.

We next specify the value of of being a follower, $F_{t,s}$. All the followers are symmetric, so the value for any one particular follower is identical. Let j denote one of the $N_t - 1$ followers. The value of being a follower at state (t,s) is

$$F_{t,s} = \left(1 - \lambda \frac{N_t}{N}\right) F_{t',s} + \frac{\lambda}{N} F_{t,s}^{\text{leader play}} + \frac{\lambda}{N} F_{t,s}^{j \text{ play}} + \frac{\lambda(N_t-2)}{N} F_{t,s}^{\text{follower play}} \tag{9}$$

Followers do not receive prizes. At time t, there are four cases. First, with probability $1-\lambda\frac{N_t}{N}$ none of the players who have entered the contest can play, so followers remain as followers and receive the continuation payoff $F_{t',s}$. Second, with probability $\frac{\lambda}{N}$ the leader can play, in which case followers receive the continuation payoff $F_{t,s}^{\text{leader play}}$. Third, follower j can play with probability $\frac{\lambda}{N}$, in which case she receives the continuation payoff $F_{t,s}^{j \text{ play}}$. Fourth, with probability $\frac{\lambda(N_t-2)}{N}$ one of the other $N_t - 2$ followers (not j) can play, in which case follower j receives the continuation payoff $F_{t,s}^{\text{follower play}}$.

The value $F_{t,s}^{\text{leader play}}$ is given by

$$F_{t,s}^{\text{leader play}} = p_{t,s}^L \left(q_s F_{t',s'} + (1-q_s) F_{t',s}\right) + (1-p_{t,s}^L) F_{t',s} \tag{10}$$

When the leader can play, she makes a submission and increases the maximum score with probability $p_{t,s}^L q_s$, in which case the followers receive $F_{t',s'}$. With probability $p_{t,s}^L(1-q_s)$, the leader makes a submission but fails to increase the maximum score, so followers receive $F_{t',s}$. With probability $1-p_{t,s}^L$ the leader chooses not to play, so the followers receive $F_{t',s}$.

The value $F_{t,s}^{j \text{ play}}$ is given by

$$F_{t,s}^{j \text{ play}} = E_c \left[\max\{q_s L_{t',s'} + (1-q_s) F_{t',s} - c, F_{t',s}\}\right]. \tag{11}$$

When follower j can play at state (t, s) chooses between making a submission or not after observing cost of making a submission, c. Conditional on c, follower j makes a submission if only if

$$q_s(L_{t',s'} - F_{t',s}) \geq c, \tag{12}$$

so the value of becoming the leader must be sufficiently larger than the value of remaining a follower. From here, the probability that a follower plays is

$$p^F_{t,s} = K\left(q_s(L_{t',s'} - F_{t',s})\right). \tag{13}$$

The value $F^{\text{follower play}}_{t,s}$ is given by

$$F^{\text{follower play}}_{t,s} = p^F_{t,s}\left(q_s F_{t',s'} + (1 - q_s)F_{t',s}\right) + (1 - p^F_{t,s})F_{t',s} \tag{14}$$

When a follower other than player j can play at state (t, s), follower j always remains a follower but the maximum score can change. The follower other than j plays with probability $p^F_{t,s}$ and increases the maximum score with probability q_s. In that case, follower j gets $F_{t',s'}$. In any other case, only time progresses and follower j receives $F_{t',s}$.

4.2 Score-contingent Prize: Simple Benchmark

We now consider score-contingent prizes, where the first player who achieves a score larger than B receives the prize π. With this design, all players are "followers" until one of them gets a score $s' > B$, in which case she wins the contest and the game ends. Thus, the value of player j at state $s < B$ is

$$F_{t,s} = \left(1 - \lambda\frac{N_t}{N}\right)F_{t',s} + \frac{\lambda}{N}F^{j\text{ play}}_{t,s} + \frac{\lambda(N_t - 1)}{N}F^{\text{follower play}}_{t',s} \tag{15}$$

Players do not receive a prize when $s < B$. At time t, there are three cases. First, with probability $1 - \lambda\frac{N_t}{N}$ none of the players who have entered the contest can play, so each player receives the continuation payoff $F_{t',s}$. Second, with probability $\frac{\lambda}{N}$ player j can play and receives the continuation payoff $F^{j\text{ play}}_{t,s}$, defined in (16). Third, with probability $\frac{\lambda(N_t-1)}{N}$ one of the other $N_t - 1$ rivals of player j can play, in which case follower j receives the continuation payoff $F^{\text{follower play}}_{t,s}$, defined in (18). We have

$$F^{j\text{ play}}_{t,s} = E_c\left[\max\{q_s(1\{s' = B\}\pi + 1\{s' < B\}F_{t',s+\varepsilon}) + (1 - q_s)F_{t',s} - c, F_{t',s}\}\right] \tag{16}$$

Player j at state (t, s) chooses between making a submission or not after observing cost of making a submission, c. Conditional on c, the she makes a submission if only if

$$q_s(1\{s' = B\}\pi + 1\{s' < B\}(F_{t',s'} - F_{t',s})) \geq c. \tag{17}$$

If the s is low ($s \leq B - 2\varepsilon$), there is no chance of collecting the prize π by playing. Thus, the incentive to play comes only from the difference in value from increasing the score. If player j increases the score, it is more likely for everyone else to collect the prize in the future. This increases the incentive of all players to play, so when player j has another opportunity to play, it is more likely that she is at state $(t, B - \varepsilon)$.

The value $F_{t,s}^{\text{follower play}}$ is given by

$$F_{t,s}^{\text{follower play}} = p_{t,s}\left(q_s F_{t',s'} + (1 - q_s)F_{t',s}\right) + (1 - p_{t,s})F_{t',s} \tag{18}$$

When a follower other than player j can play at state (t, s), follower j remains a follower but the maximum score can change. The follower other than j plays with probability $p_{t,s}$ and increases the maximum score with probability q_s. In that case, follower j gets $F_{t',s'}$. In any other case, only time progresses and follower j receives $F_{t',s}$.

Directly from Equation 17 we get

$$p_{t,s} = K(q_s(1\{s' = B\}\pi + 1\{s' < B\}F_{t',s'}) - F_{t',s})) \tag{19}$$

Lastly, the terminal values are $F_{t,s} = 0$ for any $s \geq B$ and $F_{T,s} = 0$ for any $s < B$.

4.3 Hybrid Contest: Benchmark Bonus plus Prize at the End

We also study a design a hybrid contest design, where a prize π_B is allocated to the first player that surpasses the benchmark score B and a prize $\pi - \pi_B$ to the leader at the end of the contest. For the sake of exposition, we present the equations of the value functions for this design in Appendix B.

4.4 Discussion of Modelling Assumptions

Our model hinges on several assumptions to facilitate estimation. Some of these assumptions simplify the behavior of players while others reduce the choice set of the contest designer.

Behavior of Players and Institutional Details.

1. *Learning and experimentation.* One motivation to participate in a Kaggle competition is the opportunity for players to learn and experiment. Even experienced players can benefit from learning from their performance on earlier submissions. Our model accommodates one form of learning: the probability of replacing the leader depends on current scores. That is, the function $q(s)$ can capture that all players get new ideas when they see the score increase. This learning effect collides with the effect that higher scores are harder to surpass (i.e., current competition effect). Which of these effects dominates is an empirical question.

   A different way of modelling learning would be to allow for players' performance to depend on the number of their past submissions. This approach increases the state space by m^N, where m is the maximum number of submissions and N is the number of players (e.g., with 10 submissions per player and 10 players, our state-space increases by 10 billion). We refrain from such approach for tractability.

2. *Time to build.* Data-science competitions are unique in that players spend time building a baseline prediction model. Once this model is ready, a player can tweak the model to make multiple submissions. We do not observe what players are working on. Therefore, we assume that each player builds one model, and their entry time is the time when their model is ready to be tweaked for the rest of the competition.

3. *Incentives to withhold solutions.* A strategic concern is that a good score may encourage rivals to exert effort, creating more competition. To prevent encouraging rivals, players could decide to withhold their solutions and send them closer to the end of the competition. First, players benefit from submitting their solutions as soon as possible to receive feedback, which allows them to improve their current solutions. Second, there is a limit on the number of submissions players can send each day. Third, Lemus and Marshall (2021) use the sample of contests as in this article and do not find empirical evidence suggesting strategic withholding.

4. *Leader, followers, and one final prize.* At each instant during the contest, Kaggle's leaderboard shows the best score for each player. Furthermore, multiple players receive prizes at the end of the contest (usually, three prizes). We assume one leader and one prize to reduce the dimensionality of the state space. Keeping track of each player's scores at each point in time increases our state space by $|S|^{N-1}$, where $|S|$ is the number of possible scores. Furthermore, the players' decisions are more involved as they assess the probability of ending the contest in different positions.

5. *Teamwork.* Kaggle permits players to form self-organized teams. Lemus and Marshall (2022) study the impact of team formation on performance. Different contest designs affect the formation of teams, which can impact the contest outcomes. Allowing players to form teams would allow players to improve their performance, which is a form of asymmetry that we abstract away from for tractability.

6. *Player heterogeneity.* Our model assumes players are symmetric. In our estimation, we focus on top performers, who are arguably more symmetric than two randomly-selected players in the contest, which alleviates heterogeneity concerns.

Contest design choices.

1. *Size of the prize pool.* We normalize the size of the prize to 1. The literature has documented that larger prizes elicit higher effort.

2. Open or restricted entry. Kaggle contests are typically open contests (anyone can participate). We do not model entry. Lemus and Marshall (2021) discuss at length the shortcomings of assuming exogenous entry. We further restrict the analysis to the top performers in each contest both to alleviate concerns about player heterogeneity and focus on the players who are likely to influence the contest outcomes.

3. *Evaluation measure.* Kaggle's evaluation metric is objective (e.g., root-mean-square deviation) but the public leaderboard displays a noisy signal of players' performance. The reason is that Kaggle wants players to develop robust models that work well for out-of-sample prediction, i.e., Kaggle wants to prevent solutions that overfit the data. Allowing for noisy leaderboards requires us to keep track of each player's scores and positions, which we cannot do for computational reasons discussed above.

4. *Information disclosure.* We assume that players receive public real-time feedback following Lemus and Marshall (2021). They show that displaying a real-time leaderboard improves contest outcomes on average.

5. *Length and number of stages.* Most Kaggle contests are single-stage contests with a fixed length. We normalize the length of the contest to make meaningful comparison across contests. Dividing the contest into multiple stages can reduce discouragement. In our setting, there is no notion of being "close" to the leader because each player that is not a leader is a follower. Furthermore, if the contest was divided in stages, players could always resubmit their last solution to restore the leaderboard at the beginning of

each stage. Thus, in our setting, it is unfeasible to reduce discouragement by splitting the contest in multiple stages.

4.5 Estimation

Given that only the state is payoff relevant, we use the Markov-perfect equilibrium concept. Computationally, we find it using backwards induction.

The full set of primitives for a given contest include i) the probability that a player can play at time t, λ; ii) the entry times of each player; iii) the function q_s, which indicates the probability of advancing the maximum score given that the current maximum score is s; and iv) the distribution of submission costs, $K(c;\sigma) = c^\sigma$, where $\sigma > 0$ and the support of the distribution is the interval $[0,1]$.[8] We allow these primitives to vary at the contest level.

We use a two-step procedure to estimate the primitives of each contest. In the first step, we recover or estimate primitives i)-iii) without using the full structure of the model. In the second step, we use the estimates of these primitives to estimate the cost distribution using a generalized method of moments (GMM) estimator.

We make use of a feature of the platform Kaggle to estimate the probability that a player can play at a given time period, λ. Specifically, players face a cap on the number of daily submissions, which in conjunction with the length of the contest, gives us a player's maximum number of submissions during the competition (i.e., daily cap $\cdot$ competition length (days)). We then set λ to be $N \cdot$ daily cap $\times$ competition length (days)$/T$, where N is the number of players and T is the number of time periods in the model.

The entry times of each player are assumed exogenous in the model and we recover them directly from the data. Next, we specify the function q_s as

$$q_s = \exp\{\beta_0 + \beta_1 s\}/(1 + \exp\{\beta_0 + \beta_1 s\}),$$

and we estimate β_0 and β_1 using a maximum-likelihood estimator, using data on whether each submission increased the maximum score as well as the maximum score at the time of each submission (s). Because in some competitions the maximum score is rather constant, we pool the data from all competitions to gain power in estimating the parameter β_1, which we constrain to be uniform across contests. We allow β_0 to vary across contests.

[8] Recall that we normalize the size of the prize to be 1 for every contest.

In the second step, we estimate the parameter σ of the cost distribution, $K(c;\sigma) = c^\sigma$, where $\sigma > 0$, using a GMM estimator, where the moments are based on comparisons of the number of observed submissions and the number of submissions predicted by the model. Specifically, we divide the length of each contest in five periods of equal length (henceforth, time quintiles) and we compute the number of submissions observed in the data and predicted by the model in each time quintile k: $m_k(\sigma) = \text{submissions}_k^{data} - \text{submissions}_k^{model}$. We also include a sixth moment that compares the overall number of submissions in the data and predicted by the model. The GMM estimator is then given by

$$\hat{\sigma} = \arg\min_{\sigma} \hat{\mathbf{m}}(\sigma)'\mathbf{W}\hat{\mathbf{m}}(\sigma),$$

where $\mathbf{W}$ is a weighting matrix. We present bootstrapped standard errors.

We use the full-solution method to compute the moments for a given value of σ. That is, for a given σ, we compute the equilibrium of the game using backward induction to obtain the matrices of conditional-choice probabilities (CCPs) $\mathbf{p}^L$ (leader) and $\mathbf{p}^F$ (followers) of dimensions $S \times T$ (S is the size of the set of possible scores and T is the number of periods) where element (s,t) of $\mathbf{p}^j$ is $p_{t,s}^j$.[9] Using the CCPs, we can also compute the equilibrium distribution of maximum scores at every period of time, $\mathbf{G}$ (of dimensions $S \times T$), where column t gives the distribution of maximum scores at time t.[10] Element-wise multiplication of $\lambda(\mathbf{p}^L + (N-1)\mathbf{p}^F)/N$ (i.e., the probability of play when a player is chosen at random) and $\mathbf{G}$, followed by summation of the product over the first dimension, gives us a $1 \times T$ vector with the expected probabilities of a submission at every moment of time, which we use to compute the moments.

Lastly, we restrict the sample to the top 10 players in each contest (measured by the ranking of players at the end of the competition), i.e., $N = 10$. We make this choice for two reasons: i) this is the set of players achieving scores that trigger changes in the top positions of the leaderboard, and ii) this group of players is less heterogeneous than the entire pool of players,

[9] In the estimation, we set $T = 10,000$ and S varies across contests. In a given contest, the set of scores is set to include all unique maximum scores in the competition as well as the values $-2, -1.5, -1, -0.5$, and $\bar{s} + [0.001 : 0.001 : 0.045]$, where $\bar{s}$ is the highest observed score in the competition.

[10] To compute the equilibrium distribution of scores at every moment of time, we start from period 1 where the maximum score is the lowest value in the set of scores, s_1 (i.e., nobody has played yet). At $t = 2$, the maximum score will be s_1 with probability $1 - \lambda + \lambda(1 - p_{1,s_0}) + \lambda p_{1,s_0}(1 - q_{s_0})$ (i.e., in the events when nobody is summoned to play in $t = 1$, when somebody is summoned to play in $t = 1$ but the player chooses not to make a submission, and when some player makes a submission but fails to advance the score) and s_2 with probability $\lambda p_{1,s_0} q_{s_0}$ (i.e., some player makes a submission and succeeds in advancing the score). The same logic can be used to compute the distribution of maximum scores for every subsequent period using a recursion.

which allows us to abstract away from modelling player heterogeneity.

4.6 Estimation Results and Model Fit

Table 2 reports the GMM estimates for a partial list of contests (see Table B.2 in the Online Appendix for the full list).

Regarding the goodness of fit of the model, Figure 3.A and Figure 3.B plot the actual versus the predicted maximum score and number of submissions for every contest. The figures show that the model estimates are able to replicate the data in both cases. Figure 3.C plots the actual and predicted number of submissions over time, averaged across contests, where time is divided in 10 periods of equal length. The figure shows that the model estimates are able to replicate the submission dynamics in the data without systematically under- or over-predicting the observed values.

4.7 Prize Structure and Contest Outcomes

How does the prize structure impact contest outcomes? Using our model estimates, we use counterfactual simulations to provide an answer. We consider six counterfactual designs. The first three designs consider k prizes, awarded at the end of time periods of equal length, T/k, with the size of each of the k prizes chosen optimally and $k \in \{2, 4, 6\}$. The fourth design ("2 timed prizes") allocates two prizes, one at the end of the competition and the other at an optimally-chosen time, with the magnitude of both prizes chosen optimally. The fifth design ("benchmark") awards the full prize pool to the first player surpassing an optimally-chosen benchmark score B, provided that this submission arrives before the end of the contest. The last design ("hybrid/benchmark") provides a prize to the first player surpassing a benchmark score and then a prize to the leader at the end of the competition, where the benchmark and the magnitude of both prizes are optimally chosen.

Table 3 presents a comparison of equilibrium outcomes. For each contest, we compute the optimal prize structure within each counterfactual design. The columns labeled "Optimal" compare the equilibrium number of submissions and maximum score (in expectation) under the optimal prize structure relative to the equilibrium outcomes under the baseline design (i.e., when all the prize money is awarded at the end of the contest). The table shows that awarding intermediate prizes can increase submissions by up to an average of 36.5 percent and the maximum score by 0.047 standard deviations (the case with the hybrid prize structure).

The prize structure with 6 intermediate prizes achieves the second highest gains, with an average increase in the number of submissions and maximum score of 30.5 percent and 0.039 standard deviations, respectively. The hybrid prize structure achieves the greatest gains in 82 percent of the competitions, while the 6-prize structure is optimal in the remaining 18 percent.[11] These results suggest that flexible prize structures can increase incentives to make submissions by economically significant magnitudes.

To further illustrate the impacts of contingent prizes, Figure 4 shows how the average number of submissions changes over time when implementing a 4-prize prize structure (measured relative to the baseline design). As the figure shows, the 4-prize structure boosts incentives early in the competition, especially around the times the interim prizes are given, and decreases incentives near the end. Although the decrease near the end is small, incentives to participate are greatest near the end (see Figure 3.C). Nevertheless, the total number of submissions increases on average by 8.9 percent.

Figure 5 presents details about the parameters governing the optimal prize structures within each prize class. Panels B and C show that the 4- and 6-prize optimal designs are similar across contests in that the prizes increase over time and the ranges of each of the prizes are somewhat narrow. Panel D shows that in the design with 2 timed prizes, most of the mass of the distribution of the optimal time of the first prize is in between times 0.6 and 0.9. Where we do see more heterogeneity are in the optimal benchmark scores in the benchmark and hybrid designs (Panels E and F), which reflects underlying differences in the score distributions and the probability of increasing the maximum score across contests.

We note that to compute the optimal prize structure within each counterfactual design for a contest, the contest designer needs to know all the primitives of the model. The designer may not have that information before the contest, which motivates us to ask: are the gains in contest outcomes similar if the prize structure is constrained to be uniform across all contests? We answer this question in the columns of Table 3 that are labeled "Uniform", where we compute the gains for each contest using a prize structure that was chosen by optimizing the average maximum score across contests subject to the prize structure being identical for all contests.[12] The table reveals that when a uniform prize structure is imposed on all contests, the gains of intermediate prizes are generally not too different from when prize structures

[11] Table B.3 in the Online Appendix presents estimates of a probit model for the probability of the hybrid prize structure being optimal for a contest as a function of contest primitives. The table shows an increase in the frequency of submission opportunities (i.e., a measure of the future competition effect), increases the probaility of the hybrid design being optimal.

[12] Table B.4 in the Online Appendix presents the parameters of the optimal uniform prize structure for each design.

are optimized contest by contest, with the exceptions being the benchmark and hybrid prize structures. For example, in the design with 6 prizes, the gains with a uniform prize structure are less than 1 percent smaller than those with the contest-by-contest optimal prize structure. The uniform prize structures perform worse in the benchmark and hybrid prize structures because of the heterogeneity in the optimal benchmark scores across contests we show in Figure 5.

Lastly, to measure the gains of using the optimal design in US dollars, we perform the following exercise. For every competition, we compute the equilibrium of the game using the optimal prize structure but with a prize pool that is scaled down so that the equilibrium outcomes match the outcomes observed in the data. We do this using the optimal hybrid prize structure for each contest and using the optimal 6-prize structure under the constraint of uniformity across contests, as these are the designs that perform best in each column of Table 3. The results suggest that the contest designer could achieve the same outcomes in the data and save an average of $14,894 if they used the optimal hybrid structure or $14,355 if they used the uniform 6-prize structure. That is, the contest designer could achieve the same outcomes while saving nearly half of the prize money by better managing the discouragement effects with more flexibly prize structures.

Combined, these results suggest that both discouragement effects (i.e., future competition and current competition) impact participation incentives. Contingent prizes can boost incentives by counteracting the discouragement effects (in particular, the future competition effect) and cause economically significant gains in contest outcomes.

5 Experimental Evidence

5.1 Description of the Experiment

To complement our model-based evidence, we recruited University of British Columbia (UBC) and University of Illinois at Urbana-Champaign (UIUC) students for a randomized control trial, which we ran on Kaggle.[13] We exploit experimental variation in prize structure to measure the impact of the prize structure on contest outcomes and competition dynamics.

We recruited 405 students (both undergraduates and graduates) via emails, department

[13] Approval from the University of Illinois Human Subjects Committee, IRB22154, and the University of British Columbia's Behavioral Research Ethics Board, H21-01835.

newsletters, and flyers. Registration required participants to create a Kaggle account and complete a short survey, which we use as our baseline survey. The survey asked participants whether they had participated in an online data science competition prior to the study and whether they had statistics and machine learning skills.

In the experiment, we created 81 groups (or contests) of 5 participants. Each group was randomly assigned to one of three prize structures and all other aspects of the contests were identical (e.g., difficulty, reward budget, duration, number of participants). We ran the competitions simultaneously. In the competitions, player had 11 days to solve a simple prediction problem: interpolate a function (see Online Appendix C for details). Players could submit up to 10 sets of predictions per day and all competitions displayed a real-time leaderboard providing information about the performance of all participants. The objective of the competition was to achieve prediction accuracy, as measured by RMSE. The prize pool in every competition, regardless of the prize structure, was \$100 (in Amazon gift cards).

As mentioned, each group was randomly assigned to one of three prize structures. In the first, the leader at the end of the competition received \$100 (control). In the second, the leader at 80 percent of the competition time—at the end of day 9 (out of 11 days)—received \$30, and the leader at the end of the competition received \$70 (treatment "2 prizes"). Lastly, the third one awarded \$30 to the first player to surpass a benchmark score and \$70 to the leader at the end of the competition (treatment "hybrid/benchmark"). We set the benchmark score at 0.15, which was the median score of the winning submission in an experiment that we ran in the past (Lemus and Marshall, 2021), where different participants had to solve the same problem.

Table 4 shows the outcome of the randomization. For every covariate in the baseline survey, we ran an OLS regression with indicators for every treatment assignment, where the control group is the omitted category. Column 1 reports the average value of the covariates in the control group and columns 2 to 5 report the coefficients on the treatment indicators as well as the p-values from statistical significance tests. Column 6 reports the p-value from a joint test of statistical significance of both indicators. The table shows that about 10 percent of participants had prior experience in online data science competitions, 75 percent had knowledge of statistical tools, whereas only about half reported knowing machine learning techniques. There are no statistically significant differences in these covariates across treatment groups. The table also reports that 27 out of 81 competitions featured UBC students, with the remaining 54 being composed of UIUC students.

5.2 Results

Table 5 reports the main results on the impacts of the prize structure on contest outcomes. We consider two contest-level outcome variables: the minimum score (i.e., the best score) and the number of submissions. For every outcome variable we run an OLS regression with indicators for each treatment assignment. The first two columns exclude controls, whereas the second two include the covariates in Table 4 as controls.[14]

Columns 1 and 3 suggest that the hybrid/benchmark prize structure caused the minimum score to decrease by 0.05, which is about a third of the mean of the dependent variable. These columns also suggest no statistical difference in the average minimum score between the control group and the contests assigned to a 2-prize design.[15] To shed light on heterogeneity, Figure 6 plots the minimum scores across all contests, by treatment assignment. Figure 6.A shows that the cumulative distribution functions of the control and the 2-prize groups cross each other, whereas Figure 6.B shows that the cumulative distribution function of minimum scores in the control groups first-order stochastically dominates that of the hybrid/benchmark contests.

Columns 2 and 4 of Table 5 show that at the end of the competition there are no statistical differences across treatment assignments in the average number of submissions. This comparison, however, obscures how the prize structure shapes incentives to exert effort throughout the competition. Figure 7 plots the difference in the average number of submissions between treatment X and the control group, by day. Figure 7.A shows that early in the competition, the 2-prize contests had an average number of submissions that was greater than that of the control contests, with the difference peaking at day 8 (a day before the first of the two prizes was awarded). The difference becomes negative in the last three days of the competition, which is as expected: after the first of the prizes is awarded, the control contests have a greater continuation prize (all else equal), which should lead to greater effort provision in those contests.[16] A similar pattern is observed in Figure 7.B, where the difference in submissions was greatest in the first four days, which reflects the effort of participants to surpass the benchmark score in the hybrid/benchmark contests.

[14]The table has 80 observations because 1 contest received no submissions.

[15]We had one outlier in our experiment. After removing this outlier, the 2-prizes design features lower minimum scores on average.

[16]Along these lines, Table B.5 in the Online Appendix shows that conditional on the minimum score at day 9, the number of submissions was on average lower in the 2-prizes and hybrid/benchmark contests.

6 Discussion and Implications

We study dynamic contests with public, real-time performance feedback. These type of contests are widely used in practice, so it is valuable to understand simple ways in which a contest designer can improve outcomes, on a fixed the budget. We identify two central forces governing incentives to play at any point during the competition: the future-competition effect (i.e., plays that are yet to unfold) and current-competition effect (i.e., current maximum score).

Contingent-prizes, rather than prizes only at the end of the competition, can affect the balance of these effects and, therefore, can improve outcomes. To empirically shed light on this issue, we measure the performance of various contest designs featuring time- or score-contingent prizes using both an empirical structural model and an experiment. Our model estimates and experimental results show that a combination of score- and time-contingent prizes generate significant gains relative to the baseline design where all the prize money is awarded to the leader at the end of the competition. Furthermore, we characterize the optimal prize structure for each contest in the data. Moroever, we find parameters for each prize structure that the designer can use when she does not know the primitive parameters of a contest, and we show that this "robust" design captures a large portion of the gains from using the optimal design in each contest.

7 References

Alshech, Shahar and Aner Sela (2021) "The optimal allocation of prizes in two-stage contests," *Economics Letters*, 203, 109850.

Aycinena, Diego and Lucas Rentschler (2019) "Entry in contests with incomplete information: Theory and experiments," *European Journal of Political Economy*, 60, 101803.

Benkert, Jean-Michel and Igor Letina (2020) "Designing dynamic research contests," *American economic journal: Microeconomics*, 12 (4), 270–89.

Cason, Timothy N, William A Masters, and Roman M Sheremeta (2020) "Winner-take-all and proportional-prize contests: theory and experimental results," *Journal of Economic Behavior & Organization*, 175, 314–327.

Chowdhury, Subhasish M and Sang-Hyun Kim (2017) ""Small, yet Beautiful": Reconsidering the optimal design of multi-winner contests," *Games and Economic Behavior*, 104, 486–493.

Clark, Derek J and Tore Nilssen (2020) "Creating balance in dynamic competitions," *International Journal of Industrial Organization*, 69, 102578.

Ehrenberg, Ronald G and Michael L Bognanno (1990) "Do tournaments have incentive effects?" *Journal of political Economy*, 98 (6), 1307–1324.

Ely, Jeffrey, George Georgiadis, Sina Moghadas Khorasani, and Luis Rayo (2021) "Optimal feedback in contests," in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 430–431.

Feng, Xin and Jingfeng Lu (2018) "How to split the pie: optimal rewards in dynamic multi-battle competitions," *Journal of Public Economics*, 160, 82–95.

Fu, Qiang and Jingfeng Lu (2009) "The beauty of "bigness": On optimal design of multi-winner contests," *Games and Economic Behavior*, 66 (1), 146–161.

——— (2010) "Contest design and optimal endogenous entry," *Economic Inquiry*, 48 (1), 80–88.

——— (2012) "The optimal multi-stage contest," *Economic Theory*, 51 (2), 351–382.

Gross, Daniel P (2017) "Performance feedback in competitive product development," *The RAND Journal of Economics*, 48 (2), 438–466.

Güth, Werner, René Levínskỳ, Kerstin Pull, and Ori Weisel (2016) "Tournaments and piece rates revisited: a theoretical and experimental study of output-dependent prize tournaments," *Review of Economic Design*, 20 (1), 69–88.

Harris, Christopher and John Vickers (1987) "Racing with uncertainty," *The Review of Economic Studies*, 54 (1), 1–21.

Iqbal, Hamzah and Alex Krumer (2019) "Discouragement effect and intermediate prizes in multi-stage contests: Evidence from Davis Cup," *European Economic Review*, 118, 364–381.

Kireyev, Pavel (2020) "Markets for ideas: Prize structure, entry limits, and the design of ideation contests," *The RAND Journal of Economics*, 51 (2), 563–588.

Klein, Arnd Heinrich and Armin Schmutzler (2021) "Incentives and motivation in dynamic contests," *Journal of Economic Behavior & Organization*, 189, 194–216.

Konrad, Kai A and Dan Kovenock (2009) "Multi-battle contests," *Games and Economic Behavior*, 66 (1), 256–274.

Lemus, Jorge and Guillermo Marshall (2021) "Dynamic tournament design: Evidence from prediction contests," *Journal of Political Economy*, 129 (2), 383–420.

——— (2022) "Teamwork in Contests."

Liu, Bin, Jingfeng Lu, Ruqu Wang, and Jun Zhang (2018) "Optimal prize allocation in contests: The role of negative prizes," *Journal of Economic Theory*, 175, 291–317.

Mihm, Jürgen and Jochen Schlapp (2019) "Sourcing innovation: On feedback in contests," *Management science*, 65 (2), 559–576.

Moldovanu, Benny and Aner Sela (2001) "The optimal allocation of prizes in contests," *American Economic Review*, 91 (3), 542–558.

——— (2006) "Contest architecture," *Journal of Economic Theory*, 126 (1), 70–96.

Olszewski, Wojciech and Ron Siegel (2020) "Performance-maximizing large contests," *Theoretical Economics*, 15 (1), 57–88.

Sheremeta, Roman M (2011) "Contest design: An experimental investigation," *Economic Inquiry*, 49 (2), 573–590.

Stracke, Rudi, Wolfgang Höchtl, Rudolf Kerschbamer, and Uwe Sunde (2014) "Optimal prizes in dynamic elimination contests: Theory and experimental evidence," *Journal of Economic Behavior & Organization*, 102, 43–58.

Taylor, Curtis R (1995) "Digging for golden carrots: An analysis of research tournaments," *The American Economic Review*, 872–890.

Table 1: Summary of competitions (partial list of competitions)

Competition	Total reward	Submissions	Start date	Deadline
Heritage Health Prize	500,000	2,687	04/04/2011	04/04/2013
Allstate Purchase Prediction Challenge	50,000	1,204	02/18/2014	05/19/2014
Higgs Boson Machine Learning Challenge	13,000	1,776	05/12/2014	09/15/2014
Acquire Valued Shoppers Challenge	30,000	2,347	04/10/2014	07/14/2014
Liberty Mutual Group - Fire Peril Loss Cost	25,000	1,057	07/08/2014	09/02/2014
Driver Telematics Analysis	30,000	1,619	12/15/2014	03/16/2015
Crowdflower Search Results Relevance	20,000	1,645	05/11/2015	07/06/2015
Caterpillar Tube Pricing	30,000	1,938	06/29/2015	08/31/2015
Liberty Mutual Group: Property Inspection Prediction	25,000	1,271	07/06/2015	08/28/2015
Coupon Purchase Prediction	50,000	631	07/16/2015	09/30/2015
Springleaf Marketing Response	100,000	1,567	08/14/2015	10/19/2015
Homesite Quote Conversion	20,000	2,557	11/09/2015	02/08/2016
Prudential Life Insurance Assessment	30,000	818	11/23/2015	02/15/2016
Santander Customer Satisfaction	60,000	1,138	03/02/2016	05/02/2016
Expedia Hotel Recommendations	25,000	436	04/15/2016	06/10/2016

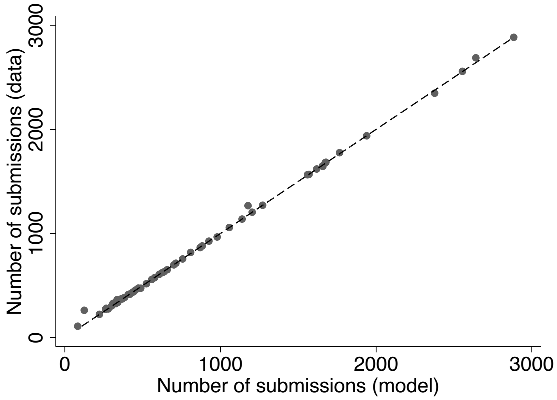
Notes: The table only considers submissions by the top 10 teams of each competition. The total reward is measured in US dollars at the moment of the competition. See Table B.1 in the Online Appendix for the complete list of competitions.

Table 2: GMM estimates of model parameters (partial list of competitions)

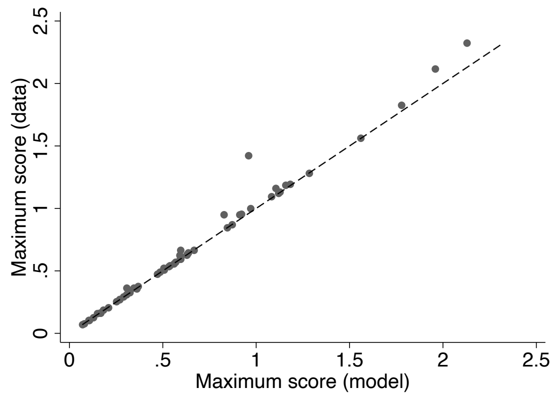
Competition	σ	SE	λ	β_0 (q)	SE	β_1 (q)	SE	Obj. Fun	N
Heritage Health Prize	0.0279	0.0001	0.7306	-3.5412	0.0461	-1.3618	0.1938	0.0138	2687
Allstate Purchase Prediction Challenge	0.0555	0.0001	0.4519	-3.1441	0.0473	-1.3618	0.1938	0.0006	1204
Higgs Boson Machine Learning Challenge	0.0648	0.0001	0.6307	-3.3141	0.0994	-1.3618	0.1938	0.0042	1776
Acquire Valued Shoppers Challenge	0.0148	0.0003	0.4759	-2.3999	0.1707	-1.3618	0.1938	0.0045	2347
Liberty Mutual Group - Fire Peril Loss Cost	0.0367	0.0001	0.2825	-2.075	0.1316	-1.3618	0.1938	0.0011	1057
Driver Telematics Analysis	0.0735	0.0001	0.4571	-4.2049	0.1214	-1.3618	0.1938	0.0018	1619
Crowdflower Search Results Relevance	0.0129	0.0001	0.2806	-2.8256	0.105	-1.3618	0.1938	0.0012	1645
Caterpillar Tube Pricing	0.0172	0.0001	0.3166	-3.0875	0.0277	-1.3618	0.1938	0.0004	1938
Liberty Mutual Group: Property Inspection Prediction	0.0233	0.0001	0.2667	-3.1092	0.06	-1.3618	0.1938	0.0009	1271
Coupon Purchase Prediction	0.1222	0.0003	0.3848	-2.0006	0.2093	-1.3618	0.1938	0.0027	631
Springleaf Marketing Response	0.0326	0.0001	0.332	-3.0862	0.0708	-1.3618	0.1938	0.0009	1567
Homesite Quote Conversion	0.0223	0.0001	0.4559	-3.1196	0.0355	-1.3618	0.1938	0.0009	2557
Prudential Life Insurance Assessment	0.0749	0.0006	0.4219	-3.0026	0.0449	-1.3618	0.1938	0.002	818
Santander Customer Satisfaction	0.0218	0.0001	0.3059	-2.2694	0.0479	-1.3618	0.1938	0.0021	1138
Expedia Hotel Recommendations	0.1011	0.0001	0.2814	-1.7786	0.0485	-1.3618	0.1938	0.0005	436

Notes: The table reports the GMM estimates with bootstrapped standard errors. See Table B.2 in the Online Appendix for the complete list of competitions.

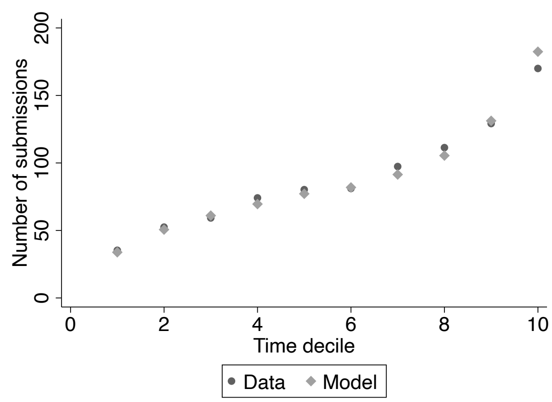
Figure 3: Model fit



A) Number of submissions



B) Maximum score



C) Number of submissions over time

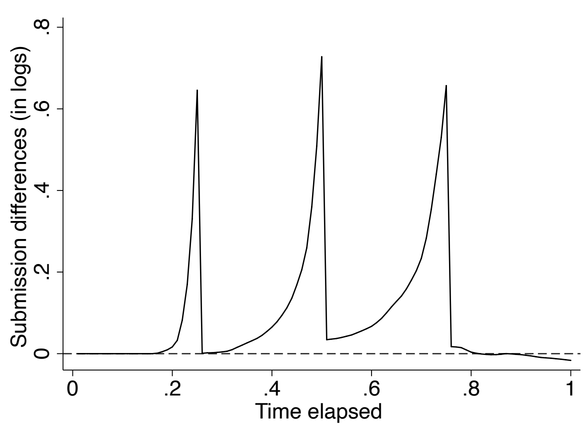
Notes: Panels A and B are scatter plots of the number of submissions and the maximum score in every contest, where the data is on the y-axis and the model predictions on the x-axis. Panel C plots the average number of submissions across contests for every time decile in the data and predicted by the model.

Table 3: Equilibrium outcomes under alternative prize structures

	Change in # of submissions (in %)		Change in max score (in st. dev.)	
	Optimal	Uniform	Optimal	Uniform
2 prizes	13.378	12.985	0.017	0.016
	(1.676)	(1.668)	(0.004)	(0.004)
4 prizes	25.061	24.742	0.032	0.032
	(2.353)	(2.338)	(0.008)	(0.007)
6 prizes	30.535	30.389	0.039	0.039
	(2.67)	(2.67)	(0.008)	(0.008)
2 timed prizes	21.096	15.368	0.03	0.02
	(2.185)	(1.805)	(0.008)	(0.005)
Benchmark	30.802	-84.503	0.04	-0.168
	(3.282)	(13.715)	(0.009)	(0.052)
Hybrid	36.528	8.45	0.047	0.01
	(3.149)	(2.542)	(0.009)	(0.003)

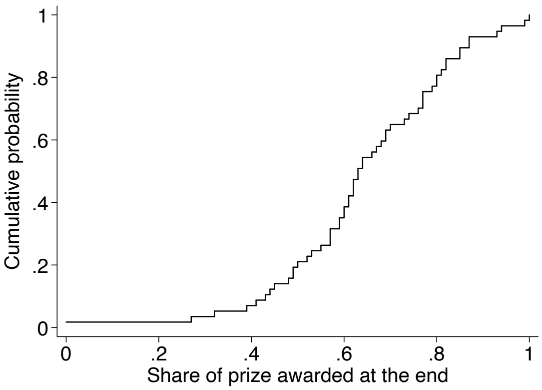
Notes: The table reports the average change in the number of submissions and maximum score when using alternative prize structures. The outcome differences compare the optimal design for each prize structure class. The column "Optimal" reports results for when the optimal design is computed for each contest, whereas the column "Uniform" for when the optimal design is constrained to be uniform across all contests.

Figure 4: Submission dynamics in a 4-prize prize structure relative to the baseline design

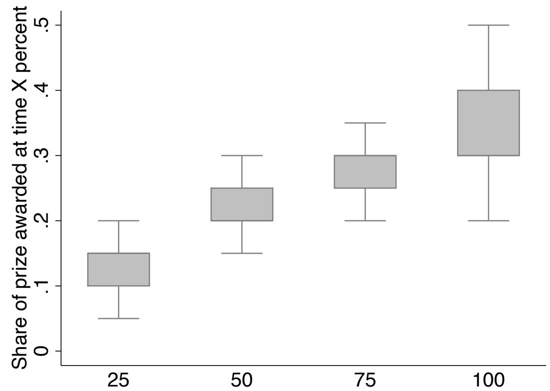


Notes: The figures present the change in the equilibrium number of submissions predicted by the model (in logs) when implementing the optimal 4-prize prize structure (relative to the baseline design).

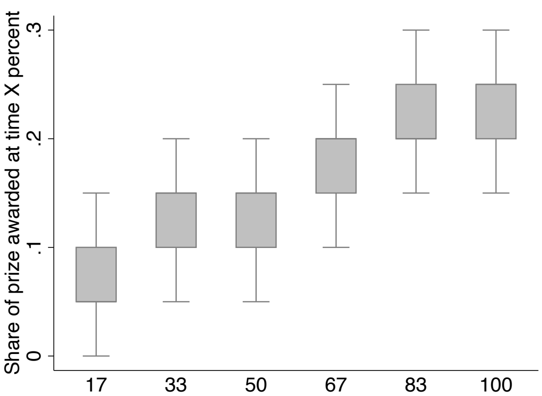
Figure 5: Optimal prize structures, by prize class



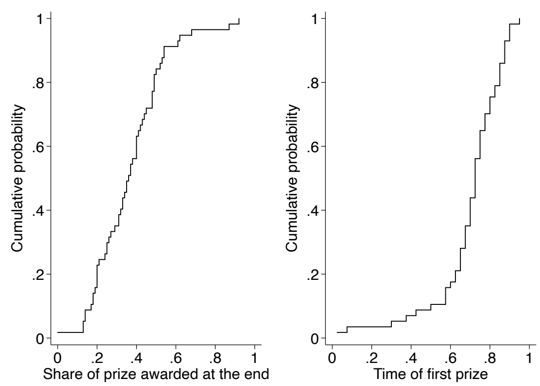
A) 2 equally-spaced prizes



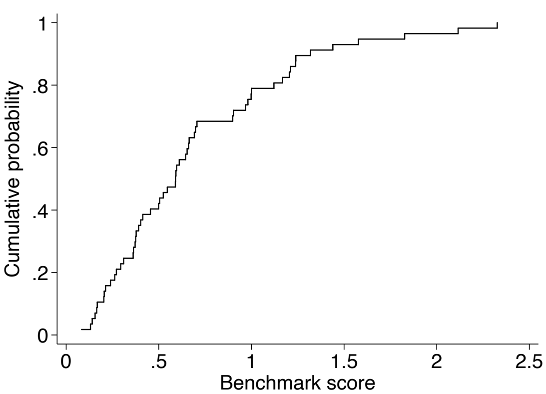
B) 4 equally-spaced prizes



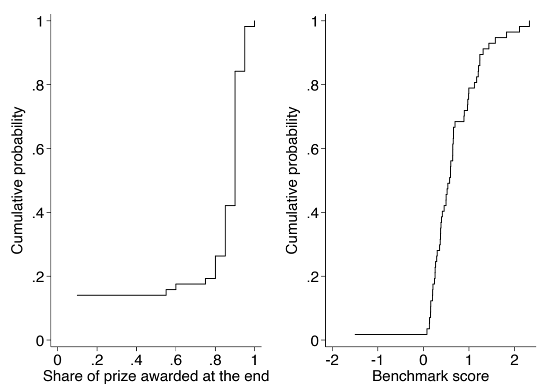
C) 6 equally-spaced prizes



D) 2 prizes with an optimally-timed first prize



E) Benchmark prize



F) Hybrid benchmark prize

Notes: The different panels present the distribution of the parameters of the optimal prize structures within each prize class. In every plot, a data point is a contest.

Table 4: Baseline summary statistics and test of balance

Variable	Control ($N = 27$)	2 prizes ($N = 27$)		Hybrid/benchmark ($N = 27$)		F-test
	Mean	Coeff.	p-value	Coeff.	p-value	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
participated_past	0.096	0.027	0.504	0	1	0.723
machine_learning	0.467	-0.036	0.569	0.022	0.727	0.562
stat_tools	0.733	-0.018	0.775	0.037	0.533	0.636
UBC	0.407	0	-	0	-	-
UIUC	0.593	0	-	0	-	-

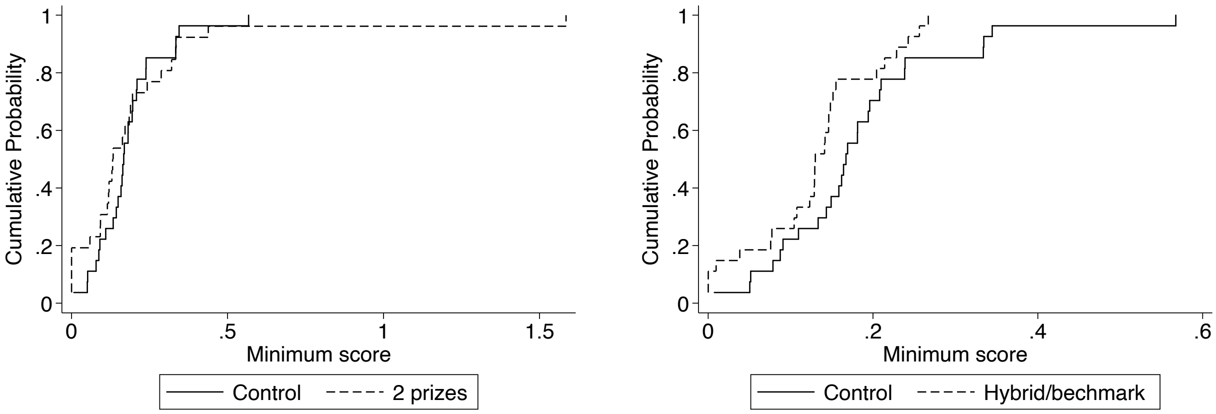
Notes: An observation is a contest. All variables are defined at the contest level as follows: 'participated_past' is the share of players in the contest who have participated in a prediction contest in the past, 'mach_learning' is the share of players in the contest who have machine learning skills, 'stat_tools' is the share of players in the contest who have learned statistics, and 'UBC' ('UIUC') is an indicator for whether the contest features UBC (UIUC) students. Columns 2-6 report the coefficients and p-values from OLS regressions of each covariate on two indicators: '2 prizes' and 'hybrid/benchmark'. Column 7 reports the p-value from a joint test of statistical significance of both indicators.

Table 5: Prize structure impacts on contest outcomes

	Min. score	# of submissions	Min. score	# of submissions
	(1)	(2)	(3)	(4)
2 prizes	0.026	4.268	0.026	5.375
	(0.063)	(10.229)	(0.064)	(10.529)
Hybrid/benchmark	-0.054**	6.000	-0.055**	6.492
	(0.026)	(10.811)	(0.027)	(11.072)
Controls	No	No	Yes	Yes
N	80	80	80	80
R^2	0.030	0.005	0.082	0.042
Mean dep. variable	0.174	39.375	0.174	39.375
Std. dev. dep. variable	0.191	37.549	0.191	37.549

Notes: An observation is a contest. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Controls include all covariates reported in Table 4.

Figure 6: Distribution of minimum scores, by prize structure

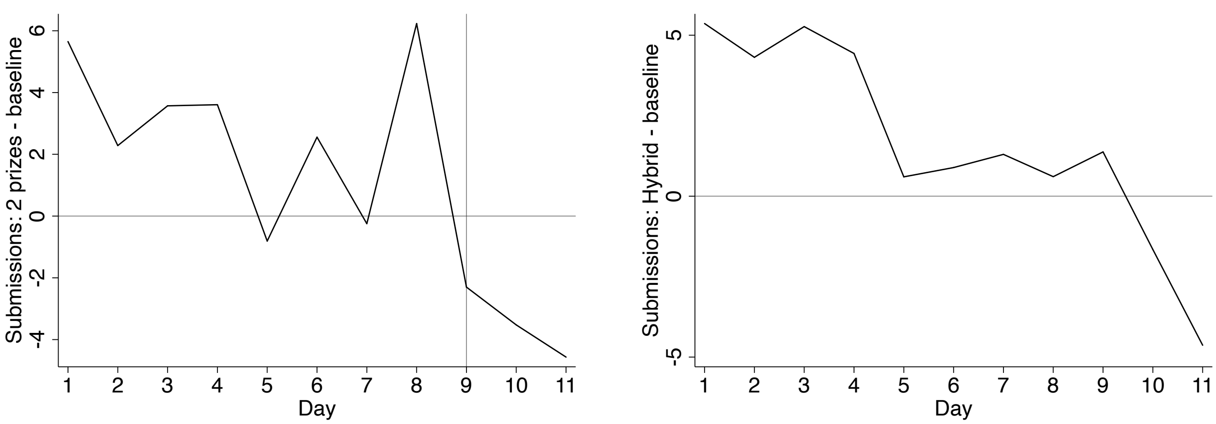


A) 2 prizes

B) Hybrid/benchmark

Notes: The figures present the cumulative distribution functions of the maximum score for the different prize structure. A data point in every figure is a contest.

Figure 7: Submissions over time, by prize structure



A) 2 prizes

B) Hybrid/benchmark

Notes: The figures present the average difference between the daily number of submissions in treatment X and the control contests. An observation is a contest. The vertical line in Panel A marks the day in which the first prize was awarded in the 2-prize contests.

Supplemental Material – Intended for Online Publication

Contingent Prizes in Dynamic Contests

Jorge Lemus and Guillermo Marshall

A Omitted Proofs

Proof of Lemma 1

Proof. If the leader has the last opportunity to play, there is no benefit in paying a cost to play, so she does not play. If the follower has the last opportunity to play, the benefit is to replace the leader at the end of the competition. If the leader's score is s, and the follower's playing cost is c, the expected benefit of playing is $q(s)\pi_2$ and the cost is c. $\square$

Proof of Lemma 2

Proof. At this stage, player 1 is the leader with score $\bar{s}$. Player 2 observes the cost she will incur by playing, c, and chooses whether to play or not. Let $E_H[c] = E[c|c < q_H\pi_2]$. Player 2's expected payoff from playing is

$$q_H \underbrace{[\pi_1 + (1-\alpha)\pi_2 + \alpha\{Q_L(1-q_L) + 1 - Q_L\}\pi_2]}_{\text{Leader}} + (1-q_H) \underbrace{(1-\alpha)Q_H(q_H\pi_2 - E_H[c])}_{\text{Follower}} - c \tag{20}$$

To explain the term 'Leader' in (20) note that when player 2 plays at stage 1, with probability q_H she becomes the leader and gets π_1. This leaves two possibilities for the next period: (a) player 2 has the last opportunity to play in the game with probability $(1-\alpha)$, in which case she does not play (because she is the leader) and wins π_2; (b) player 1, who now is the follower, has the the last opportunity to play in the game with probability α. She decides to exercise this option and play with probability Q_L, but she does not defeat the current leader (player 2) with probability $(1-q_L)$, in which case player 2 receives π_2; with probability $1-Q_L$, player 1 does not exercise the opportunity to play in the second stage, so player 2 remains the leader of the competition and gets π_2.

The term 'Follower' in (20) consists of the probability that player 2 plays but fails to take the lead in the competition, $1-q_H$, times the probability she has another opportunity to play, $1-\alpha$, times the probability that she exercises the option to play, Q_H, times the expected payoff from playing and taking the leads, $q_H\pi_2 - E_H[c]$.

If player 2 does not play in stage 1, the game will move to the next stage, and scores won't change so player 2 will be the follower for sure. Thus, player 2's payoff from not playing in the first stage is

$$(1-\alpha)Q_H(q_H\pi_2 - E_H[c]) \tag{21}$$

Comparing the payoffs in (20) and (21), player 2 exercise the option to play in stage 1 when the expected payoff change is larger than the cost, i.e. $q_H(\text{Leader} - \text{Follower}) \geq c$, which simplifies to the expression in the Lemma. □

Proof of Lemma 3

Proof. Let

$$E(\alpha, \pi_2) = q_H(1 - \pi_2\alpha q_L Q_L - (1-\alpha)Q_H(q_H\pi_2 - E[c|c \leq q_H\pi_2]))$$

First, consider the comparative static in α. We have,

$$\frac{\partial E}{\partial \alpha} = q_H(-\pi_2 q_L Q_L + Q_H(q_H\pi_2 - E[c|c \leq q_H\pi_2]))$$

This term is positive when $Q_H(q_H\pi_2 - E[c|c \leq q_H\pi_2]) > \pi_2 q_L Q_L$. Second, consider the comparative static in π_2. Note that

$$Q_H E[c|c \leq q_H\pi_2] = \int_0^{q_H\pi_2} ck(c)dc,$$

where $k(\cdot)$ is the density associated to the cost distribution. Taking partial derivative with respect to π_2

$$\frac{\partial E}{\partial \pi_2} = q_H(-\alpha q_L Q_L - \alpha\pi_2 q_L^2 k(q_L\pi_2) - (1-\alpha)Q_H q_H) < 0.$$

□

Optimal Time-Contingent Prize Structure

Proposition 1. *The optimal prize at $t = 2$, $\pi_2^* \in [0,1]$, maximizes the difference $p_{\bar{s}+2\varepsilon} - p_{\bar{s}}$, where $q_H = q(\bar{s})$, $q_L = q(\bar{s}+\varepsilon)$, $Q_H = K(q_H\pi_2)$, $Q_L = K(q_L\pi_2)$, and*

$$\begin{aligned} p_{\bar{s}+2\varepsilon} &= Q_1 q_H \alpha Q_L q_L, \\ p_{\bar{s}} &= (1 - Q_1 q_H)(1 - (1-\alpha) Q_H q_H), \\ Q_1 &= K(q_H(1 - \pi_2 \alpha q_L Q_L - (1-\alpha) Q_H (q_H \pi_2 - E[c|c \leq q_H \pi_2]))). \end{aligned}$$

Proof. We normalize the total prize pool to 1, so $\pi_1 + \pi_2 = 1$. The probability that Player 2 plays in the first stage is

$$Q_1(\pi_2) = K(q_H(1 - \pi_2 \alpha q_L Q_L - (1-\alpha) Q_H (q_H \pi_2 - E[c|c \leq q_H \pi_2])))$$

The maximum score at the end of the competition is the random variable:

$$s^* = \begin{cases} \bar{s} & \text{with prob. } p_{\bar{s}} \\ \bar{s}+\varepsilon & \text{with prob. } p_{\bar{s}+\varepsilon} \\ \bar{s}+2\varepsilon & \text{with prob. } p_{\bar{s}+2\varepsilon} \end{cases}$$

where

$$\begin{aligned} p_{\bar{s}} &= [Q_1(1-q_H) + (1-Q_1)][\alpha + (1-\alpha)(Q_H(1-q_H) + 1 - Q_H)] \\ &= (1 - Q_1 q_H)(1 - (1-\alpha) Q_H q_H) \end{aligned}$$

$$\begin{aligned} p_{\bar{s}+\varepsilon} &= [Q_1(1-q_H) + (1-Q_1)][(1-\alpha) Q_H q_H] + [Q_1 q_H][1 - \alpha + \alpha(Q_L(1-q_L) + 1 - Q_L)] \\ &= (1 - Q_1 q_H)(1-\alpha) Q_H q_H + Q_1 q_H (1 - \alpha Q_L q_L) \end{aligned}$$

$$p_{\bar{s}+2\varepsilon} = Q_1 q_H \alpha Q_L q_L$$

The expected maximum score is therefore

$$\begin{aligned} E[s^*] &= \bar{s} p_{\bar{s}} + (\bar{s}+\varepsilon) p_{\bar{s}+\varepsilon} + (\bar{s}+2\varepsilon) p_{\bar{s}+2\varepsilon} \\ &= \bar{s} + \varepsilon + \varepsilon(p_{\bar{s}+2\varepsilon} - p_{\bar{s}}) \end{aligned}$$

Thus, to maximize the expected maximum score we solve

$$\max_{0 \le \pi_2 \le 1} p_{\bar{s}+2\varepsilon} - p_{\bar{s}}.$$

Which is the same as

$$\max_{0 \le \pi_2 \le 1} Q_1 q_H \alpha Q_L q_L - (1 - Q_1 q_H)(1 - (1-\alpha) Q_H q_H).$$

Ignoring constants, the objective function can be written as

$$Q_1[1 + \alpha Q_L q_L - (1-\alpha) Q_H q_H] + (1-\alpha) Q_H$$

□

Proof of Lemma 4

Proof. (1) Benchmark $B = \bar{s} + \varepsilon$.

Stage 2: We reach this stage only if the score after Stage 1 is $\bar{s}$. The player with the opportunity to play at $t = 2$ plays and advances the score to B with probability q_H in which case she gets 1. Let $Q_H = K(q_H)$.

Stage 1: Player B's expected payoff from playing is

$$q_H + (1 - q_H)(1-\alpha) Q_H (q_H - E[c|c < q_H]) - c \tag{22}$$

The expected payoff of not playing is the payoff of continuing as a follower for sure at no cost, i.e., $(1-\alpha) Q_H (q_H - E[c|c < q_H])$. Thus, Player B plays whenever

$$q_H (1 - (1-\alpha) Q_H (q_H - E[c|c < q_H])) \ge c \tag{23}$$

□

Proof of Lemma 5

Proof. (2) Benchmark $B = \bar{s} + 2\varepsilon$.

Stage 2: There is a play only if the score after Stage 1 is $\bar{s}+\varepsilon$. In that case, there is a play when $q_L > c$. Let $Q_L = K(q_L)$.

Stage 1: Player B must succeed and must have the next opportunity to play to get a positive payoff. The expected payoff from playing is

$$q_H(1-\alpha)Q_L(q_L - E[c|c < q_L]) - c \tag{24}$$

The expected payoff of not playing is 0, so Player B plays whenever

$$q_H(1-\alpha)Q_L(q_L - E[c|c < q_L]) \geq c \tag{25}$$

□

Optimal Simple Benchmark Score

Proposition 2. *Comparing the expected maximum score, we can find the optimal simple-benchmark score*

Proof. First, consider a low benchmark score $s+\varepsilon$.

Let $Q_1 = K(q_H(1-(1-\alpha)Q_H(q_H - E[c|c < q_H])))$. Note that the contest always end before reaching score $\bar{s}+2\varepsilon$, so $p_{\bar{s}+2\varepsilon} = 0$. The probability of ending with a score $\bar{s}$ is

$$p_{\bar{s}} = (1-Q_1+Q_1(1-q_H))(1-Q_H+Q_H(1-q_H)) = (1-Q_1q_H)(1-Q_Hq_H).$$

Now consider a high benchmark score of $s+2\varepsilon$. Let $Q_1 = K(q_H(1-\alpha)Q_L(q_L - E[c|c < q_L]))$. The probability of ending with a score $\bar{s}$ is

$$p_{\bar{s}} = (1-Q_1+Q_1(1-q_H)) = (1-Q_1q_H).$$

The probability of ending with a score $\bar{s}+2\varepsilon$ is

$$p_{\bar{s}+2\varepsilon} = Q_1q_HQ_Lq_L.$$

We find the optimal benchmark score by comparing the difference $p_{\bar{s}+2\varepsilon} - p_{\bar{s}}$ in both cases. □

Analysis Hybrid Benchmark Case

Proof. (1) Benchmark $B = \bar{s} + \varepsilon$.

Stage 2: At this stage the score is either $\bar{s}$ or B. Consider the player with the opportunity to play at $t = 2$.

If the current score is $\bar{s}$ and the leader does not play she gets $1 - \pi_B$ for being the leader. If she plays and advances the score to B, she gets 1. Thus, the leader chooses to play when $q_H + (1 - q_H)(1 - \pi_B) - c \geq 1 - \pi_B$ or, equivalently,

$$q_H \pi_B \geq c.$$

Let $Q_H^{\text{leader}} = K(q_H \pi_B)$. The follower gets 0 if she does not play and $q_H - c$ if she plays. Thus, the probability that follower plays is $Q_H = K(q_H)$.

If the current score is B, then the leader has no incentives to play and gets $1 - \pi_B$ for sure. If the follower has the opportunity to play, she competes for being the leader at the end of the competition and plays whenever $q_L(1 - \pi_B) \geq c$. Let $Q_L = K(q_L(1 - \pi_B))$.

Stage 1: Player B's expected payoff from playing is

$$\begin{aligned} q_H[\underbrace{[\pi_B + (1-\alpha)(1-\pi_B) + \alpha\{Q_L(1-q_L) + 1 - Q_L\}(1-\pi_B)]}_{\text{Leader}}] + \\ (1-q_H)\underbrace{(1-\alpha)Q_H(q_H - E_H[c])}_{\text{Follower}} - c \end{aligned} \tag{26}$$

The expected payoff of not playing is the payoff of continuing as a follower for sure at no cost, i.e., $(1 - \alpha)Q_H(q_H - E_H[c])$. Thus, Player B plays whenever

$$q_H[1 - \alpha Q_L q_L(1 - \pi_B) - (1 - \alpha)Q_H(q_H - E_H[c])] \geq c \tag{27}$$

Let $Q_1 = K(q_H[1 - \alpha Q_L q_L(1 - \pi_B) - (1 - \alpha)Q_H(q_H - E_H[c])])$. The probability of ending with a score $\bar{s}$ is

$$p_{\bar{s}} = (1 - Q_1 + Q_1(1 - q_H))(\alpha(1 - Q_H^{\text{leader}} + Q_H^{\text{leader}}(1 - q_H)) + (1 - \alpha)(1 - Q_H + Q_H(1 - q_H))).$$

The probability of ending with a score $\bar{s} + 2\varepsilon$ is

$$p_{\bar{s}+2\varepsilon} = Q1 q_H Q_L q_L.$$

We can find the optimal benchmark prize π_B to maximize the expected maximum score.

(2) Benchmark $B = \bar{s} + 2\varepsilon$.

Stage 2: With this high benchmark prize we can incentivize the leader to play at period 2 when the score is $\bar{s} + \varepsilon$. With time-contingent prizes the leader never plays at the last stage (there is no "future competition" effect). Now, however, the leader gets only $1 - \pi_B$ if the score is lower than B. So, when the score is $\bar{s} + \varepsilon$ the leader plays in the last stage when $q_L \pi_B > c$. Let $Q_L^{\text{lead}} = K(q_L \pi_B)$, $Q_L = K(q_L(1 - \pi_B))$ and $Q_H = K(q_H(1 - \pi_B))$.

When the score is $\bar{s}$ in the last stage, the benchmark is unattainable, so the leader never plays and receives a prize of $1 - \pi_B$. Thus, there is a punishment from not increasing the score at stage 1, which incentivizes player B to play at stage 1.

Stage 1: Player B's expected payoff from playing is

$$q_H[\underbrace{(1-\alpha)[(1-\pi_B) + Q_L^{\text{lead}}(q_L\pi_B - E[c|c < q_L\pi_B])] + \alpha\{Q_L(1-q_L) + 1 - Q_L\}(1-\pi_B)}_{\text{Leader}}] + (1-q_H)\underbrace{(1-\alpha)Q_H(q_H(1-\pi_B) - E[c|c < q_H(1-\pi_B)])}_{\text{Follower}} - c \tag{28}$$

The expected payoff of not playing is the payoff of continuing as a follower for sure at no cost, i.e., $(1-\alpha)Q_H(q_H(1-\pi_B) - E_H[c])$. Thus, Player B plays whenever

$$q_H[1 - \pi_B + (1-\alpha)Q_L^{\text{lead}}(q_L\pi_B - E_L[c]) - \alpha(1-\pi_B)Q_L q_L - (1-\alpha)Q_H(q_H(1-\pi_B) - E_H[c])] \geq c \tag{29}$$

Let $Q_1 = K(q_H[1 - \pi_B + (1-\alpha)Q_L^{\text{lead}}(q_L\pi_B - E_L[c]) - \alpha(1-\pi_B)Q_L q_L - (1-\alpha)Q_H(q_H(1-\pi_B) - E_H[c])])$. The probability of ending with a score $\bar{s}$ is

$$p_{\bar{s}} = (1 - Q_1 + Q_1(1-q_H))(\alpha + (1-\alpha)(1 - Q_H + Q_H(1-q_H))).$$

The probability of ending with a score $\bar{s} + 2\varepsilon$ is

$$p_{\bar{s}+2\varepsilon} = Q_1 q_H((1-\alpha)Q_L^{\text{lead}} q_L + \alpha Q_L q_L).$$

We can find the optimal benchmark prize π_B to maximize the expected maximum score. □

B Value Functions Hybrid Contest

First player to reach the benchmark score B earns a prize π_B. The player that is the leader of the competition at the end receives a prize of $1 - \pi_B$.

Terminal value functions:

$$
\begin{aligned}
L_{0,T,s} &= 1 - \pi_B & (30)\\
F_{0,T,s} &= 0 & (31)\\
L_{1,T,s} &= 1 - \pi_B & (32)\\
F_{1,T,s} &= 0 & (33)\\
& & (34)
\end{aligned}
$$

The value of the leader at state $(0, t, s)$. There are three cases: nobody is summoned to play; the leader is summoned to play; one of the followers is summoned to play.

$$
\begin{aligned}
L_{0,t,s} =& \pi_t + \left(1 - \lambda \frac{N_t}{N}\right) L_{0,t',s} + \frac{\lambda}{N} L^{active}_{0,t,s} + \\
& \frac{\lambda(N-1)}{N} p^F_{0,t,s} \left(q_s(1\{s+\varepsilon = B\}F_{1,t',s+\varepsilon} + 1\{s+\varepsilon < B\}F_{0,t',s+\varepsilon}) + (1-q_s)L_{0,t',s}\right) + \\
& \frac{\lambda(N-1)}{N}(1 - p^F_{0,t,s})L_{0,t',s} \qquad (35)
\end{aligned}
$$

with

$$
L^{active}_{0,t,s} = E_c[\max\{q_s(1\{s+\varepsilon = B\}(\pi_B + L_{1,t',s+\varepsilon}) + 1\{s+\varepsilon < B\}L_{0,t',s+\varepsilon}) + (1-q_s)L_{0,t',s} - c, L_{0,t',s}\}] \qquad (36)
$$

$$
\begin{aligned}
L_{1,t,s} =& \pi_t + \left(1 - \lambda \frac{N_t}{N}\right) L_{1,t',s} + \frac{\lambda}{N} L^{active}_{1,t,s} + \frac{\lambda(N-1)}{N} p^F_{1,t,s} \left(q_s F_{1,t',s+\varepsilon} + (1-q_s)L_{1,t',s}\right) + \\
& \frac{\lambda(N-1)}{N}(1 - p^F_{1,t,s})L_{1,t',s} \qquad (37)
\end{aligned}
$$

with

$$
L^{active}_{1,t,s} = E_c\left[\max\{q_s L_{1,t',s+\varepsilon} + (1-q_s)L_{1,t',s} - c, L_{1,t',s}\}\right] \qquad (38)
$$

The value of a follower at state (t, s). There are four cases: nobody is summoned to play;

the follower is summoned to play; one of the other followers is summoned to play; the leader is summoned to play.

$$
\begin{aligned}
F_{0,t,s} = & \left(1-\lambda\frac{N_t}{N}\right)F_{0,t',s}+\frac{\lambda}{N}F^{active}_{0,t,s}+\\
& \frac{\lambda(N-2)}{N}p^F_{0,t,s}\left(q_s(1\{s+\varepsilon=B\}F_{1,t',s+\varepsilon}+1\{s+\varepsilon<B\}F_{0,t',s+\varepsilon})+(1-q_s)F_{0,t',s}\right)+\\
& \frac{\lambda(N-2)}{N}(1-p^F_{0,t,s})F_{0,t',s}+\\
& \frac{\lambda}{N}p^L_{0,t,s}\left(q_s(1\{s+\varepsilon=B\}F_{1,t',s+\varepsilon}+1\{s+\varepsilon<B\}F_{0,t',s+\varepsilon})+(1-q_s)F_{0,t',s}\right)+\\
& \frac{\lambda}{N}(1-p^L_{0,t,s})F_{0,t',s}
\end{aligned}
\tag{39}
$$

with

$$
F^{active}_{0,t,s} = E_c[\max\{q_s(1\{s+\varepsilon=B\}(\pi_B+L_{1,t',s+\varepsilon})+1\{s+\varepsilon<B\}L_{0,t',s+\varepsilon})+(1-q_s)F_{0,t',s}-c,F_{0,t',s}\}]
\tag{40}
$$

and

$$
\begin{aligned}
F_{1,t,s} = & \left(1-\lambda\frac{N_t}{N}\right)F_{1,t',s}+\frac{\lambda}{N}F^{active}_{1,t,s}+\\
& \frac{\lambda(N-2)}{N}p^F_{1,t,s}\left(q_sF_{1,t',s+\varepsilon}+(1-q_s)F_{1,t',s}\right)+\frac{\lambda(N-2)}{N}(1-p^F_{1,t,s})F_{1,t',s}+\\
& \frac{\lambda}{N}p^L_{1,t,s}\left(q_sF_{1,t',s+\varepsilon}+(1-q_s)F_{1,t',s}\right)+\frac{\lambda}{N}(1-p^L_{1,t,s})F_{1,t',s}
\end{aligned}
\tag{41}
$$

with

$$
F^{active}_{1,t,s} = E_c[\max\{q_sL_{1,t',s+\varepsilon}+(1-q_s)F_{1,t',s}-c,F_{1,t',s}\}]
\tag{42}
$$

The probability with which a follower and the leader choose to play when summoned to play

$$
p^L_{0,t,s} = \Pr\left(q_s(1\{s+\varepsilon=B\}(\pi_B+L_{1,t',s+\varepsilon})+1\{s+\varepsilon<B\}L_{0,t',s+\varepsilon}-L_{0,t',s})>c\right)
\tag{43}
$$

$$
p^L_{1,t,s} = \Pr\left(q_s(L_{1,t',s+\varepsilon}-L_{1,t',s})>c\right)
\tag{44}
$$

$$p^F_{0,t,s} = \Pr\left(q_s(1\{s+\varepsilon = B\}(\pi_B + L_{1,t',s+\varepsilon}) + 1\{s+\varepsilon < B\}L_{0,t',s+\varepsilon} - F_{0,t',s}) > c\right) \qquad (45)$$

$$p^F_{1,t,s} = \Pr\left(q_s(L_{1,t',s+\varepsilon} - F_{1,t',s}) > c\right) \qquad (46)$$

C Additional Tables and Figures

Table B.1: Summary of competitions

Competition	Total reward	Submissions	Start date	Deadline
Predict Grant Applications	5,000	371	12/13/2010	02/20/2011
RTA Freeway Travel Time Prediction	10,000	386	11/23/2010	02/13/2011
Deloitte/FIDE Chess Rating Challenge	10,000	342	02/07/2011	05/04/2011
Heritage Health Prize	500,000	2,687	04/04/2011	04/04/2013
Wikipedia's Participation Challenge	10,000	338	06/28/2011	09/20/2011
Allstate Claim Prediction Challenge	10,000	338	07/13/2011	10/12/2011
dunnhumby's Shopper Challenge	10,000	304	07/29/2011	09/30/2011
Give Me Some Credit	5,000	413	09/19/2011	12/15/2011
Don't Get Kicked!	10,000	880	09/30/2011	01/05/2012
Algorithmic Trading Challenge	10,000	442	11/11/2011	01/08/2012
What Do You Know?	5,000	371	11/18/2011	02/29/2012
Photo Quality Prediction	5,000	223	10/29/2011	11/20/2011
Benchmark Bond Trade Price Challenge	17,500	456	01/27/2012	04/30/2012
KDD Cup 2012, Track 1	8,000	1,267	02/20/2012	06/01/2012
KDD Cup 2012, Track 2	8,000	864	02/20/2012	06/01/2012
Predicting a Biological Response	20,000	651	03/16/2012	06/15/2012
Online Product Sales	22,500	418	05/04/2012	07/03/2012
EMI Music Data Science Hackathon - July 21st - 24 hours	10,000	109	07/21/2012	07/22/2012
Belkin Energy Disaggregation Competition	25,000	607	07/02/2013	10/30/2013
Merck Molecular Activity Challenge	40,000	415	08/16/2012	10/16/2012
U.S. Census Return Rate Challenge	25,000	272	08/31/2012	11/11/2012
Amazon.com - Employee Access Challenge	5,000	755	05/29/2013	07/31/2013
The Marinexplore and Cornell University Whale Detection Challenge	10,000	326	02/08/2013	04/08/2013
See Click Predict Fix - Hackathon	1,000	262	09/28/2013	09/29/2013
KDD Cup 2013 - Author Disambiguation Challenge (Track 2)	7,500	623	04/19/2013	06/12/2013
Influencers in Social Networks	2,350	281	04/13/2013	04/14/2013
Personalize Expedia Hotel Searches - ICDM 2013	25,000	517	09/03/2013	11/04/2013
StumbleUpon Evergreen Classification Challenge	5,000	328	08/16/2013	10/31/2013
Personalized Web Search Challenge	9,000	275	10/11/2013	01/10/2014
See Click Predict Fix	4,000	575	09/29/2013	11/27/2013
Allstate Purchase Prediction Challenge	50,000	1,204	02/18/2014	05/19/2014
Higgs Boson Machine Learning Challenge	13,000	1,776	05/12/2014	09/15/2014
Acquire Valued Shoppers Challenge	30,000	2,347	04/10/2014	07/14/2014
The Hunt for Prohibited Content	25,000	966	06/24/2014	08/31/2014
Liberty Mutual Group - Fire Peril Loss Cost	25,000	1,057	07/08/2014	09/02/2014
Tradeshift Text Classification	5,000	714	10/02/2014	11/10/2014
Driver Telematics Analysis	30,000	1,619	12/15/2014	03/16/2015
Diabetic Retinopathy Detection	100,000	698	02/17/2015	07/27/2015
Click-Through Rate Prediction	15,000	1,679	11/18/2014	02/09/2015
Otto Group Product Classification Challenge	10,000	926	03/17/2015	05/18/2015
Crowdflower Search Results Relevance	20,000	1,645	05/11/2015	07/06/2015
Avito Context Ad Clicks	20,000	558	06/02/2015	07/28/2015
ICDM 2015: Drawbridge Cross-Device Connections	10,000	364	06/01/2015	08/24/2015
Caterpillar Tube Pricing	30,000	1,938	06/29/2015	08/31/2015
Liberty Mutual Group: Property Inspection Prediction	25,000	1,271	07/06/2015	08/28/2015
Coupon Purchase Prediction	50,000	631	07/16/2015	09/30/2015
Springleaf Marketing Response	100,000	1,567	08/14/2015	10/19/2015
Truly Native?	10,000	474	08/06/2015	10/14/2015
Rossmann Store Sales	35,000	1,684	09/30/2015	12/14/2015
Homesite Quote Conversion	20,000	2,557	11/09/2015	02/08/2016
Prudential Life Insurance Assessment	30,000	818	11/23/2015	02/15/2016
BNP Paribas Cardif Claims Management	30,000	1,648	02/03/2016	04/18/2016
Home Depot Product Search Relevance	40,000	2,884	01/18/2016	04/25/2016
Santander Customer Satisfaction	60,000	1,138	03/02/2016	05/02/2016
Expedia Hotel Recommendations	25,000	436	04/15/2016	06/10/2016
Avito Duplicate Ads Detection	20,000	1,564	05/06/2016	07/11/2016
Draper Satellite Image Chronology	75,000	475	04/29/2016	06/27/2016

Note: The table only considers submissions by the top 10 teams of each competition. The total reward is measured in US dollars at the moment of the competition.

Table B.2: GMM estimates of model parameters

Competition	σ	SE	λ	β_0 (q)	SE	β_1 (q)	SE	GMM Obj. Fun	N
Predict Grant Applications	0.0799	0.0005	0.1391	-2.6669	0.0909	-1.3618	0.1938	0.0085	371
RTA Freeway Travel Time Prediction	0.1001	0.0004	0.1658	-0.9835	0.1434	-1.3618	0.1938	0.0013	386
Deloitte/FIDE Chess Rating Challenge	0.054	0.0001	0.1733	-1.5192	0.0541	-1.3618	0.1938	0.0036	342
Heritage Health Prize	0.0279	0.0001	0.7306	-3.5412	0.0461	-1.3618	0.1938	0.0138	2687
Wikipedia Participation Challenge	0.1496	0.0009	0.1686	-3.7402	0.0637	-1.3618	0.1938	0.0149	338
Allstate Claim Prediction Challenge	0.0956	0.0004	0.1837	-0.5781	0.2179	-1.3618	0.1938	0.0012	338
dunnhumbys Shopper Challenge	0.0919	0.0004	0.1263	-1.6724	0.1829	-1.3618	0.1938	0.009	304
Give Me Some Credit	0.0621	0.0003	0.1749	-2.6001	0.0454	-1.3618	0.1938	0.0048	413
Dont Get Kicked!	0.052	0.0001	0.2917	-2.2272	0.107	-1.3618	0.1938	0.0022	880
Algorithmic Trading Challenge	0.0452	0.0001	0.1165	-3.0026	0.0133	-1.3618	0.1938	0.0009	442
What Do You Know?	0.0707	0.0001	0.2062	-2.256	0.0923	-1.3618	0.1938	0.0123	371
Photo Quality Prediction	0.0168	0.0003	0.0444	-1.7916	0.0974	-1.3618	0.1938	0.0008	223
Benchmark Bond Trade Price Challenge	0.0653	0.0002	0.1899	-2.8435	0.0187	-1.3618	0.1938	0.003	456
KDD Cup 2012, Track 1	0.0873	0.0001	1	-1.6961	0.1663	-1.3618	0.1938	0.0177	1267
KDD Cup 2012, Track 2	0.1245	0.0003	1	-2.134	0.1743	-1.3618	0.1938	0.0005	864
Predicting a Biological Response	0.0382	0.0002	0.1824	-3.432	0.0259	-1.3618	0.1938	0.005	651
Online Product Sales	0.0552	0.0001	0.1202	-2.9822	0.0333	-1.3618	0.1938	0.0062	418
EMI Music Data Science Hackathon - July 21st - 24 hours	0	0.0001	0.023	-1.3017	0.1215	-1.3618	0.1938	0.0913	109
Belkin Energy Disaggregation Competition	0.0605	0.0002	0.2418	-1.7264	0.0767	-1.3618	0.1938	0.0021	607
Merck Molecular Activity Challenge	0.0337	0.0384	0.1222	-1.8376	0.1146	-1.3618	0.1938	0.0011	415
U.S. Census Return Rate Challenge	0.0599	0.0001	0.1435	-2.1342	0.0548	-1.3618	0.1938	0.0497	272
Amazon.com - Employee Access Challenge	0.0158	0.0001	0.1263	-2.5885	0.0863	-1.3618	0.1938	0.0008	755
The Marinexplore and Cornell University Whale Detection Challenge	0.1225	0.0004	0.236	-2.0254	0.0968	-1.3618	0.1938	0.0022	326
See Click Predict Fix - Hackathon	0	0.0001	0.0183	-2.3371	0.1248	-1.3618	0.1938	0.1398	262
KDD Cup 2013 - Author Disambiguation Challenge (Track 2)	0.0095	0.0001	0.1081	-1.7064	0.0846	-1.3618	0.1938	0.0005	623
Influencers in Social Networks	0	0.0001	0.04	-2.1596	0.0896	-1.3618	0.1938	0.0046	281
Personalize Expedia Hotel Searches - ICDM 2013	0.0213	0.0001	0.1246	-1.2842	0.1194	-1.3618	0.1938	0.0043	517
StumbleUpon Evergreen Classification Challenge	0.0857	0.0001	0.1523	-2.5777	0.096	-1.3618	0.1938	0.0496	328
Personalized Web Search Challenge	0.2166	0.0007	0.9134	-2.1926	0.0646	-1.3618	0.1938	0.002	275
See Click Predict Fix	0.0046	0.0001	0.1199	-1.5334	0.0946	-1.3618	0.1938	0.0002	575
Allstate Purchase Prediction Challenge	0.0555	0.0001	0.4519	-3.1441	0.0473	-1.3618	0.1938	0.0006	1204
Higgs Boson Machine Learning Challenge	0.0648	0.0001	0.6307	-3.3141	0.0994	-1.3618	0.1938	0.0042	1776
Acquire Valued Shoppers Challenge	0.0148	0.0003	0.4759	-2.3999	0.1707	-1.3618	0.1938	0.0045	2347
The Hunt for Prohibited Content	0.032	0.0001	0.2731	-2.7914	0.0649	-1.3618	0.1938	0.0043	966
Liberty Mutual Group - Fire Peril Loss Cost	0.0367	0.0001	0.2825	-2.075	0.1316	-1.3618	0.1938	0.0011	1057
Tradeshift Text Classification	0.0372	0.0005	0.197	-1.9588	0.0297	-1.3618	0.1938	0.0018	714
Driver Telematics Analysis	0.0735	0.0001	0.4571	-4.2049	0.1214	-1.3618	0.1938	0.0018	1619
Diabetic Retinopathy Detection	0.0889	0.0001	0.8012	-1.6264	0.1974	-1.3618	0.1938	0.0114	698
Click-Through Rate Prediction	0.0395	0.0001	0.4162	-3.2616	0.0314	-1.3618	0.1938	0.0012	1679
Otto Group Product Classification Challenge	0.0232	0.0001	0.187	-2.5233	0.0229	-1.3618	0.1938	0.0006	926
Crowdflower Search Results Relevance	0.0129	0.0001	0.2806	-2.8256	0.105	-1.3618	0.1938	0.0012	1645
Avito Context Ad Clicks	0.0612	0.0001	0.2814	-2.2102	0.0146	-1.3618	0.1938	0.0004	558
ICDM 2015: Drawbridge Cross-Device Connections	0.0684	0.0001	0.1687	-0.9735	0.192	-1.3618	0.1938	0.0051	364
Caterpillar Tube Pricing	0.0172	0.0001	0.3166	-3.0875	0.0277	-1.3618	0.1938	0.0004	1938
Liberty Mutual Group: Property Inspection Prediction	0.0233	0.0001	0.2667	-3.1092	0.06	-1.3618	0.1938	0.0009	1271
Coupon Purchase Prediction	0.1222	0.0003	0.3848	-2.0006	0.2093	-1.3618	0.1938	0.0027	631
Springleaf Marketing Response	0.0326	0.0001	0.332	-3.0862	0.0708	-1.3618	0.1938	0.0009	1567
Truly Native?	0.0526	0.0002	0.2073	-2.5242	0.0917	-1.3618	0.1938	0.0067	474
Rossmann Store Sales	0.0185	0.0001	0.3775	-3.3451	0.0452	-1.3618	0.1938	0.0015	1684
Homesite Quote Conversion	0.0223	0.0001	0.4559	-3.1196	0.0355	-1.3618	0.1938	0.0009	2557
Prudential Life Insurance Assessment	0.0749	0.0006	0.4219	-3.0026	0.0449	-1.3618	0.1938	0.002	818
BNP Paribas Cardif Claims Management	0.026	0.0001	0.3757	-3.2909	0.0187	-1.3618	0.1938	0.0013	1648
Home Depot Product Search Relevance	0.0147	0.0001	0.4918	-2.9788	0.0701	-1.3618	0.1938	0.002	2884
Santander Customer Satisfaction	0.0218	0.0001	0.3059	-2.2694	0.0479	-1.3618	0.1938	0.0021	1138
Expedia Hotel Recommendations	0.1011	0.0001	0.2814	-1.7786	0.0485	-1.3618	0.1938	0.0005	436
Avito Duplicate Ads Detection	0.0242	0.0001	0.3346	-2.4246	0.0583	-1.3618	0.1938	0.0006	1564
Draper Satellite Image Chronology	0.0857	0.0002	0.1188	-3.6358	0.0019	-1.3618	0.1938	0.0025	475

Notes: The table reports the GMM estimates with bootstrapped standard errors.

Table B.3: Probability of the hybrid prize structure being the optimal structure: Probit estimates

	hybrid (1)
σ (submission cost parameter)	-17.906**
	(7.713)
λ (frequency of submission opportunities)	6.135**
	(2.528)
β_0 (q_s function)	0.806**
	(0.407)
Observations	57
R^2	

Notes: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This probit model is for the probability that the hybrid prize structure is the optimal prize structure for a contest (with the alternative being that the optimal prize structure is the one with 6 prizes). The covariates are parameter estimates. An observation is a contest.

Table B.4: Uniform prize structure

2 prizes	30% of prize at $t = 0.5$, 70% of prize at $t = 1$
4 prizes	10%, 20%, 30%, and 40% of prize at times
	$t = .25$, $t = .5$, $t = .75$, and $t = 1$, respectively
6 prizes	5%, 10%, 15%, 20%, 25%, and 25% of prize at times
	$t = 1/6$, $t = 2/6$, $t = 3/6$, $t = 4/6$, $t = 5/6$, and $t = 1$, respectively
2 timed prizes	25% of prize at $t = 0.68$ and 75% of prize at $t = 1$
Benckmark	100% of the prize to the first player
	who surpasses the benchmark score 1.175
Hybrid	70% of prize at $t = 1$, 30% to the first player
	who surpasses the benchmark score 0.375

Notes: The table summarizes the optimal prize (within each class) subject to the constraint that all contests have the same prize structure.

Table B.5: Number of submissions in the last two days of the contests (i.e., days 10 and 11)

	# of submissions	
	(1)	(2)
Min. score at day 9	-9.835*	-11.819*
	(5.601)	(6.274)
2 prizes	-8.381**	-7.578*
	(4.063)	(4.095)
Hybrid/bechmark	-9.002**	-9.586**
	(4.124)	(4.217)
Controls	No	Yes
N	78	78
R^2	0.095	0.134

Notes: An observation is a contest. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Controls include all covariates reported in Table 4.

D Description of the Experiment

In this section, we reproduce the instructions given to all contest participants, regardless of the group they were randomly assigned to.

Description of the Competition

A large restaurant chain owns restaurants located along major highways. The average revenue of a restaurant located at distance x from the highway is $R(x)$. For simplicity, the variable distance to the highway is normalized to be in the interval [1,2]. The function $R(x)$ is unknown. The goal of this competition is to predict the value of $R(x)$ for several values of distances to the highway. Currently, the restaurant chain is located at 40 different locations. You will have access to

$$\{(x_i, R(x_i))\}_{i=1}^{30},$$

i.e., the distance to the highway and average revenue for 30 of these restaurants. Using these data, you must submit a prediction of average revenue for the remaining 10 restaurants, using their distances to the highway.

You will find the necessary datasets in the Data tab. You can send up to 10 different submission each day until the end of the competition. The deadline of the competition is Sunday September 26th at 23:59:59.

Evaluation

We will compare the actual revenue and the revenue predictions for each value

$$(x_j)_{j=31}^{40}.$$

The score will be calculated according to the Root Mean Square Deviation:

$$\text{RMSD} = \sqrt{\frac{\sum_{j=31}^{40}(\hat{R}(x_j) - R(x_j))^2}{10}},$$

which is a measure of the distance between your predictions and the actual values $R(x)$. The deadline of the competition is Sunday September 27th at 23:59:59.

Note. Following the convention used throughout the paper, we multiplied the $RSMD$ scores by minus one to make the winning score maximize private score in the competition.

Description of the Data

The goal of this competition is to predict the value of $R(x)$ for a number of values of distance to the highway. The csv file "train" contains data on the distance to the highway and average revenue for 30 restaurants

$$\{(x_i, R(x_i))\}_{i=1}^{30},$$

You can use these data to create predictions of average revenue for the remaining 10 restaurants. For these 10 restaurants you only observe their distances to the highway in the csv file "test." You can find an example of how your submission must look like in the csv file "sample_submission."

File descriptions:

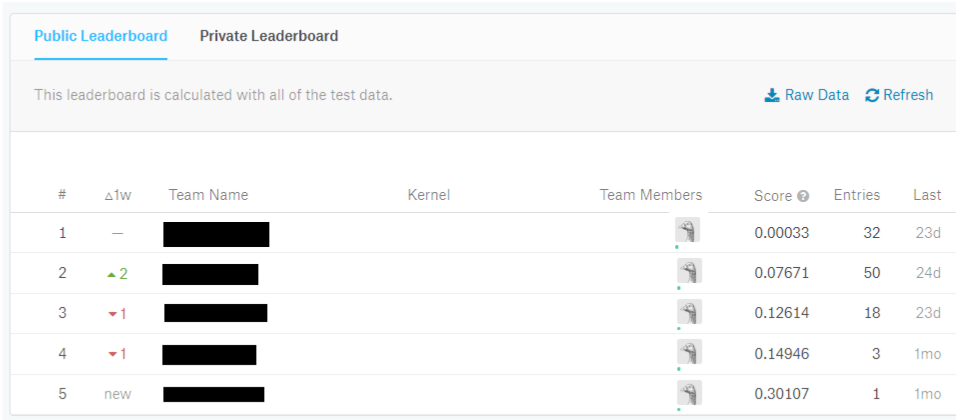
- **train.csv** - the training set
- **test.csv** - the test set
- **sample_submission.csv**- an example of a submission file in the correct format

Submission File:
The submission file must be in csv format. For every distance to the highway of the 10 restaurants, your submission files should contain two columns: distance to the highway (x) and average revenue prediction (R). The file should contain a header and have the following format:

x	R
1.047579	34.43375
1.926801	36.83077
etc.	

A correct submission must be a csv file with one row of headers and 10 rows of numerical data, as displayed above. To ensure that you are uploading your predictions in the correct format, we recommend that you upload your predictions by editing the sample submission file. There is a limit of 10 submissions per day.

Figure C.1 shows a screenshot of the leaderboard in one of our student competitions hosted on Kaggle.



Public Leaderboard Private Leaderboard

This leaderboard is calculated with all of the test data.

Raw Data Refresh

#	△1w	Team Name	Kernel	Team Members	Score	Entries	Last
1	—	[redacted]			0.00033	32	23d
2	▲2	[redacted]			0.07671	50	24d
3	▼1	[redacted]			0.12614	18	23d
4	▼1	[redacted]			0.14946	3	1mo
5	new	[redacted]			0.30107	1	1mo

Figure C.1: Snapshot of the leaderboard in one of our competitions with a leaderboard. Names are hidden for privacy reasons.