Contingent Prizes in Dynamic Contests*

Jorge Lemus[†] Guillermo Marshall[‡]

March 1, 2024

Abstract

Firms and government agencies increasingly use online contests to find solutions to problems. These contests allow players to make multiple attempts, and their progress is tracked on a real-time public leaderboard. Although the contests are *dynamic*, contest sponsors miss the opportunity to dynamically shape players' incentives, as they award the entire prize pool based only on players' final rankings. We propose the use of *contingent prizes* that depend on the game's history to incentivize players throughout the contest. We present evidence from two methodologies—a structural model and an RCT—showing 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, IIOC (2022), CMiD (Singapore), UBC (Sauder), University of Reading, 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 using online contests to procure solutions to problems. In particular, platforms hosting online data-science competitions have gained popularity over the last decade.[1] These competitions are *dynamic*: players try many attempts before a pre-determined deadline, and a public leaderboard discloses their performance in real time. Notably, these competitions only reward players based on their final rankings, which we argue is a missed opportunity to shape incentives during the contest and improve participants' performance. We study *contingent* prizes—pre-announced prize allocation rules based on a contest's *history*—which give the contest designer additional tools for motivating players throughout the contest.[2] Specifically, we ask, should *interim leaders* in a contest receive prizes? Should the contest sponsor set *milestones*? How much money is left on the table by only awarding prizes based on final rankings?

Our contribution is to empirically investigate the impact of contingent prizes on performance, measured as the best submission in the contest. Our results suggest that contest sponsors are wasting resources by awarding prizes based only on final rankings, overlooking the role of contingent prizes in shaping players' incentives. We quantify the gains from using simple contingent prizes—e.g., splitting the budget to reward interim leaders or rewarding the first player who surpasses a milestone—and show that they can achieve the same performance as a prize based only on final rankings, using merely half of the budget.

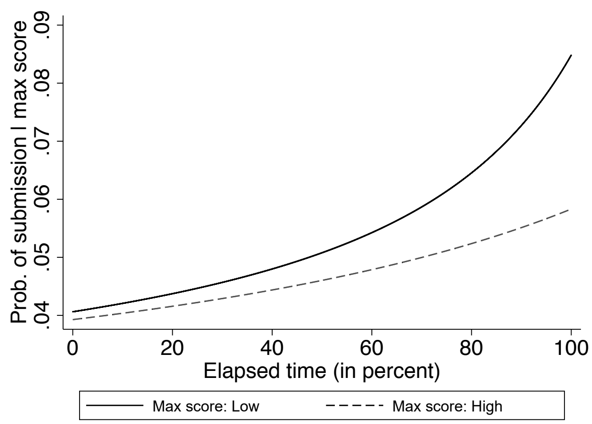
Contingent prizes can dynamically motivate players to make costly submissions by managing two salient economic forces influencing players' incentives during the competition. First, it may be unlikely to become the competition's leader at the current time after making a submission (the "current-competition" effect). Second, even if a player becomes the competition leader at the current time, she anticipates that future submissions by her rivals will threaten her lead (the "future-competition" effect). The current-competition effect is most discouraging near the end of the competition, when most opportunities to improve scores have been exhausted. In contrast, the future-competition effect is most discouraging at the beginning of the competition, when many rival submissions are yet to come.

[1] Many platforms allow firms or government agencies to sponsor online competitions, including Kaggle, DrivenData, CodaLab, InnoCentive, AIcrowd, and Topcoder. As of October 2023, Kaggle has over 15 million registered users, has hosted over 5,000 competitions, and paid millions of dollars in prizes to participants in competitions sponsored by Google, American Express, Expedia, General Eletric, Intel, and Airbus.

[2] Some dynamic contests award prizes contingent on milestones or to interim leaders. For instance, the XPRIZE Carbon Removal competition, the largest incentive prize in history, splits a $100 million budget over time: After one year of competition, up to 15 competitors will receive 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.

[Figure 1](#) illustrates these forces in a contest with a prize structure that only rewards the competition leader based on the final ranking (henceforth, a "final-ranking" prize). The figure shows the probability of making a submission at different times, while keeping a constant maximum score throughout the competition. The solid (dashed) line corresponds to the probability of a submission when the maximum score is fixed at a low (high) score. The solid line lies above the dashed line because the probability of a submission at any given time is higher when the current maximum score is lower. This difference captures the discouragement effect of the maximum score, i.e., the current-competition effect. Moving along either curve, the probability of a submission increases over time; both curves slope upwards. Part of this effect is explained by future competition, which discourages players from making submissions early 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).

Altering the importance of current- and future-competition effects throughout the contest affects the equilibrium number of submissions. The decision to make a costly submission hinges on the relative importance of the current- and future-competition effects. For instance, let us compare a final-ranking prize with an alternative time-contingent prize structure, one that allocates 50% of the budget to the interim leader in the first half of the competition and the remaining 50% to the leader based on final rankings. The future-competition effect induces a strong discouragement earlier in the contest with a final-ranking prize but milder discouragement with a time-contingent prize structure. With a final-ranking prize, the interim leader at half of the competition must endure the future competition for the remaining half of the contest before receiving a prize. Instead, with a time-contingent prize, the interim leader at the halfway point receives 50% of the budget regardless of the number of submissions in the second half of the contest. However, because players have more incentives to play early on with a time-contingent prize, the current-competition effect will be stronger than with a final-ranking prize. In addition to balancing the relative impact of current- and future-competition effects through the prize structure, the contest designer must also consider the budget constraint. Allocating a large prize early in the contest limits the maximum reward available for the remainder of the competition.

We combine two empirical methodologies to investigate the impact of contingent prizes on performance: structural estimation and experimental evidence. In the first part of our analysis, we estimate a structural model with observational data from Kaggle.com, the largest platform for hosting data science competitions. In these competitions, players can make multiple submissions that are scored based on an objective criterion (e.g., prediction accuracy). A public leaderboard displays these scores in real time, and the leader, determined by the final ranking, receives a prize. Thus, these data allow us to estimate the structural parameters of a dynamic contest where a prize is allocated based on final rankings only. Using these estimates, we simulate the equilibrium of each contest under counterfactual prize structures.

Our analysis focuses on *simple* prize structures, which include rewards for interim leaders at predetermined times, the first player to achieve a milestone, or the first player to achieve a milestone before a predetermined time. In particular, a final-ranking prize *is* a simple prize structure, as it only rewards the leader based on rankings at one specific time (at the end of the contest). Simple prize structures are both computationally manageable and relatively easy to implement in practice. Finding optimal, fully flexible, budget-constrained, contingent-prize structures, however, requires solving an optimization problem with millions of variables and an exponentially large number of constraints, making it computationally unfeasible.[3] Our decision to focus on simple prize structures is motivated by their tractability and evidence showing that they can approximate the benefits of fully flexible contingent prize structures. Specifically, we use the estimates of our structural model to simulate "short contests" and find that simple prize structures capture a large fraction of the performance of the optimal contingent prize structure in short contests. Armed with these findings, we evaluate simple prize structures using the sample of contests in our data (i.e., the full-scale contests). In this exercise, we compute the equilibrium of each contest under counterfactual simple prize structures and compare it to the observed equilibrium of each contest.

More specifically, using our model estimates, for each contest, we consider seven classes of counterfactual designs, each one corresponding to a "simple" prize structure. The first

[3]Specifically, there are $T(T+1)$ variables to optimize over and 2^T constraints (histories), where T is the number of periods in the empirical model, set at 10,000 in our baseline estimate.

three classes of designs ("time-contingent" prizes) consider k prizes to the interim leaders at k equally spaced times, with the size of each of the k prizes chosen optimally and $k \in \{2, 4, 6\}$. The fourth design ("2 timed prizes") allocates one prize of size π to the leader of the competition at time τ, and the rest of the budget to the leader based on final rankings. The fifth design ("milestone") awards the full prize pool to the first player who surpasses a milestone score. The sixth design ("hybrid") awards a prize to the first player who surpasses a milestone score and another prize to the leader based on final rankings. The last design ("elimination") eliminates a number of players at the middle of the contest. For each contest, we find the prize structure within each class that maximize the expected maximum score given the contest's primitives. For time-contingent prizes, we optimize over the size of the prizes, π_k. For 2 timed prizes, we optimize over both π and τ. For milestones, we find the optimal milestone score. For hybrid, we optimize over both the milestone score and the final prize. Lastly, for elimination contests, we find the optimal number of players to eliminate.

Finding optimal, simple prize structures for each contest rely on the designer knowing the contest's primitives, such as the contest's difficulty, the players' submission cost distribution, and the distribution of scores for a given submission. However, a contest designer may not know these primitives. To address this issue, for each class of prize structure, we find parameters that maximize the average maximum score across all of the contests in our data. We refer to this set of parameters as "uniform," meaning that they are not tailored to one specific contest but rather optimized for best performance on average. For example, the uniform parameters for the class of two time-contingent prizes, equally spaced over time, correspond to allocating 30% of the budget in the middle of the contest and 70% at the end. Meanwhile, the uniform parameters for the class of 2 timed prizes correspond to allocating 25% of the budget when 68% of the contest time has elapsed and the remaining 75% of the budget at the end of the contest.

Our estimates show that uniform parameters can achieve a significant portion of the benefits obtained by customizing parameters for each contest. In other words, even when contest primitives are heterogeneous, using uniform parameters can substantially enhance contest outcomes. Furthermore, we found that by utilizing only 52 percent of the total budget, an optimal, simple prize structure can achieve the same performance, on average, as a final-ranking prize. This result suggests that contest sponsors, who have awarded millions of dollars in prizes, might be missing out on potential gains by employing a sub-optimal prize structure.

In the second part of our analysis, we complement our model-based analysis with a randomized control experiment, serving a dual purpose. First, it provides evidence of improved contest outcomes when using contingent prizes. This experimental finding does not hinge on our modelling choices. Second, it allows us to test whether particular contingent-prize structures improve contest outcomes in line with our model predictions. In our experiment, we organized competitions for students from the University of British Columbia and the University of Illinois. The students competed in groups of up to five in a prediction competition on Kaggle.com. We randomly assigned each group to one of three conditions: (1) the leader at the end of the competition received the full prize pool (baseline); (2) the leader two days before the end of the competition received 30% of the prize pool, and the leader at the end of the competition received the remaining 70% of the prize pool, (time-contingent prizes); (3) the first player to surpass a milestone score received 30% of the prize pool, and the leader at the end of the competition received the remaining 70% of the prize pool (hybrid prizes). We designed these experimental conditions on optimal designs derived from our structural estimation.

Our model-based and experimental evidence align in showing that contingent prizes can significantly improve contest outcomes relative to a prize based on final standings. According to our model estimates, a hybrid design—providing one prize to the first player to reach a milestone score and another prize to the leader at the end of the competition—produces the best outcomes among seven counterfactual simple prize structures. Furthermore, our experimental results show that the hybrid prize structure generates similar gains in contest outcomes but with larger magnitudes than our empirical model predicts.

A contingent prize can encourage players early on by reducing the future competition effect but can discourage them by increasing the current competition effect. Our results show that simple contingent prizes, which are easy to implement in practice, resolve this tradeoff favorably for the designer and achieve better performance than a prize based on final rankings.

A contingent prize can motivate players in the early stages of the competition by reducing the future competition effect, yet it may discourage them by intensifying the current competition effect. Our results show that simple contingent prizes, which are easy to implement in practice, resolve this tradeoff better than a prize based on final rankings only. Thus, contest designers should carefully consider the use of contingent prizes if their goal is to achieve the best performance.

Related Literature. Our work contributes to the recent empirical literature on dynamic contest design. For instance, Bhattacharya (2021) studies alternative designs for the U.S. Department of Defense's multi-stage contests. Gross (2017) and Lemus and Marshall (2021) study the impact of dynamic feedback on outcomes. Our work focuses on one aspect of contest design, *contingent* prizes, fixing other design dimensions, including the number of players, contest length, evaluation rule, and information provision.

In contrast to most of the literature, we employ two complementary methodologies: structural estimation using observational data from large contests and a randomized control trial that provides model-free estimates. The theoretical literature has investigated several elements of contest design, including the number of participants that should enter a contest (Taylor, 1995; Fu and Lu, 2010; Aycinena and Rentschler, 2019), or earn a prize (Moldovanu and Sela, 2001; Olszewski and Siegel, 2020; Kireyev, 2020). In dynamic settings, researchers have studied whether contests should be divided into multiple stages (Moldovanu and Sela, 2006; Sheremeta, 2011) and whether participants should receive performance feedback during the contest (Mihm and Schlapp, 2019). Moldovanu and Sela (2006), Fu and Lu (2009), Fu and Lu (2012), Chowdhury and Kim (2017), and Clark and Nilssen (2020), among others, study the design of both the number of stages and prizes.

As a dynamic competition unfolds, players can experience discouragement (see, e.g., Harris and Vickers, 1987; Konrad and Kovenock, 2009). In multi-stage contests, Feng and Lu (2018) theoretically show that optimal time-contingent prizes depend on how effort impacts the probability of winning. Similarly, Alshech and Sela (2021) find the optimal rank-and-time-contingent prize structure, where designer chooses a prize for the player who wins both stages, in addition to a prize for the leader of each stage. Prize allocations have also been studied theoretically and experimentally by Stracke et al. (2014) (under risk aversion), Cason et al. (2020) (under noisy performance), and Güth et al. (2016) (under milestone prizes). In a two-period model, Klein and Schmutzler (2021) show that the symmetric equilibrium with a single prize based on final rankings achieves higher total effort than two equal prizes at the end of each stage. They confirm this prediction in a laboratory experiment. Our model allows for multiple periods and asymmetric players (stochastic submission costs).

Other settings have theoretically examined score-contingent prizes. For instance, in Ely et al. (2021), the designer observes the players' progress and chooses 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 an exogenous milestone score (i.e., not set by the designer). In Benkert and Letina (2020) players privately observe their progress and choose when to report it to the principal. The optimal prize structure makes interim transfers to all players while the competition has not ended and rewards the first player who reveals success upon which the contest ends. Unlike these papers, we empirically evaluate contingent prizes in contests with a fixed duration and a public leaderboard.

2 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 an average prize of $30,489. A partial list of competition characteristics is 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 training 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 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 objectively scores each submission based on its out-of-sample performance and posts its score on a public leaderboard. Prizes are awarded to top players based on the leaderboard ranking at the end of the contest.[6]

We have contest-level information on all submissions, 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 time for every contest. Using the same approach as in Lemus and Marshall (2021), we standardize the score distribution to have a mean of zero and a standard deviation of one.

3 Empirical Model

There are N contest participants. Player i enters the contest at an exogenously given time $0 \leq t_i \leq T$, at which point the player is able to make submissions. Players have perfect foresight regarding the entry times of rivals and stay until the end of the contest. At time

[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.

[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, so the public score is a noisy signal of performance. The correlation between public and private scores is 0.99, which motivates us to abstract away from the noise in the public signals for tractability.

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 A.1 in the Online Appendix for the complete list of competitions.

t, N_t players have already entered. Making a submission when the current maximum score is s increases the maximum score to $s' = s + \varepsilon$ with probability q_s and leaves the maximum score at s with probability $1 - q_s$. It becomes increasingly difficult to increase the maximum score as the current maximum score increases, i.e., q_s decreases in s.[7] As a consequence, players know that submitting today can deter future submissions. 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 undiscounted.

Players publicly observe and keep track of four state variables: the current period (t), the maximum score in the previous period (s_P), the current maximum score (s), and the identity of the current leader (ℓ). At period $t = 0, ..., T-1$, nature selects with probability $\lambda \in (0, 1)$ one randomly selected player, and with probability $1-\lambda$ no one is selected. The chosen player is the only one who can choose to submit ("play") in period t. Submissions are instantaneous and cost c, where c is a random variable distributed according to $K(\cdot)$ and i.i.d. across players and time.[8] Players observe their cost realization *before* choosing whether to play. The state space is

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

[7]The assumption of small increments of size ε is motivated by the data.

[8]Non-pecuniary benefits are implicitly accounted for. Let $(\text{Prize} + B)\Pr(\text{win}) + W > \hat{c}$, where B and W are non-pecuniary benefits from winning and submitting, and $\hat{c}$ is the actual submission cost. Then, $c = \frac{\hat{c}-W}{\text{Prize}+B}$ is the submission cost net of non-pecuniary benefits. We cannot separately identify W, B and $\hat{c}$.

The contest awards **contingent prizes**. For tractability, we restrict to "Markovian" contingent prizes rather than prizes that depend on the whole history of the contest. That is, the leader of the competition in period t, receives a prize $\pi(s,t|s_P)$, where s_P is the previous period's maximum score, and s is the current period's maximum score. Contingent prize structures allow for different prizes for reaching the same maximum score in the current period depending on the maximum score in the previous period. That is, we can have $\pi(s,t|s_P) \neq \pi(s,t|s'_P)$, when $s_P \neq s'_P$.

At time t, a player is either the leader or one of the $N_t - 1$ followers. Let L_{t,s,s_P} be the value of being the leader at time t when the current maximum score is s and the previous maximum score is s_P. Let $F_{t,s}$ be the value of being one of the followers. At time T, when the competition ends, and the maximum score is s_T, the leader gets $\pi(s_T, T|s_{T-1})$, and the followers get 0. We denote $t' = t + \delta$ and $s' = s + \varepsilon$.

The expected payoff of a leader with score s at $t = 0, 1, ..., T-1$ is

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

That is, when the maximum score in the previous and current periods are s_p and s, the leader at time t receives the prize $\pi(s,t|s_P)$. There are three continuation payoffs. 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, the score continues to be s, and the leader receives the continuation payoff $L_{t',s,s}$. Second, with probability $\frac{\lambda}{N}$ the current leader can choose to play, in which case she receives the continuation payoff $L_{t,s}^{\text{own play}}$, defined in equation (2). Third, with probability $\frac{\lambda(N_t-1)}{N}$ one of the $N_t - 1$ followers can choose to play, in which case the current leader receives the continuation payoff $L_{t,s}^{\text{rival play}}$, defined in equation (5).[9]

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

$$L_{t,s}^{\text{own play}} = E_c \left[\max\{q_s L_{t',s',s} + (1 - q_s) L_{t',s,s} - c, L_{t',s,s}\} \right], \tag{2}$$

If selected by nature, the leader 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.,

$$q_s(L_{t',s',s} - L_{t',s,s}) \geq c. \tag{3}$$

[9]Nature selects players with a probability that is independent of the number of players that have entered at time t, N_t. That is, nature selects each player with probability $1/N$, where N is the time-independent number of players that will participate in the competition. Empirically, N_t converges quickly to N.

From this condition, the probability that the leader plays is

$$p_{t,s}^{L} = K(q_s(L_{t',s',s} - L_{t',s,s})), \tag{4}$$

where $K(\cdot)$ is the distribution function of the submission cost. The last term in equation (1), $L_{t,s}^{\text{rival play}}$, 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

$$L_{t,s}^{\text{rival play}} = p_{t,s}^{F}\left(q_s F_{t',s'} + (1-q_s)L_{t',s,s}\right) + (1-p_{t,s}^{F})L_{t',s,s}. \tag{5}$$

If a follower is selected by nature, she plays 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 plays but fails to replace the leader, so the current leader remains the leader, obtaining $L_{t',s,s}$. With probability $1-p_{t,s}^{F}$, the follower chooses not to play, so the current leader remains the leader, obtaining $L_{t',s,s}$. The probability $p_{t,s}^{F}$ is an equilibrium object that we derive later on.

We next specify the value of being a follower, $F_{t,s}$, which depends on the current score and time but not on the previous maximum score. 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 is

$$F_{t,s} = \left(1 - \lambda\frac{N_t}{N}\right)F_{t',s} + \frac{\lambda}{N}F_{t,s}^{\text{leader play}} + \frac{\lambda}{N}F_{t,s}^{j\text{ play}} + \frac{\lambda(N_t-2)}{N}F_{t,s}^{\text{follower play}} \tag{6}$$

Followers do not receive prizes. In our model, prizes for followers are suboptimal because there is no entry margin, and all followers have the same probability of becoming the leader. 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 followers and receive the continuation payoff $F_{t',s}$. Second, nature selects the leader with probability $\frac{\lambda}{N}$, in which case followers receive the continuation payoff $F_{t,s}^{\text{leader play}}$. Third, nature selects follower j with probability $\frac{\lambda}{N}$, in which case she receives the continuation payoff $F_{t,s}^{j\text{ play}}$. Fourth, nature selects one of the other $N_t - 2$ followers (not j) with probability $\frac{\lambda(N_t-2)}{N}$, 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

$$F_{t,s}^{\text{leader play}} = p_{t,s}^{L}\left(q_s F_{t',s'} + (1-q_s)F_{t',s}\right) + (1-p_{t,s}^{L})F_{t',s} \tag{7}$$

When nature selects the leader, she plays and increases the maximum score with probability $p^L_{t,s} q_s$, in which case the followers receive $F_{t',s'}$. With probability $p^L_{t,s}(1-q_s)$, the leader plays but fails to increase the maximum score, so followers receive $F_{t',s}$. With probability $1-p^L_{t,s}$, the leader chooses not to play, so the followers receive $F_{t',s}$.

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

$$F^{j\text{ play}}_{t,s} = E_c \left[\max\{ q_s L_{t',s',s} + (1-q_s) F_{t',s} - c, F_{t',s} \} \right]. \tag{8}$$

If selected by nature, follower j chooses between playing or not after observing the cost of making a submission, c, playing if and only if

$$q_s (L_{t',s',s} - F_{t',s}) \geq c. \tag{9}$$

The condition above means that the expected marginal gain from becoming the leader must be sufficiently larger than the cost. From here, the probability that a follower makes a submission is

$$p^F_{t,s} = K \left(q_s (L_{t',s',s} - F_{t',s}) \right). \tag{10}$$

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

$$F^{\text{follower play}}_{t,s} = p^F_{t,s} \left(q_s F_{t',s'} + (1-q_s) F_{t',s} \right) + (1-p^F_{t,s}) F_{t',s} \tag{11}$$

When nature selects a follower other than player j, 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 $F_{t',s}$.

Key Driving Forces. Conditional on the current score and time, leaders and followers have different incentives to play. For the leader, the expected marginal benefit of a successful play is $L_{t',s',s} - L_{t',s,s}$, whereas, for a follower, the benefit is $L_{t',s',s} - F_{t',s}$. Let's first consider the leader's incentive to make a costly submission. By increasing the score, the leader collects a prize $\pi(t', s'|s)$ instead of $\pi(t', s, |s)$. This effect could be positive, negative, or zero, depending on the contingent prize structure; if it is positive, the prospect of a higher prize motivates the leader to play. Now, consider a follower's incentive to make a costly submission. By increasing the score, the follower collects a prize $\pi(t', s'|s)$ instead of 0 (a weakly positive effect) and becomes the leader. Both the leader and the followers also anticipate that a higher maximum has a deterrence effect on future submissions because $q_{s'} < q_s$. As time goes on, however, the future competition effect softens, encouraging submissions.

Objective of the Contest Designer. The contest designer's goal is to choose the prize structure that maximizes the expected maximum score at the end of the contest.[10] The set of possible histories in our game is

$$\mathcal{H} = \{(a_1, ..., a_T) \,:\, a_t \in \{0, \varepsilon\}\},$$

with $|\mathcal{H}| = 2^T$. In history $h = (a_1^h, ..., a_T^h) \in H$, a_t^h indicates the score change at time t, and the maximum score at time t is $s_t^h = \sum_{j=1}^{t} a_j^h$. Markovian contingent prizes depend on the maximum score at $t-1$, the maximum score at t, and the current time, t. Thus, they are determined by $T(T+1)$ variables, $\{\pi(s', t|s)\}$ for $t = 1, ..., T$, $0 \leq s \leq T\varepsilon$ and $s' \in \{s, s+\varepsilon\}$. A *feasible* contingent prize structure must satisfy that the sum of prizes in history $h \in H$ is less than or equal to the budget (normalized to 1). These are 2^T constraints, one for each history $h \in H$, with

$$\pi(s_1^h, 1) + \pi(s_2^h, 2|s_1^h) + ... + \pi\left(s_t^h, t|s_{t-1}^h\right) + ... + \pi\left(s_T^h, T|s_{T-1}^h\right) \leq 1 \qquad \text{(BC-}h\text{)}$$

To maximize the expected maximum score at the end of the contest, the designer solves

$$\max_{\{\pi(\cdot)\}} \int s_T^h \, dF(h|\pi(\cdot)) \text{ subject to (BC-}h\text{) for all } h \in H. \qquad (12)$$

Contingent prize structures change the balance of different histories by dynamically manipulating competitors' incentives to make costly submissions throughout the competition.

3.1 Computational Burden

Solving problem (12) is computationally demanding. It requires choosing $T(T+1)$ variables subject to 2^T budget constraints. For small instances, we can find the optimal solution of (12) by "brute force," evaluating all feasible combinations. In fact, in Section 5.1, we use our structural model estimates to simulate "short contests" instead of the actual longer duration of the contests. There, we compare optimal structure with *simple* prize structures. We show that optimal simple prize structures capture a large portion of the fully optimal prize structures in these short contests.

[10] Implicitly, we assume a risk-neutral contest designer with monotone preferences for the maximum score.

Below, we present two additional methods to assess the value of using an optimal prize structure. First, we provide an upper bound on the gains from employing one. Second, we solve an unconstrained problem (removing 2^T constraints) at the expense of requiring us to take a stance on the contest designer's value of achieving different scores.

Upper Bound on Gains. Even though we cannot explicitly compute the optimal prize structure when T is large, we can find an upper bound for the value of using it. To this end, we use duality theory to provide an upper bound for the expected maximum score. Specifically, for any vector $\mu = (\mu_h)_{h \in H}$, with $\mu_h \geq 0$, an upper bound to (12) is

$$\sup_{\{\pi(\cdot)\}} \left\{ \int s_T^h \, dF(h|\pi(\cdot)) - \sum_{h \in H} \mu_h \left(\sum_{t=1}^{T} \pi(s_t^h, t | s_{t-1}^h) - 1 \right) \right\}. \tag{13}$$

The smaller upper bound is found by minimizing (13) over $(\mu_h)_{h \in H}$, which is again computationally unfeasible because it entails solving an unconstrained problem with 2^T variables. However, fixing any vector of positive numbers $(\mu_h)_{h \in H}$ and solving (13) over $T(T+1)$ variables, which is computationally feasible, we obtain an upper bound. This is useful because it informs the designer about the performance of simple prize structures relative to the optimal contingent prize structure. If contest outcomes with a simple prize structure are not too far from the upper bound, then simple prize structures (feasible to compute) capture a large portion of the gains from using the optimal prize structure (unfeasible to compute).

Finding the Optimal Budget. The main computational burden of finding the optimal solution is the large number of constraints. Instead of solving the constrained problem for a fixed budget, we could solve an alternative unconstrained problem. Let the designer's value of obtaining a maximum score s be $V(s)$. Consider the relaxed problem, where the designer chooses the prize structure that solves

$$\max_{\{\pi(\cdot)\}} \int V(s_T^h) \, dF(h|\pi(\cdot)) - \int \sum_{t=1}^{T} \pi(s_t^h, t | s_{t-1}^h) \, dF(h|\pi(\cdot)) \tag{14}$$

While this unconstrained problem is computationally tractable ($T(T+1)$ variables), it requires us to take a stand on the shape of $V(\cdot)$.

To show that the use of flexible prize structures—rather than a final-ranking prize—is optimal for different designer's preferences, we solve problem (14) for different contests using two functional forms of $V(\cdot)$: $V(x) = x^2$ and $V(x) = \sqrt{x}$.[11] We choose these two functional forms to see whether a final-ranking prize is optimal for a convex function (representing that the

[11] The simulated contests make use of the same parameters estimated for the main model (Table 3).

designer is “risk loving” on the maximum score) or for a concave function (representing that the designer is “risk averse” on the maximum score). We find that flexible prize structures dominate a final-ranking prize in both cases. Hence, contingent prize structures can improve outcomes even if the designer is not budget-constrained but dislikes paying higher prizes.

3.2 Simple Prize Structures

To address the challenge of high dimensionality in finding optimal prize structures, we investigate different classes of *simple* prize structures. In these classes, prizes are zero for most states and positive for a small subset of states, reducing the number of variables and constraints. Simple prize structures are appealing not only for being computationally tractable but also for being easy to implement in practice.[12] We focus on several classes of simple prize structures, including time-contingent, score-contingent, hybrid, and elimination.

Time-contingent prizes. The prize structure is time-contingent if the interim leader at time t receives a prize regardless of the current or the previous maximum score. That is, $\pi(s_t, t|s_{t-1}) = \pi_t$ for all $s_t \in \{s_{t-1}, s_{t-1} + \varepsilon\}$ and $s_{t-1} \in \{0, \varepsilon, ..., T\varepsilon\}$.

Score-contingent prizes (milestones). The prize structure is score-contingent if the first player to reach a milestone receives a prize regardless of when the milestone was reached. That is, $\pi(s_t, t|s_{t-1}) = 0$ for $s_t = s_{t-1}$ and $\pi(s_t, t|s_{t-1}) \geq 0$ for $s_t = s_{t-1} + \varepsilon$. Furthermore, for any $t' \neq t$ and $s' = s + \varepsilon$, $\pi(s', t|s) = \pi(s', t'|s)$.

Hybrid prizes. The hybrid structure combines milestones and time-contingent prizes. Here, the time at which a milestone is reached matters, and there can be a final-ranking prize even if the milestones are not reached. Here, $\pi(s', t|s) \geq 0$ when s belongs to some predetermined set of scores and $s' = s + \varepsilon$, or $t = T$. Otherwise, $\pi(s', t|s) = 0$.

3.3 Discussion of Modeling 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.

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

[12] In regulation, simple menus can capture large gains of optimal menus (Rogerson, 2003) .

benefit from learning from their performance on earlier submissions. Our model accommodates a specific form of learning: each player has an equal probability of increasing the maximum score, which decreases as the maximum score increases. In other words, the function $q(s)$ captures that all players learn in the same manner when the score increases, as playing gives them an equal chance of increasing the maximum score.

A different way of modelling learning would be to make players' performance (or cost) depend on the number of past submissions. Under this form of learning, prize structures that encourage early participation would allow the designer to benefit from enhanced performance or lower costs due to learning. However, this approach significantly increases the size of the state space, enlarging it by m^N, where m is the number of "types" associated with different levels of learning, and N is the number of players. If, for instance, learning gives rise to two types, our state space would be $2^{10} = 1,024$ times larger. We refrain from such an approach for tractability.

2. *Incentives to withhold submissions.* A player could be concerned about increasing the maximum score because it may inform rivals that are "stuck" that "something else" is possible, encouraging them to exert effort. In such cases, players could withhold their submissions and send them near the end of the competition to prevent encouraging their rivals. We argue that at least three facts alleviate this concern. 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 same sample of contests as in this article and do not find empirical evidence suggesting strategic withholding.

3. *Leader, followers, and prizes.* At each instant during the contest, Kaggle's leaderboard displays the best score for each player. Furthermore, multiple players receive prizes at the end of the contest based on the final ranking (usually three prizes). In our model, there is no notion of "being close" to the leader, as each follower is symmetric. We also assume that only the leader receives a prize. These assumptions are for convenience, reducing the dimension 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 and N is the number of players. Also, the number of variables for contingent prizes would increase by m, where m is the number of "places" that receive an award. Qualitatively, the optimal contingent prize structures in this augmented model should be similar to what we find. In our setting, the incentive to play is driven by the

prize $\pi(s', t|s)$ for becoming the leader.[13] With multiple prizes, the players' decisions are more involved as they assess the probability of ending the contest in a different ranking, which determines their expected reward. With m prizes, players' incentives are driven by the expected prize $\sum_{j=1}^{m} \pi(s, t, j|s') \times \Pr(\text{rank } j \text{ at time } t)$. We speculate that our results would not dramatically change under this modification.

4. *Intensive Margin.* Effort in our model is a binary choice. We refrain from modelling an intensive margin for three reasons. First, effort is not directly observable in our data. Therefore, we would need to model effort based on, for example, time between submissions, which is a noisy measure. Second, scores evolve smoothly over time, with only a few submissions increasing the maximum score by a relatively large amount.[14] Third, if there was an intensive effort choice, players would choose more effort towards the end of the competition, when the marginal return to become the competition leader is higher. This effect is already captured by in our setting by the increased number of submissions toward the end of the competition.

5. *Player heterogeneity.* Our model assumes players are homogeneous in ability. In our estimation, we focus on top performers, who have arguably more similar abilities than two randomly selected players in the contest, which alleviates heterogeneity concerns. A simple test reveals that the top 10 players are fairly homogeneous. Specifically, less than 20 percent of the variance of scores can be explained by competition and team-fixed effects. Moreover, most of the coefficients on players' fixed effects are statistically insignificant. These facts combined suggest that most of the differences in performance among this group of players are random.

6. *Open or restricted entry.* Kaggle contests are typically open contests (anyone can participate). We do not model entry or exit. Lemus and Marshall (2021) discuss at length the exogenous entry assumption. Here, we analyze the strategic behavior of 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.

7. *Length and number of stages.* Most Kaggle contests do not allocate intermediate prizes, nor are they split into separate "stages." We normalize the length of the contests to facilitate meaningful comparisons across them. Dividing the contest into multiple stages

[13] While the leader and followers have different incentives, leapfrogging is common: conditional on the maximum score changing, a follower becomes the leader with a probability greater than 80 percent.

[14] It is plausible that most of the effort is done before the first submission, and the majority of players' subsequent submissions comprise "tweaking" their initial algorithms. In that case, each "tweak" would likely require a similar amount of effort.

can reduce discouragement by bringing players "closer together" at the beginning of each stage. In our setting, there is no notion of being "close" to the leader because every player who is not a leader is a follower. Furthermore, if a contest were divided into stages, and there is no information exchanged among players, they could always resubmit their last solution to reset the latest leaderboard at the beginning of each new stage. Hence, in our setting, splitting the contest into multiple stages is unfeasible.

4 Estimation

Given that only the state is payoff relevant, we use the Markov-perfect equilibrium concept. Computationally, we find it using backward 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 $[0, 1]$.[15] 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 calibrate 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 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. This expression gives a measure of the fraction of periods in which a player can play, which is what we intend to capture with λ.

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

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

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

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 changes infrequently, 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.

In the second step, we estimate the parameter σ of the cost distribution, $K(c;\sigma) = c^{\sigma}$, where $\sigma > 0$. We leverage revealed preference for identification: if σ is larger (smaller), participants find making submissions costlier (cheaper), resulting in fewer (more) submissions predicted by the model. In the estimation, we search for the value of σ that rationalizes the observed number of submissions.

In practice, we use 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 into 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 $S^2 \times T$ (S is the size of the set of possible scores and T is the number of periods) where element (s, t, s^P) of $\mathbf{p}^j$ is p^j_{s,t,s^P}.[16] Using the CCPs, we can also compute the equilibrium distribution of maximum scores at every period of time, $\mathbf{G}$ (of dimensions $S^2 \times T$), where column t gives the distribution of maximum scores at time t. 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 a summation of the product over the first dimension, gives us a $1 \times T$ vector with the expected probabilities of a submission at every instant of time, which we use to compute the moments.

[16]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.

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	0.0020	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 A.2 in the Online Appendix for the complete list of competitions.

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) we suspect that this group of players is less heterogeneous than the entire pool of players, which allows us to abstract away from modeling player heterogeneity.

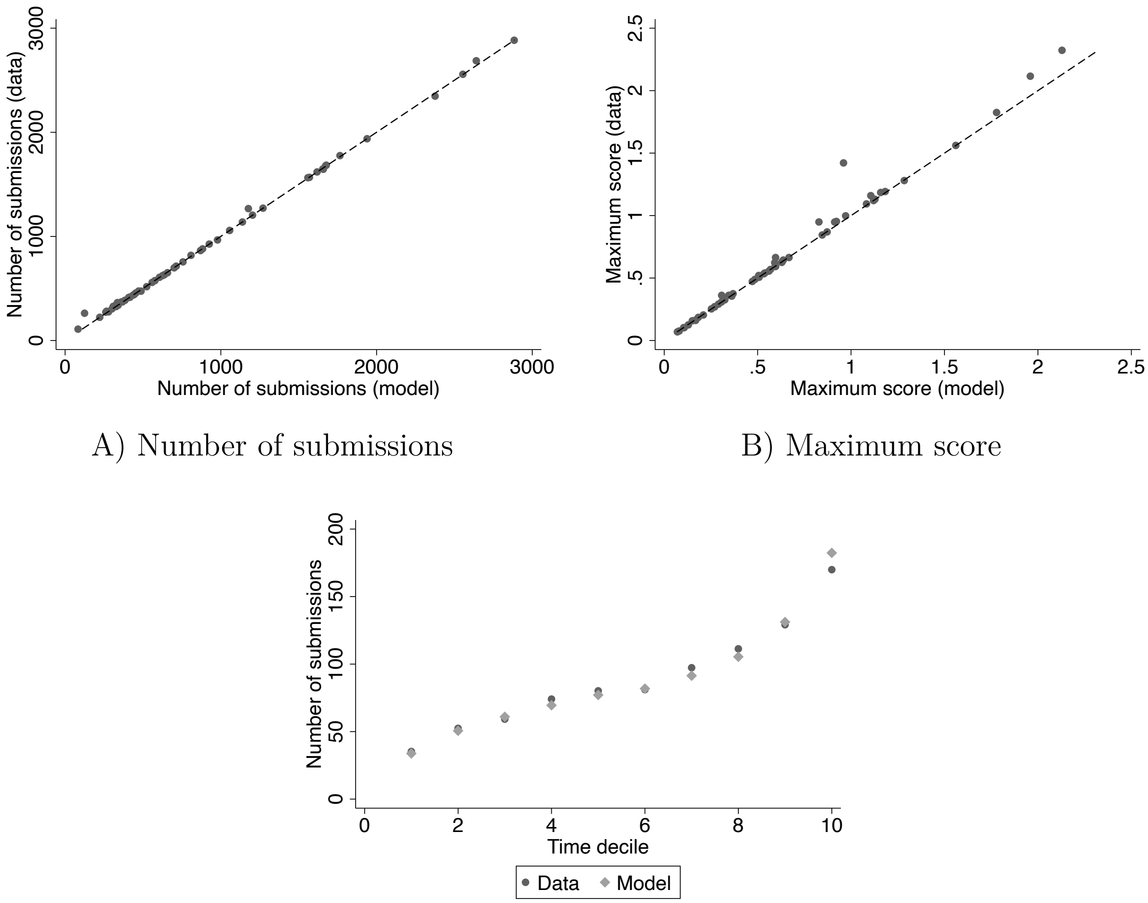
4.1 Estimation Results and Model Fit

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

Regarding the goodness of fit of the model, Figure 2.A and Figure 2.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 2.C plots the actual and predicted number of submissions over time, averaged across contests, where time is divided into 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.

Figure A.1.A in the Online Appendix plots the cross-contest distribution of the expected cost of making a submission. Given the distribution of costs that we use in our model, i.e., $K(c; \sigma) = c^{\sigma}$, the expected cost is given by $\sigma/(1 + \sigma)$. Since the model normalizes the value of the prize pool to 1, we have to scale this up by the size of the prize in order to translate

Figure 2: 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.

costs to dollars. The figure shows that the median expected cost of making a submission is \$642 dollars and 75 percent of contests have an expected cost that is less than \$1,280. If a player needs 10 hours to prepare to make a submission, the median hourly cost is \$64.2 dollars.

Note, however, that these averages represent the unconditional distribution of costs. When a player decides to make a submission, they weigh the benefit against the cost. A submission occurs only when the benefit outweighs the cost, creating selection: making a submission is indicative of a low cost draw. We approximate the expected conditional cost of making a submission for every contest by computing the expected conditional mean at time zero.[17] We

[17]Specifically, we compute $E[c|c < B_0] = \frac{B_0^{\sigma+1}\sigma}{\sigma+1}$, where $B_0 = \frac{q_0}{N} \times$ Prize.

Table 3: Percentage change in the expected number of score increments under alternative prize structures in simulated contests (relative to a prize structure where all the prize money is awarded at the end of the competition)

	$\bar{s}=0$			$\bar{s}=0.4$			$\bar{s}=0.8$		
	Optimal	SC	2-parameter SC	Optimal	SC	2-parameter SC	Optimal	SC	2-parameter SC
$T=5$	2.52	2.3	2.25	2.76	2.61	2.58	3.01	2.91	2.89
	(0.25)	(0.23)	(0.23)	(0.27)	(0.26)	(0.26)	(0.29)	(0.28)	(0.28)
$T=10$	1.87	1.5	1.41	2.1	1.79	1.73	2.38	2.14	2.1
	(0.19)	(0.16)	(0.16)	(0.21)	(0.19)	(0.19)	(0.23)	(0.22)	(0.22)
$T=15$	1.68	1.29	1.14	1.83	1.47	1.38	2.09	1.78	1.73
	(0.17)	(0.15)	(0.13)	(0.19)	(0.16)	(0.16)	(0.21)	(0.19)	(0.19)

Notes: The table reports the average percentage change in the expected number of score increments (i.e., the number of times the maximum score increases during the contest) under alternative prize structures relative to a final-ranking prize. The optimal column captures the gains from using the optimal prize structure, the 'SC' shows the gains of a hybrid prize structure, and the '2-parameter SC column' shows the gains of an optimally calibrated milestone in combination with a prize to the leader at the end. These outcomes were computed based on simulated contests that make use of the model estimates in Table 2. In these simulated contests, we vary two parameters relative to the "full" model: T (number of periods) and $\bar{s}$ (the score at the beginning of the contest). The parameter ε is fixed at 0.1. Standard errors are in parentheses.

report these values in Figure A.1.B and find that the average cost of a submission, conditional on choosing to make a submission, is $26.6 dollars, highlighting the role of selection.

5 Prize Structure and Contest Outcomes

How do different prize structures impact contest outcomes? Using our model estimates, we use counterfactual simulations to provide an answer using two exercises. In the first exercise, we use the model estimates in Table 2 to simulate contests that are shorter than the ones observed in our sample. The shorter length of these simulated contests allows us to solve for the optimal prize structure and compare contest outcomes under the optimal prize structure with outcomes under alternative prize structures. Because the dimensionality of the optimal prize structure explodes as we increase the length of the contest, we use this exercise to explore whether simple prize structures can approximate the gains of the optimal prize structure. In the second exercise, we use our model estimates and compare the observed equilibrium of the contest in our sample with the equilibria of these contests under alternative prize structures, where our focus is on simple prize structures.

5.1 Optimal Prize Structure

We begin our analysis by finding the optimal prize structure in short contests. We simulate a contest for each set of primitives in Table 2 with different T (number of periods), $\bar{s}$ (the score at the beginning of the contest), and ε (the size of the score increments). We fix ε at 0.1, and vary T and $\bar{s}$ to consider the sensitivity of our results to these parameters.

For every combination of parameters, we compare the performance of a final-ranking prize with the performance of three alternative prize structures: (1) the optimal prize structure, which solves problem (12); (2) the optimal score-contingent prize structure ("SC") that rewards players who advance the maximum score (i.e., we allow for $\pi(s_t = s, t|s_{t-1}) = \pi_s \geq 0$ for every s) as well as the player who leads the competition at the end; (3) a "2-parameter SC" that awards one prize from reaching a milestone and one prize for the leader at the end of the contest; the size of the milestone and the size of the two prizes are chosen optimally. We measure the performance of each prize structure by the expected number of score increments during the contest (i.e., the number of times the maximum score increases during the contest). We use this variable because it contains all the information needed to compute the final maximum score (together with ε and $\bar{s}$) and is unaffected by the scale of scores.

Table 3 shows that the optimal prize structure always increases performance relative to a final-ranking prize in short contests. Specifically, it increases the number of score increments between 1.68 percent (when $T = 15$ and $\bar{s} = 0$) and 3.01 percent (when $T = 5$ and $\bar{s} = 0.8$) on average.

The table shows two salient patterns related to the current- and the future-competition effects. We capture the intensity of the current competition effect by the initial score: A higher value of $\bar{s}$ implies that the probability of increasing the maximum score is lower. We capture the intensity of the future-competition effect by the length of the contest: a higher value of T means that the future-competition effect early in the contest is stronger.

First, the gains from using the optimal prize structure are larger when the current competition effect is stronger. For instance, when $T = 5$, the gain increases from 2.52 percent to 3.01 percent. Intuitively, a stronger current-competition effect discourages players, so with a final-ranking prize, the leader can "rest on her laurels" and wait for the competition to end to receive a prize. The optimal structure must reward players who increase the score but do not necessarily lead the competition at the end. We can see that rewarding score increments (column 'SC') capture a large fraction of the gains of the optimal prize structure (between 77 and 97 percent). Perhaps more surprisingly, a milestone prize in combination with a prize to the leader based on final standings (column '2 parameter SC') achieves between 68 and 96 percent of the gains of the optimal prize structure. This is surprising because the '2 parameter SC' is very sparse. For example, in a contest with 15 periods, the optimal prize structure can award up to 210 prizes, whereas a 2-parameter SC awards only 2 prizes.

Second, the gains from using the optimal prize structure are smaller when the future competition effect is stronger. For instance, when $\bar{s} = 0$ the gain decreases from 2.52 percent to 1.68 percent. Here, the optimal prize structure can use contingent prizes to motivate players early on. However, it cannot reward early plays too much; otherwise, the remaining budget will be too small, discouraging players later on. One concern regarding the gains from the optimal structure when T becomes very large is that they could converge to zero since there is a decreasing pattern. In the next section, we show this is not the case. While we cannot compute the optimal prize structure for large values of T, we show that the gains from simple structures are bounded away from zero and, in fact, can be substantial.

In summary, studying short contests suggests that simple prize structures can approximate the gains of the optimal prize structure despite being sparse.

5.2 The Gains of Simple Prize Structures

We next turn to compare the contests we observe in our sample (i.e., the observed equilibria) with the equilibrium of each of these contests under alternative prize structures. We consider seven counterfactual designs, each of which has a simple structure. The first three designs consider k prizes to the interim leaders at k equally spaced times, 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 of size $\pi \in [0, 1]$ at time $\tau \in [0, T]$, and one of size $1 - \pi$ at the end of the competition. Both π and τ are chosen optimally. The fifth design ("milestone") awards the full prize pool to the first player surpassing an optimally chosen milestone score B. The sixth design ("hybrid") awards one prize to the first player surpassing a milestone score and one prize to the leader of the competition based on final standings, where the milestone and the magnitude of both prizes are optimally chosen. The last design ("elimination") eliminates k followers (chosen at random) in the middle of the contest, where k is optimally chosen.

Table 4 presents a comparison of equilibrium outcomes. For each contest, we compute the optimal set of parameters for each prize structure class. The columns labelled "Optimal" compare the equilibrium number of submissions and maximum score (in expectation) under the within-class optimal prize structure relative to the equilibrium outcomes under the base-

Table 4: 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)
Milestone	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)
Elimination	10.539	10.539	0.013	0.013
	(2.319)	(2.319)	(0.005)	(0.005)

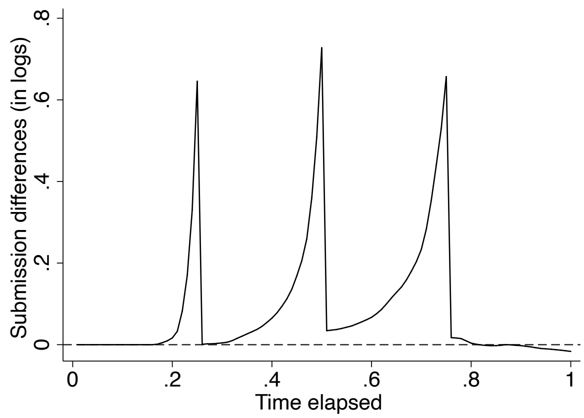
Notes: The table reports the average change in the number of submissions and maximum score when using alternative prize structures. The outcome differences compare the optimal set of parameters for each prize structure class. The column "Optimal" finds optimal parameters for each contest and averages the gains. The column "Uniform" finds a single set of parameters that maximize the average maximum score across contests.

line design (i.e., when all the prize money is awarded based on the final ranking). 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 third highest gains, with an average increase in the number of submissions and a maximum score of 30.5 percent and 0.039 standard deviations, respectively. The milestone prize structure is the second highest performer, with an average increase in the number of submissions and a maximum score of 30.8 percent and 0.04 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.[18] 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 3 shows how the average num-

[18] Table A.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 that the hybrid design is more likely to be optimal when the frequency of submission opportunities is higher (i.e., stronger future competition).

Figure 3: Submission dynamics in a 4-prize prize structure relative to the baseline design



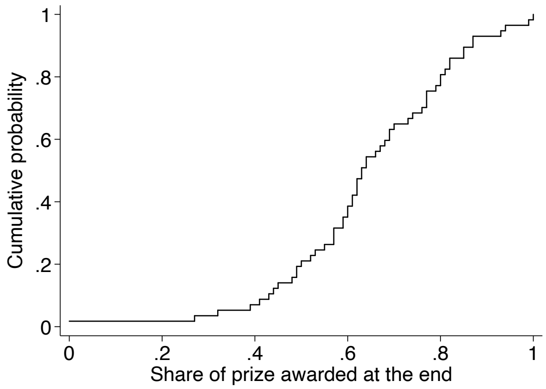
Notes: The figures present the change in the equilibrium number of submissions predicted by the model (in logs) when implementing the optimal 4-prize prize structure (relative to the baseline design).

ber 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 2.C). Nevertheless, the total number of submissions increases on average by 25.1 percent relative to a prize structure based only on final standings.

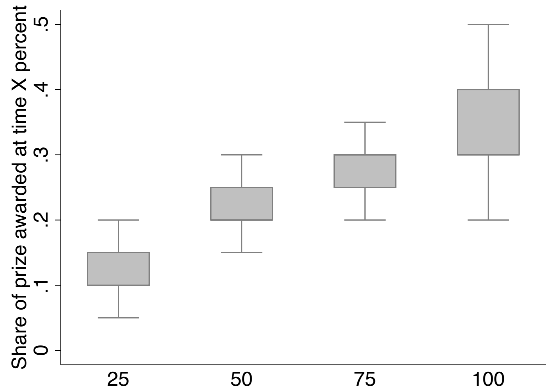
Figure 4 presents details about the parameters governing the optimal prize structures within each prize class. Panels B and C show that the 4- and 6-prize optimal designs are similar across contests in that the prizes increase over time, and the ranges of each of the prizes are somewhat narrow. Panel D shows that in the design with 2-timed prizes, about 80 percent of the probability mass of the distribution of the optimal time of the first prize is in between times 0.6 and 0.8. The optimal milestone score and hybrid designs (Panels E and F) feature more heterogeneity across contests, which reflects underlying differences in the score distributions and the probability of increasing the maximum score across contests. Omitted in the figure is the elimination design, as we find that it is optimal to eliminate 1 player in the middle of the competition in every contest. Although we find that the punishment of eliminating players boosts the incentive to make submissions early in the competition, the cost of having fewer players able to make submissions in the second half is too great to eliminate more than one player.

We note that to compute the optimal prize structure within each counterfactual design for

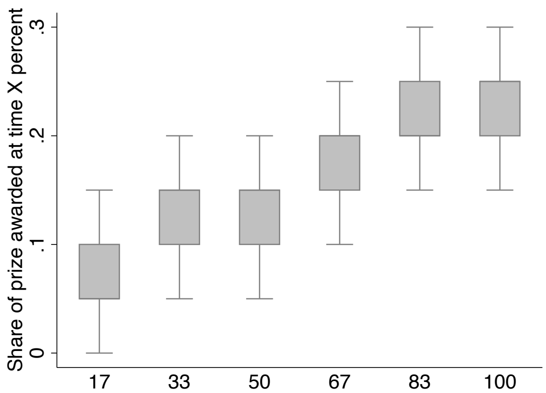
Figure 4: Optimal prize structures, by prize class



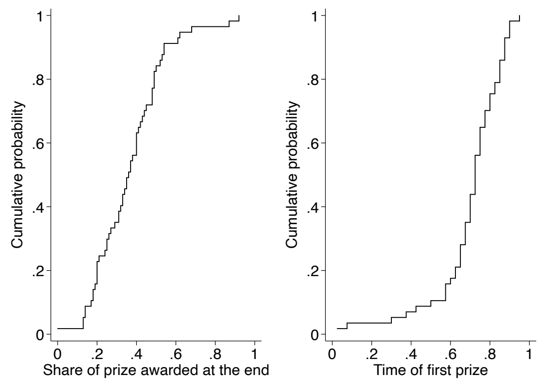
A) 2 equally-spaced prizes



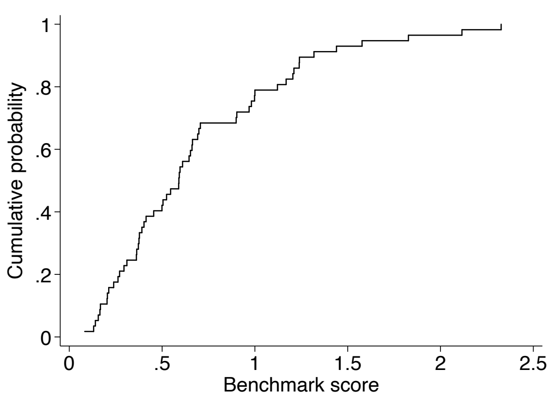
B) 4 equally-spaced prizes



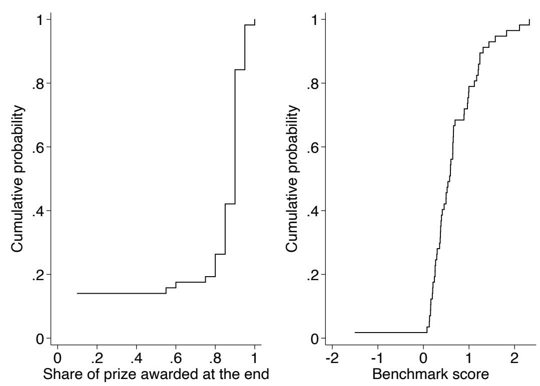
C) 6 equally-spaced prizes



D) 2 prizes with an optimally-timed first prize



E) Milestone prize



F) Hybrid 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.

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 4 that are labelled "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.[19] 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 are optimized contest by contest, with the exceptions being the milestone 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 milestone and hybrid prize structures because of the heterogeneity in the optimal milestone scores across contests we show in Figure 4.

Lastly, to measure the gains of using optimal designs in US dollars, we perform the following exercise. For every competition, we compute the equilibrium of the game using the optimal simple prize structure for a given *class* of simple prize structure. However, we compute this optimal structure scaling down the prize pool up to the point of matching equilibrium outcomes with observed outcomes. Specifically, we find the optimal hybrid prize structure and 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 4. 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 flexible 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.

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

6 Experimental Evidence

This section presents experiment evidence, which is valuable as a complement to our structural estimates for two reasons. First, it allows us to investigate whether contingent prizes can improve contest outcomes without relying on model assumptions. Estimates from our randomized-controlled trial are "model-free" estimates. Second, it gives us an opportunity to test some of the predictions of the empirical model regarding the performances of particular prize structures. To this end, we use our structural estimates to guide our choice of experimental treatments.

6.1 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.[20] We exploit experimental variation in prize structure to measure the impact of the prize structure on contest outcomes and competition dynamics. This exercise is meant to provide model-free evidence of the value of contingent prizes.

We recruited 405 students (both undergraduates and graduates) via emails, department 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, had statistics knowledge, or had machine learning skills.

In the experiment, we created 81 groups (or contests) of 5 participants. Of these groups, 27 out of 81 featured UBC students, with the remaining 54 being composed of UIUC students. 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, and number of participants). We ran the competitions simultaneously. In the competitions, players had 11 days to solve a simple prediction problem: interpolate a function (see Online Appendix C for details). Players were allowed to 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. Unlike the rest of the paper, a better submission here is one with a

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

Table 5: Baseline summary statistics and test of balance

Variable	Control ($N = 27$) Mean (1)	2 prizes ($N = 27$) Coeff. (2)	2 prizes ($N = 27$) p-value (3)	Hybrid ($N = 27$) Coeff. (4)	Hybrid ($N = 27$) p-value (5)	F-test p-value (6)
Past participation	0.096	0.027	0.504	0	1	0.723
Knows machine learning	0.467	-0.036	0.569	0.022	0.727	0.562
Uses statistical tools	0.733	-0.018	0.775	0.037	0.533	0.636

Notes: An observation is a contest. All variables are defined at the contest level as follows: 'Past participation' is the share of players in the contest who have participated in a prediction contest in the past, 'Knows machine learning' is the share of players in the contest who have machine learning skills, and 'Uses statistical tools' is the share of players in the contest who have learned statistics. Columns 2-6 report the coefficients and p-values from OLS regressions of each covariate on two indicators: '2 prizes' and 'hybrid'. Column 7 reports the p-value from a joint test of statistical significance of both indicators.

smaller score (i.e., lower 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 predetermined milestone score and $70 to the leader at the end of the competition (treatment "hybrid"). We set the milestone 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 5 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, and only about half reported knowing machine learning techniques. There are no statistically significant differences in these covariates across treatment groups.

Table 6: Prize structure impacts on contest outcomes

	Min. score (1)	# of submissions (2)	Min. score (3)	# of submissions (4)
2 prizes	0.026	4.268	0.026	5.375
	(0.063)	(10.229)	(0.064)	(10.529)
Hybrid	-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 5.

6.2 Results

Table 6 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. We run an OLS regression for every outcome variable with indicators for each treatment assignment. The first two columns exclude controls, whereas the second two include the covariates in Table 5 as controls.[21]

Columns 1 and 3 suggest that the hybrid prize structure caused the minimum score to decrease by 0.05, about a third of the mean of the dependent variable.[22] 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.[23] To shed light on heterogeneity, Figure 5 plots the minimum scores across all contests by treatment assignment. Figure 5.A shows that the cumulative distribution functions of the control and the 2-prize groups cross each other, whereas Figure 5.B shows that the cumulative distribution function of minimum scores in the control groups first-order stochastically dominates that of the hybrid contests.

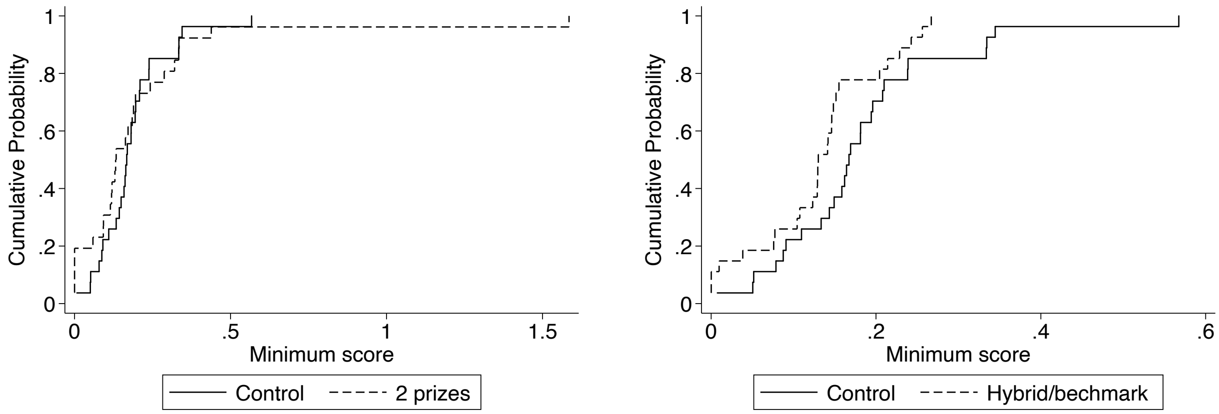
Columns 2 and 4 of Table 6 show that at the end of the competition, the 2-prize and the hybrid contests had a greater number of submissions than the control group, although the estimates

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

[22] We replicate these results using quantile regression for the median in Table A.5 in the Online Appendix.

[23] We had one outlier: In one of the 2-prize contests, there was a single submission (minimum score ≥ 1.5 in Figure 5). The 2-prizes design would feature lower minimum scores on average had we removed that outlier.

Figure 5: Distribution of minimum scores, by prize structure



A) 2 prizes

B) Hybrid

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

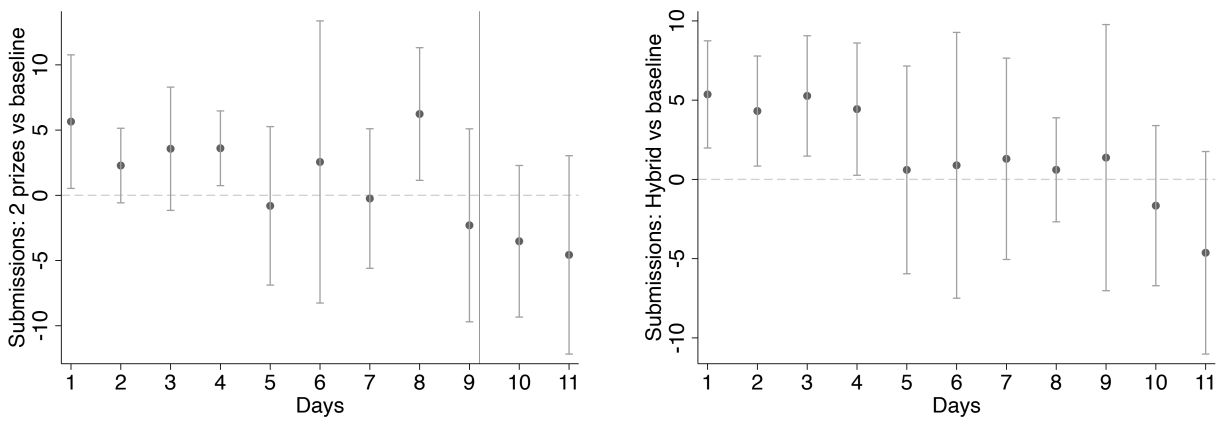
are noisy. This comparison, however, obscures how the prize structure shapes incentives to exert effort throughout the competition. Figure 6 plots the difference in the average number of submissions between treatment X and the control group, by day. Figure 6.A shows that early in the competition, the 2-prize contests had an average number of submissions that was greater than that of the control contests, with the difference peaking on 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 (although the estimates are noisy), 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.[24] A similar pattern is observed in Figure 6.B, where the difference in submissions was greatest in the first four days, reflecting participants' effort to surpass the milestone score in the hybrid contests.

7 Summary

We study dynamic contests with public, real-time performance feedback. These types 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 budget. We identify two central forces governing incentives to play at any point during the competition: the future competition

[24] Along these lines, Table A.6 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 contests.

Figure 6: Submissions over time, by prize structure



A) 2 prizes

B) Hybrid

Notes: The figures show the difference in average daily submissions in treatment (2-prizes or hybrid) and control contests. Each contest–day combination is an observation. In Panel A, the vertical line indicates the day of the first prize award in the 2-prize contests.

effect (i.e., plays that are yet to unfold) and the current-competition effect (i.e., current maximum score).

Contingent prizes, rather than prizes based on the final ranking only, can affect the balance of these effects and, therefore, can improve outcomes. To shed light on this issue empirically, we measure the performance of various contest designs featuring time- or score-contingent prizes using 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. We characterize the optimal prize structure among a set of simple prize structures for each contest in the data. We also 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 "uniform" design captures a large portion of the gains from using a tailored prize structure in each contest.

8 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.

Bhattacharya, Vivek (2021) "An empirical model of R&D procurement contests: An analysis of the DOD SBIR program," *Econometrica*, 89 (5), 2189–2224.

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.

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.

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.

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.

Rogerson, William P (2003) "Simple menus of contracts in cost-based procurement and regulation," *American Economic Review*, 93 (3), 919–926.

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.

Supplemental Material – Intended for Online Publication

Contingent Prizes in Dynamic Contests

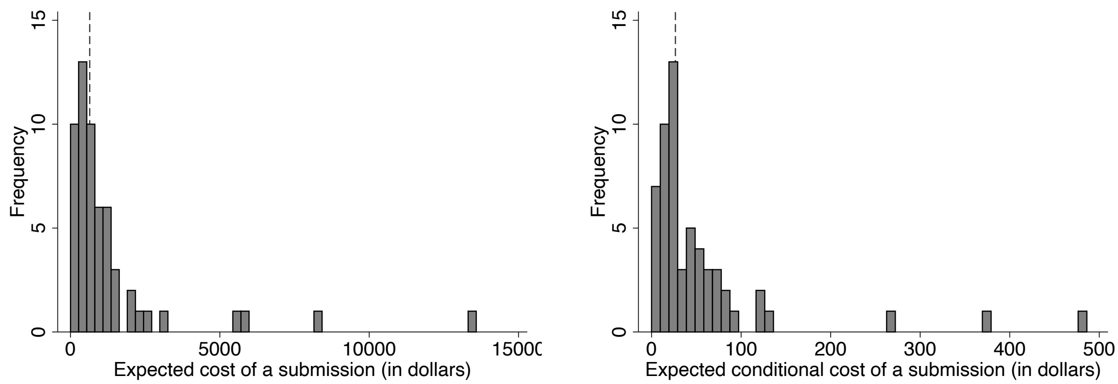
Jorge Lemus and Guillermo Marshall

A Additional Tables and Figures

Table A.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.



A) Unconditional B) Conditional

Figure A.1: Estimates for the cost of making a submission

Note: An observation is a contest. Panel A: "Cost of a submission" is the expected cost of making a submission (i.e., given the distribution of costs we use in our model, the average cost is given by $\sigma/(1+\sigma)$). Since the model normalizes the value of the prize pool to 1, we have to scale this up by the size of the prize in order to translate costs to dollars. The median across contests (dashed vertical line) is \$642 (USD). Panel B: "Conditional cost of a submission" is the expected cost conditional on choosing to make a submission. Computing this moment is computationally intensive. A middle-ground result is an approximation at time 0: Initially players are symmetric, the gains from a submission are bounded by $B = q(s_0) \cdot \text{Prize}/(\text{number of players})$, where $q(s_0)$ is the probability of surpassing the maximum score evaluated at the initial score. Thus, we can compute the conditional mean at time zero explicitly: $E[c|c < B] = B^{\sigma+1} \cdot \sigma/(\sigma + 1)$. We report this value for every contest. The median is \$26.6 (USD).

Table A.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	0.0001	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	0.0001	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	0.0001	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 A.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 A.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 milestone score 1.175
Hybrid	70% of prize at $t = 1$, 30% to the first player
	who surpasses the milestone 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 A.5: Prize structure impacts on contest outcomes

	Min. score			
	OLS	Quant. Reg.	OLS	Quant. Reg.
	(1)	(2)	(3)	(4)
2 prizes	0.026	-0.034	0.026	-0.043
	(0.063)	(0.027)	(0.064)	(0.034)
Hybrid	-0.054**	-0.037*	-0.055**	-0.040*
	(0.026)	(0.019)	(0.027)	(0.022)
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	0.174	0.174	0.174
Std. dev. dep. variable	0.191	37.549	0.191	37.549

Notes: An observation is a contest. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. "Quant. Reg." is a quantile regression for the median. Controls include all covariates reported in Table 5.

Table A.6: 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 5.

B 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

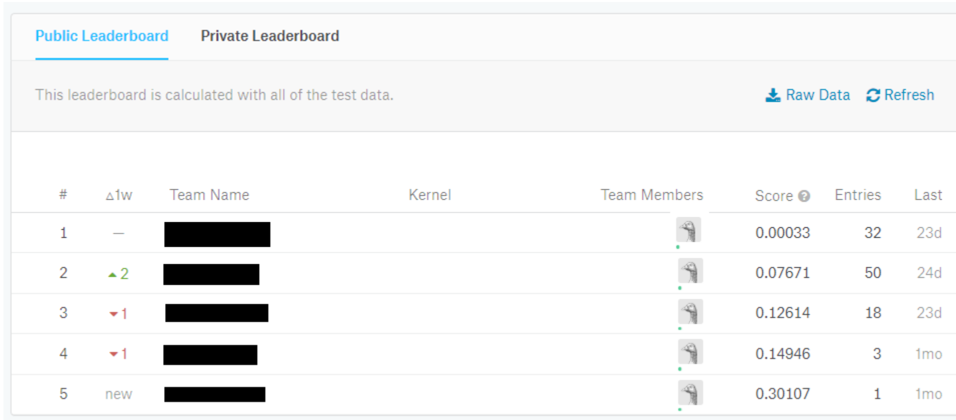
Submission File:

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.