EC8855 Empirical Methods Final Project

Ying Wang

May 18, 2021

The goal of this final project is to replicate the results from the paper "Dynamic Tournament Design: Evidence from Prediction Contests" by Lemus and Marshall (2021).

Information disclosure can play an important role in online contests outcomes. In practice, it can be implemented by displaying a real-time public leaderboard with information about the performance of all participants. In theory, the information disclosure of participants' progress has two opposing effects. On the one hand, it can make participants more optimistic about the state of the world (encouraging effect); on the other hand, it may negatively affect effort provision for the laggard since their winning probability becomes smaller (competition effect). Thus, the net impact of real-time public leaderboard on players' decision and contest outcomes can be ambiguous in a dynamic setting.

Lemus and Marshall (2021) investigates the impact of a leaderboard on contest outcomes using two complementary approaches. First, they estimate a dynamic model using observational data of 57 large competitions from Kaggle, an online platform that hosts prediction contests, and compare the equilibrium with and without a leaderboard. Second, they present experimental evidence from student competition. The results imply that a leadeboard on average improves competition outcomes. For simplicity, I will use data from one competition "Predict Grant Applications" to replicate some of the empirical analysis results in Lemus and Marshall (2021).

I will first present some preliminary observations and facts using data from one competition following the order of Lemus and Marshall (2021).

Figure 1 contains a histogram that partitions all the submissions into time interval based on their submission time. In contrast with Lemus and Marshall (2021), Figure 1 shows that the number of submissions does not monotonically increasing over time based on data from one competition. Specifically, the number of submissions first increases for the first 40% of contest time, and then decreases from 40% to 60% of contest time. For the remaining 40% of contest time, the number of submissions keeps increasing.

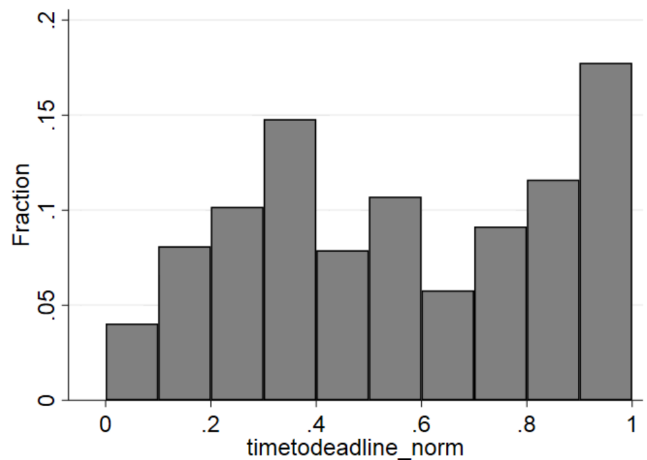

Figure 1: Figure 1

Figure 2 plots the result of local polynomial regression of the number of teams with one or more submissions as a function of time. It shows the entry of new teams into the competition. The figure shows that the rate of entry is roughly constant over time, with about 20% of teams making their first submission when 20% of the contest time remains.

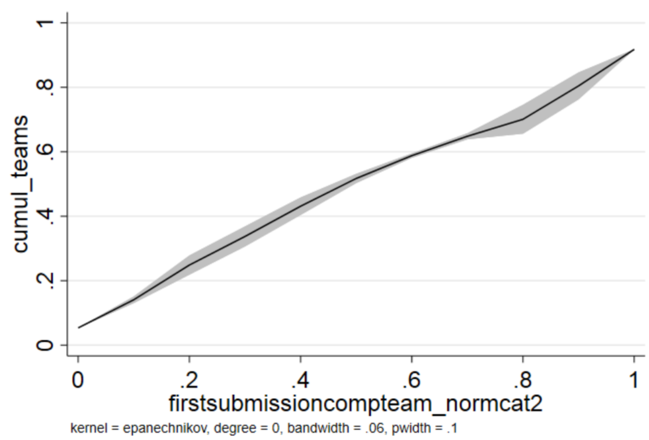

Figure 2: Figure 2

Figure 3 plots the results of local polynomial regression of the time between submissions as a function of time. In contrast with Lemus and Marshall (2021), Figure 3 shows that the average time between submissions does not monotonically increasing over time based on data from one competition. It implies that there might exist learning effect. Teams might become more experienced and make submissions at a faster pace when they have made large enough number of submissions. Moreover, the pattern of Figure 3 can help to explain the U-shaped part of Figure 1, since teams are able to make more submissions if it takes less time to complete one submission.

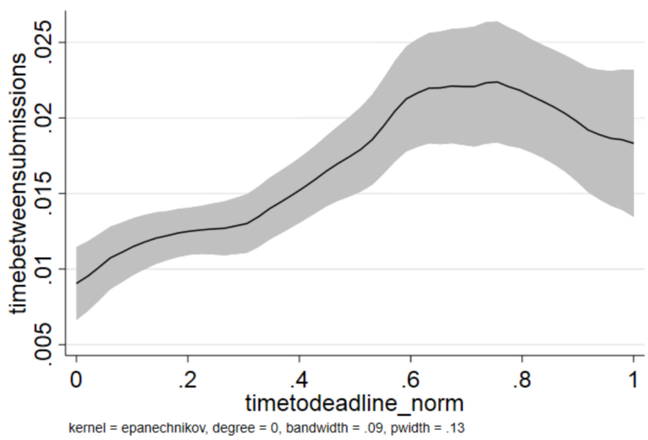

Figure 3: Figure 3

In Lemus and Marshall (2021), they claim that lower scores are attributed to participants who may be not trying to win it but instead are participating for nonpecuniary motives. Since they are interested in modeling competitive players whoa re affected by the design of the competition and are trying to win, they group teams into ”competitive” and ”noncompetitive” categories. Competitive teams are defined as teams that obtain scores above the 75th percentile of the score distribution in a competition.

The following table decomposes the variance of public scores using a regression analysis. The dependent variable is the public score, and the regressor is the submission number, which is defined at the competition-team-submission level and measures the number of submissions made by a team up to the time of a submission. Column 1 and 2 state the results based on all teams, while column 3 and 4 consider only submissions from competitive teams. In column 1, 30% of the variation in public score is between-team variation, suggesting that teams systematically differ in the scores that they achieve. Column 2 controls for the number of submissions that a team has submitted up to that time of each submission, which allows us to capture whether learning can explain some of the variation in scores. However, in contrast with Lemus and Marshall (2021), the coefficient is negatively significant and only 1% of the variance in scores is explained by this control. Column 3 and 4 repeat the analysis for competitive teams. Now the coefficient becomes positive, suggesting that learning may be present for competitive teams. Moreover, in this restricted sample, teams are more homogeneous, so team fixed effects explain less of the variation when compared with the whole sample.

	(1) pubscore_n~l	(2) pubscore_n~l	(3) pubscore_n~l	(4) pubscore_n~l
sub_number		-0.00736***		0.00167
		(0.000977)		(0.000920)
_cons	-0.000576	0.143***	0.459***	0.421***
	(0.0107)	(0.0222)	(0.0141)	(0.0226)
N	2776	2776	1388	1388
R-sq	0.701	0.711	0.629	0.630

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

To understand how the public leaderboard shapes incentives to participate, Lemus and Marshall (2021) regress an indicator for whether a given submission was a team's last submission on the distance between the team's best public score up to that time and the best public score across all teams up to that time. The following table states the regression results. Column 1 shows that it is more likely for teams to drop out of the competition when they start falling behind in the public score leaderboard. A standard deviation increase in a team's deviation from the maximum public score at time t is associated with a 2.35% increase in the likelihood of a team dropping out of the competition at time t. Column 2 explores whether this result is heterogeneous between competitive and noncompetitive teams and shows that competitive teams are less discouraged to quit the competition when they are falling behind, compared with noncompetitive teams.

	(1) lastsubmis~m	(2) lastsubmis~m
dev_leader~t	0.0235***	0.0186*
	(0.00605)	(0.00778)
inter_dev_75		-0.0283***
		(0.00813)
inc75		-0.0366**
		(0.0119)
_cons	0.0584***	0.0822***
	(0.00540)	(0.0104)
N	2800	2800
R-sq	0.008	0.016

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Next I am going to discuss more details about empirical model estimation. I will skip the model description and focus on model estimation.

The full set of parameters for a given contest include:

- the distribution of new player arrival times, which assumes to follow an exponential distribution with parameter μ.
- the distribution of submission arrival times, which assumes to follow an exponential distribution with parameter λ.
- the distribution of private score conditional on public score, $H(\cdot|p^{public})$, which we assume is given by $p^{private} = \alpha + \beta p^{public} + \epsilon$, with ϵ distributed according to a double exponential distribution.
- the type-specific cdf of public score, which we assume is given by the standard normal distribution, $Q_i(x) = \Phi[(x - \theta_j^{mean})/\theta_j^{st.dev}]$, for type θ_j
- the distribution of types, κ, which we assume is a discrete distribution over the set of player types Θ.
- the distribution of submission costs, which we assume has a cdf given by $K(c;\sigma) = c^{\sigma}$.

I summarize some of the key points of the estimation procedure used to estimate the primitives of the model stated above.

- **Distribution of entry times**

  Given the vector of entry times for the set of players I in a given contest, $\{t_i\}_{i \in I}$, estimate μ by using MLE.

  $$\hat{\mu} = argmaxlogL(\mu) = argmax \sum_{i \in I} log(\mu) - \mu t_i = \frac{1}{\bar{t}}$$

  where $\frac{1}{\bar{t}} = \sum_{i \in I} \frac{t_i}{|I|}$

- **Distribution of entry time between submissions**

  Given the vector of times between submissions for the set of players I in a given contest, $\{t_{i,m}\}_{m \in M_i, i \in I}$, estimate λ by using MLE.

  $$\hat{\lambda} = argmaxlogL(\lambda) = argmax \sum_{i \in I} \sum_{m \in M_i} log(\lambda) - \lambda t_{i,m} = \frac{1}{\bar{t}}$$

  where $\frac{1}{\bar{t}} = \sum_{i \in I \; m \in M_i} \frac{t_{i,m}}{|\{t_{i,m}\}_{m \in M_i, i \in I}|}$

- **Conditional distribution of the private scores**

  Given the pairs of scores for all M submissions in a contest, $\{p_m^{public}, p_m^{private}\}_{m \in M}$, estimate (α, β) using MLE.

  $$(\hat{\alpha}, \hat{\beta}) = argmaxlogL(\alpha, \beta) = argmax \sum_{m \in M} -\epsilon_m + exp\{-\epsilon_m\}$$

  where $\epsilon_m = p_m^{private} - \alpha - \beta p_m^{public}$

- **Type-specific distribution of scores & Distribution of types**

  - Use the expectation-maximization (EM) algorithm to estimate the MLE of θ and κ.
  - Assume there are $k = 2$ player types.
  - The EM algorithm estimates the k Gaussian distributions that best predict the observed distribution of public score.

- **Distribution of submission costs - Challenging**
  - Players decide to build a new submission if the benefits are greater than the cost of building a new submission.
  - The distribution of costs is identified based on how variation in the benefits of building a new submission changes the decision of a player to build a new submission.
  - Compute the benefits of building a new submission via simulation.
  - Use these simulated benefits to compute the MLE of cost distribution.
  - Simulating the benefits of building a new submission require estimates of all primitives of the model except for the distribution of costs as well as the equilibrium conditional choice probabilities (CCPs).
  - CCPs are estimated using data on decisions as well as data on the state variables faced by the players at the moment of making those decisions (e.g., scores on the leaderboard).
  - Then simulate the expected payoffs associated with each action in the choice set given the state variables faced by the players and use the estimates of the CCPs to capture the players' beliefs about how all other players would behave in response to the player's action.

I have gone through the code provided by the author and will discuss what I have learned as well as some of the challenges I faced during the estimation procedure.

- I learned how to code up EM algorithm from the authors. Each iteration consists of computing an expectation and then maximizing the expectation with respect to θ and κ. These steps are repeated until convergence of the vector parameters. Note that when implement the EM algorithm, the authors only use three scores for each player in a given contest: median score, 75th percentile score and maximum score.

- The most challenging part is to estimate the distribution of cost. This step is computationally intensive and it took me several days to get the estimate results. One of the technical issues that took me a while to understand is how they compute the benefits of building a new submission. For every (s, t, θ), where s is a vector of state variables,

t is time and θ is player type, they simulate $NS = 200$ continuation histories of the game under two cases: the player chooses to build a new submission (d = Play) and the player chooses not to build a new submission (d = Not Play). Histories are simulated using estimates for all primitives of the model and the CCPs. For every simulated history, they compute the payoff of the player at the end of the game.

All in all, it is a great learning experience following the estimation procedure proposed by the authors and understanding how they wrote the code. Since I am interested in the design of contest, the replication exercise of this paper provides me with more detailed knowledge preparation both theoretically and empirically for future research.

References

[1] Lemus, Jorge, and Guillermo Marshall. "Dynamic tournament design: Evidence from prediction contests." Journal of Political Economy 129.2 (2021): 383-420.