Measuring Performance Of Peer Prediction Mechanisms Using Replicator Dynamics

Victor Shnayder
edX; Harvard SEAS
shnayder@seas.harvard.edu

Rafael M. Frongillo
CU Boulder
raf@colorado.edu

David C. Parkes
Harvard SEAS
parkes@eecs.harvard.edu

Abstract

Peer prediction is the problem of eliciting private, but correlated, information from agents. By rewarding an agent for the amount that their report "predicts" that of another agent, mechanisms can promote effort and truthful reports. A common concern in peer prediction is the multiplicity of equilibria, perhaps including high-payoff equilibria that reveal no information. Rather than assume agents counterspeculate and compute an equilibrium, we adopt *replicator dynamics* as a model for population learning. We take the size of the basin of attraction of the truthful equilibrium as a proxy for the robustness of truthful play. We study different mechanism designs, using models estimated from real peer evaluations in several massive online courses. Among other observations, we confirm that recent mechanisms present a significant improvement in robustness over earlier approaches.

1 Introduction

Peer prediction formalizes the challenge of eliciting information from agents in settings without verification. Whereas scoring rules [Gneiting and Raftery, 2007] and prediction markets [Hanson, 2003; Chen *et al.*, 2007] can be used to elicit beliefs about observable events (e.g., the outcome of the U.S. Presidential election), peer prediction addresses settings without direct access to the ground truth. Consider, for example, eliciting information about noise in a restaurant, about the quality of an e-commerce search algorithm, or the suggested grade for a student's assignment in an online course, where obtaining ground truth is either not possible or costly.

The theory of peer prediction has developed rapidly in recent years. From the simple approach of output agreement [von Ahn and Dabbish, 2004; Waggoner and Chen, 2014], the field has moved to scoring-rule based approaches with varying knowledge requirements on the part of the designer [Miller *et al.*, 2005; Witkowski and Parkes, 2012a], later relaxing the requirement of a common prior [Witkowski and Parkes, 2012b; Radanovic and Faltings, 2013; Kamble *et al.*, 2015]. These early mechanisms all had *uninformative* equilibria, where agents could make reports without looking at their assigned task, and yet get a higher score than by being truthful. Several recent papers propose mechanisms that ensure that truthfulness is not only a strict correlated equilibrium, but has higher payoff than certain other strategies.

Jurca and Faltings [2009] discourage strategies where all agents report identically, by rewarding near-agreement rather than complete agreement with peers. Radanovic and Faltings [2015] present the logarithmic peer truth serum, with a large population and many peers performing each task, comparing an agent's agreement with their peers to their agreement with the population as a whole. Dasgupta and Ghosh [2013] propose a *multi-task* approach, where each agent completes multiple tasks, and compare agreement on overlapping tasks to expected agreement on non-overlapping ones, showing that truthfulness is optimal for settings with binary reporting. Shnayder et al. [2016] extend this method to settings with more than two possible reports.

There is also experimental work on peer prediction. One study [Gao *et al.*, 2014] showed that Mechanical Turk workers are able to coordinate on an uninformative equilibrium in some peer prediction mechanisms, while behaving in an unpredictable way in a design inspired by Jurca and Faltings [2009]. A second experimental study is more positive, showing that simple scoring mechanisms can encourage effort, and that workers do not seem to coordinate on uninformative equilibria [Faltings *et al.*, 2014].[1]

We adopt replicator dynamics as a model of population learning in peer prediction mechanisms. Our interest is to understand the robustness of different designs when, rather than pre-computing equilibria, participants adjust their behavior via a simple dynamic. Learning is widely used to study behavior in games, giving a useful measure of the likelihood that various equilibria emerge in repeated play of a mechanism, as well as the stability of those equilibria. Intuitively, these dynamics capture how players may adjust their behavior slightly each round depending on the success of their previous actions. While truthfulness may be an equilibrium of the game, if learning dynamics steer away from it, one may not expect to see (long-lasting) truthful behavior in practice.

Analyzing models derived from peer evaluation data in

[1] A possible reason for the difference in results is that the environment in this second study had many possible reports, making it harder to coordinate.

several massive online courses, we confirm concerns about uninformative equilibria in early peer prediction mechanisms: despite the existence of a truthful equilibrium, learning dynamics move toward uninformative equilibria in these mechanisms. The learning dynamics still tend toward all participants adopting the same uniformed report in the approach of Jurca and Faltings [2009]. In contrast, the multi-task mechanisms do better, with a larger basin of attraction of the truthful equilibrium. Truthfulness is most stable under the *correlated agreement* mechanism [Shnayder *et al.*, 2016], which generalizes the method of Dasgupta and Ghosh [2013], while the logarithmic peer truth serum [Radanovic and Faltings, 2015] does not work well unless each task is performed by a comparatively large number of agents.

1.1 Case study: Peer grading

To choose realistic parameters for our experimental study, we use data from peer evaluation in several Massive Open Online Courses (MOOCs). Organizations such as edX, Coursera, and many others around the world are scaling online learning to tens of thousands of students per course without a corresponding expansion in course staff. A key challenge is to scalably teach topics that are difficult to automatically assess, such as writing, judgement, or design. Peer evaluation is a promising tool—students submit assignments, which are evaluated by several peers using an instructor-created rubric. Peers provide scores as well as written feedback.

In today's systems, the evaluators are not scored, though participation can be coupled with being able to see feedback from their peers. This means that students can (and do) submit minimal feedback without giving it much thought.[2] This setting fits the peer prediction model—it is expensive for staff to make "ground truth" evaluations by grading submissions, and because several peers evaluate each submission, their assessments are naturally correlated and can be compared.

Other research on scalable peer evaluation evaluates students' assessment skills, identifies and compensates for their biases [Piech *et al.*, 2013], and helps students self-adjust for bias [Kulkarni *et al.*, 2013]. The Mechanical TA project [Wright and Leyton-brown, 2015] aims to reduce TA workload in high-stakes peer grading.

1.2 Background on replicator dynamics

We use one of the simplest models of evolutionary population dynamics, which were first introduced to study evolution [Smith, 1972; Sandholm, 2009; Gintis, 2009]. Such models track segments of a population, gradually adjusting behavior in response to feedback. Evolutionary dynamics have been used in many applications besides evolutionary biology. For example, Erev and Roth [1998] show that learning dynamics can capture key features of human behavior in economic games, and they have many applications in multi-agent systems [Bloembergen *et al.*, 2015].

Replicator dynamics track a continuous population of agents playing a game over time, with each agent adopting a pure strategy and probabilistically switching to higher-payoff strategies in proportion to the gain in expected payoff. Nash equilibria are known to be fixed points of replicator dynamics, but the converse need not hold [Easley and Kleinberg, 2010, Thm 12.6]. These dynamics also provide an appealing model for learning at the individual level, as they are a continuous-time limit of the *multiplicative-weights* learning algorithm, and guarantee no regret [Hofbauer *et al.*, 2009, Prop 4.1 and Prop 6.2]. See Arora *et al.* [2012] for more about the multiplicative weights algorithm.

Replicator dynamics have been used to compute the symmetric, mixed equilibria in empirical game theory [Reeves *et al.*, 2005]. Recently, replicator dynamics have been applied to assess the likelihood or stability of various equilibria in games [Panageas and Piliouras, 2014] (see also [Kleinberg *et al.*, 2011; 2009]). We employ this latter interpretation; specifically, we adopt the basin of attraction of the truthful equilibrium, meaning the set of strategy profiles leading eventually to the equilibrium, as a proxy for how likely and how stable truthfulness would be under repeated play.

2 Model

There is a continuum of agents, representing a distribution over strategies observed in the population. At each time t, finite groups of agents are sampled from this distribution, and each group is assigned to a particular task (e.g., label an image, evaluate a particular homework submission, judge the mood of a video clip, etc), which has a hidden type $h \in H$.

Each agent i privately observes a signal $s_i \in S=\{0, 1, \ldots, n-1\}$, identically and independently distributed, conditioned on type h. Let $\Pr(h)$ denote the type prior and let $\Pr(s|h)$ denote the signal distribution conditioned on type. For simplicity, we assume that the number of types is equal to the number of signals. For example, in a peer evaluation setting, the hidden type would be the "true" quality of a submission, and the signal a student's assessment of the quality, both on a scale of e.g. 0, 1, or 2. We assume that $\Pr(h)$ and $\Pr(s|h)$ are the same for all tasks and all agents, though the methodology extends to heterogeneous agent populations with non-identical signal models.

Once the agents observe their signals, they use a strategy, θ, to compute a report $r_i = \theta(s_i)$. In general, θ can be randomized, but we focus on deterministic strategies, relying on the random sampling from the population for mixing. A peer-prediction mechanism, without knowing the hidden type or the observed signals, computes a score σ_i for each agent based on reports. This score can depend on the reports of *peer* agents who did the same task, as well as on the overall set of reports across all tasks. A good scoring rule leads agents to maximize expected score by truthfully revealing their signals, and is robust to alternate equilibria as well as misreports or noise from other agents. A special concern is to prevent high-payoff, *uninformed equilibria*, where agents adopt signal-independent strategies; e.g., "always report 1."

We represent the *population strategy profile* as a distribution $x = (x_1, \ldots, x_m)$, where x_k is the fraction of agents who adopt strategy θ_k, and m is the total number of strategies. Let $U(k, x)$ denote the expected payoff from strategy

[2] This is a well-known issue in on-campus peer-evaluation settings as well, though there, instructors can review the feedback and intervene. In MOOCs, that kind of oversight may not be scalable.

θ_k given population profile x. The average population payoff is defined as $A(x) = \sum_{k=1}^{m} x_k U(k, x)$, leading to the replicator dynamics differential equation:

$$\dot{x}_k = x_k(U(k, x) - A(x)). \quad (1)$$

We numerically solve this equation for particular starting strategy profiles to predict whether the population will tend toward the all-truthful profile.

2.1 Peer prediction mechanisms

We focus on *strictly proper* peer prediction mechanisms, where truthful reporting is a strict correlated equilibrium.

Single-task mechanisms. We first define mechanisms that only depend on the reports for a single task.

(1) **Output Agreement (OA)** [von Ahn and Dabbish, 2004]. The system picks a reference agent j for each agent i, and defines $\sigma_i(r_i, r_j) = \mathbf{1}_{(r_i=r_j)}$, where $\mathbf{1}_{x=y}$ is 1 if x=y, 0 otherwise. The OA mechanism is only strictly proper when observing a signal s makes s the most likely signal for a reference agent as well (see Frongillo and Witkowski [2016] for an elaboration). A useful property of OA is that it is *detail-free*, requiring no knowledge of the probabilistic model of the world.

(2) **MRZ**. The peer prediction method [Miller *et al.*, 2005] (MRZ), which uses proper scoring rules [Gneiting and Raftery, 2007] to achieve strict properness. In MRZ, the system gets a report r_i, picks a reference peer j, and uses a proper scoring rule R based on the likelihood of r_j given r_i. By the properties of proper scoring rules, this makes truthful reporting a strict correlated equilibrium. In our experiments, we use the *log scoring rule* $R(\gamma, o) = \log(\gamma_o)$, where γ is a probability distribution over outcomes, and o is the observed outcome. MRZ is not detail-free, as computing γ requires knowledge of the world model.

(3) **JF09.** A problem with both OA and MRZ is that they also have uninformative, pure-strategy symmetric Nash equilibria, one of which always results in the highest possible payoff [Jurca and Faltings, 2005]. The JF09 [Jurca and Faltings, 2009] mechanism removes these (pure) Nash equilibria in binary settings, relying on four or more peers doing a single task. To evaluate a report r_i in a binary signal setting ($S = \{0, 1\}$), the mechanism picks three reference agents, defines z_i as the total number of 1 reports among them, and gives score $\sigma_i(r_i, z_i) = M[r_i, z_i]$, where M is the matrix $\left(\begin{smallmatrix} 0 & \alpha & 0 & \epsilon \\ \epsilon & 0 & \beta & 0 \end{smallmatrix}\right)$. α and β are set based on the world parameters to preserve strict properness, while the form of the payoff matrix ensures that if all agents coordinate on 0 or 1, they get score 0.[3] JF09 is not detail free because the designer needs the world model to compute the score matrix.

Multitask mechanisms. The next two mechanisms are *strong truthful*, meaning that all agents being truthful is an equilibrium with higher payoff than any other strategy profile, with the inequality strict except for signal permutations.

(4) **RF15**. The RF15 [Radanovic and Faltings, 2015] mechanism scores an agent based on the statistical significance of the agent's report compared to the reports of their peers and the distribution of reports in the entire population across multiple tasks. Given report r_i and the fractions $z_{\text{peer}}, z_{\text{global}}$ of reference peers and global population respectively reporting r_i, the agent's score is $\sigma_i = \log(z_{\text{peer}}/z_{\text{global}})$. As the number of reference peers goes to infinity, this approaches $\log(\Pr(r_{peer} = r_i)/\Pr(r_i))$.[4] RF15 is detail-free.

(5) **DG13**. The DG13 mechanism [Dasgupta and Ghosh, 2013] is detail-free and multi-task, so each agent reports on several tasks. It is defined for binary signals. An agent is rewarded for being more likely to match the reports of peers doing the same task than the reports of peers doing other tasks.

We present a slightly generalized form, parametrized by a score matrix Λ. The mechanism is described, w.l.o.g., for two agents, 1 and 2:

1. Assign the agents to three or more tasks, with each agent to two or more tasks, including at least one overlapping task. Let M_s, M_1, and M_2 denote the shared, agent-1 and agent-2 tasks, respectively.
2. Let r_1^k denote the report received from agent 1 on task k (and similarly for agent 2). The payment to both agents for a shared task $k \in M_s$ is

$$\sigma_i = \Lambda(r_1^k, r_2^k) - \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \Lambda(i, j) \cdot h_{1,i} \cdot h_{2,j},$$

   where $\Lambda : \{0, \ldots, n-1\} \times \{0, \ldots, n-1\} \to \mathbb{R}$ is a score matrix, $h_{1,i} = \frac{|\{\ell \in M_1 | r_1^\ell = i\}|}{|M_1|}$ is the empirical frequency with which agent 1 reports signal i in tasks in set M_1, and $h_{2,j} = \frac{|\{\ell \in M_2 | r_2^\ell = j\}|}{|M_2|}$ is the empirical frequency with which agent 2 reports signal j in tasks in set M_2.
3. The total payment to an agent is the sum of the payments across all shared tasks.

In the DG13 mechanism, Λ is the identity matrix ('1' for agreement, '0' for disagreement.) For binary signals and positive correlation between signals, DG13 is strong truthful.

(6) **DGMS.** Shnayder et al. [2016] extend the DG13 mechanism in two ways. The first is DGMS, the direct extension of DG13 to multiple signals, using the identity matrix for scoring. DGMS is strong truthful when the world satisfies a *categorical* property, where, given an agent's signal, the likelihood of peers having any other signal goes down: $\Pr(s'|s) < \Pr(s')$ for all $s' \neq s$; this property holds trivially for binary signal models with positive correlation.

(7) **Correlated Agreement (CA)**. The second extension of DG13 yields the CA mechanism, which adopts a different scoring rule. Rather than the identity matrix, CA sets $\Lambda(i, j) = 1$ if $\Pr(s_j|s_i) > \Pr(s_i)$, and 0 otherwise; it rewards agreement on positively correlated signals. CA reduces to DGMS in categorical settings. In general settings, it is proper (not strictly), and *informed truthful*. The payoff for truthfulness is weakly higher than any other strategy profile,

[3] There are no results about mixed equilibria. Our analysis in Section 3 shows them to be problematic.

[4] Because log is non-linear, the expected score with a finite number of reference peers is lower than this limit, even in a continuous population, and this affects the attractiveness of different strategies. We examine this effect in Section 3.

and strictly higher than any uninformed, signal-independent reporting strategy. CA only requires that the designer know the direction of correlation between pairs of signals, not the entire world model.

2.2 Strategy selection

To fully define the replicator dynamics, we need to instantiate a finite set of strategies available to the population. In mechanisms where agents do multiple tasks per round, each agent uses the same strategy for each task. We omit permutation strategies, which exchange the names of signals in a 1-to-1 mapping, from our analysis. These are unnatural in practice, and do not give higher payoffs than the remaining strategies in the mechanisms we study.

With two signals, the remaining pure strategies are *const0, const1, T*, corresponding to agents always reporting *0, 1*, or being truthful, respectively. For three or more signals, there are more strategies possible, and we include the monotonic strategies that overreport, underreport, or merge adjacent signals, using *const0, const1, const2, merge01, merge12, bias+, bias-, T*. *merge01* reports 0 for signals 0 and 1. *merge12* reports 1 for signals 1 and 2. *bias+* over-reports, mapping signal i to $\min(i+1, n-1)$. *bias-* maps i to $\max(i-1, 0)$. For four signals, we add *mergeAdj*, which reports 0 for signals 0 and 1, and 2 for signals 2 and 3. For five signals, we add *mergeEach3* which rounds down to the nearest multiple of three. The merging strategies lose information and increase the frequency of agreement, and are an intermediate step between truthfulness and constant reports.

As a simple model of effort, we distinguish between *informed* and *uninformed* strategies. An informed strategy depends on the agent's signal. In contrast, constant strategies such as *const1* are uninformed. The distinction reflects that it takes effort to obtain a signal, so informed misreporting strategies are less appealing to agents than uninformed ones.

For certain strategy profiles and mechanisms, there may be multiple strategies with equal payoff. When there is a tie between truthfulness and another informed strategy, we believe it is natural for agents to be truthful— it is simpler because it does not require strategic reasoning, while the effort of signal acquisition is needed either way. To model this, we add a tiny cost to the expected payoffs for non-truthful informed strategies, so as to break such ties in favor of truthfulness.[5]

2.3 World models

Our initial qualitative analysis compares the mechanisms in four world models, selected to illustrate common scenarios; the worlds vary the correlation between agent signals and include bias toward particular values (Figure 1).

3 Replicator dynamics of peer prediction

Starting with the single-task mechanisms, we show that in OA, MRZ, and JF09, non-truthful equilibria are attractors of replicator dynamics and the basin of attraction of truthfulness is small.

[5]The consequence for replicator dynamics is that areas of the strategy simplex where the derivative between truthful and another informed strategy was exactly zero now tend toward truthful.

World	$\Pr(h)$	$\Pr(s\|h)$	Description
W2a	[0.5, 0.5]	$\begin{pmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{pmatrix}$	Strong correlation
W2b	[0.5, 0.5]	$\begin{pmatrix} 0.4 & 0.6 \\ 0.1 & 0.9 \end{pmatrix}$	Bias toward 1
W3a	[0.3, 0.3, 0.4]	$\begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$	Unbiased noise
W3b	[0.3, 0.3, 0.4]	$\begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.4 & 0.5 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$	0 and 1 correlated

Figure 1: Our manually selected world models.

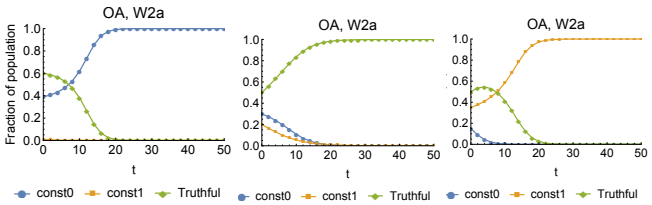

Figure 2: Replicator dynamics in OA for different initial strategy distributions. Even when a large fraction of the population starts out truthful, the dynamics can converge to all-ones or all-zeros uninformative equilibria.

3.1 Single task mechanisms

We first look at the replicator dynamic for OA in the W2a world (Figure 2). This illustrates replicator dynamics for different initial values. At least half the population starts out truthful in each plot, but the dynamics can still converge to an uninformative strategy where all agents say 0 or 1.

It is difficult to understand the overall dynamics of a mechanism from plots of strategies versus time, because each only shows a particular starting point. A better visualization for analyzing convergence is a flow plot of the derivatives of the replicator equation, as shown in Figure 3. The area in green shows the *basin of attraction*, the set of starting points from which the dynamics converge to truthful play (the bottom-left corner). From the plots for W2b, we see that OA is not strictly proper, and that the all-0 and all-1 corners are much stronger attractors than truthfulness for MRZ.

From the JF09 plots, we can clearly see that even though the $(1, 0)$ and $(0, 1)$ corners are not equilibria, there are still attractors very nearby. This illustrates how replicator dynamics complement equilibrium analysis, showing that the pure-strategy-only theoretical guarantees of JF09 are not robust.

3.2 Multi-task mechanisms

Multi-task mechanisms leverage reports across multiple tasks to make coordination on uninformed behavior less attractive to agents. We first confirmed that constant reporting is longer an attractive strategy in replicator dynamics under DG13 and RF15 for any binary world with correlated signals, including W2a and W2b. Instead, the basin of attraction of truthful play covers the entire strategy simplex.

However, for RF15, this is in the large-population limit, as both the total population and the number of reference peers for each task go to infinity. Figure 4 shows what happens

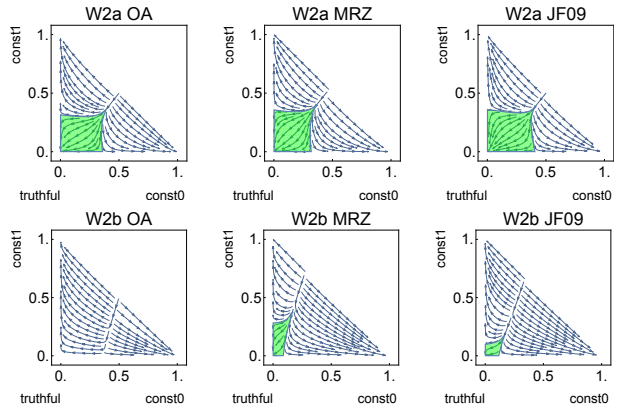

Figure 3: Flow plots of the derivative of the replicator equation (Eqn 1) for W2a and W2b, with OA, MRZ, and JF09. The all-truthful strategy profile is at $(0, 0)$, with its basin of attraction shown by the green shaded area. OA is not truthful for W3. Even when the mechanisms are truthful, the basins of attraction can be quite small.

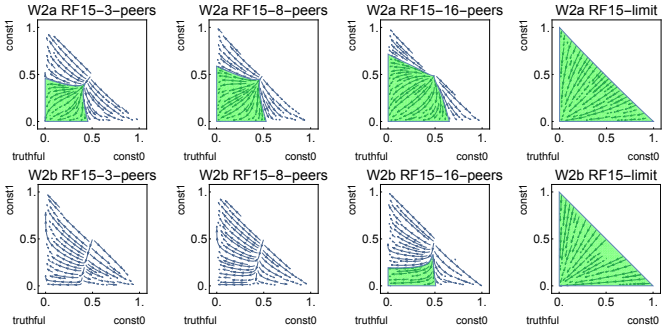

Figure 4: RF15 with finite sets of peers. The non-linearity of the log function makes RF15 far less robust with small numbers of peers, with much smaller basins of attraction for truthfulness.

when the population is large (formally, a continuum), but the number of peers per task is finite. We see that a large group of reference peers is needed for RF15 to behave as in its limit—even with 16 peers, the non-linearity of the log function in the definition of the score rule makes constant reporting attractive if enough of the population agrees. Going forward, when using RF15 with a finite number of reference peers we fix this number to three and study *RF15-3-peer*; for motivation, consider that it is typical for 3-5 students to assess a peer's work for peer assessment in online courses.

We now look at settings with more than two signals, and examine the recent extensions of DG13 to multi-signal settings. The strategy space quickly grows, so we cannot visualize the full basin of attraction in the same way. Instead, we first consider T along with two non-truthful strategies at a time, looking to develop qualitative understanding through representative examples. We will then adopt a quantitative metric, which estimates the basin size for more than three

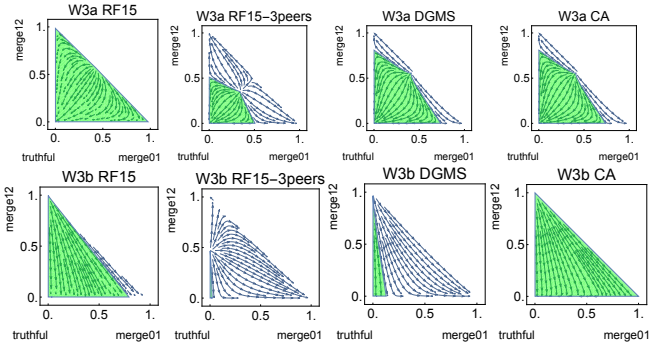

Figure 5: Flow plots for W3a, a categorical world, and W3b, a non-categorical one. *merge01* is the highest payoff strategy under DGMS, and is a strong attractor.

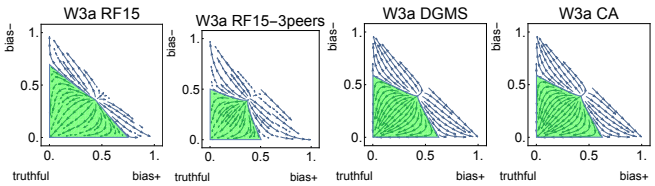

Figure 6: Flow plots for W3a, now using the *bias+* and *bias-* strategies. The difference in the basins of attraction compared to the top row of Figure 5 shows the limitations of two-dimensional flow plots in a many-strategy setting.

strategies by sampling.

First we compare W3a, a categorical three-signal model, and W3b, a non-categorical model, showing dynamics for *merge01, merge12*, and T (Figure 5). For W3a, the CA and DGMS mechanisms are identical, and both converge to truthfulness from a large set of starting values. For W3b, *merge01* has higher payoff than T under DGMS, and the dynamics converge to *merge01* from almost the whole space.[6] Figure 6 parallels the W3a plots just discussed, but now adopting different strategies. Here, the basins of attraction for truthfulness are smaller. This illustrates the need to examine many combinations of strategies to understand a mechanism's behavior.

4 Peer assessment in MOOCs

Our qualitative analysis suggests that the RF15 and CA are robust across a range of strategies and models, while non-truthful strategies can be attractors for OA, MRZ, and JF09. We now examine these patterns quantitatively on realistic world models. We study 325,523 peer assessments from 17 courses from a major MOOC platform. These comprise 104 questions, each with a minimum of 100 evaluations. There are 9, 67, 25, and 3 questions with 2, 3, 4, and 5 signals, respectively. We use maximum likelihood estimation to gen-

[6]CA gives equal payoff for T and *merge01*. The tiny boost to truthfulness described in Section 2 breaks the tie toward the truthful corner.

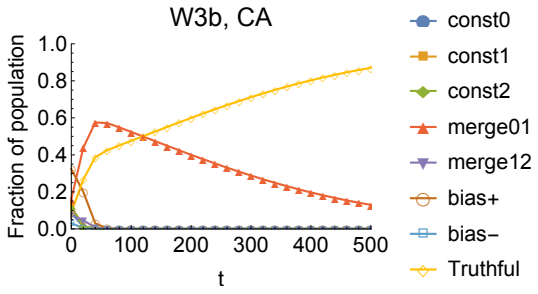

Figure 7: Dynamics with many strategies. We cannot easily visualize the many-dimensional simplex, but can sample to estimate the size of the basin of attraction of an equilibrium.

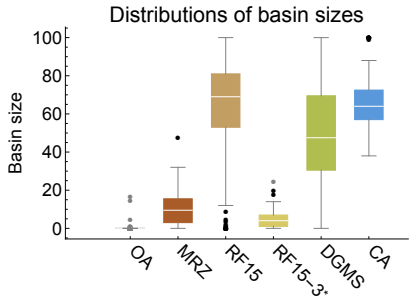

Figure 8: For each mechanism on the horizontal axis, and each of 104 MOOC-based world models, we estimate the size of the basin of attraction of T (as a fraction of the full space, on the vertical axis) for that world. The result is a distribution of 104 basin sizes for each mechanism, which we illustrate with a box plot above each mechanism name. The results match our earlier qualitative analyses.

erate a probabilistic model, $\Pr(h)$ and $\Pr(s|h)$, for each question.

We base our model fit on student reports, not the unobservable signals, which are not available. For the current work, in the absence of better data sets, we will simply stipulate that these are representative of true world models. This gives us a set of observed, non-hand-selected distributions, and provides a systematic way to compare the performance of the various mechanisms. Our analysis remains robust as long as the observed reports do not vary too much from the true signals learners would get if they all invested effort. We believe that as MOOCs start to provide valuable credentials based on peer-assessed work, there will be more incentive to cheat, and this condition may no longer hold without explicit credit mechanisms for peer assessment.

To ensure that our earlier observations were not specific to the particular strategies chosen for each plot, we look at dynamics with many strategies at once. For a qualitative example, see Figure 7, which shows an example for W3b and CA, now with eight strategies. Despite the small fraction of the population starting out truthful, the dynamics converge to the truthful equilibrium.

To quantitatively compare the mechanisms, we estimate the size of the basin of attraction of truthfulness for each question and mechanism pair: we choose 100 starting strategy profiles uniformly at random in the strategy simplex, and measure the percentage for which the dynamics converge to truthful. We exclude JF09 because it is only defined for binary signals while the MOOC models have up to five signals. For each model, we use the corresponding strategy set from Section 2.[7]

This gives us a distribution of 104 basin sizes for each mechanism, shown as box plots in Figure 8. DGMS basin sizes span a large range because many of the estimated models are non-categorical. The CA and RF15 mechanisms have the most robust performance. However, recall that RF15 is defined in the limit as the number of peers per task grows large, and is thus not a good fit for this domain. RF15-3-peer, which is a better match for the domain, does not do as well.

[7]Due to computational limitations in simulating RF15-3-peer, we do not include the full strategy set in its analysis, using only *const0,const1,mergeAdj,bias-,T*. Our comparison thus favors RF15-3-peer, as other potentially attractive strategies are excluded.

The CA mechanism appears promising in terms of its ability to robustly promote the convergence of population learning strategies to informed, truthful play.

5 Conclusions

Replicator dynamics provide a good complement to equilibrium analysis and experiments for studying peer prediction mechanisms. Our analysis confirms that single-task mechanisms such as OA, MRZ, and JF09 can be very unstable with a learning population, even when being truthful is an equilibrium.Newer, multi-task mechanisms (DG13, DGMS and CA), on the other hand, are much better at avoiding uninformative equilibria. The analysis also supports the need for large peer-group sizes with RF15, a point already made in Radanovic and Faltings [2015].

We show that CA, in particular, is a promising candidate for real applications. Given over 100 distributions from peer assessment data, we can have some confidence that our findings will generalize; although distributions for other applications will differ, the size of the differences between the mechanisms suggest that our qualitative findings will be robust.

This analysis can be extended in various directions. We assume the same world model over many rounds of learning, for example, which may not apply if the types of tasks change over time (imagine different homework assignments during a course). In addition, replicator dynamics ignores variance and small-population effects. Also of interest is the behavior of peer-prediction mechanisms with more complex models of human learning from behavioral economics or cognitive neuroscience. Finally, there is a need to validate these results with real people, in the lab, in real online courses, or in other crowdsourcing applications.

Acknowledgments

This research is supported in part by a grant from Google, the SEAS TomKat fund, and NSF grant CCF-1301976. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors alone.

References

[Arora *et al.*, 2012] Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[Bloembergen *et al.*, 2015] Daan Bloembergen, Karl Tuyls, Daniel Hennes, and Michael Kaisers. Evolutionary dynamics of multi-agent learning: A survey. *Journal of Artificial Intelligence Research*, 53:659–697, 2015.

[Chen *et al.*, 2007] Yiling Chen, Lance Fortnow, Evdokia Nikolova, and David M. Pennock. Betting on permutations. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, pages 326–335, 2007.

[Dasgupta and Ghosh, 2013] Anirban Dasgupta and Arpita Ghosh. Crowdsourced Judgement Elicitation with Endogenous Proficiency. In *WWW13*, pages 1–17, 2013.

[Easley and Kleinberg, 2010] David Easley and Jon Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.

[Erev and Roth, 1998] Ido Erev and Alvin E. Roth. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *The American Economic Review*, 88(4):848–881, 1998.

[Faltings *et al.*, 2014] Boi Faltings, Pearl Pu, and Bao Duy Tran. Incentives to Counter Bias in Human Computation. In *The Second AAAI Conference on Human Computation & Crowdsourcing (HCOMP 2014)*, pages 59–66, 2014.

[Frongillo and Witkowski, 2016] Rafael Frongillo and Jens Witkowski. A geometric method to construct minimal peer prediction mechanisms. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*, February 2016.

[Gao *et al.*, 2014] Xi Alice Gao, Andrew Mao, Yiling Chen, and Ryan P Adams. Trick or Treat : Putting Peer Prediction to the Test. In *The Fifteenth ACM Conference on Economics and Computation (EC'14)*, 2014.

[Gintis, 2009] Herbert Gintis. Chapter 12 – Evolutionary Dynamics. In *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*, chapter 12, pages 271–297. 2009.

[Gneiting and Raftery, 2007] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

[Hanson, 2003] R. Hanson. Combinatorial Information Market Design. *Information Systems Frontiers*, 5(1):107–119, 2003.

[Hofbauer *et al.*, 2009] Josef Hofbauer, Sylvain Sorin, and Yannick Viossat. Time average replicator and best-reply dynamics. *Mathematics of Operations Research*, 34(2):263–269, 2009.

[Jurca and Faltings, 2005] Radu Jurca and Boi Faltings. Enforcing truthful strategies in incentive compatible reputation mechanisms. In *Proceedings of the 1st International Workshop on Internet and Network Economics (WINE05),*, volume 3828 LNCS, pages 268–277, 2005.

[Jurca and Faltings, 2009] Radu Jurca and Boi Faltings. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research*, 34(1):209–253, 2009.

[Kamble *et al.*, 2015] Vijay Kamble, Nihar Shah, David Marn, Abhay Parekh, and Kannan Ramachandran. Truth Serums for Massively Crowdsourced Evaluation Tasks. 2015.

[Kleinberg *et al.*, 2009] Robert Kleinberg, Georgios Piliouras, and Éva Tardos. Multiplicative updates outperform generic no-regret learning in congestion games. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 533–542. ACM, 2009.

[Kleinberg *et al.*, 2011] Robert D Kleinberg, Katrina Ligett, Georgios Piliouras, and Éva Tardos. Beyond the Nash equilibrium barrier. In *ICS*, pages 125–140, 2011.

[Kulkarni *et al.*, 2013] Chinmay Kulkarni, Koh Pang Wei, Huy Le, Daniel Chia, Kathryn Papadopoulos, Justin Cheng, Daphne Koller, and Scott R. Klemmer. Peer and self assessment in massive online classes. *ACM Transactions on Computer-Human Interaction*, 20(6):1–31, December 2013.

[Miller *et al.*, 2005] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Management Science*, 51:1359–1373, 2005.

[Panageas and Piliouras, 2014] Ioannis Panageas and Georgios Piliouras. From pointwise convergence of evolutionary dynamics to average case analysis of decentralized algorithms. *arXiv preprint arXiv:1403.3885*, 2014.

[Piech *et al.*, 2013] Chris Piech, Jonathan Huang, Zhenghao Chen, Chuong Do, Andrew Ng, and Daphne Koller. Tuned Models of Peer Assessment in MOOCs. *Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee*, 2013.

[Radanovic and Faltings, 2013] Goran Radanovic and Boi Faltings. A Robust Bayesian Truth Serum for Non-Binary Signals. In *AAAI13*, pages 833–839, 2013.

[Radanovic and Faltings, 2015] Goran Radanovic and Boi Faltings. Incentive Schemes for Participatory Sensing. In *AAMAS 2015*, 2015.

[Reeves *et al.*, 2005] Daniel M. Reeves, Michael P. Wellman, Jeffrey K. MacKie-Mason, and Anna Osepayshvili. Exploring bidding strategies for market-based scheduling. *Decision Support Systems*, 39(1):67–85, March 2005.

[Sandholm, 2009] William H Sandholm. Evolutionary game theory. In *Encyclopedia of Complexity and Systems Science*, pages 3176–3205. Springer, 2009.

[Shnayder *et al.*, 2016] Victor Shnayder, Arpit Agarwal, Rafael Frongillo, and David C. Parkes. Informed truthfulness in multi-task peer prediction. *arXiv preprint arXiv:1603.03151*, 2016.

[Smith, 1972] J Maynard Smith. Game theory and the evolution of fighting. *On evolution*, pages 8–28, 1972.

[von Ahn and Dabbish, 2004] Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 319–326, New York, NY, USA, 2004. ACM.

[Waggoner and Chen, 2014] Bo Waggoner and Yiling Chen. Output Agreement Mechanisms and Common Knowledge. In *Second AAAI Conference on Human Computation*, 2014.

[Witkowski and Parkes, 2012a] Jens Witkowski and David C Parkes. A Robust Bayesian Truth Serum for Small Populations. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI 2012).*, 2012.

[Witkowski and Parkes, 2012b] Jens Witkowski and David C Parkes. Peer Prediction Without a Common Prior. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC2012)*, 2012.

[Wright and Leyton-brown, 2015] James R Wright and Kevin Leyton-brown. Mechanical TA : Partially Automated High-Stakes Peer Grading. In *SIGSCE'15*, 2015.