

Trick or Treat: Putting Peer Prediction to the Test

XI ALICE GAO, Harvard University
ANDREW MAO, Harvard University
YILING CHEN, Harvard University
RYAN P. ADAMS, Harvard University

Collecting truthful subjective information from multiple individuals is an important problem in many social and online systems. While *peer prediction* mechanisms promise to elicit truthful information by rewarding participants with carefully constructed payments, they also admit uninformative equilibria where coordinating participants provide no useful information. To understand how participants behave towards such mechanisms in practice, we conduct the first controlled online experiment of a peer prediction mechanism, engaging the participants in a multiplayer, real-time and repeated game. Using a hidden Markov model to capture players' strategies from their actions, our results show that participants successfully coordinate on uninformative equilibria and the truthful equilibrium is not focal, even when some uninformative equilibria do not exist or are undesirable. In contrast, most players are consistently truthful in the absence of peer prediction, suggesting that these mechanisms may be harmful when truthful reporting has similar cost to strategic behavior.

Categories and Subject Descriptors: J.4 [Social and Behavioral Sciences]: Economics

Additional Key Words and Phrases: peer prediction; online behavioral experiment; hidden Markov models

1. INTRODUCTION

Businesses and organizations often face the challenge of gathering accurate and informative feedback or opinion from multiple individuals. Notably, community-based websites such as Yelp, IMDb, Amazon, TripAdvisor, and Angie's List that review products and services are largely dependent on voluntary feedback contributed by their users. The proliferation of online labor markets has also created an opportunity for outsourcing simple tasks, such as classifying images and identifying offensive content on the web, to a readily available online workforce.

In all of these settings, there are some significant barriers in getting participants to honestly reveal their information. First, it is often costly to formulate and share an honest opinion—for example, when evaluating qualities of books or restaurants—whereas uninformative contributions require little or no effort. More importantly, it may be difficult or impossible to verify individual contributions against an observable ground truth, because either the information is inherently subjective or it is too costly to be verified.

These difficulties led to the development of *peer prediction* mechanisms, pioneered by Miller et al. [2005]. To incentivize truthful reports, these mechanisms leverage the

This work was partially supported by the NSF under grant CCF-0953516. Xi Alice Gao is partially supported by the Canadian NSERC Postgraduate Scholarship and the Siebel Scholarship. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors alone.

Authors' address: School of Engineering and Applied Sciences, Harvard University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EC'14, June 8–12, 2014, Stanford, CA, USA.

ACM 978-1-4503-2565-3/14/06 ...\$15.00.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

<http://dx.doi.org/10.1145/2600057.2602865>

stochastic correlation of participants' information and reward each participant by comparing the participant's report with those of his peers. Existing theory on peer prediction focuses on designing such monetary rewards to induce a *truthful Bayesian Nash equilibrium* among all participants—it is in a participant's best interest to report his information truthfully if he believes all other participants will also be truthful.

Despite the existence of the truthful equilibrium, peer prediction theory provides little assurance that participants will adhere to it in practice. This is because peer prediction mechanisms also induce *uninformative equilibria* where participants can coordinate to make a set of reports that are independent of their information. Moreover, such uninformative equilibria are unavoidable in general [Jurca and Faltings 2009; Waggoner and Chen 2013]. Although Miller et al. [2005] argue that the truthful equilibrium is likely focal due to limited communication or ethical preferences of participants, there is generally little theoretical or empirical support for this conjecture.

In this work, we aim to understand how participants will behave towards the peer prediction mechanisms in the presence of multiple equilibria. Specifically, we address the following question:

Will the participants play one of the multiple equilibria of the peer prediction mechanisms? If they do, which equilibrium will they choose and why?

To tackle this question, we tested participants' behavior towards the Jurca and Faltings [2009] (JF) mechanism by engaging them in a multiplayer, real-time and repeated game through a controlled online experiment on Amazon's Mechanical Turk (MTurk) [Mason and Suri 2012]. We allow the participants to repeatedly interact with the JF mechanism, revealing their behavioral dynamics over time. Our work is the first empirical evaluation of participants' behavior towards a peer prediction mechanism in terms of convergence to game-theoretic equilibria.

Our results show that the truthful equilibrium is not focal and that participants clearly favor the uninformative equilibria when paid by peer prediction. In contrast, a majority of the participants are consistently truthful when peer prediction is not used. Specifically, for payment rules where the uninformative equilibria have higher payoffs, a majority of the players coordinated on an uninformative equilibrium. Moreover, for payment rules where the symmetric uninformative equilibria do not exist or have less payoff than the truthful equilibrium, we successfully deterred the players from choosing the uninformative equilibria but still did not motivate truthful reports. Hence, our results suggest that adopting peer prediction mechanisms may be harmful in scenarios where the cost of being truthful is similar to that of acting strategically.

In answering our question about equilibrium play, we needed to infer the players' intended strategies, which are not directly observed in an experiment. This is a common challenge in experimental studies of game-theoretic mechanisms. Researchers typically try to detect equilibrium play by comparing the frequencies of players' actions with equilibrium or other strategies. In this work, we use the novel approach of a hidden Markov model (HMM) to capture players' strategies as latent variables. The HMM is ideal for game-theoretic modeling because it allows us to infer strategies and equilibrium play from observed actions, while allowing the strategies to evolve over time. Our analytical results show that there is great potential in using probabilistic latent variable models to describe game-theoretic and other experimental data.

Related Work. There have been considerable theoretical developments on peer prediction mechanisms. The first peer prediction mechanism was proposed by Miller et al. [2005] and later generalized by Jurca and Faltings [2009] to counter collusions. Subsequently, several mechanisms [Prelec 2004; Witkowski and Parkes 2012b,a; Radanovic and Faltings 2013] were proposed to relax the strong common knowledge assumption.

tions of the earlier mechanisms. Beyond incentivizing truthful information revelation, recent work proposed mechanisms to incentivize participants to put in high efforts [Witkowski et al. 2013; Dasgupta and Ghosh 2013]. Among these, we chose the one proposed by Jurca and Faltings [2009] for the simplicity of its payment rules.

Very few experimental work has adopted peer prediction mechanisms. Among them, the work by John et al. [2012] is the only one with the goal of evaluating the incentives of a peer prediction mechanism, specifically the Bayesian truth serum [Prelec 2004]. This motivated us to experimentally evaluate the incentives of other peer prediction mechanisms. John et al. [2012] asked psychologists to report their engagements in questionable research practices in an anonymous survey study, using BTS to incentivize truth telling. They showed that the self admission rate was higher in the BTS group than in the control group. Several experiments adopted but did not evaluate peer prediction mechanisms [Prelec and Seung 2006; Shaw et al. 2011; Gao et al. 2012]. Prelec and Seung [2006] used BTS to infer the ground truth even if most participants' subjective judgements are wrong. Shaw et al. [2011] used the description of BTS as the contextual manipulation for one financial incentive tested in their online experiment. Gao et al. [2012] used the Witkowski and Parkes [2012b] mechanism to score judges on their evaluations of the quality of short tourism ads.

2. THE JURCA AND FALTINGS MECHANISM

We describe the JF mechanism [Jurca and Faltings 2009], which include the first peer prediction mechanism [Miller et al. 2005] as a special case.

We are interested in an item for which some people have opinions or subjective information, given by their private signals, and we aim to design monetary payments so that participants will truthfully reveal their signals about the item. The item has a type $\omega \in \Omega$; we focus on the binary case with $|\Omega| = 2$. Among $n \geq 3$ participants providing their opinions, participant i 's private opinion is a binary signal $s \in S$, $S = \{s_1, s_2\}$ that only he observes.

The JF mechanism assumes that the relationship between the participants' signals and the type of the item is common knowledge. This assumption allows inference about the likelihood of one participant's signal given another participant's signal. The participants' signals are assumed to be conditionally independent given the type of the item. Specifically, the type of the item is drawn by nature according to a probability distribution $\Pr(\omega)$, $\omega \in \Omega$ where $\sum_{\omega \in \Omega} \Pr(\omega) = 1$, and each participant's signal s is independently drawn according to the conditional probability distribution $\Pr(s | \omega)$, $s \in S$, $\omega \in \Omega$. Moreover, the mechanism assumes that the prior $\Pr(\omega)$ and the conditional distribution $\Pr(s | \omega)$ are common knowledge for all participants and the mechanism.

With their realized signals, each participant makes a report $r \in S$ to the mechanism. The mechanism pays each participant according to the participant's report r and n_f reports of his peers ($1 \leq n_f \leq n - 1$), denoted the participant's *reference reports*. For example, participant i 's reference report could be one other random participant's report ($n_f = 1$) or all other participants' reports ($n_f = n - 1$).

The payment rule of the JF mechanism is designed to induce a truthful Bayesian Nash equilibrium among rational and risk-neutral participants: a participant maximizes his expected payment by truthfully reporting his private signal if he believes that all other participants are reporting their signals truthfully. A payment rule that supports the truthful equilibrium must satisfy the following linear constraints:

$$\sum_{0 \leq m \leq n_f} \Pr(m | s_i) (u(s_i, m) - u(s_{3-i}, m)) \geq 0, \forall i = 1, 2$$

where $u(r, m)$ ($r \in S$, $m \in [1, n_f]$) denotes a participant's payment if he reports r and m out of his n_f reference reports are s_1 and $\Pr(m | s)$ denotes the probability that given a signal of $s \in S$, m out of n_f signals are s_1 . We use this linear program with additional constraints to derive payment rules in our experiment.

Although the JF mechanism supports the truthful equilibrium by design, it always induces other *uninformative* pure strategy equilibria. Furthermore, the uninformative equilibria may yield higher payoffs than the truthful equilibrium, making it questionable whether participants will choose the truthful equilibrium in practice. For example, for the payment rule in Table Ia, for every $s \in S$, it is an equilibrium for all participants to report s regardless of their signals. These equilibria yield higher payoff (1.20 or 1.50) than the truthful equilibrium (0.91).

	$r = s_1$	$r = s_2$
$r_f = s_1$	1.50	0.30
$r_f = s_2$	0.10	1.20

(a) Example 1.

	$r = s_1$	$r = s_2$
$m = 0$	0.90	0.80
$m = 1$	0.10	1.50
$m = 2$	1.50	0.10
$m = 3$	0.80	0.90

(b) Example 2.

	$r = s_1$	$r = s_2$
$m = 0$	0.10	0.15
$m = 1$	0.10	0.90
$m = 2$	1.50	0.15
$m = 3$	0.15	0.10

(c) Example 3.

Table I: Payment rule examples. In (a), each cell gives a player's payoff if he reports r and his reference report is r_f . In (b) and (c), each cell gives a player's payoff if he reports r and m out of the n_f reference reports are s_1 .

Fundamentally, the existence of such uninformative equilibria is unavoidable as participants' reports are only compared with one another and not with the ground truth. Although we cannot make truthful reporting the unique pure strategy equilibrium, it is possible to modify the payment rule to remove the symmetric pure strategy coordinating equilibria entirely or make them less desirable than the truthful equilibrium. Either can be achieved if we have at least 4 participants and a participant's payment depends on all other participants' reports [Jurca and Faltings 2009].

Table Ib shows a payment rule without any symmetric coordinating equilibrium. If all participants make the same report, then any participant can improve his payoff from 0.80 to 0.90 by reporting the other signal. However, this payment rule has asymmetric coordinating equilibria where 3 participants always report s_1 and the remaining participant always reports s_2 (or vice versa). These asymmetric coordinating equilibria also reveal no information but may be harder to reach if the participants cannot communicate with one another.

Alternatively, we could make the symmetric coordinating equilibria less desirable than the truthful equilibrium (Table Ic). This is possible because a participant's payment is maximized when his report agrees with the majority but not all of his reference reports. At either symmetric coordinating equilibrium, every participant gets 0.15, which is much less than the expected payoff of 0.50 at the truthful equilibrium.

3. EXPERIMENT DESIGN AND SETUP

Among all peer prediction mechanisms, we chose to test the binary version of the JF mechanism due to its simplicity. Its payment rule can be presented in a concise table and the binary setting avoids the potential difficulty of estimating and reporting continuous values.

We evaluated the JF mechanism in a repeated setting, although it was defined and analyzed as a one-shot mechanism. In practice, it is natural to expect participants to learn and adapt to the mechanism by providing their opinions for different products and services. Hence, a repeated setting best captures the intended uses of peer pre-

diction and the participants' learning dynamics. Such studies of one-shot games in a repeated setting are typical in experimental economics [Shachat et al. 2012].

We conduct our experiment on MTurk. Peer prediction mechanisms are naturally suited for online experiments, as their intended application is for web-based or impersonal settings. Online participants enjoy greater anonymity and are not subject to the social norms typical of real-world interaction. Additionally, online labor markets, such as MTurk, provide immediate access to a large and diverse subject pool, which is not readily available through alternative means. We recruited over 3000 unique participants in our experiment. Recent studies have shown that online experiments are not only viable, but can be advantageous for large-scale studies of behavior [Horton et al. 2011; Mason and Suri 2012; Rand 2012].

Trick or treat story. In our experiment, we described the mechanism using a simple and fun story about trick or treating on Halloween night:

A group of kids are trick or treating on Halloween night. There are two types of houses, A and B, giving out two types of candies, M&M's and gummy bears, in different proportions. The kids randomly choose a house to go trick or treating. The chosen house may be one of the two types with equal chance but the kids don't know which type of house was chosen. Each kid secretly receives one randomly selected candy from the chosen house. A clown shows up and asks each kid to tell him the type of candy received, promising a payment in return. Each kid may claim to have either type of candy to the clown. The clown collects reports from all the kids, and then rewards each kid based on the kid's claim and the other kids' claims according to a payment rule.

This story can be easily mapped to the setting of the JF mechanism. Each player is a kid, the house is the type of the item, the candies are the signals and reports, and the clown is the mechanism.

This story serves several purposes. First, it makes the mechanism accessible to participants who may be unfamiliar with economic theory. Second, it satisfies the mechanism's common knowledge assumptions, which are required for their theoretical properties to hold. The story also highlights the conditional independence of the signals given the type of the item by emphasizing that the proportions of candies at each house remains the same regardless of the candies given out.

The story also explains that misreporting one's private information is perfectly acceptable. Although this is a common concept in game theory, it can be unintuitive for MTurk workers since they often associate misbehavior with rejection or punishment. To counter this bias, we stated that each player's candy is obtained in secret and is not observed by anybody else. We used the clown as a neutral character to represent the mechanism. The action of "making a claim" is a neutral phrasing in lieu of words such as "lying" or "cheating" with negative connotations. We also emphasized that each player can claim to have either type of candy, and that the clown cannot verify whether the player's claim matches the player's candy.

Setup. Our experiment engaged the participants in a real-time, multiplayer, and repeated game through MTurk. A fixed number of players are matched together and play a game repeatedly for a fixed number of rounds. We use TurkServer [Mao et al. 2012], a framework and API for real-time interaction between experiment participants, in our experiment.

We offer a \$1.00 base payment for completing the task. Each player also receives a bonus payment equal to his average reward in the game (ranging from \$0.10 to \$1.50). We chose a large bonus relative to the base payment to motivate workers to focus on their performance in the game.

We controlled our subject pool in several ways. Each worker may participate only once in the experiment, so that no worker has prior experience with the game. We also restricted our tasks to US workers, for two reasons. Having US workers minimizes the likelihood of connection issues since we require synchronous connections to a US-based server. Second, controlling for geography avoids unexpected behavior if people from other regions have different behavioral norms or a language barrier in understanding the instructions. Additionally, we used common qualifications ($> 95\%$ HITs approved, > 100 HITs completed) to ensure that workers are familiar with the MTurk system.

Task. Users progress through the task in several sections. The first page describes some general information and requires consent, followed by an 11-page tutorial consisting mainly of pictures. The tutorial describes the trick or treat story and the game interface. After the tutorial, the participant must pass a short quiz testing their understanding of the task. Each participant has 3 attempts to pass the quiz with a score of at least 80%. If they fail all 3 attempts, they are permanently blocked from our experiment. After passing the quiz, participants wait in a virtual lobby for enough players to start a new game. When enough players arrive, a 'READY' button appears for each player, starting a new game when enough players press this button; this ensures that all players are paying attention when the experiment starts. We explain the game interface in the next section. After the game, the participants are asked to describe the strategies and reasoning they used in a short exit survey.

Interface. The game interface (Figure 1) describes some general information such as the number of rounds and the number of other players at the top, the steps of the current round on the left, and the history of game play on the right. Players cannot communicate with one another during the task, other than viewing reports from previous rounds.

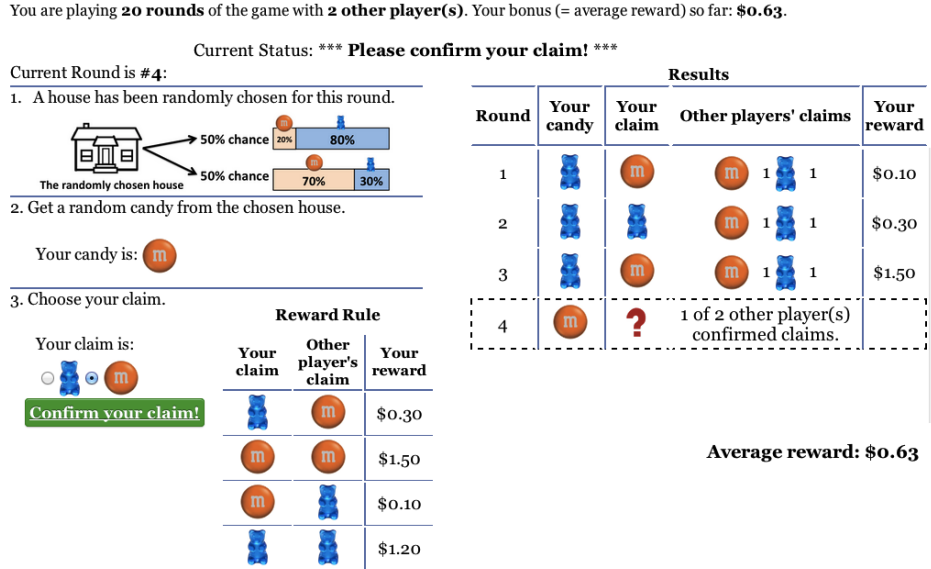


Fig. 1: The Game Interface

When designing the table showing the history of play, we carefully considered how to show the optimal amount of information. With an abundance of information, a participant may become distracted or confused, and pay less attention to the task. However, with insufficient information, a participant may not be able to learn or improve his strategy by observing other participants' actions. After initial trials of our experiment,

	T1	T2	T3	T4	T5
Number of players per game	3	3	4	4	1
Number of rounds	20	20	20	30	20
Number of games without expelled players	103	104	103	103	411

Table II: Treatments

we ultimately chose to display the other participants' claims as an aggregate summary, which is a concise yet still informative representation.

To control for position biases on the game interface, we randomize the row and column order of the payment rule once for each participant and show this randomized table throughout the task. We also randomize the order of questions on the quiz and the order of the radio buttons for choosing claims.

Dealing with disconnections. Due to the synchronicity of the experiment, disruptions from connection issues are possible. To ensure that a game progresses smoothly when such issues occur, we expel a participant from the game if he is disconnected for at least 1 minute (a reasonable threshold since a typical game lasts 5 minutes). An expelled player cannot reconnect to the game, and the server will choose truthful reports on behalf of the expelled player. This ensures that other players experience the game as normal. Our analysis excludes 37 games with expelled players (4% of the games).

4. TREATMENTS

We designed and conducted five treatments in sequence (Table II). Because we had no *a priori* prediction of players' empirical behavior, we used the results of the earlier treatments to design the later treatments. We allowed each participant to participate only once in any treatment, and our quiz showed that participants have similar comprehension of the task in all treatments.

For each game, we recruited a small number of players and allowed them to play for a large number of rounds. Having a small number of players minimizes players' waiting times and potential connection issues, as well as making it easier for each player to reason about other players' actions. With a large number of rounds, we hoped to give players sufficient time to explore and improve their strategies.

For all payment rules, we chose the maximum payment of 1.50 to make it more attractive than the 1.00 base payment, and the minimum payment of 0.10 to prevent extreme behavior resulting from attempting to avoid a payment of zero.

Let $\Omega = \{A, B\}$ denote the two types of houses and let $S = \{MM, GB\}$ be the common signal and report space. We use the MM (or GB) equilibrium to refer to the symmetric coordinating equilibrium where all players always report MM (or GB), and let $xMMyGB$ denote an asymmetric coordinating equilibrium for a $(x + y)$ -player game where x players always report MM and y players always report GB.

The prior. A fixed prior is used for all treatments, as shown in equation (1).

$$\Pr(A) = 0.5, \quad \Pr(MM \mid A) = 0.2, \quad \Pr(MM \mid B) = 0.7 \quad (1)$$

This prior has the nice property that if one player receives a given signal, other players are more likely to have received the same signal. This reasoning may motivate players to be truthful if they believe other players are also truthful.

Treatment 1. For treatments 1 and 2, we used payment rules where each player has one reference report, chosen randomly from all other reports. For treatment 1 (Table IIIa), all four values in the payment rule are distinct and neither report strictly dominates the other. At the truthful equilibrium, every player obtains 0.91 in expectation. Moreover, both the MM and GB equilibria yield higher payments than the truth-

ful equilibrium. Every player gets the maximum payment of 1.50 at the MM equilibrium, making it the highest-paying choice among all equilibria.

	$r = \text{MM}$	$r = \text{GB}$
$r_f = \text{MM}$	1.50	0.30
$r_f = \text{GB}$	0.10	1.20

(a) Treatment 1 payment rule.

	$r = \text{MM}$	$r = \text{GB}$
$r_f = \text{MM}$	1.50	0.10
$r_f = \text{GB}$	0.10	1.50

(b) Treatment 2 payment rule.

Table III: The cell at (r, r_f) gives a player's payment if the player reports r and his reference report is r_f .

Treatment 2. For treatment 1, it's easy to identify the MM equilibrium as having the highest payment. Thus, in treatment 2, we modified the payment rule of treatment 1 to be symmetric (Table IIIb), such that the MM and GB equilibria have the same payments. We hypothesized that such a payment rule may deter the players from reaching either the MM or GB equilibrium, especially without direct communication.

Treatment 3. For payment rules where each player has only one reference report, we are inevitably limited in incentivizing truthful reporting because one of the MM and GB equilibria will always yield higher payment than the truthful equilibrium [Jurca and Faltings 2009]. To overcome this limitation we tested 4-player payment rules in treatments 3 and 4 where each player's payment depends on all other players' reports. We aim to either eliminate the MM and GB equilibria altogether or make these equilibria yield worse payment than the truthful equilibrium.

The payment rule for treatment 3 (Table IVa) does not have the MM and GB equilibria. However, it does support the 3MM1GB and 1MM3GB equilibria and an equilibrium where every player always reports the signal he did not receive. Compared to the truthful equilibrium, the 3MM1GB or 1MM3GB equilibria seem more attractive. They give 3 players the maximum payment of 1.50 and the remaining player 0.90, which is comparable to the expected payment of 0.91 at the truthful equilibrium.

	$r = \text{MM}$	$r = \text{GB}$
$n_f = 0$	0.90	0.80
$n_f = 1$	0.10	1.50
$n_f = 2$	1.50	0.10
$n_f = 3$	0.80	0.90

(a) Treatment 3 payment rule.

	$r = \text{MM}$	$r = \text{GB}$
$n_f = 0$	0.10	0.15
$n_f = 1$	0.10	0.90
$n_f = 2$	1.50	0.15
$n_f = 3$	0.15	0.10

(b) Treatment 4 payment rule.

Table IV: The cell at (r, f_f) gives a player's payment if the player reports r and n_f of his reference reports are MM .

Treatment 4. For treatment 4's payment rule, the MM and GB equilibria pay very little, as shown in Table IVb. At the either MM or GB equilibrium, a player obtains a small payment of 0.15, which is close to the minimum payment of 0.10. In contrast, every player receives 0.50 in expectation at the truthful equilibrium.

We chose this payment rule for two reasons. First, it is impossible to eliminate all of the MM, GB, 3MM1GB and 1MM3GB equilibria for any such 4-player payment rule. Moreover, by supporting the MM and GB equilibria instead of the 3MM1GB and 1MM3GB equilibria, it is possible to make the payments from the coordinating equilibria much less than those of the truthful equilibrium.

Non Peer-Prediction Treatment. For comparison, we would like to understand how players behave when they are not paid by any peer prediction mechanism. In this treatment, each player is paid 0.90 in every round, which is comparable to a player's

expected payment at the truthful equilibrium in treatments 1 and 3. Also, each player plays the game alone, without observing other players' reports. We believe that this setting is closest to how such a constant payment would be used in practice.

5. RESULTS

We collected results on a fairly large scale, recruiting 3533 unique subjects over 65 days for both the pilot and the actual experiment. In the pilot experiment, 705 workers passed the quiz and 542 of them played a total of 181 games. In the actual experiment, 2031 workers passed the quiz and 1988 of them played a total of 861 games.

We received generally positive feedback about the design of our task, suggesting that peer prediction mechanisms can be made accessible to lay participants. Many participants remarked that the game was easy to understand, quick, smooth and enjoyable. 79% of all workers who attempted the quiz eventually passed it, suggesting that the quiz was of appropriate difficulty and the participants had adequate understanding of the mechanism.

5.1. Summary of Data

Figure 2 shows a summary of our experimental data: the percentage of players receiving a particular signal and making a specified report for each round in each treatment. In treatments 1 and 4, the percentage of players with GB signals and MM reports increases, whereas in treatment 2, the percentage of players with MM signals and GB reports increases. The total percentage of misreporting is smallest in treatment 5.

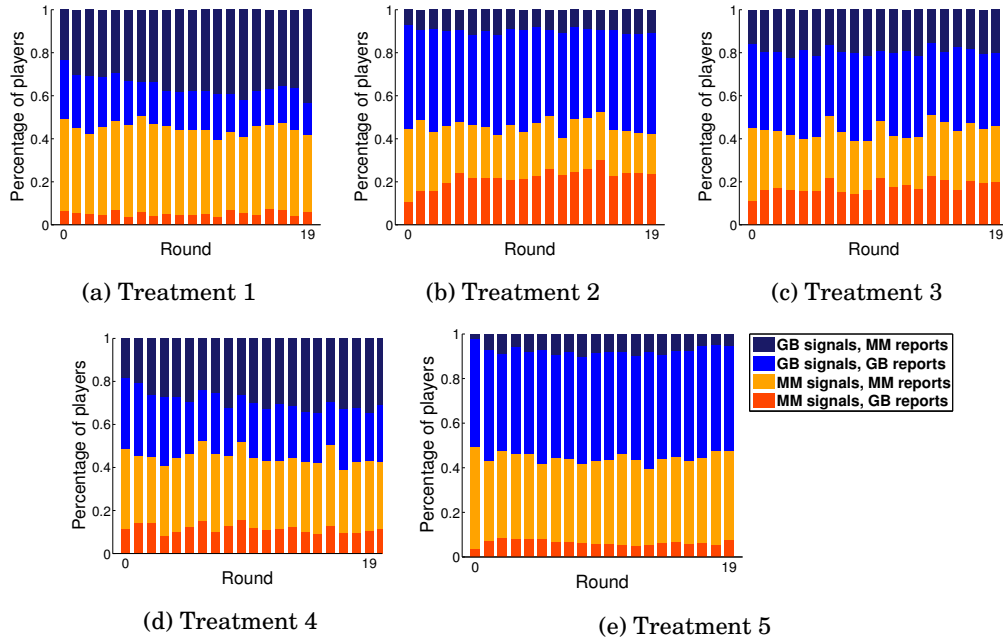


Fig. 2: Percentage of players with the specified signal and report

Table V compares the players' average payoffs in the game with their expected payoff at the truthful equilibrium in treatments 1-4. Compared to a player's expected payoff at the truthful equilibrium, the players' average payoff is higher in treatments 1, 2, and 4, and lower in treatment 3.

	T1	T2	T3	T4
Actual payoff	1.13	1.05	0.87	0.57
Expected payoff at truthful equilibrium	0.91	0.98	0.90	0.50

Table V: Comparison of actual payoff with expected payoff at truthful equilibrium

5.2. Learning a hidden Markov model

Our main goal is to answer the following question:

Will the players reach one of the multiple equilibria of the game, and if so which equilibrium will the players choose and why?

While equilibrium concepts are defined based on players' strategies, we only observe the players' signals and reports in our experiment. In fact, characterizing unobserved strategies is a common challenge for experimentally testing game-theoretic predictions. A common approach is to test whether a player's actions are consistent with an equilibrium strategy, but this heuristic assumes that a player's actions are drawn from a stationary distribution [Selten and Chmura 2008]. Alternatively, one might directly elicit mixed strategies from players in the form of a probability distribution over actions [Noussair and Willinger 2011], but this invasive elicitation method may significantly influence players' behavior.

We use the hidden Markov model [Rabiner 1989] to accurately capture the players' strategies in repeated games. Widely used in speech recognition, natural language processing, and computational biology, the HMM allows us to infer strategies from actions and analyze how they evolve over time without fixing the strategies *a priori*. Recent work has adopted similar probabilistic models to detect latent behavior in repeated games [Shachat et al. 2012; Ansari et al. 2012].

We model the players' behavior as follows. There are K latent states in the HMM, each corresponding to a mixed strategy

$$(\mu_j(\text{MM}, \text{MM}), \mu_j(\text{GB}, \text{MM}))$$

where $\mu_j(s, r)$ is the probability of making a report r when receiving signal s for strategy $j \in \{1, \dots, K\}$. Each player $i \in \{1, \dots, N\}$ chooses his starting strategy from the distribution

$$\pi = (\pi_1, \dots, \pi_K)$$

where π_j is the probability that a player adopts strategy j in the first round. At each time step, players change strategies according a stochastic matrix \mathbf{A} , with

$$\mathbf{A}(j, j') = \Pr(\psi_{t+1}^i = j' \mid \psi_t^i = j)$$

where ψ_t^i is player i 's strategy in round t and $\mathbf{A}(j, j')$ is the probability that a player switches from strategy j to j' from round t to $t + 1$.

This HMM makes several assumptions about players' behavior. A player only chooses strategies among the K states of the HMM. Moreover, a player's distribution over strategies in round $t + 1$ is Markovian and depends only on her strategy in round t . Thus, each player changes his strategy stochastically according to a fixed transition distribution, and this may not capture the scenario when players change their strategies in response to their peers' actions. However, by considering each player independently, the HMM describes the population as a whole instead of capturing the intentions of individual players. It is thus a natural first step in studying the evolution of unobserved strategies.

Our experimental observations are pairs (s_t^i, r_t^i) , corresponding to player i 's signal s_t^i and report r_t^i in round t . The HMM defines the following probability distribution over

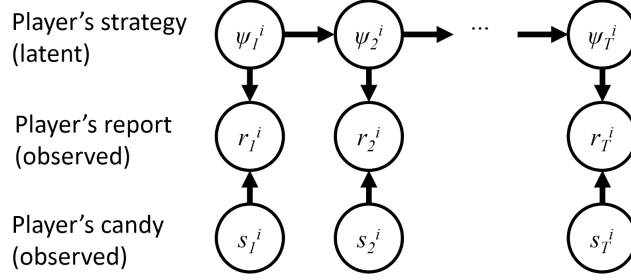


Fig. 3: The graphical model for each player i implied by the HMM.

players' reports based on their signals and model parameters:

$$\Pr(\mathbf{r} \mid \mathbf{s}, \boldsymbol{\pi}, \boldsymbol{\mu}, \mathbf{A}) = \prod_{i=1}^N \left(\prod_{t=2}^T \mu_{\psi_t^i}(s_t^i, r_t^i) \mathbf{A}(\psi_{t-1}^i, \psi_t^i) \right) \mu_{\psi_1^i}(s_1^i, r_1^i) \cdot \boldsymbol{\pi}(\psi_1^i) \quad (2)$$

As shown in Figure 3, this model differs from the canonical HMM as both the observed signal s_t^i and the hidden state ψ_t^i influence the observed action r_t^i , but can be estimated using the same methods. We maximize equation (2) over $\boldsymbol{\pi}$, $\boldsymbol{\mu}$, and \mathbf{A} using the Baum-Welch Expectation-Maximization (EM) algorithm, obtaining maximum likelihood values of the parameters. These parameters reveal the set of strategies, the initial and the long-run distribution over strategies.

Robustness. The likelihood function for the HMM is not log-concave and so the local optimum found by the EM algorithm depends on the initial parameters. To get closer to the global optimum, we ran the EM algorithm with 100,000 restarts with random initial parameters and chose the parameters with the highest log likelihood. As many of these restarts produced equivalent “best” solutions, we feel confident that we found solutions equivalent to the global optimum.

Model selection. The number of states (strategies) K for the HMM may significantly impact our equilibrium convergence analysis. As K increases, there is a diminishing return on the increase in the log likelihood. A common criterion for model selection is the Bayesian information criterion (BIC) [Schwarz 1978]. To maximize the BIC, we chose $K = 4$ for all treatments.

	GB	MM	Truthful	Mixed	Mixed 2
Treatment 1	(0.13, 0.09)	(1.00, 0.99)	(0.99, 0.04)	(0.82, 0.45)	
Treatment 2	(0.01, 0.00)	(0.99, 0.99)	(0.96, 0.01)	(0.60, 0.32)	
Treatment 3	(0.02, 0.03)	(0.87, 0.97)	(0.97, 0.05)	(0.54, 0.42)	
Treatment 4		(0.96, 0.97)	(0.96, 0.06)	(0.73, 0.61)	(0.34, 0.37)
Constant payment	(0.02, 0.00)		(0.98, 0.02)	(0.16, 0.96)	(0.68, 0.34)

Table VI: Each tuple gives the estimated strategy $(\mu_j(\text{MM}, \text{MM}), \mu_j(\text{GB}, \text{MM}))$. All numbers are rounded to 2 decimal places.

Estimated HMM parameters. The 4 states in each HMM are shown in Table VI. The majority of the estimated strategies are close enough to one of the pure strategies (MM, GB or truthful) and are therefore interpreted correspondingly. Notably, these strategies emerged as a result of estimating our model on the data without *a priori* restrictions. Every state in each estimated HMM has a large self-transition probability (≥ 0.9). This suggests that players rarely switch between strategies, and they play each strategy for 10 rounds on average before switching to a different strategy.

When estimating each HMM, the number of strategies K affects how the strategies capture the players' noisy behavior. When K increases, the pure strategies in the HMM become less noisy and closer to their theoretical definitions. For instance, as K increases from 3 to 4, the truthful strategy in treatment 1 changes from (0.91, 0.22) to (0.99, 0.04). When K is small, the strategies are less pure because they must incorporate players' exploratory behavior. As K increases, the pure strategies become less noisy because the noisy behavior can be captured using additional strategies.

Equilibrium convergence using HMM analysis. The set of strategies in each HMM only describes the population in aggregate. To better understand the actions of each player, we use the Viterbi [1967] algorithm to estimate the most likely sequence of strategies used by each player over the multiple rounds of the game. Formally, for given parameters μ , π , and \mathbf{A} , the Viterbi algorithm computes

$$\psi^{i*} = (\psi_1^{i*}, \dots, \psi_T^{i*}) = \underset{\psi^i}{\operatorname{argmax}} \left(\prod_{t=2}^T \mu_{\psi_t^i}(s_t^i, r_t^i) \mathbf{A}(\psi_{t-1}^i, \psi_t^i) \right) \mu_{\psi_1^i}(s_1^i, r_1^i) \cdot \pi(\psi_1^i) \quad (3)$$

Having a sequence of most likely strategies for each player allows us to characterize equilibrium convergence for each game—here, we are referring to convergence to game-theoretic equilibria and not to the stationary distribution of the Markov chain for the HMM. In the following results, we use the values $\psi^{i*}, i \in \{1, \dots, N\}$, to describe strategies and equilibria in games.

Treatment 1 Results. Treatment 1's results are shown in Figure 4. Strikingly, many players adopted more profitable strategies and converged to the uninformative equilibria during the game. In particular, they started with the truthful or mixed strategy but switched to and stayed with the MM or GB strategy until the end (Figure 4a). The MM and GB strategies are that are close to absorbing under the HMM interpretation.

Moreover, players particularly favored the MM equilibrium, which yields the highest payoff. 1/3 of the players used the MM strategy throughout the game and another 1/3 of the players switched to playing the MM strategy during the game (Figure 4a). The percentage of games playing the MM equilibrium increased dramatically from 8% to around 46%, whereas the total percentage of games playing the GB or truthful equilibria remained less than 10% (Figure 4b).

Unsurprisingly, few players are truthful due to the high payoffs of the MM and GB equilibria. Around 15% of the players are truthful throughout the game, but only 5% of the games converged to the truthful equilibrium by the end.

Treatment 2 Results. In treatment 2, we aimed to deter the players from choosing either the MM or GB equilibrium by giving them the same payoff (Figure 5). This was unsuccessful as 36% of the games converged to the GB equilibrium (Figure 5b).

Compared to treatment 1, this treatment is better at promoting truthful behavior in the short term, but not in the long run. More players in this treatment adopted the truthful strategy at the beginning—32% of the players in this treatment compared to 16% in treatment 1 (Figures 4a and 5a). We hypothesize that this happened because it is harder to coordinate on the GB equilibrium than to coordinate on the MM equilibrium in treatment 1. However, by the end of the game, less than 4% of the games in either treatment converged to the truthful equilibrium (Figures 4b and 5b).

Interestingly, players clearly favored the GB equilibrium over the MM equilibrium although they yield the same payoffs. One reason for this seems to be a property of the prior: the probability of the GB signal (55%) is higher than that of the MM signal. If players start by being truthful, this property alone could naturally lead them to the GB equilibrium. Indeed, we observe that nearly all players starting with the truthful strategy who switched, changed to the GB rather than the MM strategy (Figure 5a).

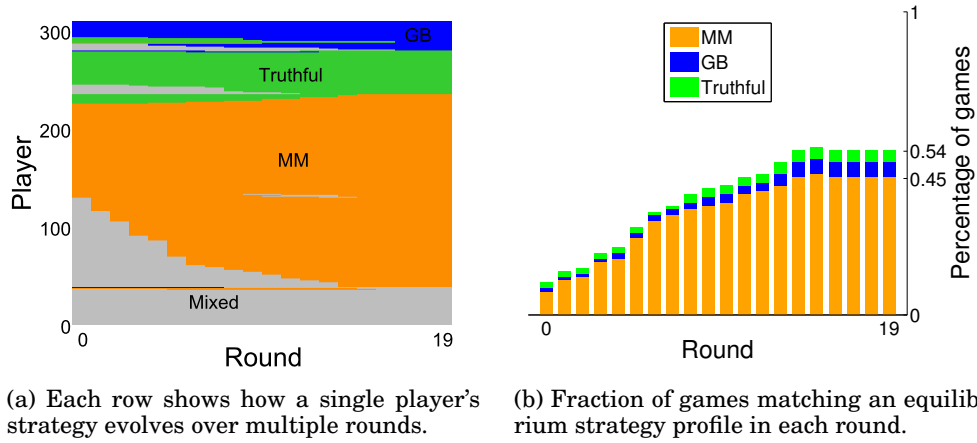


Fig. 4: Treatment 1 Results.

Moreover, players' exit survey answers revealed that they deliberately coordinated on the GB equilibrium once they recognized that other players were more likely to receive the GB candy. These suggest that players have a basic understanding of the prior and can use it to determine their strategies.

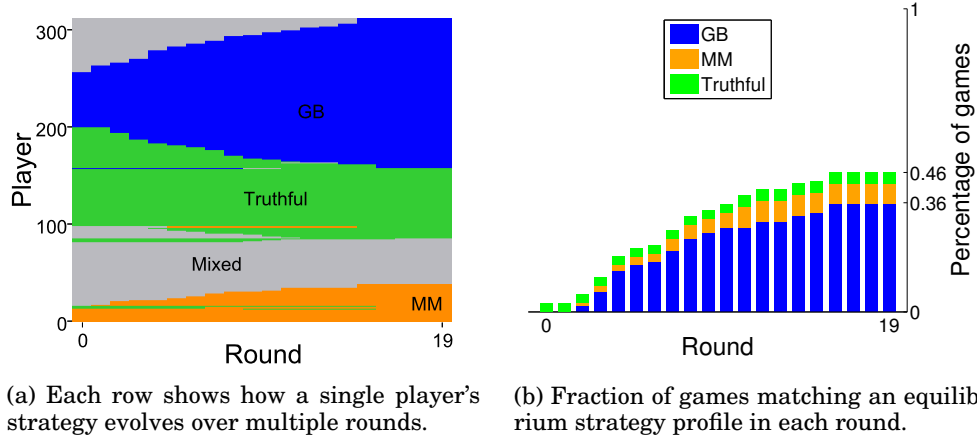


Fig. 5: Treatment 2 Results.

Treatment 3 Results. In this treatment, we aimed to promote truthful behavior by eliminating the symmetric coordinating equilibria, although there still exist asymmetric coordinating equilibria (Figure 6). As expected, it was empirically much more difficult for players to find and reach the asymmetric equilibria versus the symmetric equilibria. Less than 5% of the games converged to an asymmetric equilibrium (Figure 6b) whereas more than $1/3$ of the games converged to a symmetric equilibrium in the first two treatments (Figures 4b and 5b).

However, the increased difficulty of finding a coordinating equilibrium did not promote truthful reporting. Not a single game converged to the truthful equilibrium (Figure 6b) despite 20% of players being truthful over the entire game (Figure 6a).

The mixed strategy $(0.54, 0.42)$ used by more than half of the players may capture the players' random exploration because they are unable to decide on a strategy for this

complex payment rule. Alternatively, this strategy is very close to the random strategy $(0.50, 0.50)$, which is part of a symmetric mixed strategy equilibrium, and 11% of games appear to reach this mixed strategy equilibrium (Figure 6b).

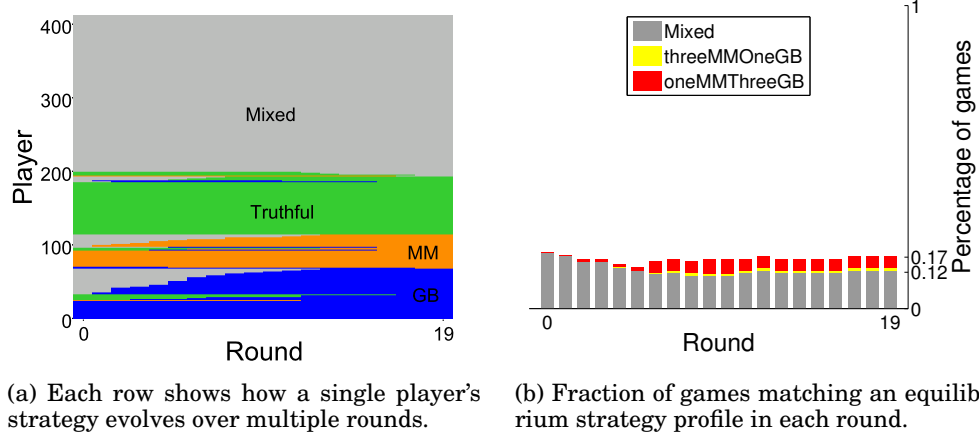


Fig. 6: Treatment 3 Results.

Treatment 4 Results. In this treatment, we made the MM and GB equilibria to have very low payoffs. Similar to treatment 3, this effectively deterred the players from choosing them (Figure 7). Less than 2% of the games reached the MM equilibrium and no game reached the GB equilibrium (Figure 7b).

However, this did not promote truthful behavior. No game reached the truthful equilibrium (Figure 7b) although 16% of the players were truthful during the entire game (Figure 7a).

Players strongly favored the MM strategy over the GB strategy, possibly because of the higher payments in the MM column (Figure IVb). A large percentage (26%) of the players adopted the MM strategy by the end of the game (Figure 7a). In contrast, they seemed to not consider the GB strategy at all as none of the 4 estimated strategies (Table VI) is close to the GB strategy (This is true even for $K = 6$).

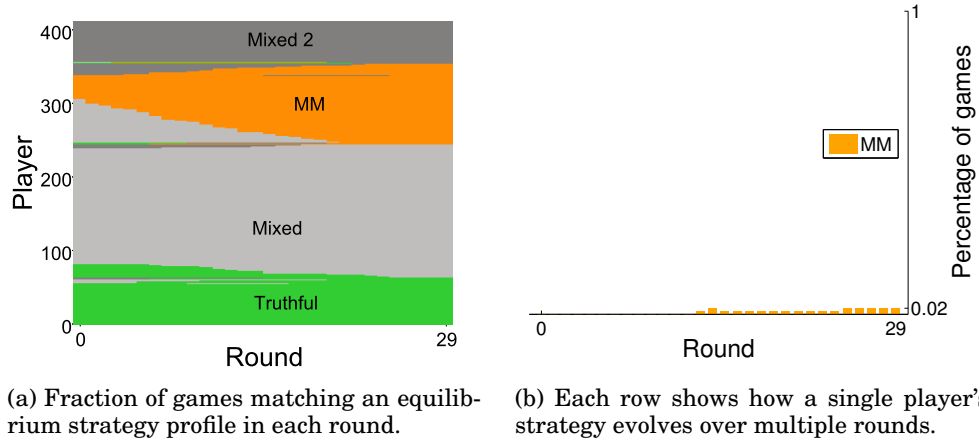


Fig. 7: Treatment 4 Results.

5.3. Simple method for classifying pure strategy equilibria

In contrast to the HMM analysis, a simpler approach to detect equilibrium convergence is to compare players' actions with an equilibrium strategy. Using this simple approach, described below, we classified pure strategy equilibrium convergence and the results are almost identical to those of our HMM analysis.

Consider a game g with a set of players I and T rounds. For each pure strategy equilibrium $e \in E$ of game g with strategy profile s_e , let s_{ei} be player i 's strategy at equilibrium e . For each round, player i is playing strategy s_{ei} if i 's report is observed with positive probability given strategy s_{ei} and i 's signal. A player's report in each round may be consistent with multiple pure strategies. For each player i , we find the earliest round t_{ei} such that player i is playing strategy s_{ei} from round t to the last round. Then, the players in game g are said to have converged to equilibrium e at round $\max_{i \in I} t_{ei}$.

Players in a game may converge to different equilibria at different rounds. Hence, we find the equilibrium for which the convergence occurred earliest in the game, that is the players converged at round $t^* = \min_{s_e} (\max_{i \in I} t_{ei})$ to the equilibrium with strategy profile $s^* = \arg \min_{s_e} (\max_{i \in I} t_{ei})$. If convergence occurred sufficiently early ($t^* \leq (T - 5)$), we classify the game to have converged to the equilibrium with strategy profile s^* . Otherwise, the game remains unclassified.

5.4. Non peer-prediction treatment

Remarkably, in the absence of peer prediction methods, this treatment was much more successful in incentivizing truthful reports than all other treatments. 2/3 of the players reported truthfully during the entire game (Figure 8), despite having no explicit incentive to do so. We conjecture that the reason for these observations is that truthful reporting is the easiest choice for players who are offered a constant payment, since the cost of exploring and adopting alternative strategies may be greater than that of being truthful in the design of our trick or treat game. When peer prediction is used in the other treatments, it may prompt and even motivate the players to explore non-truthful strategies, leading to the strategic play observed. This implies that peer prediction mechanisms may be better suited for scenarios where behaving truthfully is much more costly than acting strategically.

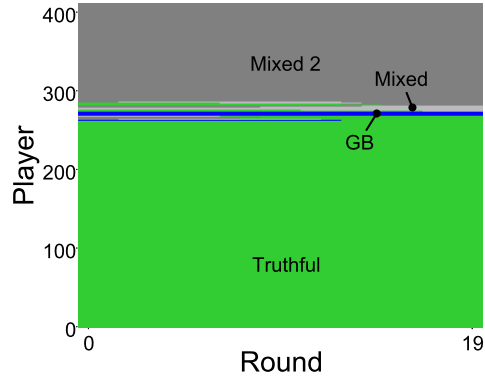


Fig. 8: Each row shows how a single player's strategy evolves over multiple rounds.

6. EXPERIMENTAL CHALLENGES

In designing an online behavioral experiment to test an economic mechanism, we overcame some unique challenges. We believe that awareness of these issues is important for future experimental work.

Coordinating simultaneous participation of multiple workers. Our experiment requires multiple unique workers to participate in each game simultaneously. This is unusual for MTurk, where workers typically participate independently, without interaction, and complete many similar tasks (HITs) at a time.

To coordinate the simultaneous participation of several workers, we built a virtual lobby, where workers wait for enough others to start the game [Mason and Suri 2012]. However, simple posting such tasks on MTurk is not sufficient to ensure a smooth experience for the workers and to collect high quality data. In our pilot experiment, workers experienced a long wait in the lobby because they accepted our tasks at very different times. This resulted in games with workers not paying attention or accidentally disconnecting from the server, causing further frustration for other workers.

We solved this problem by mimicking the recruitment process for lab experiments. In a separate recruitment task, we described our experiment and paid \$0.10 for workers to consent to being contacted. We invite the recruits via email to participate in our experiment at specified times. Each time, we posted the tasks during a 30-minute window to encourage the timely arrival of the participants and to minimize their waiting time. This recruitment process worked extremely well: we completed the experiment significantly more quickly and collected higher quality data, compared to our pilot experiment. By establishing prior history with a worker, we can compensate a worker for his time using a bonus payment if he could not finish the experiment due to technical issues. Many of these features are now part of the TurkServer [Mao et al. 2012] platform, and we encourage its use for deploying similar experiments.

Ensuring attention and comprehension. Using unique participants, we face the challenge of ensuring workers’ attention and comprehension with only a single interaction. Lab experiments can easily ensure participants’ attention with location and time constraints. In contrast, getting online workers’ full attention can be a luxury due to the distractions in their environment. We addressed this issue in several ways. First, we went through several design iterations to ensure that our task was clear, enjoyable and smooth. We also promised a generous bonus payment contingent on workers’ performance in the task, motivating them to pay attention. Finally, we used a built-in tutorial and quiz to ensure that workers have adequate understanding of the task.

The MTurk meta-environment. A platform such as MTurk is a small part of a much bigger online environment. Workers constantly communicate through various forums¹, review requesters on sites such as TurkOpticon², and use third party services like TurkAlert to monitor requester activities. These systems can significantly influence the results of online experiments [Chandler et al. 2013].

While many such communities have policies to protect the integrity of research data, workers may still unwittingly disclose details of our task that we do not want revealed, or they may speculate and reach incorrect conclusions about the purpose of our experiment. They may also share their confusion and frustration with other workers without notifying the requester. The unique worker requirement also makes the task less attractive than other tasks that allow repeated participation. Furthermore, a requester’s reputation on review sites such as TurkOpticon can seriously influence his ability to recruit participants in the future.

As a result, we carefully designed our task and conducted our experiments to provide workers with a positive experience. We promptly responded to email communications from workers at all times, especially during the time window when tasks were posted. We made sure our payments compared fairly to market wages, and paid workers promptly upon task completion. We also extensively monitored workers’ activities on the various online communities, advertising our presence to prevent intentional discussion of our task, and responding quickly to workers’ questions. This also allowed

¹Popular sites include <http://www.turkernation.com>, <http://www.reddit.com/r/mturk>, <http://www.mturkforum.com>, <http://www.cloudmebaby.com>, and <http://www.mturkgrind.com>.

²<http://turkopticon.ucsd.edu>

us to recruit workers from the broader online community. Despite only able to participate once, many workers left positive TurkOpticon reviews for us and said that they would be happy to work on our tasks again.

7. DISCUSSION AND FUTURE WORK

Our results suggest a failure of peer prediction mechanisms in incentivizing truthful reports. Players easily converged to the uninformative equilibria, and hampering these equilibria did not induce truthful behavior. In contrast, players are generally truthful in the absence of economic incentives.

This observation, however, may be due to the similar costs of truthful reporting and behaving strategically in our setting. An interesting direction is to evaluate peer prediction mechanisms when truthful reporting is much more costly than behaving otherwise. Moreover, in our experiment, we assign players to fixed groups and let them play the game repeatedly with the same peers. This essentially tests the peer prediction as a repeated game. While there is an equilibrium of the repeated game where every player in every round adopts his strategy at an equilibrium of the stage game (i.e. the one-shot peer prediction game), the repeated game may have other equilibria where players' behavior in each round is different from their behavior at any equilibrium of the stage game. For future work, to alleviate the repeated game effects, we would like to test peer prediction by recruiting a large number of players and randomly rematching the players for each round of the game. We are also excited about exploring other social, psychological and economic techniques for motivating truthful behavior.

Our work shows the promise of evaluating game theoretic mechanisms through online behavioral experiments, especially for mechanisms inherently designed for an online or crowdsourced setting. Online infrastructures allow for conducting experiments at a much greater scale, as we show in recruiting over 3000 participants for our experiment. In addition, our results also motivate the general use of probabilistic models for analyzing game-theoretic and other experimental data, demonstrating the potential of greatly improved explanatory power of participants' behavior over existing techniques.

ACKNOWLEDGMENTS

The authors are very grateful to Paul Resnick and Grant Schoenebeck for helpful discussions and email exchanges regarding possibly inducing more truthful behavior by making the symmetric coordinating equilibria less attractive.

REFERENCES

- Asim Ansari, Ricardo Montoya, and Oded Netzer. 2012. Dynamic learning in behavioral games: A hidden Markov mixture of experts approach. *Quantitative Marketing and Economics* 10, 4 (2012), 475–503.
- Jesse Chandler, Pam Mueller, and Gabriele Paolacci. 2013. Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods* (2013). DOI: <http://dx.doi.org/10.3758/s13428-013-0365-7>
- Anirban Dasgupta and Arpita Ghosh. 2013. Crowdsourced Judgement Elicitation with Endogenous Proficiency. *ACM International World Wide Web Conference (WWW)* (2013).
- Xi Alice Gao, Yoram Bachrach, Peter Key, and Thore Graepel. 2012. Quality Expectation-Variance Tradeoffs in Crowdsourcing Contests. In *AAAI*.
- John J Horton, David G Rand, and Richard J Zeckhauser. 2011. The online laboratory: conducting experiments in a real labor market. *Experimental Economics* 14, 3 (Feb. 2011), 399–425.

- L K John, G Loewenstein, and Drazen Prelec. 2012. Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science* 23, 5 (May 2012), 524–532.
- R Jurca and B Faltings. 2009. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research* 34, 1 (2009), 209–253.
- Andrew Mao, Yiling Chen, Krzysztof Z. Gajos, David Parkes, Ariel D. Procaccia, and Haoqi Zhang. 2012. TurkServer: Enabling Synchronous and Longitudinal Online Experiments. In *Proceedings of the 4th Workshop on Human Computation (HCOMP'12)*.
- W. Mason and S. Suri. 2012. Conducting Behavioral Research on Amazon's Mechanical Turk. *Behavior Research Methods* 44, 1 (2012), 1–23.
- Nolan Miller, Paul Resnick, and Richard Zeckhauser. 2005. Eliciting Informative Feedback: The Peer-Prediction Method. *MANAGEMENT SCIENCE* 51, 9 (1 Sept. 2005), 1359–1373. DOI: <http://dx.doi.org/10.1287/mnsc.1050.0379>
- Charles Noussair and Marc Willinger. 2011. *Mixed strategies in an unprofitable game: an experiment*. Technical Report.
- Drazen Prelec. 2004. A Bayesian truth serum for subjective data. *Science* 306, 5695 (Oct. 2004), 462–466.
- Drazen Prelec and HS Seung. 2006. An algorithm that finds truth even if most people are wrong. (2006).
- Lawrence R Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 2 (1989), 257–286.
- Goran Radanovic and Boi Faltings. 2013. A Robust Bayesian Truth Serum for Non-binary Signals. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI'13)*. 833–839.
- David G Rand. 2012. The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology* 299, C (April 2012), 172–179.
- Gideon Schwarz. 1978. Estimating the Dimension of a Model. *The Annals of Statistics* 6, 2 (1978), 461–464. DOI: <http://dx.doi.org/10.2307/2958889>
- Reinhard Selten and Thorsten Chmura. 2008. Stationary concepts for experimental 2x2-games. *The American Economic Review* 98, 3 (2008), 938–966.
- Jason Shachat, J Swarthout, and Lijia Wei. 2012. A hidden Markov model for the detection of pure and mixed strategy play in games. *Available at SSRN* (2012).
- AD Shaw, JJ Horton, and DL Chen. 2011. Designing incentives for inexperienced human raters. *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (2011), 275–284.
- Andrew Viterbi. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *Information Theory, IEEE Transactions on* 13, 2 (1967), 260–269.
- Bo Waggoner and Yiling Chen. 2013. Information Elicitation Sans Verification. In *Proceedings of the 3rd Workshop on Social Computing and User Generated Content (SC13)*.
- Jens Witkowski, Yoram Bachrach, Peter Key, and David C. Parkes. 2013. Dwelling on the Negative: Incentivizing Effort in Peer Prediction. In *Proceedings of the 1st AAAI Conference on Human Computation and Crowdsourcing (HCOMP'13)*.
- Jens Witkowski and David C. Parkes. 2012a. A Robust Bayesian Truth Serum for Small Populations. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI'12)*.
- Jens Witkowski and David C Parkes. 2012b. Peer Prediction without a Common Prior. *EC 2012* (April 2012), 1–18.