Learning to Incentivize: Eliciting Effort via Output Agreement

Yang Liu and **Yiling Chen**

Harvard University, Cambridge MA, USA

{yangl,yiling}@seas.harvard.edu

Abstract

In crowdsourcing when there is a lack of verification for contributed answers, output agreement mechanisms are often used to incentivize participants to provide truthful answers when the correct answer is hold by the majority. In this paper, we focus on using output agreement mechanisms to elicit effort, in addition to eliciting truthful answers, from a population of workers. We consider a setting where workers have heterogeneous cost of effort exertion and examine the data requester's problem of deciding the reward level in output agreement for optimal elicitation. In particular, when the requester knows the cost distribution, we derive the optimal reward level for output agreement mechanisms. This is achieved by first characterizing Bayesian Nash equilibria of output agreement mechanisms for a given reward level. When the cost distribution is unknown to the requester, we develop sequential mechanisms that combine learning the cost distribution with incentivizing effort exertion to approximately determine the optimal reward level.

1 Introduction

Our ability to reach an unprecedentedly large number of people via the Internet has enabled crowdsourcing as a practical way for knowledge or information elicitation. For instance, crowdsourcing has been widely used for getting labels for training samples in machine learning. One salient characteristic of crowdsourcing is that a requester often cannot verify or evaluate the collected answers, because either the ground truth doesn't exist or is unavailable or it is too costly to be practical to verify the answers. This problem is called *information elicitation without verification* (IEWV) [Waggoner and Chen, 2014].

In the past decade, researchers have developed a class of economic mechanisms, collectively called the *peer prediction* mechanisms [Prelec, 2004; Miller *et al.*, 2005; Jurca and Faltings, 2006; 2009; Witkowski and Parkes, 2012a; 2012b; Radanovic and Faltings, 2013; Frongillo *et al.*, 2015], for IEWV. The goal of most of these mechanisms is to design payment rules such that participants truthfully report their information at a game-theoretic equilibria. Each of these mechanisms makes some restriction on the information structure of the participants. Under the restriction, truthful elicitation is then achieved by rewarding a participant according to how his answer compares with those of his peers. Within this class, *output agreement mechanisms* are the simplest and they are often adopted in practice [von Ahn and Dabbish, 2004]. In a basic output agreement mechanism, a participant receives a positive payment if his answer is the same as that of a random peer and zero payment otherwise. When the majority of the crowd hold the correct answer, output agreement mechanisms can truthfully elicit answers from the crowd at an equilibrium.

Most of these works on peer prediction mechanisms, with the exception of Dasgupta and Ghosh [2013] and Witkowski *et al.* [2013], assume that answers of participants are exogenously generated, that is, participants are equipped with their private information. However, in many settings, participants can exert more effort to improve their information and hence the quality of their answers is endogenously determined. Recent experiments [Yin and Chen, 2015; Ho *et al.*, 2015] have also shown that the quality of answers can be influenced by the magnitude of contingent payment in settings where answers can be verified.

In this paper, we study eliciting efforts as well as truthful answers in output agreement mechanisms. Taking the perspective of a requester, we ask the question of how to optimally set the payment level in output agreement mechanisms when the requester cares about both the accuracy of elicited answers and the total payment.

Specifically, we focus on binary-answer questions and binary effort levels. We allow workers to have heterogeneous cost of exerting effort. Such a cost is randomly drawn from a distribution that is common knowledge to all participants. We consider two scenarios. In the first scenario, a static setting, the requester is assumed to know the cost distribution of the participants. Her objective is to set the payment level in output agreement mechanisms such that when a game-theoretic equilibrium is reached, her expected utility is maximized. In the second scenario, a dynamic setting, the data requester doesn't know the cost distribution of the participants but only knows an upper bound of the cost. Here, the requester incorporates eliciting and learning the cost distribution into incentivizing efforts in output agreement mechanisms when she

repeatedly interacts with the set of participants over multiple tasks. The ultimate goal of the requester is to learn to set the optimal payment level in this sequential variant of output agreement mechanism for each interaction so that when participants reach a game-theoretic equilibrium of this dynamic game, the data requester minimizes her regret on expected utility over the sequence of tasks.

We summarize our main contributions as follows:

- Since the quality of answers is endogenously determined, a requester's utility depends on the behavior of participants. Optimizing the payment level requires an understanding of the participant's behavior. We characterize Bayesian Nash equilibria (BNE) for two output agreement mechanisms with any given level of payment and show that at equilibrium there is a unique threshold strategy for positive effort exertion.

- For the static setting where the requester knows the cost distribution, when the cost distribution satisfies certain conditions, we show that the optimal payment level in the two output agreement mechanisms is a solution to a convex program and hence can be efficiently solved.

- For the dynamic setting where the requester doesn't know the cost distribution, we design a sequential mechanism that combines eliciting and learning the cost distribution with incentivizing effort exertion in a variant of output agreement mechanism. Our mechanism ensures that participants truthfully report their cost of effort exertion when asked, in addition to following the same strategy on effort exertion and answer reporting as that in the static setting for each task. We further prove performance guarantee of this mechanism in terms of the requester's regret on expected utility.

All omitted proofs can be found in [Liu and Chen, 2016].

1.1 Related work

The literature on peer prediction mechanisms hasn't addressed costly effort until recently. Dasgupta and Ghosh [2013] and Witkowski *et al.* [2013] are the two papers that formally introduce costly effort into models of information elicitation without verification. Dasgupta and Ghosh [2013] design a mechanism that incentivizes maximum effort followed by truthful reports of answers in an equilibrium that achieves maximum payoffs for participants. Witkowski *et al.* [2013] focuses on simple output agreement mechanisms as this paper. They study the design of payment rules such that only participants whose quality is above a threshold participate and exert effort. Both Dasgupta and Ghosh [2013] and Witkowski *et al.* [2013] assume that the cost of effort exertion is fixed for all participants and is known to the mechanism designer. This paper studies effort elicitation in output agreement mechanisms but allow participants to have heterogeneous cost of effort exertion drawn from a common distribution. Moreover, we consider a setting where the mechanism designer doesn't know this cost distribution, which leads to an interesting question of learning to optimally incentivize effort exertion followed by truthful reports of answers in repeated interactions with a group of participants.

Roth and Schoenebeck [2012] and Abernethy *et al.* [2015]

consider strategic data acquisition for estimating the mean and statistical learning in general respectively. Both works do not consider costly effort but participants may have stochastic and heterogeneous cost for revealing their data and need to be appropriately compensated. Moreover, these two works all assume that workers won't misreport their obtained answers.

2 Problem formulation

2.1 Our mechanisms

A data requester has a set of tasks that she wants to obtain answers from a crowd $\mathcal{C} = \{1, ..., K\}$ of $K \geq 2$ candidate workers. In this paper, we consider binary-answer tasks, for example, identifying whether a picture of cells contains cancer cells, and denote the answer space of each task as $\{0, 1\}$. The requester assigns each task to N randomly selected workers, with $N \geq 2$ being potentially much less than K.[1] Such a redundant assignment strategy, when combined with some aggregation method (e.g. majority voting), has been found effective in obtaining accurate answers [Sheng *et al.*, 2008; Liu and Liu, 2015].

The requester cannot verify the correctness of contributed answers for a task, either because ground truth is not available or verification is too costly and defies the purpose of crowdsourcing. Thus, in addition to a base payment, each worker is rewarded with a contingent bonus that is determined by how his answer compares with those of other workers for completing a task. Specifically:

1. The requester assigns a task to a randomly selected subset $\mathcal{U} \subseteq \mathcal{C}$ of workers, where $|\mathcal{U}| = N$. She announces a base payment $b > 0$ and a bonus $B > 0$, as well as the criteria for receiving the bonus. The criteria of receiving the bonus is specified by an output agreement mechanism, which we will introduce shortly

2. Each worker $i \in \mathcal{U}$ independently submits his answer $L_i \in \{0, 1\}$ to the requester.

3. After collecting the answers, the requester pays base payment b to every worker who has submitted an answer and a bonus B to those who met the specified criteria.

The criteria for receiving bonus B is specified by an output agreement mechanism. *Output agreement* is a term introduced by von Ahn and Dabbish [2008] to capture the idea of "rewarding agreement" in their image labeling game, the ESP game [von Ahn and Dabbish, 2004]. We define two variants of output agreement mechanisms:

Peer output agreement (PA): For each worker $i \in \mathcal{U}$, the data requester randomly selects a *reference worker* $j \neq i$ and $j \in \mathcal{U}$. If $L_i = L_j$, worker i receives bonus B. Note worker j's reference worker could be different from i.

Group output agreement (GA): For each worker $i \in \mathcal{U}$, the data requester compares L_i with the majority answer of the rest of the workers, L_M, where $L_M = 1$ if $\frac{\sum_{j \in \mathcal{U}, j \neq i} L_j}{N-1} >$

---

[1] We assume N is fixed, though how to optimally choose N could be an interesting future direction.

0.5, $L_M = 0$ if $\frac{\sum_{j \in \mathcal{U}, j \neq i} L_j}{N-1} < 0.5$ and $L_M = \{0,1\}$ if $\frac{\sum_{j \in \mathcal{U}, j \neq i} L_j}{N-1} = 0.5$. If $L_i \in L_M$, worker i receives bonus B.

2.2 Agent models

A worker can decide how much effort to exert to complete a task and the quality of his answer stochastically depends on his chosen effort level. Specifically, a worker can choose to either exert or not exert effort. If a worker exerts effort, then with probability $P_H \leq 1$ his answer is correct. If a worker does not exerts effort, with probability P_L, where $P_L < P_H$, he will provide the correct answer. We further assume $P_L \geq 0.5$, that is, when no effort is exerted the worker can at least do as well as random guess. This assumption is also used by Dasgupta and Ghosh [2013] and Karger *et al.* [2011, 2013]. For now, we assume P_L and P_H are the same for all workers.

Since workers can choose their effort level, the quality of an answer is endogenously determined. Let $e_i \in \{0,1\}$ represents the chosen effort level of worker i, with 0 corresponding to not exerting effort and 1 corresponding to exerting effort. The accuracy of worker i can be represented as $p_i(e_i) = P_H e_i + P_L(1 - e_i)$.

Workers have heterogeneous abilities, which are reflected by their cost of exerting effort. When worker i doesn't exert effort, he incurs zero cost. A cost of $c_i \geq 0$ is incurred if agent i chooses to exert effort on a task. c_i is randomly generated according to a distribution with pdf $f(c)$ and cdf $F(c)$ for each pair of (worker, task). We further assume this distribution stays the same[2] across all workers and all tasks, and it has a bounded support $[0, c_{\max}]$. Moreover we enforce the following assumption on $F(c)$:

Assumption 2.1. $F(c)$ is strictly concave on $c \in [0, c_{\max}]$.

This assumption is stating that the probability of having a larger cost c_i is decreasing. Several common distributions, e.g. exponential and standard normal (positive side), satisfy this assumption. Throughout this paper, we assume $F(c), c \in [0, c_{\max}]$ is common knowledge among all workers.[3] Nevertheless each realized cost c_i is private information, that is each worker i observes his own realized cost c_i, but not the one for others. In Section 3, we assume the requester also has full knowledge of $F(\cdot)$, but we relax this assumption in Section 4.

Given that the cost of not exerting effort is zero, the positive base payment b ensures that every worker will provide an answer for a task assigned to him. We focus on understanding how to determine the bonus B in output agreement mechanisms to better incentivize effort in this paper. The base payment b doesn't enter our analysis directly but it allows us to not worry about workers' decisions on participation. When reporting their answer to the data requester, workers can choose to report truthfully, or to mis-report. Denote this decision variable for each worker i as $r_i \in \{0,1\}$, where $r_i = 1$ represents worker i truthfully reporting his answer, and $r_i = 0$ represents worker i mis-reporting (reverting the

---

[2]Realization for each (worker, task) pair can be very different.

[3]In practice each worker can estimate such distribution based on their past experiences.

answer in our case). Then the accuracy of each worker i's report is a function of (e_i, r_i):

$$p_i(e_i, r_i) = p_i(e_i)r_i + (1 - p_i(e_i))(1 - r_i).$$

When each worker $j \in \mathcal{U}$ takes actions (e_j, r_j), we denote the probability that worker $i \in \mathcal{U}$ receives bonus B as $P_{i,B}(\{(e_j, r_j)\}_j)$. In the PA mechanism, this quantity is

$$P_{i,B}(\{(e_j, r_j)\}_j) = \frac{\sum_{j \neq i} P(L_i = L_j)}{N - 1}$$

In the GA mechanism, it is $P_{i,B}(\{(e_j, r_j)\}_j) = P(L_i = L_M)$. Then, the utility for worker i is:

$$u_i(\{(e_j, r_j)\}_j) = b - e_i c_i + B \cdot P_{i,B}(\{(e_j, r_j)\}_j).$$

2.3 Requester model

The data requester has utility function U_D, which in theory can be of various forms balancing accuracy of elicited answers and total budget spent. In this paper, we assume that the requester uses majority voting to aggregate elicited answers and has utility function

$$U_D(B) = P^c(N, B) - b N - B N_e(B),$$

where $P^c(N, B)$ is the probability that the majority answer is correct, and $N_e(B)$ is the number of workers who receive the bonus. Data requester's goal is to find a B^* s.t.

$$B^* \in \text{argmax}_{B \in \mathbf{R}^+} P^c(N, B) - b N - B E[N_e(B)]. \quad (1)$$

$P^c(N, B)$ and $E[N_e(B)]$ depend on workers' strategy towards effort exertion and answer reporting. The equilibrium analysis in the next section will help us define these quantities. The data requester is then hoping to choose a reward level that maximizes the expected utility at an equilibrium.

3 Optimal bonus strategy with known cost distribution

In this section we set out to find the optimal bonus strategy when the data requester knows workers' cost distribution. Because the requester's utility depends on the behavior of workers, we first characterize symmetric Bayesian Nash Equilibria (BNE) for the two output agreement mechanisms for an arbitrary bonus level B. Then based on workers' equilibrium strategies, we show the optimal B^* can be calculated efficiently for certain cost distributions. Note that due to the independence of tasks, this is a static setting and we only need to perform the analysis for a single task.

3.1 Equilibrium characterization

For any given task, we have a Bayesian game among workers in $\mathcal{U}$. A worker's strategy in this game is a tuple $(e_i(c_i), r_i(e_i))$ where $e_i(c_i) : [0, c_{\max}] \rightarrow \{0, 1\}$ specifies the effort level for worker i when his realized cost is c_i and $r_i(e_i) : \{0, 1\} \rightarrow \{0, 1\}$ gives the reporting strategy for the chosen effort level, with $r_i(e_i) = 1$ representing reporting truthfully and $r_i(e_i) = 0$ representing misreporting.

We first argue that at any Bayesian Nash equilibrium (BNE) of the game, $e_i(c_i)$ must be a threshold function. That is, there is a threshold c_i^* such that $e_i(c_i) = 1$ for all $c_i \leq c_i^*$

and $e_i(c_i) = 0$ for all $c_i > c_i^*$ at any BNE. The reason is as follows: suppose at a BNE worker i exerts effort with cost c_i. Since the other workers' outputs do not depend on c_i (due to the independence of reporting across workers), worker i's chance of getting a bonus will not change when he has a cost $c_i' < c_i$ and only obtains a higher expected utility by exerting effort. This allow us to focus on threshold strategies for effort exertion. We restrict our analysis to symmetric BNE where every worker has the same threshold for effort exertion, i.e. $c_i^* = c^*$. In the rest of the paper, we often use $(c^*, \cdot)$ to denote that a worker playing an effort exertion strategy with threshold c^*. In addition, we use $r_i \equiv 1$ to denote the reporting strategy that $r_i(1) = r_i(0) = 1$, i.e. always reporting truthfully for either effort level.

Before we present our main results, we note we have a set of zero-effort exertion equilibriums ($c^* = 0$). One set of such equilibriums is no one exerting effort combined with truthful or non-truthful reporting, or a mix of the two. Nonetheless these equilibriums return strictly less expected utility for each worker. Another one is when workers collude to always report the same uninformative signal.[4] In this paper we mainly focus on positive effort exertion equilibriums, that is $c^* > 0$.

PA: We have the following results for the PA mechanism.

Lemma 3.1. *The strategy profile $\{(c^*, r_i \equiv 1)\}_{i \in \mathcal{U}}$ is a symmetric BNE for the PA game if*

$$2(P_H - P_L)F(c^*) + 2P_L - 1 = \frac{c^*}{(P_H - P_L)B}. \quad (2)$$

Denote $B_{PA} := \frac{c_{\max}}{(2P_H - 1)(P_H - P_L)}$, the minimum bonus level needed to induce full effort exertion. With above lemma, we have the following equilibrium characterization.

Theorem 3.2. *When $P_L > 0.5$, there exists a unique threshold $c^* > 0$ s.t. $(c^*, 1)$ is a symmetric BNE for the PA game:*

- *When $B \geq B_{PA}$, $c^* = c_{\max}$.*
- *O.w. c^* is the unique solution to Eqn. (2).*

When $P_L = 0.5$, we can prove similar results for the uniqueness of $c^* > 0$ such that $(c^*, 1)$ is a symmetric BNE: but under certain condition non-effort exertion is the only equilibrium. We would like to note that always mis-reporting ($r_i \equiv 0$) combined with the same threshold c^* for effort exertion as in Theorem 3.2 is also a symmetric BNE when $P_L > 0.5$. This equilibrium gives workers the same utility as the equilibrium in Theorem 3.2. This phenomenal has also been observed by Dasgupta and Ghosh [2013] and Witkowski *et al.* [2013]. Dasgupta and Ghosh [2013] argue that always mis-reporting is risky, and workers may prefer breaking the tie towards always truthful reporting.

GA: For GA, directly calculating the probability term for matching a majority voting is not easy; but if we adopt a

Chernoff type approximation for it, and suppose such approximation is common knowledge[5], we can prove similar results.

Similar to Lemma 3.1, we can show that the strategy profile $\{(c^*, r_i \equiv 1)\}_{i \in \mathcal{U}}$ is a symmetric BNE for the GA game if

$$1 - 2[(\alpha - 1)F(c^*) + 1]^{N-1} = c^*/(B(P_H - P_L)). \quad (3)$$

where $\alpha := e^{-2(P_H - P_L)^2}$.

Denote $B_{GA} := \frac{c_{\max}}{(1 - 2\alpha^{N-1})(P_H - P_L)}$ we have:

Theorem 3.3. *When $P_L > 0.5$, there always exists a unique threshold $c^* > 0$ such that $\{(c^*, r_i \equiv 1)\}_{i \in \mathcal{U}}$ is a symmetric BNE for the GA game:*

- *When $B \geq B_{GA}$, $c^* = c_{\max}$.*
- *O.w., c^* is the unique solution to Eqn. (3).*

We note the reward level and the total expected payment is lower in GA than in PA for eliciting the same level of efforts:

Lemma 3.4. *Denote the smallest bonus level corresponding to an arbitrary equilibrium threshold $c^* > 0$ for PA and GA as $B_{PA}(c^*)$ and $B_{GA}(c^*)$ respectively. Then $B_{PA}(c^*) > B_{GA}(c^*)$, when $N \geq \frac{-\log(1 - P_H)}{2(P_L - 0.5)^2} + 1$ (sufficiently large). Furthermore, the total payment in GA is lower than that in PA, i.e., GA leads to a higher requester utility.*

Heterogeneity of P_L and P_H. So far we have assumed that P_L and P_H are the same for all workers. If workers have heterogeneous accuracy $\{P_L^i, P_H^i\}_i$ that are generated from some distribution with mean P_L, P_H, we can show that the above results hold in a similar way, with more involved arguments.

3.2 Optimal solution for data requester

Now consider the optimization problem stated in Eqn. (1) for the requester's perspective. For each $B > 0$, denote $\{(c, r_i \equiv 1)\}_{i \in \mathcal{U}}$ as the corresponding strategy profile at equilibrium. $P^c(N, B)$ can then be calculated based on $c, F(c)$ (controlling how much effort can be induced), and P_L, P_H. Same can be done for $E[N_e(B)]$. Denote the optimization problem in (1) with above calculation as (PB). Directly investigating the two objective functions may be hard. We seek to relax the objectives. First of for PA, we will be omitting the $2P_L - 1$ term as when P_L is only slightly larger than 0.5, this quantity is close to 0. Also for both PA and GA, we again use the Chernoff type of approximation for calculating P_B^c. We further introduce three conditions: (i) $f(c)$ is twice differentiable and $\partial^2 f(c)/\partial^2 c \geq 0$. (ii) $cF(c)$ is convex on $c \in [0, c_{\max}]$. (iii) $G(c) := 1 - [(\alpha - 1)F(c^*) + 1]^{N-1}$ satisfies that $\partial^3 G(c)/\partial^3 c$ exists and being non-negative.

Lemma 3.5. *If (i) and (ii) hold, the objective function of (PB) is concave if we adopt PA. When (ii) and (iii) hold, the objective function of (PB) is concave if we adopt GA.*

For example, exponential distribution ($\exp(\lambda)$) for $c_{\max} \leq 2/\lambda$ satisfies (i)&(ii) for PA; and $\exp(\lambda)$ for $c_{\max} = \frac{-\log \alpha}{\lambda}$ satisfies (ii)&(iii) for GA. It is worth to note above results hold for a wide range of other $U_D(\cdot)$s: for instance the ones with a linear combination of $P^c(N, B)$ and $E[N_e(B)]$.

---

[4]Interested readers can find more discussions and solutions for the colluding scenario in [Dasgupta and Ghosh, 2013].

[5]This is not entirely unreasonable as in practice this Chernoff type bounds are often used to estimate such majority voting probability term, when the number of variables (population) is large.

4 Learning the optimal bonus strategy

In this section we propose a sequential mechanism for learning the optimal bonus strategy, when the requester has no prior knowledge of the cost distribution but only knows $c_{\max}$. This assumption can be further relaxed by assuming knowing an upper bound of $c_{\max}$ instead of knowing $c_{\max}$ precisely. Also similar as last section, P_L, P_H are known. In reality these two quantities can be estimated through a learning procedure by repeated sampling and output matching as shown in [Liu and Liu, 2015], via setting bonus level $B := 0$ and $B := B_{\mathrm{PA(GA)}}$[6] respectively (to induce effort level corresponding to P_L, P_H). In this work we focus on learning the cost functions, which is a more challenging task when the workers are strategic. We are in a dynamic setting where the requester sequentially ask workers to complete a set of task. *(P1):* We start our discussions with a simpler case. When asked to report their cost, workers maximize their collected utility from a set of data elicitation tasks and are not aware of the potential influence of their reports on calculating optimal bonus levels for any future tasks. The data requester's goal is to elicit cost data to estimate cost distribution and then the optimal bonus level $\tilde{B}^*$, such that when $\tilde{B}^*$ is applied to a *newly arrived* task we can bound $|U_D(\tilde{B}^*) - U_D(B^*)|$. where B^* is the optimal bonus level if the cost distribution is known. *(P2):* We then consider the case when workers are forward looking and are aware of that their reported cost on a task will be utilized to calculate optimal bonus strategy for future tasks. We form a sequential learning setting, where we separate the stages for task assignment into two types: one for data elicitation, which we also refer as *exploration*, and the other for utility maximization, which we refer as *exploitation*. The data requester's objective in this case is to minimize the *regret* defined as follows:

$$R(T) = \sum_{t=1}^{T} E|U_D(\{B_i(t)\}_{i \in \mathcal{U}}) - U_D(B^*_{\mathrm{GA}})|, \quad (4)$$

where B^*_{GA} is the optimal bonus level for GA when cost distribution is know,[7] and $\{B_i(t)\}_{i \in \mathcal{U}}$ is the bonus bundle offered at time t. Note $U_D(\cdot)$ is mechanism dependent: both $P^c(N, B)$ and $N_e(B)$ depends on not only the bonus level, but also the equilibrium behavior in a particular mechanism.

For simplicity of presentation, throughout this section we consider $P_L > 0.5$: this is to remove the ambiguity introduced in by the trivial equilibrium $c^* = 0$. Also we assume with the same expected utility, workers will favor truthful reporting $r_i \equiv 1$.

4.1 (M_Crowd) for (P1)

Suppose the data requester allocates T tasks to elicit the cost data sequentially, and exactly one of them is assigned to the workers at each time step $t = 1, 2, \cdots, T$. For simplicity of analysis we fix the set of N workers we will be assigning tasks to. Denote worker i's realized cost for the t-th task as $c_i(t)$. We propose mechanism (M_Crowd):

---

[6]Calculating $B_{\mathrm{PA(GA)}}$ only requires knowing $c_{\max}$, not $F(\cdot)$.

[7]Since GA is more cost efficient, we define the regret w.r.t. the optimal utility that can be obtained via GA.

Mechanism 1 (M_Crowd)

For each step t:

1. Assign the task, and workers then report costs. Denote the reports as $(\tilde{c}_1(t), ..., \tilde{c}_N(t))$. This is a voluntary procedure. A worker i can choose to not report his cost, in which case, the requester sets $\tilde{c}_i(t) := c_{\max}$.

2. Data requester randomly selects a threshold $c^*(t)$ *uniformly* from the support $[0, c_{\max}]$, such that only the workers who reported $\tilde{c}_i(t) \leq c^*(t)$ will be considered for bonus following PA; others will be given a bonus according to a probability that is independent of workers' report (see Remarks for details).

3. The requester estimates a bonus level $\tilde{B}_i(t)$ for each worker i that corresponds to the threshold level $c^*(t)$ under PA, using *only* the data collected from user $j \neq i$, and from *all* previous stages. This is done via estimating $F(\cdot)$ first and then plugging it into Eqn.(2). Then the requester adds a positive perturbation $\delta(t)$ that is time dependent to $\tilde{B}_i(t)$: $B_i(t) := \tilde{B}_i(t) + \delta(t)$.

4. The data requester will then announce the bonus bundle $[B_1(t), ..., B_N(t)]$.

Remarks: 1. When a worker, say worker i, reports higher than the selected threshold, his probability of receiving a bonus will be calculated using the following experiment, which is independent of his output: suppose out of N workers, there are $N(t)$ of them reported lower than $c^*(t)$. Then we will "simulate" N workers' reports with the following coin-toss procedure: toss $N(t)$ P_H-coin and $N - N(t)$ P_L-coin. Assign a P_L-coin toss to worker i, and select a reference answer from the rest of the tosses, and compare their results. If there is a match, worker i will receive a bonus. Simply put, the probability for receiving a bonus can be calculated as the matching probability in the above experiment. 2. Since we have characterized the equilibrium equation for PA with a clean and simple form, this set of equilibriums is good for eliciting workers' data. 3. After estimating the bonus level for each worker $\tilde{B}_i(t)$, the data requester will add a positive perturbation term $\delta(t)$ to each of them. This is mainly to remove the bias introduced by (i) imperfect estimation due to finite number of collected samples, and (ii) the (possible) mis-reports from workers. Such term will become clear later in the stated results. 4. The fact that we can use collected cost data to estimate B depends crucially on the assumption that the cost distribution is the same for all tasks.

4.2 Equilibrium analysis for (M_Crowd)

We present the main results for characterizing how workers report their cost at an equilibrium. Even though we are in a dynamic game setting, we use BNE rather than Perfect Bayesian equilibrium (PBE) as our general solution concept as they do not have material differences in our setting. Due to the independence of tasks, workers make decisions on effort exertion at each stage just as they are in a static game with corresponding reward level. Because a worker's decision on effort exertion can not be directly observable, a deviation from an equilibrium effort exertion strategy cannot

be detected by others, making the notion of "off-equilibrium path" not meaningful in this setting. This suggests that every BNE of the game is also a PBE. We will adopt ϵ-approximate BNE as our exact solution concept, defined as follows:

Definition 4.1. *A set of reporting strategy* $\{\tilde{\mathbf{c}}_{\mathbf{i}} := \{\tilde{c}_i(t)\}_{t=1,\dots,T}\}_{i \in \mathcal{C}}$ *is ϵ-BNE if for any i, $\forall \tilde{\mathbf{c}}_{\mathbf{i}}' \neq \tilde{\mathbf{c}}_{\mathbf{i}}$ we have*

$$\sum_{t=1}^{T} (E[\max_{e_i, r_i} u_i^t(\tilde{\mathbf{c}}_{\mathbf{i}}, \tilde{\mathbf{c}}_{-\mathbf{i}})] - E[\max_{e_i, r_i} u_i^t(\tilde{\mathbf{c}}_{\mathbf{i}}', \tilde{\mathbf{c}}_{-\mathbf{i}})])/T \geq \epsilon.$$

We explicitly denote the expected utility for each worker as a function of $\{\tilde{\mathbf{c}}_{\mathbf{i}}\}_{i \in \mathcal{C}}$. This is a short-hand notation, as u_i^ts also depend on the effort exertion and reporting strategies. The $\max_{e_i, r_i}$ term allows worker i to optimize his effort exertion and reporting procedure based on their cost reporting.

Theorem 4.2. *With (`M_Crowd`), set $\delta(t) := O(\sqrt{\frac{\log t}{t}})$, let $\gamma > 0$ being arbitrarily small, there exists a $O(\frac{\log^2 T}{T})$-BNE for each worker i with reporting $\tilde{c}_i(t)$ at time t such that*

$$\max\{c_i(t) - \epsilon_1(t), 0\} \leq \tilde{c}_i(t) \leq \min\{c_i(t) + \epsilon_2(t), c_{\max}\},$$

where $0 \leq \epsilon_1(t) = o(\sqrt{\log t/t})$, $0 \leq \epsilon_2(t) = o(1/t^{2-\gamma})$.

The effort exertion game at each step looks alike the static game introduced in Section 3 with the following difference: instead of workers who have cost $c_i(t) \leq c^*$ will exert effort, now it is the workers who reported $\tilde{c}_i(t) \leq c^*$ will exert effort. This is mainly due to the addition of the perturbation term to the estimated bonus level. Meanwhile the mechanism excludes workers who reported higher than the threshold from exerting effort by offering bonus with a probability that is independent of worker's output. Nevertheless, we can bound the fraction of workers whose actions are different for the above two games.

4.3 Performance of (`M_Crowd`)

With this set of collected data, we bound the performance loss in offering optimal bonus level B for an incoming task (or task $T + 1$). Suppose we adopt GA, where the optimal bonus level with known cost distribution is given by B_{GA}^*, and the estimated optimal solution is given by $\tilde{B}_{GA}^*$. We will have the following lemma: (similar results hold for PA)

Lemma 4.3. *With probability being at least $1 - \eta$,*

$$|U_D(\tilde{B}_{GA}^*) - U_D(B_{GA}^*)| = o(\sqrt{\frac{\log 2/\eta}{2NT}} + \sqrt{\frac{\log T}{T}}) .$$

When we chose $\eta = O(1/T^2)$, the above regret term is roughly on the order of $\sqrt{\log T/T}$.

4.4 (`RM_Crowd`) for (P2)

We propose a (`RM_Crowd`) for **(P2)**:

Remarks: 1. The dependence on T is to simplify the presentation and our algorithm design. This can be easily extended to a T-independent one. 2. At exploitation phases

---

Mechanism 2 (`RM_Crowd`)

Specify a constant $0 < z < 1$, and initialize $t = 1$. Define $p(t) := \min\{1, \frac{\log T}{t^{1-z}}\}$.
At time t, assign the task; workers then report costs.
Toss a $p(t)$-coin.
When HEAD, algorithm enters *exploration* phase:

- Follow same steps as in (`M_Crowd`).

When TAIL, algorithm enters *exploitation* phase,

1. estimate the optimal $B_i(t)$ and its corresponding threshold $c_i^*(t)$ for each worker i with GA, using the cost data collected *only* from the *exploration* phases. Only workers who reported $c_i(t) \leq c_i^*(t)$ will be given bonus according to GA; others receive bonus with a probability that is independent of her report.
2. Follow rest steps in (`M_Crowd`).

---

we assume there exists a solver that can find the optimal solution with a noisy estimation of $F(\cdot)$. In practice search heuristics can help achieve the goal. 3. We adopted different bonus mechanisms for different phases. When we calculate the bonus level according to a particular mechanism (PA or GA), we will also adopt it for evaluating workers' answers. 4. When using GA, the independent probability for giving out bonus when a report is higher than the threshold will be adjusted to a probability of matching a majority voting of the experiment we presented for (`M_Crowd`).

Theorem 4.4. *With (`RM_Crowd`), set $\delta(t) := O(zt^{-z/2})$, let $z > 1/3$ and $\gamma > 0$ being arbitrarily small, there exists a $O(\frac{z^2}{T^z})$-BNE for worker i reporting $\tilde{c}_i(t)$ at time t that:*

$$\max\{c_i(t) - \epsilon_1(t), 0\} \leq \tilde{c}_i(t) \leq \min\{c_i(t) + \epsilon_2(t), c_{\max}\},$$

where $0 \leq \epsilon_1(t) = o(z/t^{z/2})$, $0 \leq \epsilon_2(t) = o(1/t^{3z-1-\gamma})$.

We have similar observations for the effort exertion game in (`RM_Crowd`) as we made for (`M_Crowd`). Further we prove the following regret results:

Lemma 4.5. $R(T) \leq O(T^z \log T + T^{1-z/2})$.

Order-wise, the best z is when $1 - z/2 = z \Rightarrow z = \frac{2}{3}$, which leads to a bound on the order of $O(T^{2/3} \log T)$.

5 Conclusion

In this paper we focus on using output agreement mechanisms to elicit effort, in addition to eliciting truthful answers, from crowd-workers when there is no verification of their outputs. Workers' cost for exerting efforts are stochastic and heterogeneous. We characterize the symmetric BNE for workers' effort exertion and reporting strategies for a given bonus level, and show data requester's optimal bonus strategy at equilibrium is a solution to a convex program for certain cost distribution. Then a learning procedure is introduced to help the requester learn the optimal bonus level via eliciting cost data from strategic workers. We bound the mechanism's performance loss w.r.t. offering the best bonus bundle, compared to the case when workers' cost distribution is known a priori.

Acknowledgement

References

[Abernethy *et al.*, 2015] Jacob Abernethy, Yiling Chen, Chien-Ju Ho, and Bo Waggoner. Actively purchasing data for learning. 2015.

[Dasgupta and Ghosh, 2013] Anirban Dasgupta and Arpita Ghosh. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd international conference on World Wide Web*, pages 319–330. International World Wide Web Conferences Steering Committee, 2013.

[Frongillo *et al.*, 2015] Rafael M. Frongillo, Yiling Chen, and Ian A. Kash. Elicitation for Aggregation. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI'15)*, 2015.

[Ho *et al.*, 2015] Chien-Ju Ho, Aleksandrs Slivkins, Siddharth Suri, and Jennifer Wortman Vaughan. Incentivizing high quality crowdwork. In *Proceedings of the 24th International Conference on World Wide Web*, pages 419–429. International World Wide Web Conferences Steering Committee, 2015.

[Jurca and Faltings, 2006] R. Jurca and B. Faltings. Minimum payments that reward honest reputation feedback. In *Proceedings of the 7th ACM conference on Electronic commerce*, EC '06, pages 190–199. ACM, 2006.

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

[Karger *et al.*, 2011] David R Karger, Sewoong Oh, and Devavrat Shah. Iterative learning for reliable crowdsourcing systems. In *Advances in neural information processing systems*, pages 1953–1961, 2011.

[Karger *et al.*, 2013] David R Karger, Sewoong Oh, and Devavrat Shah. Efficient crowdsourcing for multi-class labeling. In *ACM SIGMETRICS Performance Evaluation Review*, volume 41, pages 81–92. ACM, 2013.

[Liu and Chen, 2016] Yang Liu and Yiling Chen. Learning to incentivize: Eliciting effort via output agreement. *arXiv preprint arXiv: 1604.04928*, April 2016.

[Liu and Liu, 2015] Yang Liu and Mingyan Liu. An online learning approach to improving the quality of crowdsourcing. In *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '15, pages 217–230, New York, NY, USA, 2015. ACM.

[Miller *et al.*, 2005] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Management Science*, 51(9):1359 – 1373, 2005.

[Prelec, 2004] Dražen Prelec. A bayesian truth serum for subjective data. *Science*, 306(5695):462–466, 2004.

[Radanovic and Faltings, 2013] G. Radanovic and B. Faltings. A robust bayesian truth serum for non-binary signals. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, AAAI '13, 2013.

[Roth and Schoenebeck, 2012] Aaron Roth and Grant Schoenebeck. Conducting truthful surveys, cheaply. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 826–843. ACM, 2012.

[Sheng *et al.*, 2008] Victor S Sheng, Foster Provost, and Panagiotis G Ipeirotis. Get another label? improving data quality and data mining using multiple, noisy labelers. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 614–622. ACM, 2008.

[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. ACM, 2004.

[von Ahn and Dabbish, 2008] Luis von Ahn and Laura Dabbish. Designing games with a purpose. *Communications of the ACM*, 51(8):58–67, 2008.

[Waggoner and Chen, 2014] Bo Waggoner and Yiling Chen. Output agreement mechanisms and common knowledge. In *Proceedings of the 2nd AAAI Conference on Human Computation and Crowdsourcing*, HCOMP'14, 2014.

[Witkowski and Parkes, 2012a] J. Witkowski and D. Parkes. A robust bayesian truth serum for small populations. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, AAAI '12, 2012.

[Witkowski and Parkes, 2012b] J. Witkowski and D.C. Parkes. Peer prediction without a common prior. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 964–981. ACM, 2012.

[Witkowski *et al.*, 2013] Jens Witkowski, Yoram Bachrach, Peter Key, and David C. Parkes. Dwelling on the Negative: Incentivizing Effort in Peer Prediction. In *Proceedings of the 1st AAAI Conference on Human Computation and Crowdsourcing (HCOMP'13)*, 2013.

[Yin and Chen, 2015] Ming Yin and Yiling Chen. Bonus or Not? Learn to Reward in Crowdsourcing. In *In the Proc. of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15), 2015*, 2015.