Parametric Prediction from Parametric Agents

Yuan Luo[1], Nihar B. Shah[2], Jianwei Huang[1], and Jean Walrand[2]

[1]The Chinese University of Hong Kong
[2]UC Berkeley*

Abstract

We consider a prediction problem based on opinions elicited from multiple agents. Making an accurate prediction with a minimal cost requires a *joint* design of incentive mechanisms and prediction algorithms. Such a problem lies at the interface of statistical learning theory and game theory, and arises in many domains such as the consumer surveys and mobile crowdsourcing. Under a fairly general problem setup, we jointly design a mechanism and algorithm to achieve the *optimal* system objective. Our results offer several valuable engineering insights. When the costs incurred by the agents are linear in the exerted effort, then the optimal mechanism calls for a "crowd-tender", where the principal only employs the agent with the lowest bid. When the costs are quadratic, then it is optimal to use a "crowd-sourcing" mechanism that employs multiple agents at the same time. Synthetic simulations demonstrate the gains achieved under our proposed mechanism, as compared to those that do not account for the heterogeneity of the agents.

1 Introduction

Prediction algorithms are often designed under the assumption that the training data is provided to the algorithm, and that the algorithm has no control over the quality of the training data. In many situations, however, the training data is collected by surveying people, for instance, in the prediction of the future demand for a product by surveying a number of potential customers [Hay98], or the prediction of the winner of an election by surveying potential voters [WR11]. Collecting data from people is much cheaper, easier and faster today due to the emergence of several commercial crowdsourcing platforms such as Amazon Mechanical Turk and others. In such situations, it is possible to monetarily incentivize the respondents to provide higher quality inputs.

In any realistic setup, the responses obtained from people ("the agents") are noisy: one cannot expect a naive customer to gauge the sales of a product accurately. Moreover, every individual has a different expertise and ability, and will likely react differently to the amount of money paid per task. For example, some people may be active users of the surveyed product, therefore have a better understanding of its anticipated usage. We assume that the surveyor ("the principal") has no knowledge of the behavior of individual agents. It is therefore important to design an appropriate incentive mechanism for the prediction procedure that exploits the heterogeneity of the agents, motivating them to participate and exert suitable levels of effort. An appropriate incentive will provide higher quality data and as result, a superior prediction performance. This requirement motivates the problem at the interface between statistical estimation and mechanism design considered in this paper.

As compared to problems that tackle only one of the prediction and the mechanism design problems, the problem of joint design poses a significantly greater challenge. From the statistical prediction point

*E-mail address: ly011@ie.cuhk.edu.hk, nihar@eecs.berkeley.edu, jwhuang@ie.cuhk.edu.hk, walrand@berkeley.edu

of view, the challenge is that every sample is drawn from a different distribution, whose properties are unknown apriori to the principal. From the mechanism design perspective, the challenge is that the incentivization procedure not only needs to ensure that agents report truthfully, but also needs to ensure that each agent exerts an effort that minimizes the overall prediction error. In this paper, we formulate and optimally solve a "parametric" form of this joint design problem. More specifically, the principal desires to predict a parameter of a known distribution. Each agent is modeled in a parametric fashion, with her work quality (or expertise) governed by a single parameter that is the agent's private information. While each agent aims to maximize her own expected payoff (i.e., the revenue minus the cost of effort), the principal must optimize a joint utility that trades off the prediction error and the monetary costs.[1]

Our contributions. We design a mechanism (which we call "COPE", for COst and Prediction Elicitation) that jointly optimizes the principal's payoff in terms of the payments made to the agents and the prediction error incurred. COPE provides a systematic way for the principal to incentivize all participating agents to report their estimations truthfully and exert appropriate amounts of effort based on their respective capabilities. The mechanism incorporates and exploits the heterogenity of the agents in terms of their capabilities and costs, in order to minimize the prediction error.

Our COPE mechanism operates under a fairly general framework, encompassing a wide range of distributions of the parameters for prediction as well as the cost functions of the agents. In this paper, however, we focus on two special settings to gain engineering insights towards the design of such mechanisms. We investigate the special scenario where the noise follows a Gaussian distribution, and study the impact of two specific cost functions on the principal's decision. Our results show that when the costs incurred by the agents are linear in the amount of exerted effort, the principal should conduct a *crowd-tender*, soliciting service of only the agent with the lowest reported cost. On the other hand, when the costs are quadratic in the exerted effort, the optimal mechanism is that of *crowd-sourcing*, where the principal recruits multiple agents to complete the task. We subsequently comment on the structure of the optimal mechanism in the general case, but defer the details for a future version.

Related literature. Mechanism design for truthful elicitation of agents' opinions is an extensively studied problem, most recently investigated in [CJ15, MRZ05, Pre04, SZP15, DG13, TPPZ15] in the context of modern crowdsourcing setups. In contrast to our work, this line of literature does not consider the estimation aspect, and only focuses on the elicitation problem alone. Mechanism design for truthful elicitation of agents' opinions is also studied in the context of prediction markets (e.g., see [WZ04, Con09]). These works, however, study the scenario where the agents take the responsibility of aggregating information. Our paper concerns a different setting and objective in which the principal is in charge of information gathering and making the final prediction.

A parallel line of literature considers the problem of efficiently predicting or estimating underlying parameters *given* the agents' responses. The problems addressed there include estimation of an objective ground truth [IPSW14, KOS11, ZBMP12], or inference of subjective preferences [NOS12, SBB+15]. Here, the agents are assumed to be noisy, but not strategic, and there are no monetary considerations.

The scenario turns quite different when prediction must be done taking incentives into account, and calls for the design of new procedures catering to both aspects. The recent works [FSW07, DFP08, FCK15, CDP15] address problems in this space. The work that is closest to ours is that of [CDP15] which considers a setting that involves both elicitation and estimation. There are several differences between [CDP15] and our work. First, in [CDP15], the relationship between each agent's effort and the noise in its observation is assumed to be known. Our work, however, assumes this relation is governed by a parameter that is *unknown* to the principal, which adds a non-trivial complexity to the problem. We address this issue by eliciting the cost types of the agents and ensuring that they are incentivized to report their cost types correctly. Second, it is assumed in [CDP15] that the agent always reports

[1] For ease of exposition, we refer to the principal as "he" and each agent as "she".

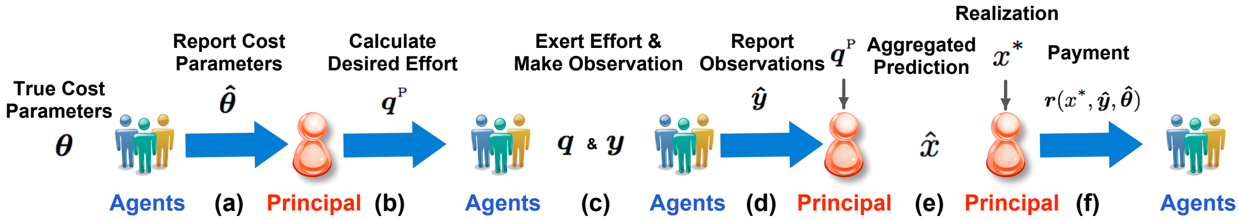


Figure 1: Sequence of interactions between the principal and the agents.

truthfully once she makes an observation. Our work, in contrast, accommodates strategic agents, and truth-telling is no longer automatically achieved. In order to address this key difficulty, we add a scoring-rule component, where we assume that the principal eventually observes the true value of the parameter. This assumption is not required in [CDP15]. Making this assumption limits the scope of the application, but makes the joint optimization of elicitation and estimation possible.

2 Problem Setting

We begin with a formal description of the problem formulation. The description requires us to set up notation to capture the behavior of the agents, the objective of the principal, the prediction problem, and the mechanism-design problem.

2.1 System Model

We consider a setting where the principal wishes to make a *parametric prediction*, that is, to form an informed estimate about a parameter $x^* \in \mathcal{X} \subseteq \mathbb{R}$. Predicting the winner of an election is a motivating example. We assume that x^* has a prior distribution that is publicly known, for instance, from the results of an earlier election. We assume that the principal will come to know the precise value of x^* sometime in the future, for instance, upon completion of the election.

Figure 1 pictorially depicts the interaction between agents and the principal. The individual components of the figure are explained in the following text.

Observation and Reporting: The principal's prediction is based on queries made by the principal to a set $\mathcal{A} = \{1, \ldots, N\}$ of N agents. When queried, an agent can put in some effort to form an "observation" whose value is known only to that agent. We assume that the distribution of the observation y_n comes from a parameterised family of distributions $\phi(x^*, q_n)$, where q_n represents the effort exerted by agent n to make observation. The higher value of q_n, the more effort agent n exerts, and thus the better quality of agent n's observation. An example that we focus on subsequently in the paper is additive Gaussian noise, with

$$y_n \sim \mathcal{N}(x^*, \frac{1}{q_n}). \tag{1}$$

Conditioned on x^*, the observations of the agents are assumed to be mutually independent. As a shorthand, we let $\boldsymbol{y} = [y_1, y_2, \ldots, y_N]^T$. We assume that agents do not collude with each other.

Effort Level and Cost Type: Every agent $n \in \mathcal{A}$'s performance is governed by two parameters, q_n and θ_n, whose values are known privately only to that agent. The parameter $q_n \geq 0$ introduced earlier represents the amount of effort exerted by agent $n \in \mathcal{A}$ in making her observation. The agent is free to choose the value of q_n, and a higher value of q_n leads to a less noisy observation (see (1) for instance). We let $\boldsymbol{q} = [q_1, q_2, \ldots, q_N]^T$ as a shorthand.

Each agent $n \in \mathcal{A}$ incurs some cost to exert effort when making an observation, and this cost is governed by the agent's cost-type parameter θ_n. The cost types of different agents are allowed to be different, capturing the heterogenity of the agents. A smaller value of θ_n implies a higher capability of agent n. Specifically, we consider a publicly known cost function $C : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ and assume that the cost incurred by an agent $n \in \mathcal{A}$ (having the cost type θ_n) when exerting an effort q_n is $C(q_n, \theta_n)$.

The cost types $\{\theta_n\}_{n\in\mathcal{A}}$ are assumed to be random, independently and identically distributed on support $[\underline{\theta}, \bar{\theta}]$ for some $0 \le \underline{\theta} < \bar{\theta} < \infty$. This distribution is assumed to be public knowledge. In this paper we focus on the case where the distribution is uniform on the interval $[\underline{\theta}, \bar{\theta}]$. We let $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_N]^T$.

Reporting Observations and Making Payments: The principal employs monetary incentives to ensure that agents make their observations and report them to the principal. In order to incentivize agents to participate the prediction task, the payment to an agent must, at the least, cover the cost incurred by that agent in putting effort to make the observation. However, since each agent's cost parameter is known only to that agent, the principal needs to ask each agent to report her own cost type (Figure 1a). The agents are strategic, and any agent $n \in \mathcal{A}$ may report a cost type $\hat{\theta}_n$ that is different from her true cost type θ_n. Let $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N]^T$.

As we will see subsequently, incentivizing different agents to put different levels of effort depending on their respective cost types allows for a significantly better prediction performance. The principal must incentivize these different effort levels, and the choice of these effort levels is based on the agents' reported cost types $\hat{\boldsymbol{\theta}}$ (Figure 1b). Let function $Q^{\text{P}} : [\underline{\theta}, \bar{\theta}]^N \to \mathbb{R}_+$ denote the effort that the principal requires an agent to exert. The function Q^{P} depends on the cost types reported by the agents: $Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$ represents the effort required from agent $n \in \mathcal{A}$, where $\hat{\boldsymbol{\theta}}_{-n} = [\hat{\theta}_1, \ldots, \hat{\theta}_{n-1}, \hat{\theta}_{n+1}, \ldots, \hat{\theta}_N]^T$ is the reported cost parameters of all agents except agent n. Here (and elsewhere in the paper), we use the superscript "P" to represent the principal. In order to simplify notation, we will henceforth use q_n^{P} as a shorthand for $Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$. We let $\boldsymbol{q}^{\text{P}} = [q_1^{\text{P}}, \ldots, q_N^{\text{P}}]^T$.

Each agent $n \in \mathcal{A}$ is strategic and may exert an effort $q_n \neq q_n^{\text{P}}$ to suit her own interests. When choosing the effort to exert, the agent may also exploit the fact that the principal cannot directly observe the actual effort exerted. Upon exerting the chosen effort q_n, the agent obtains an observation y_n (Figure 1c). The principal seeks the value of the observation y_n, but the agent may report a strategically chosen value $\hat{y}_n \neq y_n$ to the principal (Figure 1d) that suits her own interests. We adopt the shorthand $\hat{\boldsymbol{y}} = [\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_N]^T$. Based on the information obtained, the principal must make a prediction for the value of x^* (Figure 1e).

The principal makes payment to each agent once he observes the true value of x^*. Specifically, we define the payment function as $R : \mathbb{R} \times \mathbb{R} \times [\underline{\theta}, \bar{\theta}]^N \times \to \mathbb{R}_+$; the payment to agent n is $R(x^*, \hat{y}_n, \hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$, which depends on the value of x^*, the agent n's reported observation $\hat{y}_n$, and all agents' reported cost parameters $\hat{\boldsymbol{\theta}}$.

As indicated above, the model considered is a one-shot model where the principal communicates with all agents simultaneously and not sequentially.

Payoff of Agent: Given the payment function announced by the principal, each agent n's payoff $U^{\text{A}} : \mathbb{R} \times [\underline{\theta}, \bar{\theta}] \times \mathbb{R}_+ \times \mathbb{R} \times [\underline{\theta}, \bar{\theta}]^N \to \mathbb{R}_+$ is defined as the difference between the payment received from the principal and the cost incurred in making the observation, and is given as

$$U^{\text{A}}(x^*, \hat{\theta}_n, q_n, \hat{y}_n, \theta_n, \hat{\boldsymbol{\theta}}_{-n}) = R(x^*, \hat{y}_n, \hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) - C\big(q_n, \theta_n\big), \tag{2}$$

Here the superscript "A" indicates a term associated to the agents. The above equation shows that agent $n \in \mathcal{A}$'s payoff also depends on other agents' reported cost parameters $\hat{\boldsymbol{\theta}}_{-n}$. When each agent $n \in \mathcal{A}$ chooses her strategy, i.e., her cost reported value $\hat{\theta}_n$, exerted effort q_n, and the reported observation $\hat{y}_n$, to maximize her expected payoff, she assumes that all other agents report truthfully. The expected payoff of the agent n is calculated as

$$\mathbb{E}[U^{\text{A}}(x^*, \hat{\theta}_n, q_n, \hat{y}_n, \theta_n, \boldsymbol{\theta}_{-n})] = \mathbb{E}[R(x^*, \hat{y}_n, \hat{\theta}_n, \boldsymbol{\theta}_{-n})] - C(q_n, \theta_n),$$

where the expectation is taken with respect to the distributions of x^* and all agents' cost parameters $\boldsymbol{\theta}_{-n}$. Recall that each agent n only knows her own cost parameter θ_n, and only has distributional information about other agents' cost parameters.

Payoff of the Principal: Let $\ell^{\mathrm{P}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ be the loss function that characterises the penalty term for mistakes in the prediction. For instance, one could consider the squared loss $\ell^{\mathrm{P}}(x^*, \hat{x}) = (x^* - \hat{x})^2$ as the penalty for the principal. We measure the utility gained by the principal through the prediction in terms of the *Bayes risk* incurred under this loss function: If all agents report their true observations (i.e., $\hat{\boldsymbol{y}} = \boldsymbol{y}$) and exert efforts as desired by the principal (i.e., $\boldsymbol{q} = \boldsymbol{q}^{\mathrm{P}}$), then the *Bayes risk* $B^{\mathrm{P}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ is

$$B^{\mathrm{P}}(\boldsymbol{q}^{\mathrm{P}}) = \inf_{\hat{x}} \mathbb{E}[\ell^{\mathrm{P}}(x^*, \hat{x}(\boldsymbol{y}, \boldsymbol{q}^{\mathrm{P}}))],$$

where the expectation is taken with respect to x^* and $\boldsymbol{y}$. The net payoff of the principal $U^{\mathrm{P}} : \mathbb{R} \times \mathbb{R}^N \times [\underline{\theta}, \bar{\theta}]^N \to \mathbb{R}$ is then defined as the difference between his utility obtained from prediction and the monetary payments to all agents:

$$U^{\mathrm{P}}(x^*, \boldsymbol{q}^{\mathrm{P}}, \boldsymbol{y}, \hat{\boldsymbol{\theta}}) = -B^{\mathrm{P}}(\boldsymbol{q}^{\mathrm{P}}) - \sum_{n \in \mathcal{A}} R(x^*, y_n, \hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}). \tag{3}$$

Here, we assumed without loss of generality that the monetary payment and the prediction loss is normalized to be on the same scale.

2.2 Design Objective

Before we explain our objective, we begin by defining two standard game-theoretic terms (applied to our setting) that are required for subsequent discussions.

Definition 1. *(BIC: Bayesian Incentive Compatibility) A mechanism is said to be Bayesian incentive compatible (BIC) if for every agent $n \in \mathcal{A}$, her expected payoff satisfies*

$$\mathbb{E}\big[U^{\mathrm{A}}(x^*, \theta_n, q_n^{\mathrm{P}}, y_n, \theta_n, \boldsymbol{\theta}_{-n})\big] \geq \mathbb{E}\big[U^{\mathrm{A}}(x^*, \hat{\theta}_n, q_n, \hat{y}_n, \theta_n, \boldsymbol{\theta}_{-n})\big] \quad \forall (\hat{\theta}_n, q_n, \hat{y}_n) \neq (\theta_n, q_n^{\mathrm{P}}, y_n), \tag{4}$$

where the expectation is taken with respect to x^* and all other agents cost parameters $\boldsymbol{\theta}_{-n}$.

BIC means that for any agent n, reporting the true cost parameter, exerting the effort requested by the principal, and reporting true observation will maximize her expected payoff, given common knowledge about the distribution on agents cost parameters and when other agents are truthfully report their cost parameters.

Definition 2. *(BIR: Bayesian Individual Rationality) A mechanism is said to be Bayesian incentive rationality (BIR), if the expected payoff of every agent $n \in \mathcal{A}$ is non-negative, given that she reports truthfully, exerts effort as the principal desires, and assumes that all other agents report their cost parameters truthfully, that is,*

$$\mathbb{E}\big[U^{\mathrm{A}}(x^*, \theta_n, q_n^{\mathrm{P}}, y_n, \theta_n, \boldsymbol{\theta}_{-n})\big] \geq 0 \quad \forall n \in \mathcal{A}, \tag{5}$$

where the expectation is taken with respect to x^* and all other agents' cost parameters $\boldsymbol{\theta}_{-n}$.

Assuming (without loss of generality) that the payoff of an agent not participating in this process equals zero, BIR means that an agent will participate only if her expected payoff is at least as much as that of a non-participating agent.

Based on the above definitions, the problem that we want to solve is formalized as follows. The goal is to design a mechanism, say $\mathcal{M}$, that maximizes the principal's utility while ensuring truthful reports from the agents:

$$\begin{aligned} &\sup_{\mathcal{M}} \ \mathbb{E}\big[U^{\mathrm{P}}(x^*, \boldsymbol{q}^{\mathrm{P}}, \boldsymbol{y}, \hat{\boldsymbol{\theta}})\big] \\ &\text{subject to: BIC and BIR in (4) and (5),} \end{aligned} \tag{6}$$

where the expectation is taken with respect to x^*, $\boldsymbol{y}$ and $\boldsymbol{\theta}$. In words, the goal is to design a mechanism such that: (i) the principal's payoff is maximized in expectation; (ii) the principal can elicit truthful information from all agents; and (iii) the principal can incentivise suitable effort from the agents based on their respective cost parameters.

3 The COPE Mechanism

In this section, we present an optimal mechanism which we call "COPE" (COst and Prediction Elicitation) to solve (6). In this paper, we choose to focus on two specific cases, enumerated below. The reason for our choice is that these two settings illustrate all the key ideas behind the general construction, are easier to understand, and offer some concrete engineering insights. We subsequently provide a brief commentary on the mechanism in the general case; we defer the details to a subsequent version of the paper.

We consider the following pair of specific settings in this section. We consider the Gaussian case, where we assume the prior $x^* \sim \mathcal{N}(\mu_0, \sigma_0^2)$, and the observation of every agent n follows the distribution $y_n \sim \mathcal{N}(x^*, \frac{1}{q_n})$. The values of μ_0 and σ_0 are assumed to be public knowledge. We assume $\theta_n \sim \text{Uniform}[\underline{\theta}, \bar{\theta}]$, independent for every $n \in \mathcal{A}$. We consider the squared ℓ_2-loss to measure the prediction error, namely, $\ell^{\mathrm{P}}(x^*, \hat{x}) = (x^* - \hat{x})^2$. We consider two cost functions: (i) linear cost function $C(q, \theta) = q\theta$, and (ii) quadratic cost function $C(q, \theta) = \frac{1}{2}\theta q^2$.

3.1 Linear Cost Function $C(q, \theta) = q\theta$

We first consider the linear cost function $C(q, \theta) = q\theta$ and discuss the corresponding COPE mechanism. Algorithm 1 presents the higher-level structure of the mechanism; the working of the mechanism crucially relies on the careful construction of specific functions referred to in the algorithm, and these constructions are described below.

Algorithm 1: COPE

Step 1: The principal announces a payment function R
Step 2: Every agent $n \in \mathcal{A}$ independently reports a cost type $\hat{\theta}_n \in [\underline{\theta}, \bar{\theta}]$ to the principal
Step 3: The principal sends each agent $n \in \mathcal{A}$ a contract which specifies the effort level q_n^{P} along with values of functions $\pi(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$, $K(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$, and $S(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$ that comprise the function R
Step 4: Each agent $n \in \mathcal{A}$ exerts effort q_n and makes an observation y_n
Step 5: Each agent $n \in \mathcal{A}$ reports an estimate $\hat{y}_n$
Step 6: The principal makes prediction $\hat{x}$
Step 7: The true value x^* is realized
Step 8: The principal makes the payment $R(x^*, \hat{y}_n, \hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$ to every agent $n \in \mathcal{A}$

Recall that the function $Q^{\mathrm{P}} : [\underline{\theta}, \bar{\theta}] \times [\underline{\theta}, \bar{\theta}]^{N-1} \to \mathbb{R}_+$ specifies the effort that the principal requires an agent to exert, based on the cost parameters reported by all agents. In Theorem 1 subsequently, we show that when the cost function is linear, the principal requires *only one* agent to exert effort. This property is reflected in the following choice of function Q^{P}:

$$Q^{\mathrm{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) = \begin{cases} \max\{(2\hat{\theta}_n - \underline{\theta})^{-\frac{1}{2}} - \sigma_0^{-2}, 0\} & \text{if } n = \arg\min_{m\in\mathcal{A}} \hat{\theta}_m \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

The function Q^{P} is designed to strike an optimal balance between the prediction error and the monetary expenditure, accommodating the fact that the agents are heterogeneous. As a shorthand, we let $q_n^{\mathrm{P}} = Q^{\mathrm{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$ be the value of the effort that the principal requires the agent n to exert. We also define $n_0 = \arg\min_{m\in\mathcal{A}} \hat{\theta}_m$, that is, n_0 is the agent with the lowest reported cost parameter.

We now characterize the function R that governs the payment made by the principal to the agents. The payments to all agents other than agent n_0 are zero since these agents are not involved in the observation and prediction procedure. The payment made to agent n_0 is

$$R(x^*, \hat{y}_{n_0}, \hat{\theta}_{n_0}, \hat{\boldsymbol{\theta}}_{-n_0}) = \pi(\hat{\theta}_{n_0}, \hat{\boldsymbol{\theta}}_{-n_0}) - (x^* - \hat{y}_{n_0})^2 \cdot K(\hat{\theta}_{n_0}, \hat{\boldsymbol{\theta}}_{-n_0}) + S(\hat{\theta}_{n_0}, \hat{\boldsymbol{\theta}}_{-n_0}), \tag{8}$$

where

$$\begin{aligned} \pi(\hat{\theta}_{n_0}, \hat{\boldsymbol{\theta}}_{-n_0}) =& \hat{\theta}_{n_0}(2\hat{\theta}_{n_0} - \underline{\theta})^{-\frac{1}{2}} - \bar{\theta}\sigma_0^{-2} + 2[(2\bar{\theta} - \underline{\theta})^{\frac{1}{2}} - (2\hat{\theta}_n - \underline{\theta})^{\frac{1}{2}}], \\ K(\hat{\theta}_{n_0}, \hat{\boldsymbol{\theta}}_{-n_0}) =& \hat{\theta}_{n_0}(2\hat{\theta}_{n_0} - \underline{\theta})^{-1}, \quad S(\hat{\theta}_{n_0}, \hat{\boldsymbol{\theta}}_{-n_0}) = \hat{\theta}_{n_0}(2\hat{\theta}_{n_0} - \underline{\theta})^{-\frac{1}{2}}. \end{aligned}$$

Let us explain the main ideas behind this construction. The term $(x^* - \hat{y}_{n_0})^2$ in (8) ensures that the agent reports her observation truthfully. Since no other terms in (8) depend on $\hat{y}$, and $K > 0$, the agent must necessarily report a certain deterministic function of her observation y_n in order to maximize her own expected payoff. The choices of functions K and S ensure that the agent's expected payoff is maximized only when the agent chooses $q_{n_0} = q_{n_0}^{\mathrm{P}}$. This property ensures that the agent exerts an effort as desired by the principal. Now given that the agent's behavior is guaranteed to be truthful with respect to $q_{n_0}^{\mathrm{P}}$ and y_{n_0}, the expected value of R simply equals to the value of the function π. This function is designed to ensure that the agent truthfully reports her cost parameter in the first place. Interestingly, in the case of linear costs, given the identity of agent n_0, the payment function R is independent of $\hat{\boldsymbol{\theta}}_{-n_0}$. Finally, the predictor $\hat{x}$ employed by the principal is the standard Bayes estimator operating on the agents' responses $\hat{x} = \hat{y}_{n_0}$.

We prove that the proposed COPE mechanism is indeed guaranteed to work as claimed.

Theorem 1. *Under the linear cost function $C(q, \theta) = q\theta$, COPE is feasible and maximizes the principal's expected payoff.*

An important consequence of the theorem is that the optimal mechanism in the case of linear costs awards the task to the single agent with the lowest bid. When costs are linear, this result therefore suggests the practitioner to build what we call a "crowd-tender" system where all agents submit their cost parameters, and the lowest bidder is awarded the task.

3.2 Quadratic Cost Function $C(q, \theta) = \frac{1}{2}\theta q^2$

We now consider a quadratic cost function $C(q, \theta) = \frac{1}{2}\theta q^2$ and present COPE for this setting. The higher level structure of COPE is again given by Algorithm 1, and the design of the specific functions referred to in the algorithm is provided below. Under COPE, the function $Q^{\mathrm{P}} : [\underline{\theta}, \bar{\theta}] \times [\underline{\theta}, \bar{\theta}]^{N-1} \to \mathbb{R}_+$ that governs the effort that the principal requires an agent to exert is given as

$$Q^{\mathrm{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) = (2\hat{\theta}_n - \underline{\theta})^{-1}(W(\hat{\boldsymbol{\theta}}))^{-2}, \tag{9}$$

where W is the solution of the equation $[W(\boldsymbol{\theta})]^3 - \frac{1}{\sigma_0^2}[W(\boldsymbol{\theta})]^2 = \sum_{m\in\mathcal{A}} \frac{1}{2\theta_m - \underline{\theta}}$. An explicit (although cumbersome) solution of W is provided in Appendix A.2.

As in the case of linear costs, the function Q^{P} is designed to optimally harness the heterogenity of the agents in order to minimize the prediction error with a small enough payment. Observe that in contrast to the linear case (7), here the principal requires every agent to exert a positive effort.

We again adopt the shorthand of $q_n^{\text{P}} = Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})$. We define the function R that governs the payment to any agent n as

$$R(x^*, \hat{y}_n, \hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) = \pi(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) - (x^* - \hat{y}_n)^2 \cdot K(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) + S(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}),$$

where

$$\begin{aligned}\pi(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) =& \frac{1}{2}\big(\hat{\theta}_n \cdot \big[Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n})\big]^2 + \int_{\hat{\theta}_n}^{\bar{\theta}} \big[Q^{\text{P}}(z, \hat{\boldsymbol{\theta}}_{-n})\big]^2 \mathrm{d}z\big),\\ K(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) =& \big[Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) + 1/\sigma_0^2\big]^2 \hat{\theta}_n \cdot Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}),\\ S(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) =& \big[Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}) + 1/\sigma_0^2\big] \hat{\theta}_n \cdot Q^{\text{P}}(\hat{\theta}_n, \hat{\boldsymbol{\theta}}_{-n}).\end{aligned}$$

These functions have a form similar to those in the case of linear costs (8), except that these functions depend on the reported cost parameters of all N agents, whereas the corresponding functions in the linear cost setting depended only on the reported cost parameter of one agent. The remaining higher level intuition behind this construction is identical to that behind the linear-cost case described in the previous section.

The principal uses the Bayes estimate as his predictor: $\hat{x}(\hat{\boldsymbol{y}}, \boldsymbol{q}^{\text{P}}) = \frac{(1-N)\mu_0/\sigma_0^2 + \sum_{n\in\mathcal{A}} (1/\sigma_0^2 + q_n^{\text{P}}) \cdot \hat{y}_n}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q_n^{\text{P}}}$.

The following theorem establishes the optimality guarantee of COPE under quadratic costs.

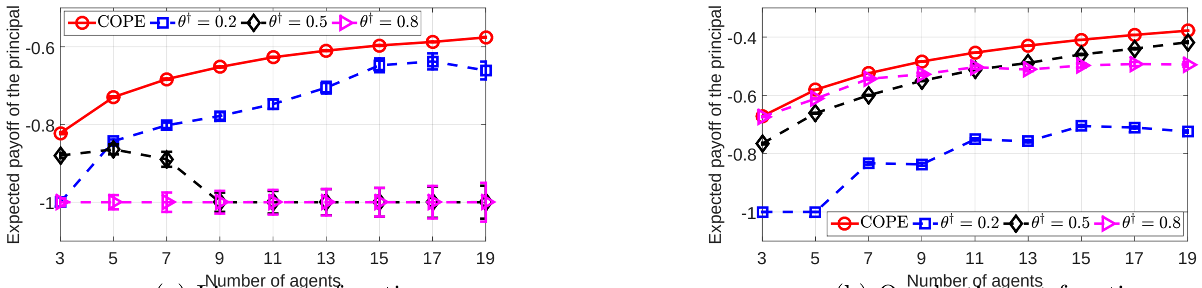
Theorem 2. *Under the quadratic cost function $C(q, \theta) = \frac{1}{2}\theta q^2$, COPE is feasible and maximizes the principal's expected payoff.*

Thus the optimal mechanism in the case of quadratic costs requires participation of all agents. Our theory thus recommends the practitioner to build a "crowd-sourcing" system when the costs of agents are believed to be quadratic.

3.3 Commentary on Interpretations and Generalizations

Our results show that interestingly, it is optimal for the principal to call for a *crowd-tender* when the cost function is linear, while it is optimal to design a *crowd-sourcing* mechanism when the cost function is quadratic. Informally, the cost function acts as a regularizer on the choice of effort levels q^{P}, and the dichotomy of these two cost functions is related to the sparsity inducing properties of the ℓ_1-regularizer, and the lack thereof of the squaredâĂž ℓ_2-regularizer.

We briefly comment upon the structure of the optimal mechanism under more general forms of the cost function, the noise distribution, the prior distribution, and the prediction loss function. Under these general conditions, the structure of the mechanism remains identical to Algorithm 1. The structure of the payment function remains identical to (8), except that under a more general distribution of noise in the agents' observations, the term $(x^* - \hat{y}_{n_0})^2$ is replaced by a loss function that permits a unique Bayes estimator under the given noise distribution. The structure of the functions π, K, and S in (8) is different from that for the linear and quadratic cases; interestingly, they depend on the distribution of $\{\theta_n\}$ only through the ratio of the c.d.f. $F(\theta_n)$ and the p.d.f. $f(\theta_n)$ of the distribution. The principal's predictor $\hat{x}$ remains the standard Bayes estimator. We have shown that COPE with these modifications to be optimal and feasible under certain regularity conditions. Given the space constraints, we defer the details for a future version.



(a) Linear cost function. (b) Quadratic cost function.

Figure 2: The principal's expected payoff under COPE and the homogenous mechanism.

4 Simulations

We conduct numerical studies to evaluate the performance of COPE. In particular, we investigate the amount of gain that can be achieved by (optimally) exploiting the heterogenity of the agents. To this end, we compare the performance of COPE to the following "homogeneous" benchmark mechanism. The homogenous mechanism assumes that all agents are identical and that the principal operates under the belief that every agent's cost parameter equals $\theta^\dagger \in [\underline{\theta}, \bar{\theta}]$. The principal then chooses a mechanism that incentivizes every agent to exert optimal effort and to report observations truthfully in a manner that maximizes his expected payoff, and then employs the predictor that leads to the smallest risk. (Please see Appendix B for details on this optimal homogeneous mechanism.) Each agent, on the other hand, knows the value of her own cost parameter, and given this payment function, exerts an effort and reports the observation that maximizes her own payoff.

In the simulations, we draw $x^* \sim \mathcal{N}(0, 1)$, and set $\underline{\theta} = 0$ and $\bar{\theta} = 1$. We vary the number of agents from $N = 3$ to $N = 19$. Each point in the plots is an average across 50000 trials. Without loss of generality, we have normalized the principal's payoff (see (3)) so that it equals zero in the ideal (unachievable) case of zero prediction error and a zero payment. Note that the principal can always achieve a payoff of -1 by not making any payments, and simply choosing the prior mean has her prediction.

Figure 2 depicts the expected payoff of the principal under COPE and under the homogeneous mechanism with different values of $\theta^\dagger$ (The error bars on many points in Figure 2 are too small to be visible). The two primary observations from the simulation results are as follows. First, under COPE, the principal's expected payoff increases with the number of agents. This is because COPE optimally exploits the presence of additional agents by making them exert different efforts based on their respective cost types. A second inference is that exploiting the heterogeneity of agents allows COPE to outperform the homogeneous mechanism consistently, and the difference depends on the principal's belief of $\theta^\dagger$.

5 Conclusions

All in all, we see that one can obtain a (much) higher payoff by eliciting and exploiting the heterogenity in the agents. The mechanism COPE proposed in this paper indeed achieves this goal, and moreover, attains the optimal payoff. In the future, we intend to use this mechanism as a building block for studying more complex problems at the interface of statistical learning theory and game theory. In particular, it remains to be seem whether COPE can be extended to non-parametric settings with minimal structural assumptions on the behavior of the agents.

References

[CDP15] Yang Cai, Constantinos Daskalakis, and Christos H Papadimitriou. Optimum statistical estimation with strategic data sources. In *Conference on Learning Theory (COLT)*, 2015.

[CJ15] Ruggiero Cavallo and Shaili Jain. Efficient crowdsourcing contests. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 677–686, 2015.

[Con09] Vincent Conitzer. Prediction markets, mechanism design, and cooperative game theory. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 101–108, 2009.

[DFP08] Ofer Dekel, Felix Fischer, and Ariel D Procaccia. Incentive compatible regression learning. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms (SODA)*, pages 884–893, 2008.

[DG13] Anirban Dasgupta and Arpita Ghosh. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd international conference on World Wide Web (WWW)*, pages 319–330, 2013.

[FCK15] Rafael M Frongillo, Yiling Chen, and Ian A Kash. Elicitation for aggregation. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, 2015.

[FSW07] Fang Fang, Maxwell Stinchcombe, and Andrew Whinston. Putting your money where your mouth is-a betting platform for better prediction. *Review of Network Economics*, 6(2), 2007.

[Hay98] Bob E Hayes. *Measuring customer satisfaction: Survey design, use, and statistical analysis methods*. ASQ Quality Press, 1998.

[IPSW14] Panagiotis G Ipeirotis, Foster Provost, Victor S Sheng, and Jing Wang. Repeated labeling using multiple noisy labelers. *Data Mining and Knowledge Discovery*, 28(2):402–441, 2014.

[KOS11] David R Karger, Sewoong Oh, and Devavrat Shah. Iterative learning for reliable crowdsourcing systems. In *Advances in neural information processing systems (NIPS)*, volume 24, pages 1953–1961, 2011.

[MRZ05] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Management Science*, 51(9):1359–1373, 2005.

[NOS12] Sahand Negahban, Sewoong Oh, and Devavrat Shah. Iterative ranking from pair-wise comparisons. In *Advances in Neural Information Processing Systems (NIPS)*, volume 22, pages 2474–2482, 2012.

[Pre04] Drazen Prelec. A Bayesian truth serum for subjective data. *Science*, 306(5695):462–466, 2004.

[SBB+15] Nihar B Shah, Sivaraman Balakrishnan, Joseph Bradley, Abhay Parekh, Kannan Ramchandran, and Martin Wainwright. Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence. In *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015.

[SZP15] Nihar B Shah, Dengyong Zhou, and Yuval Peres. Approval voting and incentives in crowdsourcing. In *International Conference on Machine Learning (ICML)*, 2015.

[TPPZ15] Panos Toulis, David C Parkes, Elery Pfeffer, and James Zou. Incentive-compatible experimental design. In *ACM Conference on Economics and Computation (EC)*, 2015.

[WR11] Justin Wolfers and David Rothschild. Forecasting elections: Voter intentions versus expectations. In *Law and Economics Workshop*, 2011.

[WZ04] Justin Wolfers and Eric Zitzewitz. Prediction markets. Technical report, National Bureau of Economic Research, 2004.

[ZBMP12] Dengyong Zhou, Sumit Basu, Yi Mao, and John C Platt. Learning from the wisdom of crowds by minimax entropy. In *Advances in Neural Information Processing Systems (NIPS)*, volume 22, pages 2195–2203, 2012.

Appendix

A Proof

A.1 Proof for Theorem 1

Proof. The proof will proceed in four steps. The first three steps show that our mechanism incentivizes the agents to be truthful, and the fourth step proves optimality of our mechanism. First, we show that irrespective of what an agent reports as her cost parameter, and irrespective of the effort she exerts, the agent is always incentivized to report her true observation. We follow this up and show that irrespective of the effort that an agent exerts, she is always incentivized to report her cost parameter correctly. The third step completes the proof of truthfulness, showing that under truthful reporting of the cost parameter and the observation, under our mechanism, an agent is always incentivized to exert precisely the effort as desired by the principal. Finally, we show that among all mechanisms that ensure truthful reports, our mechanism maximizes the principal's expected utility.

We assume that the random variables $\{\theta_n\}_{n\in\mathcal{A}}$ are independently and identically distributed on support $[\underline{\theta},\bar{\theta}]$, with a cumulative distribution function $F:[\underline{\theta},\bar{\theta}]\to\mathbb{R}_+$ and a probability density function $f:[\underline{\theta},\bar{\theta}]\to\mathbb{R}_+$. We further assume that the c.d.f. function F is continues, differentiable, and log concave in $[\underline{\theta},\bar{\theta}]$. This assumption is satisfied by a wide range of distributions, such as the uniform, gamma, and beta distributions.

Step 1. Truthful reporting of observation under COPE

We will analyze the strategies of the agent who is recruited by the principal and the agents who are not recruited by the principal, respectively.

We first study the observation reporting strategy of the agent n_0 who is recruited and rewarded by the principal, where $n_0=\arg\min_{m\in\mathcal{A}}\hat{\theta}_m$.

We will show that the agent n_0 will choose

$$\hat{y}_{n_0}=\frac{\mu_0\cdot 1/\sigma_0^2+y_{n_0}\cdot q_{n_0}}{1/\sigma_0^2+q_{n_0}} \tag{10}$$

to maximize her expected payoff given her exerting effort q_{n_0} and own observation y_{n_0}.

As shown in (8), $\pi(\hat{\theta}_{n_0})$, $K(\hat{\theta}_{n_0})$ and $S(\hat{\theta}_{n_0})$ are independent of $\hat{y}_{n_0}$. Moreover, the value of calculated by $K(\hat{\theta}_{n_0})$ is always positive. Hence, when the agent n_0 makes reporting observation strategy to maximize her expected payoff, i.e.,

$$\hat{y}_{n_0}\in\arg\max\mathbb{E}\big[\pi(\hat{\theta}_{n_0})-K(\hat{\theta}_{n_0})\cdot(x^*-\hat{y}_{n_0})^2+S(\hat{\theta}_{n_0})\big]-C\big(q_{n_0},\theta_{n_0}\big),$$

where the expectation is taken with respect to x^* and cost parameters $\boldsymbol{\theta}_{-n_0}=[\theta_1,\ldots,\theta_{n_0-1},\theta_{n_0+1},\ldots,\theta_N]^T$ except agent n_0, it is equivalent for the agent n_0 to choose reporting strategy such that

$$\hat{y}_{n_0}\in\arg\min\mathbb{E}_{x^*}[(x^*-\hat{y}_{n_0})^2]. \tag{11}$$

It is well known that only when $\hat{y}_{n_0}=\frac{\mu_0\cdot 1/\sigma_0^2+y_{n_0}\cdot q_{n_0}}{1/\sigma_0^2+q_{n_0}}$, the value of $\mathbb{E}_{x^*}[(x^*-\hat{y}_{n_0})^2]$ is minimized and the expected value is

$$\mathbb{E}_{x^*}[(x^*-\hat{y}_{n_0})^2]=\frac{1}{1/\sigma_0^2+q_{n_0}}.$$

We then study the observation reporting strategy of other agents who are not recruited and rewarded by the principal. For agent $n\in\mathcal{A},n\neq n_0$, she will put zero effort as she does not received the reward from the principal. In such case, only when reporting her observation $\hat{y}_n = \mu_0$ can minimize $\mathbb{E}_{x^*}[(x^* - \hat{y}_n)^2]$. The expected value of $\mathbb{E}_{x^*}[(x^* - \hat{y}_n)^2]$ is

$$\mathbb{E}_{x^*}[(x^* - \hat{y}_n)^2] = \frac{1}{1/\sigma_0^2}, \quad n \in \mathcal{A}, n \neq n_0.$$

Step 2. Truthful reporting of cost parameter under COPE

We first show that the agent n_0 will truthfully reveals her cost type. We first rewrite the function π, K, and S as follows.

$$\begin{aligned}
\pi(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) =& \hat{\theta}_{n_0} \cdot Q^{\text{P}}(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) + \int_{\hat{\theta}_{n_0}}^{\bar{\theta}} Q^{\text{P}}(z, \boldsymbol{\theta}_{-n_0}) \mathrm{d}z, \\
K(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) =& \big[Q^{\text{P}}(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big]^2 \cdot \hat{\theta}_{n_0}, \\
S(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) =& \big[Q^{\text{P}}(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big] \cdot \hat{\theta}_{n_0}.
\end{aligned}$$

The expected payoff of the agent whose cost type is θ_{n_0} but report $\hat{\theta}_{n_0}$ is:

$$\begin{aligned}
&\mathbb{E}_{\{x^*, y_{n_0}, \boldsymbol{\theta}_{-n_0}\}}\big[U^{\text{A}}(x^*, \hat{\theta}_{n_0}, q_{n_0}, y_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] \\
&= \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\big[\pi(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) - K(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) \cdot \frac{1}{1/\sigma_0^2 + q_{n_0}} + S(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) - q_{n_0}\theta_{n_0}\big].
\end{aligned} \tag{12}$$

For notation convenience, we define the function $U^{\text{AE}} : \mathbb{R} \times [\underline{\theta}, \bar{\theta}] \times \mathbb{R}_+ \times [\underline{\theta}, \bar{\theta}]^N \to \mathbb{R}_+$ as

$$U^{\text{AE}}(\hat{\theta}_{n_0}, q_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0}) = \big[\pi(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) - K(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) \frac{1}{1/\sigma_0^2 + q_{n_0}} + S(\hat{\theta}_{n_0}, \boldsymbol{\theta}_{-n_0}) - q_{n_0}\theta_{n_0}\big], \tag{13}$$

where $\boldsymbol{\theta}_{-n_0}$ are the random variables of all agents cost type except the agent n_0. By comparing (12) to (13), the expected payoff of the agent n is

$$\mathbb{E}_{\{x^*, y_{n_0}, \boldsymbol{\theta}_{-n_0}\}}\big[U^{\text{A}}(x^*, \hat{\theta}_{n_0}, q_{n_0}, y_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] = \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\big[U^{\text{AE}}(\hat{\theta}_{n_0}, q_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big].$$

By the mean value theorem, we have:

$$\mathbb{E}\big[U^{\text{AE}}(\theta_{n_0}, q_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] - \mathbb{E}\big[U^{\text{AE}}(\hat{\theta}_{n_0}, q_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] = \mathbb{E}\left[\frac{\partial U^{\text{AE}}(\eta, q_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})}{\partial \eta}\right] \cdot (\theta_{n_0} - \hat{\theta}_{n_0}), \tag{14}$$

where the expectation is taken with respect to $\boldsymbol{\theta}_{-n_0}$, and η lies between θ_{n_0} and $\hat{\theta}_{n_0}$.

We further have:

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\left[\frac{\partial U^{\text{AE}}(\eta, q_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})}{\partial \eta}\right] \\
&= \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\Bigg[\frac{\partial}{\partial \eta}\Bigg(\eta Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + \int_{\eta}^{\bar{\theta}} Q^{\text{P}}(z, \boldsymbol{\theta}_{-n_0}) \mathrm{d}z - \frac{\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big]^2}{1/\sigma_0^2 + q_{n_0}} \eta \\
&\qquad + \big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big]\eta - q_{n_0}\theta_{n_0}\Bigg)\Bigg] \\
&= \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\Bigg[2\eta \frac{\partial Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0})}{\partial \eta} - \frac{\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big]^2}{1/\sigma_0^2 + q_{n_0}} \\
&\qquad - 2\frac{\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big]}{1/\sigma_0^2 + q_{n_0}} \cdot \frac{\partial Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0})}{\partial \eta} \cdot \eta + \big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big]\Bigg] \\
&= \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\Bigg[\Bigg(1 - \frac{Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2}{q_{n_0} + 1/\sigma_0^2}\Bigg) \cdot \Bigg(2\eta \cdot \frac{\partial Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0})}{\partial \eta} + Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\Bigg)\Bigg].
\end{aligned} \tag{15}$$

If we have

$$-\frac{\partial Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n_0})/\big(Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n_0})+1/\sigma_0^2\big)}{\partial\theta_{n_0}/\theta_{n_0}}\geq\frac{1}{2}, \tag{16}$$

then we have

$$2\eta\cdot\frac{\partial Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n_0})}{\partial\eta}+Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n_0})+1/\sigma_0^2\leq 0.$$

Lemma 1. *If $\theta_n \sim$ Uniform$[\underline{\theta},\bar{\theta}]$, independent for every $n\in\mathcal{A}$, then* (16) *is satisfied.*

Proof. First consider the case $N=1$. Then there is only one agent, and hence the principal automatically selects that agent. So Q^{P} is simply Q of that agent:

$$Q(\theta)=1/\sqrt{\theta+F(\theta)/f(\theta)}-1/\sigma_0^2,$$

and hence

$$\begin{aligned}-\frac{\partial Q(\theta)}{\partial\theta}\frac{\theta}{Q(\theta)+1/\sigma_0^2}&=\frac{1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right)}{2[\theta+F(\theta)/f(\theta)]}\frac{1}{\sqrt{\theta+F(\theta)/f(\theta)}}\frac{\theta}{1/\sqrt{\theta+F(\theta)/f(\theta)}}\\&=\frac{1}{2}\frac{1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right)}{1+\frac{1}{\theta}\left(\frac{F(\theta)}{f(\theta)}\right)}\geq\frac{1}{2},\end{aligned}$$

where the final inequality holds for uniform distribution.

We now extend this condition to $N>1$. Observe that the calculation above will be violated only when the cost parameter of some other agent is infinitesimally close to θ_{n_0} (since in that case, $\frac{\partial Q^{\mathrm{P}}(\theta_{n_0},\boldsymbol{\theta}_{-n_0})}{\partial\theta_{n_0}}$ is different from that calculated above). However given our assumptions that the distribution of θ has a valid pdf, and the number of agents N is finite, θ_{n_0} will be well separated from the cost types of all other agents with probability 0. □

Because the agent will exert effort q_{n_0} to maximize her expected payoff. By taking the first order derivative of (12) with respect to q_{n_0} and set it to zero, we have

$$(1/\sigma_0^2+q^{\mathrm{P}}_{n_0})^2\cdot\hat{\theta}_{n_0}=(1/\sigma_0^2+q_{n_0})^2\cdot\theta_{n_0}, \tag{17}$$

where $q^{\mathrm{P}}_{n_0}$ is a shorthand for $Q^{\mathrm{P}}(\hat{\theta}_{n_0},\boldsymbol{\theta}_{-n_0})$.

Based on (17), we have (i) if $\hat{\theta}_{n_0}>\theta_{n_0}$, $q^{\mathrm{P}}_{n_0}<q_{n_0}$, (ii) if $\hat{\theta}_{n_0}<\theta_{n_0}$, $q^{\mathrm{P}}_{n_0}>q_{n_0}$, and (iii) if $\hat{\theta}_{n_0}=\theta_{n_0}$, $q^{\mathrm{P}}_{n_0}=q_{n_0}$.

Hence, If $\hat{\theta}_{n_0}>\theta_{n_0}$, the equation (15) is negative and (14) is positive. This inequality also holds for $\hat{\theta}_{n_0}<\theta_{n_0}$, by a similar argument. Therefore, agent n_0 will truthfully report her own cost parameter.

We then show that the agent $n\in\mathcal{A},n\neq n_0$ will truthfully reveals her cost type. Recall that the principal does not recruit and reward the agent $n\in\mathcal{A},n\neq n_0$. Hence, the payment to the agent $n\in\mathcal{A},n\neq n_0$ is zero. The we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{AE}}(\theta_n,q_n,\theta_n,\boldsymbol{\theta}_{-n})\big]-\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{AE}}(\hat{\theta}_n,q_n,\theta_n,\boldsymbol{\theta}_{-n})\big]=0,\forall n\in\mathcal{A},n\neq n_0,$$

which shows that there is indifference between truthfully report cost type or not in terms of expected payoff for the agent n. We assume that in such case, the agent will truthfully report their cost types.

Step 3. Incentivized agent to exert precisely the effort as desired by the principal under COPE

As we have proved in Step 2 that the agent n_0 would truthfully report her cost type ($\hat{\theta}_n = \theta_n$), then we will show that the agent n_0 exerts effort such that $q_{n_0} = q^{\text{P}}_{n_0}$ would maximize her expected payoff which is given as

$$\mathbb{E}\big[U^{\text{AE}}(\theta_{n_0}, q^{\text{P}}_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] = \pi(\theta_{n_0}, \boldsymbol{\theta}_{-n_0}) - K(\theta_{n_0}, \boldsymbol{\theta}_{-n_0})\frac{1}{1/\sigma_0^2 + q_{n_0}} + S(\theta_{n_0}, \boldsymbol{\theta}_{-n_0}) - q_{n_0}\theta_{n_0}. \quad (18)$$

where the expectation is taken with respect to $\boldsymbol{\theta}_{-n_0}$.

It can be verified that (18) is concave in q_{n_0}. Hence, by taking the first order derivative of (18) with respect to q_{n_0}, we have

$$\frac{\partial}{\partial q_{n_0}}\mathbb{E}\big[U^{\text{AE}}(\theta_{n_0}, q^{\text{P}}_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] = \left[\frac{1/\sigma_0^2 + q^{\text{P}}_{n_0}}{1/\sigma_0^2 + q_{n_0}}\right]^2 \cdot \theta_{n_0} - \theta_{n_0}. \quad (19)$$

We can verify that the value of (19) equals to zero only when $q_{n_0} = q^{\text{P}}_{n_0}$. Hence, agent n_0 will exert the effort as the principal desires to maximize her expected payoff. Then (10) is rewritten as

$$\hat{y}_{n_0} = \frac{\mu_0 \cdot 1/\sigma_0^2 + y_{n_0} \cdot q^{\text{P}}_{n_0}}{1/\sigma_0^2 + q^{\text{P}}_{n_0}}. \quad (20)$$

Because the principal knows the value of μ_0, σ_0^2, and $q^{\text{P}}_{n_0}$, he can infer the agent n_0's truth observation y_{n_0} from (20).

Step 4. Maximizes the principal's expected utility under COPE

Then we look at the expected payoff of the principal. The following lemma describes COPE is the optimal mechanism that maximizes the principal's expected utility.

Lemma 2. *The optimal predictor* $\hat{x} = \frac{\mu_0 \cdot 1/\sigma_0^2 + \sum_{n\in\mathcal{A}} y_n \cdot q^{\text{P}}_n}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}$ *defined in COPE maximizes the principal's expected utility, and the Bayes risk of the principal's prediction is* $B^{\text{P}}\big(\boldsymbol{q}^{\text{P}}\big) = \frac{1}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}$.

Proof. Recall that $q^{\text{P}}_n = Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$. Given all agents' observation $\boldsymbol{y}$ and agents' exert effort $\boldsymbol{q}^{\text{P}}$, the principal's updated belief on the realization of x^* can be expressed as

$$x^*|(\boldsymbol{y}, \boldsymbol{q}^{\text{P}}) \sim N\left(\frac{\mu_0 \cdot 1/\sigma_0^2 + \sum_{n\in\mathcal{A}} y_n \cdot q^{\text{P}}_n}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}, \frac{1}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}\right).$$

To maximize the expected utility for the prediction, the principal solves

$$\begin{aligned}
&\max_{\hat{x}} \mathbb{E}\big[v - (x^* - \hat{x})^2|(\boldsymbol{y}, \boldsymbol{q}^{\text{P}})\big] \\
&= \max_{\hat{x}} \left(v - \left\{\mathbb{E}\big[x^{*2}|(\boldsymbol{y}, \boldsymbol{q}^{\text{P}})\big] - 2\hat{x}\mathbb{E}\big[x^*|(\boldsymbol{y}, \boldsymbol{q}^{\text{P}})\big] + \hat{x}^2\right\}\right) \\
&= \max_{\hat{x}} \left(v - \left[\hat{x} - \frac{\mu_0 \cdot 1/\sigma_0^2 + \sum_{n\in\mathcal{A}} y_n \cdot q^{\text{P}}_n}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}\right]^2 - \frac{1}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}\right) \\
&\leq v - \frac{1}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}
\end{aligned}$$

The equality holds only when

$$\hat{x} = \frac{\mu_0 \cdot 1/\sigma_0^2 + \sum_{n\in\mathcal{A}} y_n \cdot q^{\text{P}}_n}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q^{\text{P}}_n}. \quad (21)$$

Hence, the optimal predictor that maximizes the principal's expected utility is

$$\hat{x}\big(\boldsymbol{y}, \boldsymbol{q}^{\mathrm{P}}\big) = \frac{\mu_0 \cdot 1/\sigma_0^2 + \sum_{n\in\mathcal{A}} y_n \cdot q_n^{\mathrm{P}}}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q_n^{\mathrm{P}}} \tag{22}$$

and the Bayes risk is

$$B^{\mathrm{P}}\big(\boldsymbol{q}^{\mathrm{P}}\big) = \inf_{\hat{x}} \mathbb{E}[(x^* - \hat{x})^2] = \frac{1}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q_n^{\mathrm{P}}},$$

where the expectation is taken with respect to x^* and $\boldsymbol{y}$.

Recall that under the linear cost function, the principal only recruits agent n_0 to exert effort, in such case, $q_n^{\mathrm{P}} = 0$, $\forall n \in \mathcal{A}$, $n \neq n_0$. Also recall that the principal can infer the true observation of the agent n_0 through the function $g: \mathbb{R} \to \mathbb{R}$ and is defined as $y_{n_0} = g(\hat{y}_{n_0}) = \hat{y}_{n_0} + (\hat{y}_{n_0} - \mu_0)/(q_{n_0}^{\mathrm{P}} \sigma_0^2)$. Then putting back to (22) we can get the conclusion.

□

We then show that the desired effort level $Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ defined in (7) and the function $\pi(\theta_n, \boldsymbol{\theta}_{-n})$ defined in (8) can maximize the principal's expected payoff and satisfy BIC and BIR condition.

Notice that the agent n_0 exerts effort such that $q_{n_0} = q_{n_0}^{\mathrm{P}}$ and reports $\hat{y}_{n_0} = \frac{\mu_0 \cdot 1/\sigma_0^2 + y_{n_0} \cdot q_{n_0}^{\mathrm{P}}}{1/\sigma_0^2 + q_{n_0}^{\mathrm{P}}}$, the expected payment function is reduced to

$$\begin{aligned}
&\mathbb{E}_{\{x^*, y_{n_0}, \boldsymbol{\theta}_{-n_0}\}}\big[R(x^*, y_{n_0}, q_{n_0}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] \\
&= \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\big[\pi(\theta_{n_0}, \boldsymbol{\theta}_{-n_0}) - K(\theta_{n_0}, \boldsymbol{\theta}_{-n_0}) \cdot \frac{1}{1/\sigma_0^2 + q_{n_0}} + S(\theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big] = \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\big[\pi(\theta_{n_0}, \boldsymbol{\theta}_{-n_0})\big].
\end{aligned} \tag{23}$$

For other agent $n \in \mathcal{A}, n \neq n_0$, as the principal does not require her to do the observation, we first assume that the expected payment to her is as follows

$$\mathbb{E}_{\{x^*, y_{n_0}, \boldsymbol{\theta}_{-n_0}\}}[R(x^*, y_n, \theta_n, \boldsymbol{\theta}_{-n})] = \mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\big[\pi(\theta_n, \boldsymbol{\theta}_{-n})\big], \forall n \in \mathcal{A}, n \neq n_0 \tag{24}$$

Later we will show that $\pi(\theta_n, \boldsymbol{\theta}_{-n}) = 0, \forall n \neq n_0$.

The expected payoff of agent $n \in \mathcal{A}$ is

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{AE}}\big(\hat{\theta}_{n_0}, q_{n_0}^{\mathrm{P}}, \theta_{n_0}, \boldsymbol{\theta}_{-n_0}\big)\big] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) - \theta_n Q^{\mathrm{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n})\big],$$

where $q_{n_0}^{\mathrm{P}}$ is the shorthand of $Q^{\mathrm{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n})$. For notation convenience, we adopt $U^{\mathrm{AE}}\big(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\mathrm{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n\big)$ in the later proof of Theorem 1, where the function U^{AE} is rewritten as

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{AE}}\big(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\mathrm{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\big] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) - \theta_n Q^{\mathrm{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n})\big]. \tag{25}$$

Correspondingly, BIC and BIR conditions, i.e., (4) and (5) can be rewritten as

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{AE}}\big(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\big] \geq \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{AE}}\big(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\mathrm{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\big],\ \forall \hat{\theta}_n \neq \theta_n \tag{26}$$

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{AE}}\big(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\big] \geq 0,\ \forall \theta_n \in [\underline{\theta}, \bar{\theta}]. \tag{27}$$

Base on Lemma 2, (23), and (24), the expected payoff of the principal is

$$\begin{aligned}
\mathbb{E}_{x^*, \boldsymbol{y}, \boldsymbol{\theta}}[U^{\mathrm{P}}(x^*, \boldsymbol{q}^{\mathrm{P}}, \boldsymbol{y}, \hat{\boldsymbol{\theta}})] &= -B^{\mathrm{P}}\big(\boldsymbol{q}^{\mathrm{P}}\big) - \mathbb{E}_{x^*, \boldsymbol{y}, \boldsymbol{\theta}}\big[\sum_{n\in\mathcal{A}} R(x^*, y_n, \theta_n, \boldsymbol{\theta}_{-n})\big] \\
&= -\frac{1}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q_n^{\mathrm{P}}} - \mathbb{E}_{\boldsymbol{\theta}}\big[\sum_{n\in\mathcal{A}} \pi(\theta_n, \boldsymbol{\theta}_{-n})\big].
\end{aligned}$$

Recall that $q_n^{\text{P}} = Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$, the principal's optimal problem defined in (6) can be rewritten as

$$\begin{aligned} \sup_{\{Q^{\text{P}}(\boldsymbol{\theta}),\pi(\boldsymbol{\theta})\},\forall\theta_n\in\boldsymbol{\theta},\forall n\in\mathcal{A}} \quad & \mathbb{E}[U^{\text{P}}(x^*, \boldsymbol{q}^{\text{P}}, \boldsymbol{y}, \hat{\boldsymbol{\theta}})], \\ \text{subject to:} \quad & \text{BIC and BIR in (26) and (27).} \end{aligned} \tag{28}$$

In the following lemmas, we characterize an equivalent formulation for the feasible region defined by BIC and BIR. Using these lemmas, we show that $Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ defined in (7) and $\pi(\theta_n, \boldsymbol{\theta}_{-n})$ defined in (8) are the optimal solution that solves the principal's problem in (28).

Lemma 3. *The solution of* (28) *is feasible if and only if it satisfies the following conditions for all* $\theta_n \in [\underline{\theta}, \bar{\theta}]$, $\forall n \in \mathcal{A}$:

- *the expected payoff of agent n is*

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\text{AE}}\big(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\Big] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[\int_{\theta_n}^{\bar{\theta}} Q^{\text{P}}(x, \boldsymbol{\theta}_{-n})\mathrm{d}x\Big] \tag{29}$$

- $Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ *is non-increasing in* θ_n.

Proof. The proof of Lemma 3 is as follows. We first show that BIC and BIR imply the condition in (29).

Notice that the first derivative of (25) is:

$$\frac{\partial \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}\big(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\big]}{\partial \theta_n} = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[-Q^{\text{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n})\big] \le 0. \tag{30}$$

Then, for any $\theta_n^1 > \theta_n^2$, we have

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\theta_n^1, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n^1, \boldsymbol{\theta}_{-n}), \theta_n^1)\big] &\le \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\theta_n^1, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n^1, \boldsymbol{\theta}_{-n}), \theta_n^2)\big] \\ &\le \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\theta_n^2, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n^2, \boldsymbol{\theta}_{-n}), \theta_n^2)\big]. \end{aligned} \tag{31}$$

where the first inequality is because (30) and the second is from the BIC condition defined in (26). In other words, for the agent $n \in \mathcal{A}$ whose cost parameter $\underline{\theta} \le \theta_n \le \bar{\theta}$, we have

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\bar{\theta}, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\bar{\theta}, \boldsymbol{\theta}_{-n}), \bar{\theta})\big] &\le \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n)\big] \\ &\le \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\underline{\theta}, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\underline{\theta}, \boldsymbol{\theta}_{-n}), \underline{\theta})\big]. \end{aligned} \tag{32}$$

Recall that the BIR condition is

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}\big(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\big] \ge 0,\ \forall \theta_n \in [\underline{\theta}, \bar{\theta}], \tag{33}$$

which implies that, for the agent $n \in \mathcal{A}$ with any value $\theta_n \in [\underline{\theta}, \bar{\theta}]$, her expected payoff should at least be zero. Then the expected payoff of the agent n with cost parameter $\bar{\theta}$ must be binding at zero. Otherwise, the principal can reduce the $\pi(\bar{\theta}, \boldsymbol{\theta}_{-n})$ by a small value of $\delta > 0$, which does not violate the constraint of (33) but raises the principal's expected payoff. Hence, we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\bar{\theta}, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\bar{\theta}, \boldsymbol{\theta}_{-n}), \bar{\theta})\big] = 0. \tag{34}$$

Let $U^{\text{AE}}(\theta_n, \boldsymbol{\theta}_{-n}) = U^{\text{AE}}\big(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n\big)$. From BIC condition, we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\theta_n, \boldsymbol{\theta}_{-n})\big] = \max_{\hat{\theta}_n} \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}\big(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n\big)\big].$$

By using the envelope theorem, we have:

$$\frac{\partial \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\theta_n,\boldsymbol{\theta}_{-n})\big]}{\partial \theta_n} = \frac{\partial \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\theta_n)\big]}{\partial \theta_n}\Bigg|_{\hat{\theta}_n=\theta_n} = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[-Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})\big], \tag{35}$$

where θ_n is a parameter. By integrating both sides from the value of θ_n to $\bar{\theta}$ and using (34) and the assumption that the random variable θ_n of the agent n is independent for every $n \in \mathcal{A}$, we get

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\big] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\bigg[\int_{\theta_n}^{\bar{\theta}} Q^{\text{P}}(x,\boldsymbol{\theta}_{-n})\mathrm{d}x\bigg] \tag{36}$$

We prove $Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_n)$ is nonincreasing in θ_n by contradiction. Let p_n as the shorthand for $\pi(\theta_n,\boldsymbol{\theta}_{-n})$. Suppose for any $\theta_n^1 > \theta_n^2$, we have $Q^{\text{P}}(\theta_n^1,\boldsymbol{\theta}_{-n}) > Q^{\text{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})$. Because

$$\frac{\partial^2 U^{\text{AE}}\big(p_n,q_n^{\text{P}},\theta_n\big)}{\partial q_n^{\text{P}}\partial\theta_n} = -1 < 0, \tag{37}$$

$$\frac{\partial^2 U^{\text{AE}}\big(p_n,q_n^{\text{P}},\theta_n\big)}{\partial {q_n^{\text{P}}}^2} = 0, \tag{38}$$

we have

$$\begin{aligned} 0 &= \frac{\partial U^{\text{AE}}\big(p_n,q_n^{\text{P}},\theta_n^1\big)}{\partial q_n^{\text{P}}}\Bigg|_{q_n^{\text{P}}=Q^{\text{P}}(\theta_n^1,\boldsymbol{\theta}_{-n})} \\ &= \frac{\partial U^{\text{AE}}\big(p_n,q_n^{\text{P}},\theta_n^1\big)}{\partial q_n^{\text{P}}}\Bigg|_{q_n^{\text{P}}=Q^{\text{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})} \\ &< \frac{\partial U^{\text{AE}}\big(p_n,q_n^{\text{P}},\theta_n^2\big)}{\partial q_n^{\text{P}}}\Bigg|_{q_n^{\text{P}}=Q^{\text{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})}, \end{aligned} \tag{39}$$

where the first equality is because of BIC when the agent n's cost parameter θ_n has the value of θ_n^1, the second equality is because of (38), and the inequality is because of (37).

However, based on the BIC condition, if the agent n's cost parameter θ_n has the value of θ_n^2, then we should have

$$\frac{\partial U^{\text{AE}}\big(p_n,q_n^{\text{P}},\theta_n^2\big)}{\partial q_n^{\text{P}}}\Bigg|_{q_n^{\text{P}}=Q^{\text{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})} = 0,$$

which holds true for all scalar value of p_n. Hence, for any $\theta_n^1 > \theta_n^2$, $Q^{\text{P}}(\theta_n^1,\boldsymbol{\theta}_{-n}) \le Q^{\text{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})$.

Then we need to prove that (29) implies BIC and BIR defined in (26) and (27).

BIR is verified by putting θ_n back to (29). Besides, by putting $\theta_n = \bar{\theta}$ back to (29), we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\text{AE}}\big(\pi(\bar{\theta},\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\bar{\theta},\boldsymbol{\theta}_{-n}),\bar{\theta}\big)\Big] = 0.$$

Then we prove that (29) implies BIC. Notice that we have:

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\text{AE}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\Big]\\
&\overset{1}{=}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[-\int_{\theta_n}^{\bar{\theta}}\frac{\partial U^{\text{AE}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),z\big)}{\partial z}\mathrm{d}z\Big]\\
&\overset{2}{=}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\text{AE}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\hat{\theta}_n\big)-\int_{\theta_n}^{\hat{\theta}_n}\frac{\partial U^{\text{AE}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),z\big)}{\partial z}\mathrm{d}z\Big]\\
&\overset{3}{=}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[\int_{\hat{\theta}_n}^{\bar{\theta}}Q^{\text{P}}(\eta,\boldsymbol{\theta}_{-n})\mathrm{d}\eta-\int_{\theta_n}^{\hat{\theta}_n}\frac{\partial U^{\text{AE}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),z\big)}{\partial z}\mathrm{d}z\Big]\\
&\overset{4}{=}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[-\int_{\bar{\theta}}^{\theta_n}Q^{\text{P}}(\eta,\boldsymbol{\theta}_{-n})\mathrm{d}\eta-\int_{\theta_n}^{\hat{\theta}_n}Q^{\text{P}}(\eta,\boldsymbol{\theta}_{-n})\mathrm{d}\eta+\int_{\theta_n}^{\hat{\theta}_n}Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n})\mathrm{d}z\Big]\\
&\overset{5}{=}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\text{AE}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)+\int_{\theta_n}^{\hat{\theta}}\Big(Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n})-Q^{\text{P}}(\eta,\boldsymbol{\theta}_{-n})\Big)\mathrm{d}\eta\Big]
\end{aligned}
$$

where the third equality and the fifth equality is obtained by (29).

If $\hat{\theta}_n>\theta_n$, then the above equation is non-positive (because $Q^{\text{P}}(\eta,\boldsymbol{\theta}_{-n})$ is non-increasing in η) and hence

$$
\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\theta_n)\big]<\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\text{AE}}(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n)\big].
$$

This inequality also holds for $\hat{\theta}_n<\theta_n$ by a similar argument. Therefore, the two condition imply BIC. □

Then based on Lemma 3, we have the following Lemma.

Lemma 4. *The optimisation problem in (28) has the following equivalent formulation:*

$$
\begin{aligned}
\max_{\{Q^{\text{P}}(\boldsymbol{\theta})\},\forall\theta_n\in\boldsymbol{\theta}}\mathbb{E}_{\boldsymbol{\theta}}\Big[&-\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})}-\sum_{n\in\mathcal{A}}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})\cdot\theta_n-\sum_{n\in\mathcal{A}}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})\cdot\frac{F(\theta_n)}{f(\theta_n)}\Big],\\
&\text{s.t. } Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n}) \text{ is nonincreasing in } \theta_n,
\end{aligned}
\tag{40}
$$

where the expectation is taken with respect to $\boldsymbol{\theta}$.

Proof. The proof of Lemma 4 is as follows. The expected payoff of the principal can be written as:

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}}\Big[-\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})}-\sum_{n\in\mathcal{A}}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})\cdot\theta_n-\sum_{n\in\mathcal{A}}U^{\text{A}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\Big]\\
&=\mathbb{E}_{\boldsymbol{\theta}}\Big[-\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})}-\sum_{n\in\mathcal{A}}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})\cdot\theta_n-\sum_{n\in\mathcal{A}}\int_{\theta_n}^{\bar{\theta}}Q^{\text{P}}(x,\boldsymbol{\theta}_{-n})\mathrm{d}x\Big]
\end{aligned}
\tag{41}
$$

where the expectation is taken with respect to $\boldsymbol{\theta}$. Notice that

$$
\begin{aligned}
\mathbb{E}_{\theta_n}\big[\int_{\theta_n}^{\bar{\theta}}Q^{\text{P}}(x,\boldsymbol{\theta}_{-n})\mathrm{d}x\big]&=\int_{\underline{\theta}}^{\bar{\theta}}\int_{z}^{\bar{\theta}}Q^{\text{P}}(x,\boldsymbol{\theta}_{-n})\mathrm{d}x\cdot f(z)\mathrm{d}z=\int_{\underline{\theta}}^{\bar{\theta}}F(z)Q^{\text{P}}(z,\boldsymbol{\theta}_{-n})\mathrm{d}z\\
&=\int_{\underline{\theta}}^{\bar{\theta}}\frac{F(z)}{f(z)}Q^{\text{P}}(z,\boldsymbol{\theta}_{-n})f(z)\mathrm{d}z=\mathbb{E}_{\theta_n}\big[\frac{F(\theta_n)}{f(\theta_n)}Q^{\text{P}}(\theta_n,\boldsymbol{\theta}_{-n})\big]
\end{aligned}
$$

where the first equation is obtained by using integration by parts. Then by applying the above equation to (41) and the fact that $\{\theta_n\}_{n\in\mathcal{A}}$ are assumed to be random, independently and identically distributed on support $[\underline{\theta}, \bar{\theta}]$, we can get the conclusion.

$\square$

Based on Lemma 4, the principal's problem thus reduces to choosing the desired effort $q_n^{\text{P}} = Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ for each agent $n \in \mathcal{A}$. We first consider the problem in (40) without the constraint. If the solution to this unconstrained problem is increasing, then it is also a solution to the constrained problem.

Lemma 5. *$Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ defined in* (7) *and $\pi(\theta_n, \boldsymbol{\theta}_{-n})$ defined in* (8) *are the optimal solution that solves the principal's problem in* (28)

Proof. We first prove that for the agent $\forall n \in \mathcal{A}$,

$$Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n}) = \begin{cases} \max\{1/\sqrt{\gamma(\theta_n)} - 1/\sigma_0^2, 0\} & \text{if } n_0 = \arg\min_{m\in\mathcal{A}} \theta_m, \\ 0, & \text{otherwise}, \end{cases}$$

is the optimal solution of (40) by contradiction.

As $q_n^{\text{P}} = Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$, the problem in (40) is equivalent to

$$\begin{aligned} &\min_{\boldsymbol{q}^{\text{P}}\geq 0} \frac{1}{1/\sigma_0^2 + \sum_{n\in\mathcal{A}} q_n^{\text{P}}} + \sum_{n\in\mathcal{A}} q_n^{\text{P}} \cdot \gamma(\theta_n), \\ &\text{s.t. } q_n^{\text{P}} \text{ is nonincreasing in } \theta_n, \end{aligned} \tag{42}$$

where $\gamma(\theta_n) = \theta_n + F(\theta_n)/f(\theta_n)$.

Without loss of generality, let $\gamma(\theta_1) \leq \gamma(\theta_2) \ldots \leq \gamma(\theta_N)$. If the principal's desired effort from all agents are positive and the solution is

$$q_1^{\text{P}} = q_1^{\text{P},*}, q_2^{\text{P}} = q_2^{\text{P},*}, \ldots, q_N^{\text{P}} = q_N^{\text{P},*} \tag{43}$$

Suppose there are another solution such that

$$\begin{cases} q_1^{\text{P},\dagger} = q_1^{\text{P}} + q_j^{\text{P}}, & \\ q_i^{\text{P},\dagger} = q_i^{\text{P}}, & \forall i \neq 1, j, i \in \mathcal{A} \\ q_j^{\text{P},\dagger} = 0, & \end{cases} \tag{44}$$

We can verify that

$$\sum_{n\in\mathcal{A}} q_n^{\text{P},\dagger} = \sum_{n\in\mathcal{A}} q_n^{\text{P}} \text{ and } \sum_{n\in\mathcal{A}} \left(q_n^{\text{P},\dagger} \cdot \gamma(\theta_n)\right) \leq \sum_{n\in\mathcal{A}} \left(q_n^{\text{P}} \cdot \gamma(\theta_n)\right).$$

Hence, (43) is not an optimal solution. Then we let $q_n^{\text{P}} = 0, \forall n > 1$, the problem in (42) becomes

$$\begin{aligned} &\min_{q_1^{\text{P}}} \frac{1}{1/\sigma_0^2 + q_1^{\text{P}}} + q_1^{\text{P}} \cdot \gamma(\theta_1), \\ &\text{s.t. } q_1^{\text{P}} \geq 0. \end{aligned} \tag{45}$$

By solving the above problem we can get that $q_1^{\text{P}} = \max\{1/\sqrt{\gamma(\theta_1)} - 1/\sigma_0^2, 0\}$. As we define $n_0 = \arg\min_{m\in\mathcal{A}} \theta_m$ and the assumption that F is log-concave in $[\underline{\theta}, \bar{\theta}]$, we have $q_{n_0}^{\text{P}} = \max\{1/\sqrt{\gamma(\theta_{n_0})} - 1/\sigma_0^2, 0\}$.

According to (29), we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[\pi(\theta_n,\boldsymbol{\theta}_{-n}) - Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\cdot\theta_n\Big] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[\int_{\theta_n}^{\bar\theta} Q^{\mathrm{P}}(x,\boldsymbol{\theta}_{-n})\mathrm{d}x\Big].$$

Then the optimal payment function given the agent n_0 and $1/\sqrt{\gamma(\theta_{n_0})} - 1/\sigma_0^2 \geq 0$ is

$$\pi(\theta_{n_0}) = \theta_{n_0}/\sqrt{\gamma(\theta_{n_0})} - \theta_{n_0}/\sigma_0^2 + \int_{\theta_{n_0}}^{\bar\theta}\left(\frac{1}{\sqrt{\gamma(z)}} - \frac{1}{\sigma_0^2}\right)\mathrm{d}z = \theta_{n_0}/\sqrt{\gamma(\theta_{n_0})} - \bar\theta/\sigma_0^2 + \int_{\theta_{n_0}}^{\bar\theta}\frac{1}{\sqrt{\gamma(z)}}\mathrm{d}z.$$

and the payment will be zero if $1/\sqrt{\gamma(\theta_{n_0})} - 1/\sigma_0^2 < 0$.

For other agents, the payments will be zero as they are not involved in the observation and prediction. □

As in the Theorem 1 , we assume that $\theta_n \sim \text{Uniform}[\underline{\theta},\bar\theta]$, independent for every $n \in \mathcal{A}$. Putting the expression of F and f back to the above equations, we can have the conclusion.

□

A.2 Proof for Theorem 2

Proof. The proof is similar to that in Section A.1. The difference is as follows.

First, the function $\pi : [\underline{\theta},\bar\theta]^N \to \mathbb{R}_+$, $K, S : [\underline{\theta},\bar\theta]^N \times \mathbb{R}_+ \to \mathbb{R}_+$ are defined as

$$\begin{aligned}\pi(\hat\theta_n,\boldsymbol{\theta}_{-n}) =& \frac{1}{2}\cdot\left[\hat\theta_n\cdot\big[Q^{\mathrm{P}}(\hat\theta_n,\boldsymbol{\theta}_{-n})\big]^2 + \int_{\hat\theta_n}^{\bar\theta}\left(\big[Q^{\mathrm{P}}(z,\boldsymbol{\theta}_{-n})\big]^2\right)\mathrm{d}z\right], &(46)\\ K(\hat\theta_n,\boldsymbol{\theta}_{-n}) =& \big[Q^{\mathrm{P}}(\hat\theta_n,\boldsymbol{\theta}_{-n}) + 1/\sigma_0^2\big]^2\hat\theta_n\cdot Q^{\mathrm{P}}(\hat\theta_n,\boldsymbol{\theta}_{-n}), &(47)\\ S(\hat\theta_n,\boldsymbol{\theta}_{-n}) =& \big[Q^{\mathrm{P}}(\hat\theta_n,\boldsymbol{\theta}_{-n}) + 1/\sigma_0^2\big]\hat\theta_n\cdot Q^{\mathrm{P}}(\hat\theta_n,\boldsymbol{\theta}_{-n}). &(48)\end{aligned}$$

Step 1. Truthful reporting of observation under COPE

We will analyze the strategies of the agent n, $\forall n \in \mathcal{A}$. We will show that the agent n will choose

$$\hat y_n = \frac{\mu_0\cdot 1/\sigma_0^2 + y_n\cdot q_n}{1/\sigma_0^2 + q_n} \tag{49}$$

to maximize her expected payoff given her exerting effort q_n and own observation y_n.

As $\pi(\hat\theta_n,\boldsymbol{\theta}_{-n})$, $K(\hat\theta_n,\boldsymbol{\theta}_{-n})$ and $S(\hat\theta_n,\boldsymbol{\theta}_{-n})$ are independent of $\hat y_n$ and the value of calculated by $K(\hat\theta_n,\boldsymbol{\theta}_{-n})$ is always positive. Hence, when the agent n makes reporting observation strategy to maximize her expected payoff, i.e.,

$$\hat y_n \in \arg\max \mathbb{E}\big[\pi(\hat\theta_n,\boldsymbol{\theta}_{-n}) - K(\hat\theta_n,\boldsymbol{\theta}_{-n})(x^* - \hat y_n)^2 + S(\hat\theta_n,\boldsymbol{\theta}_{-n})\big] - C\big(q_n,\theta_n\big),$$

where the expectation is taken with respect to x^* and cost parameters $\boldsymbol{\theta}_{-n} = [\theta_1,\ldots,\theta_{n-1},\theta_{n+1},\ldots,\theta_N]^T$ except agent n, it is equivalent for the agent n to choose reporting strategy such that

$$\hat y_n \in \arg\min \mathbb{E}_{x^*}[(x^* - \hat y_n)^2]. \tag{50}$$

The value of $\mathbb{E}_{x^*}[(x^* - \hat y_n)^2]$ is minimized when $\hat y_n = \frac{\mu_0\cdot 1/\sigma_0^2 + y_n\cdot q_n}{1/\sigma_0^2 + q_n}$. The expected value in this case is

$$\mathbb{E}_{x^*}[(x^* - \hat y_n)^2] = \frac{1}{1/\sigma_0^2 + q_n}.$$

Step 2. Truthful reporting of cost parameter under COPE

We will show that the agent n, $\forall n \in \mathcal{A}$ will truthfully reveals her cost type. The expected payoff of the agent whose cost type is θ_n but report $\hat{\theta}_n$ is:

$$\begin{aligned}&\mathbb{E}_{\{x^*, y_n, \boldsymbol{\theta}_{-n}\}}\big[U^{\text{A}}(x^*, \hat{\theta}_n, q_n, y_n, \theta_n, \boldsymbol{\theta}_{-n})\big]\\&=\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) - K(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) \cdot \frac{1}{1/\sigma_0^2 + q_n} + S(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) - \frac{1}{2}\theta_n q_n^2\big].\end{aligned} \tag{51}$$

For notation convenience, we define

$$U^{\text{A}}(\hat{\theta}_n, q_n, \theta_n, \boldsymbol{\theta}_{-n}) = \big[\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) - \frac{K(\hat{\theta}_n, \boldsymbol{\theta}_{-n})}{1/\sigma_0^2 + q_n} + S(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) - \frac{1}{2}\theta_n q_n^2\big]$$

By the mean value theorem, we have:

$$\mathbb{E}\big[U^{\text{A}}(\theta_n, q_n, \theta_n, \boldsymbol{\theta}_{-n})\Big] - \mathbb{E}\Big[U^{\text{A}}(\hat{\theta}_n, q_n, \theta_n, \boldsymbol{\theta}_{-n})\big] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[\frac{\partial U^{\text{A}}(\eta, q_n, \theta_n, \boldsymbol{\theta}_{-n})}{\partial \eta}\big](\theta_n - \hat{\theta}_n), \tag{52}$$

where the expectation is taken with respect to $\boldsymbol{\theta}_{-n}$, and η lies between θ_n and $\hat{\theta}_n$.

We further have

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[\frac{\partial U^{\text{A}}(\eta, q_n, \theta_n, \boldsymbol{\theta}_{-n})}{\partial \eta}\Big]\\&=\mathbb{E}_{\boldsymbol{\theta}_{-n_0}}\Bigg[\frac{\partial}{\partial \eta}\Big(\frac{1}{2}\cdot \eta\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})\big]^2 + \int_{\eta}^{\bar{\theta}}\big[Q^{\text{P}}(z, \boldsymbol{\theta}_{-n})\big]^2 \mathrm{d}z\\&\qquad + \big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n_0}) + 1/\sigma_0^2\big]\eta\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})\big] - \frac{\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + 1/\sigma_0^2\big]^2}{1/\sigma_0^2 + q_n}\eta\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})\big] - \frac{1}{2}\theta_n q_n^2\Big)\Bigg]\\&=\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Bigg[\Big(1 - \frac{Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + 1/\sigma_0^2}{q_n + 1/\sigma_0^2}\Big)\Big(2Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) \cdot \eta \cdot \frac{\partial Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})}{\partial \eta} + \big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + 1/\sigma_0^2\big] \cdot Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})\\&\qquad + \big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + 1/\sigma_0^2\big] \cdot \eta \cdot \frac{\partial}{\partial \eta}\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})\big]\Big)\Bigg]\end{aligned} \tag{53}$$

We can check that if

$$-\frac{\partial Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})/(Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + 1/\sigma_0^2)}{\partial \theta_n/\theta_n} \geq \frac{1}{3}, \forall n \in \mathcal{A}, \tag{54}$$

we have

$$3Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) \cdot \eta \cdot \frac{\partial Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})}{\partial \eta} + \big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + 1/\sigma_0^2\big] \cdot Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) \leq 0$$

Later we will show that $Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})$ is non-increasing in η, $\forall n \in \mathcal{A}$ (i.e., Lemma 6), hence, the term

$$3Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})\eta\frac{\partial Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})}{\partial \eta} + \big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + 1/\sigma_0^2\big]Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n}) + \frac{\eta}{\sigma_0^2}\frac{\partial}{\partial \eta}\big[Q^{\text{P}}(\eta, \boldsymbol{\theta}_{-n})\big] \leq 0.$$

Because the agent n will exert effort q_n to maximize her expected payoff. By taking the first order derivative of (51) with respect to q_n and set it to zero, we have

$$(1/\sigma_0^2 + q_n^{\text{P}})^2 \cdot q_n^{\text{P}} \cdot \hat{\theta}_n = (1/\sigma_0^2 + q_n)^2 \cdot q_n \cdot \theta_n, \tag{55}$$

where q_n^{P} is a shorthand for $Q^{\text{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n})$.

Based on (55), we have (i) if $\hat{\theta}_n > \theta_n$, $q_n^{\text{P}} < q_n$, (ii) if $\hat{\theta}_n < \theta_n$, $q_n^{\text{P}} > q_n$, and (iii) if $\hat{\theta}_n = \theta_n$, $q_n^{\text{P}} = q_n$.

Then if $\hat{\theta}_n > \theta_n$, the equation (53) is negative and (52) is positive. This inequality also holds for $\hat{\theta}_n < \theta_n$, by a similar argument. Therefore, the agent n will truthfully report her own cost type.

Step 3. Incentivized agents to exert precisely the effort as desired by the principal under COPE

Then we will show that the agent n, $\forall n \in \mathcal{A}$ exerts effort such that $q_n = q_n^{\text{P}}$ would maximize her expected payoff which is given as

$$\begin{aligned}&\mathbb{E}_{\{x^*, y_n, \boldsymbol{\theta}_{-n}\}}\left[U^{\text{A}}(x^*, q_n, y_n, \theta_n, \boldsymbol{\theta}_{-n})\right] \\ &= \mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[\pi(\theta_n, \boldsymbol{\theta}_{-n}) - K(\theta_n, \boldsymbol{\theta}_{-n}) \cdot \frac{1}{1/\sigma_0^2 + q_n} + S(\theta_n, \boldsymbol{\theta}_{-n}) - \frac{1}{2}\theta_n q_n^2\right].\end{aligned} \tag{56}$$

where the expectation is taken with respect to $\boldsymbol{\theta}_{-n}$, x^*, and y_n.

It can be verified that (56) is concave in q_n. Hence, by taking the first order derivative of (56) with respect to q_n, we have

$$\frac{\partial}{\partial q_n}\mathbb{E}\left[U^{\text{A}}(x^*, q_n, y_n, \theta_n, \boldsymbol{\theta}_{-n})\right] = \left[\frac{1/\sigma_0^2 + q_n^{\text{P}}}{1/\sigma_0^2 + q_n}\right]^2 \cdot \theta_n \cdot q_n^{\text{P}} - \theta_{n_0} \cdot q_n. \tag{57}$$

We can verify that the value of (57) equals to zero only when $q_n = q_n^{\text{P}}$. Hence, agent n will exert the effort as the principal desires to maximize her expected payoff. Then (49) is rewritten as

$$\hat{y}_n = \frac{\mu_0 \cdot 1/\sigma_0^2 + y_n \cdot q_n^{\text{P}}}{1/\sigma_0^2 + q_n^{\text{P}}}. \tag{58}$$

Because the principal knows the value of μ_0, σ_0^2, and q_n^{P}, he can infer the agent n's truth observation y_n from (58).

Step 4. Maximizes the principal's expected utility under COPE Then we look at the expected payoff of the principal. We then show that the desired effort level $Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ defined in (9) and the function $\pi(\theta_n, \boldsymbol{\theta}_{-n})$ defined in (46) can maximize the principal's expected payoff and satisfy BIC and BIR condition.

Notice that the agent n, $\forall n \in \mathcal{A}$ exerts effort such that $q_n = q_n^{\text{P}}$ and reports $\hat{y}_n = \frac{\mu_0 \cdot 1/\sigma_0^2 + y_n \cdot q_n^{\text{P}}}{1/\sigma_0^2 + q_n^{\text{P}}}$, the expected payment function is reduced to

$$\begin{aligned}&\mathbb{E}_{\{x^*, y_n, \boldsymbol{\theta}_{-n}\}}\left[R(x^*, y_n, q_n, \theta_n, \boldsymbol{\theta}_{-n})\right] \\ &= \mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[\pi(\theta_n, \boldsymbol{\theta}_{-n}) - K(\theta_n, \boldsymbol{\theta}_{-n})(x^* - \hat{y}_n)^2 + S(\theta_n, \boldsymbol{\theta}_{-n})\right] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[\pi(\theta_n, \boldsymbol{\theta}_{-n})\right].\end{aligned} \tag{59}$$

The expected payoff of the agent n is rewritten as

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[U^{\text{A}}\left(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n\right)\right] = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}) - \frac{1}{2}\theta_n \cdot \left[Q^{\text{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n})\right]^2\right], \tag{60}$$

and the BIC and BIR conditions, i.e., (4) and (5) can be rewritten as

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[U^{\text{A}}\left(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n\right)\right] \geq \mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[U^{\text{A}}\left(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n\right)\right], \ \forall \hat{\theta}_n \neq \theta_n \tag{61}$$

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[U^{\text{A}}\left(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\text{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n\right)\right] \geq 0, \ \forall \theta_n. \tag{62}$$

Base on Lemma 2 and (59), the expected payoff of the principal is

$$\begin{aligned}\mathbb{E}_{x^*,\boldsymbol{y},\boldsymbol{\theta}}[U^{\mathrm{P}}(x^*,\boldsymbol{q}^{\mathrm{P}},\boldsymbol{y},\hat{\boldsymbol{\theta}})] &= -B^{\mathrm{P}}\big(\boldsymbol{q}^{\mathrm{P}}\big) - \mathbb{E}_{x^*,\boldsymbol{y},\boldsymbol{\theta}}\big[\sum_{n\in\mathcal{A}} R(x^*,y_n,\theta_n,\boldsymbol{\theta}_{-n})\big] \\ &= -\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}} q_n^{\mathrm{P}}} - \mathbb{E}_{\boldsymbol{\theta}}\big[\sum_{n\in\mathcal{A}} \pi(\theta_n,\boldsymbol{\theta}_{-n})\big].\end{aligned}$$

Recall that $q_n^{\mathrm{P}} = Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})$, the principal's optimal problem defined in (6) can be rewritten as

$$\begin{aligned}\sup_{\{Q^{\mathrm{P}}(\boldsymbol{\theta}),\pi(\boldsymbol{\theta})\},\forall\theta_n\in\boldsymbol{\theta},\forall n\in\mathcal{A}} \quad & \mathbb{E}[U^{\mathrm{P}}(x^*,\boldsymbol{q}^{\mathrm{P}},\boldsymbol{y},\hat{\boldsymbol{\theta}})], \\ \text{subject to}: \quad & \text{BIC and BIR in (61) and (62).}\end{aligned} \tag{63}$$

In the following lemmas, we characterize an equivalent formulation for the feasible region defined by BIC and BIR. Using these lemmas, we show that $Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})$ defined in (9) and $\pi(\theta_n,\boldsymbol{\theta}_{-n})$ defined in (46) are the optimal solution that solves the principal's problem in (63).

Lemma 6. *The solution of* (63) *is feasible if and only if it satisfies the following conditions for all* $\theta_n \in [\underline{\theta},\bar{\theta}]$*:*

- *the expected payoff of the agent* n*,* $\forall n\in\mathcal{A}$ *is*

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\big] = \frac{1}{2}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[\int_{\theta_n}^{\bar{\theta}} \big[Q^{\mathrm{P}}(x,\boldsymbol{\theta}_{-n})\big]^2 \mathrm{d}x\right] \tag{64}$$

- $Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})$ *is non-increasing in* θ_n*.*

Proof. The proof of Lemma 6 is as follows. We first show that BIC and BIR imply the condition in (64).

Notice that the first derivative of (60) is:

$$\frac{\partial\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\big]}{\partial\theta_n} = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[-\frac{1}{2}\big[Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n})\big]^2\big] \le 0. \tag{65}$$

Then, for any $\theta_n^1 > \theta_n^2$, we have

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\pi(\theta_n^1,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n^1,\boldsymbol{\theta}_{-n}),\theta_n^1)\big] &\le \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\pi(\theta_n^1,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n^1,\boldsymbol{\theta}_{-n}),\theta_n^2)\big] \\ &\le \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\pi(\theta_n^2,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n^2,\boldsymbol{\theta}_{-n}),\theta_n^2)\big].\end{aligned} \tag{66}$$

where the first inequality is because (65) and the second is from the BIC condition defined in (61).

Recall that the BIR condition is

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\big] \ge 0,\ \forall\theta_n\in[\underline{\theta},\bar{\theta}], \tag{67}$$

which implies that, for the agent $n\in\mathcal{A}$ with any value $\theta_n\in[\underline{\theta},\bar{\theta}]$, her expected payoff should at least be zero. Then the expected payoff of the agent n with cost parameter $\bar{\theta}$ must be binding at zero. Otherwise, the principal can raise the $\pi(\bar{\theta},\boldsymbol{\theta}_{-n})$ by a small value of $\delta>0$, which does not violate the constraint of (67) but raises the principal's expected payoff. Hence, we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\pi(\bar{\theta},\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\bar{\theta},\boldsymbol{\theta}_{-n}),\bar{\theta})\big] = 0. \tag{68}$$

Let $U^{\mathrm{A}}(\theta_n,\boldsymbol{\theta}_{-n}) = U^{\mathrm{A}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)$. From BIC condition, we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\theta_n,\boldsymbol{\theta}_{-n})\big] = \max_{\hat{\theta}_n}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\left[U^{\mathrm{A}}\Big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\theta_n\Big)\right].$$

By using the envelope theorem, we have:

$$\frac{\partial \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\theta_n,\boldsymbol{\theta}_{-n})\big]}{\partial \theta_n} = \frac{\partial \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\theta_n)\big]}{\partial \theta_n}\Bigg|_{\hat{\theta}_n=\theta_n} = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Bigg[-\frac{1}{2}\big[Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\big]^2\Bigg],$$

where θ_n is a parameter. By integrating both sides from the value of θ_n to $\bar{\theta}$ and using (68), we get

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\mathrm{A}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\Big] = \frac{1}{2}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[\int_{\theta_n}^{\bar{\theta}}\big[Q^{\mathrm{P}}(x,\boldsymbol{\theta}_{-n})\big]^2 \mathrm{d}x\Big] \tag{69}$$

We prove $Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_n)$ is nonincreasing in θ_n by contradiction. Let p_n as the shorthand for $\pi(\theta_n,\boldsymbol{\theta}_{-n})$. Suppose for any $\theta_n^1 > \theta_n^2$, we have $Q^{\mathrm{P}}(\theta_n^1,\boldsymbol{\theta}_{-n}) > Q^{\mathrm{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})$. Because

$$\frac{\partial^2 U^{\mathrm{A}}\big(p_n,q_n^{\mathrm{P}},\theta_n\big)}{\partial q_n^{\mathrm{P}}\partial\theta_n} = -q_n^{\mathrm{P}} < 0, \quad \text{and} \tag{70}$$

$$\frac{\partial^2 U^{\mathrm{A}}\big(p_n,q_n^{\mathrm{P}},\theta_n\big)}{\partial {q_n^{\mathrm{P}}}^2} = -\theta_n \leq 0. \tag{71}$$

we have

$$\begin{aligned} 0 = & \frac{\partial U^{\mathrm{A}}\big(p_n,q_n^{\mathrm{P}},\theta_n^1\big)}{\partial q_n^{\mathrm{P}}}\Bigg|_{q_n^{\mathrm{P}}=Q^{\mathrm{P}}(\theta_n^1,\boldsymbol{\theta}_{-n})} \\ \leq & \frac{\partial U^{\mathrm{A}}\big(p_n,q_n^{\mathrm{P}},\theta_n^1\big)}{\partial q_n^{\mathrm{P}}}\Bigg|_{q_n^{\mathrm{P}}=Q^{\mathrm{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})} \\ < & \frac{\partial U^{\mathrm{A}}\big(p_n,q_n^{\mathrm{P}},\theta_n^2\big)}{\partial q_n^{\mathrm{P}}}\Bigg|_{q_n^{\mathrm{P}}=Q^{\mathrm{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})}, \end{aligned}$$

where the first equality is because of BIC when the agent n's cost parameter θ_n has the value of θ_n^1, the second equality is because of (70), and the inequality is because of (71).

However, based on the BIC condition, if the agent n's cost parameter θ_n has the value of θ_n^2, then we should have

$$\frac{\partial U^{\mathrm{A}}\big(p_n,q_n^{\mathrm{P}},\theta_n^2\big)}{\partial q_n^{\mathrm{P}}}\Bigg|_{q_n^{\mathrm{P}}=Q^{\mathrm{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})} = 0,$$

which holds true for all scalar value of p_n. Hence, for any $\theta_n^1 > \theta_n^2$, $Q^{\mathrm{P}}(\theta_n^1,\boldsymbol{\theta}_{-n}) \leq Q^{\mathrm{P}}(\theta_n^2,\boldsymbol{\theta}_{-n})$.

Then we need to prove that (64) implies BIC and BIR defined in (61) and (62). Notice that we have:

$$\begin{aligned} & \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\mathrm{A}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\Big] \\ & = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[-\int_{\theta_n}^{\bar{\theta}}\frac{\partial U^{\mathrm{A}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),z\big)}{\partial z}\mathrm{d}z\Big] \\ & = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[\frac{1}{2}\int_{\hat{\theta}_n}^{\bar{\theta}}\big[Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n})\big]^2\mathrm{d}\eta - \int_{\theta_n}^{\hat{\theta}_n}\frac{\partial U^{\mathrm{A}}\big(\pi(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n}),z\big)}{\partial z}\mathrm{d}z\Big] \\ & = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[-\frac{1}{2}\int_{\bar{\theta}}^{\theta_n}\big[Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n})\big]^2\mathrm{d}\eta - \frac{1}{2}\int_{\theta_n}^{\hat{\theta}_n}\big[Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n})\big]^2\mathrm{d}\eta + \frac{1}{2}\int_{\theta_n}^{\hat{\theta}_n}\big[Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n})\big]^2\mathrm{d}z\Big] \\ & = \mathbb{E}_{\boldsymbol{\theta}_{-n}}\Big[U^{\mathrm{A}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big) + \frac{1}{2}\int_{\theta_n}^{\hat{\theta}}\Big(\big[Q^{\mathrm{P}}(\hat{\theta}_n,\boldsymbol{\theta}_{-n})\big]^2 - \big[Q^{\mathrm{P}}(\eta,\boldsymbol{\theta}_{-n})\big]^2\Big)\mathrm{d}\eta\Big] \end{aligned}$$

where the second equality and the forth equality is obtained by (64).

If $\hat{\theta}_n > \theta_n$, then the above equation is non-positive (because $Q^{\mathrm{P}}(\eta, \boldsymbol{\theta}_{-n})$ is non-increasing in η) and hence

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\pi(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), Q^{\mathrm{P}}(\hat{\theta}_n, \boldsymbol{\theta}_{-n}), \theta_n)\big] < \mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[U^{\mathrm{A}}(\pi(\theta_n, \boldsymbol{\theta}_{-n}), Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n}), \theta_n)\big].$$

This inequality also holds for $\hat{\theta}_n < \theta_n$ by a similar argument. Therefore, the two condition imply BIC.

BIR is verified by putting θ_n back to (64). □

Then based on Lemma 6, we have the following lemma.

Lemma 7. *The optimisation problem in (63) has the following equivalent formulation:*

$$\max_{\{Q^{\mathrm{P}}(\boldsymbol{\theta})\},\forall\theta_n\in\boldsymbol{\theta}} \mathbb{E}_{\boldsymbol{\theta}}\left[-\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}}Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})} - \frac{1}{2}\sum_{n\in\mathcal{A}}\left[Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\right]^2\theta_n - \frac{1}{2}\sum_{n\in\mathcal{A}}\left[Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\right]^2\frac{F(\theta_n)}{f(\theta_n)}\right],$$
$$\text{s.t. } Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n}) \text{ is nonincreasing in } \theta_n. \tag{72}$$

Proof. The proof of Lemma 7 is as follows. The expected payoff of the principal can be written as:

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}}\left[-\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}}Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})} - \frac{1}{2}\sum_{n\in\mathcal{A}}\left[Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\right]^2\cdot\theta_n - \sum_{n\in\mathcal{A}}U^{\mathrm{A}}\big(\pi(\theta_n,\boldsymbol{\theta}_{-n}),Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n}),\theta_n\big)\right]\\
&=\mathbb{E}_{\boldsymbol{\theta}}\left[-\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}}Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})} - \frac{1}{2}\sum_{n\in\mathcal{A}}\left[Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\right]^2 - \frac{1}{2}\sum_{n\in\mathcal{A}}\int_{\theta_n}^{\bar{\theta}}\left[Q^{\mathrm{P}}(x,\boldsymbol{\theta}_{-n})\right]^2\mathrm{d}x\right]
\end{aligned}$$

Using integration by parts and Lemma 6, we can get the conclusion. □

Based on Lemma 7, the principal's problem thus reduces to choosing the desired effort $Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ for each agent $n \in \mathcal{A}$.

Let $q_n^{\mathrm{P}} = Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ and

$$M\big(q_1^{\mathrm{P}},\ldots,q_N^{\mathrm{P}}\big) = -\frac{1}{1/\sigma_0^2+\sum_{n\in\mathcal{A}}q_n^{\mathrm{P}}} - \frac{1}{2}\sum_{n\in\mathcal{A}}\left[q_n^{\mathrm{P}}\right]^2\cdot\theta_n - \frac{1}{2}\sum_{n\in\mathcal{A}}\left[q_n^{\mathrm{P}}\right]^2\frac{F(\theta_n)}{f(\theta_n)}$$

Let $G = [\partial^2 M/\partial q_i^{\mathrm{P}}\partial q_j^{\mathrm{P}}]$ is the matrix of second order derivatives and is a symmetric matrix with negative diagonal terms as

$$\begin{aligned}
\frac{\partial^2 M}{\partial q_i^{\mathrm{P}}\partial q_j^{\mathrm{P}}} &= -\frac{2}{\left[1/\sigma_0^2+\sum_{n\in\mathcal{A}}q_n^{\mathrm{P}}\right]^3}, j\neq i\\
\frac{\partial^2 M}{\partial {q_i^{\mathrm{P}}}^2} &= -\frac{2}{\left[1/\sigma_0^2+\sum_{n\in\mathcal{A}}q_n^{\mathrm{P}}\right]^3} - \theta_i - \frac{F(\theta_i)}{f(\theta_i)}
\end{aligned}$$

As we can verify that, for $k = 1, \ldots, N$, the kth leading principal minors of G alternate in sign, hence G is negative definite and M is strictly concave. Thus, the principal's desired effort level from agents $q_n^{\mathrm{P}} = Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n})$, $\forall n \in N$ is the solution of the below equations:

$$\frac{1}{\left[1/\sigma_0^2+\sum_{n\in\mathcal{A}}q_n^{\mathrm{P}}\right]^2} - Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\cdot\theta_n - Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})\cdot\frac{F(\theta_n)}{f(\theta_n)} = 0, n = 1,2,\ldots,N \tag{73}$$

Using Cramer's rule and the assumption that the c.d.f. function F is log concave in θ , we can verify that

$$\frac{\partial Q^{\mathrm{P}}(\theta_n,\boldsymbol{\theta}_{-n})}{\partial\theta_n} = -\frac{\partial^2 M/\partial q_n^{\mathrm{P}}\partial\theta_n}{\partial^2 M/\partial {q_n^{\mathrm{P}}}^2} \leq 0, \tag{74}$$

which shows that $Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n})$ derived from (73) is non increasing in θ_n, so that it is the feasible solution of (72).

The solution of (73) is

$$Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n}) = \frac{1}{\theta_n + \frac{F(\theta_n)}{f(\theta_n)}} \cdot \frac{1}{\big[W(\boldsymbol{\theta})\big]^2}, \tag{75}$$

where the function $W : [\underline{\theta}, \bar{\theta}]^N \to \mathbb{R}_+$ is the solution of the below equation:

$$\big[W(\boldsymbol{\theta})\big]^3 - \frac{1}{\sigma_0^2} \cdot \big[W(\boldsymbol{\theta})\big]^2 - \sum_{m \in \mathcal{A}} \frac{1}{\theta_m + \frac{F(\theta_m)}{f(\theta_m)}} = 0. \tag{76}$$

The real root of the above cubic equation is as follows.

$$W(\hat{\boldsymbol{\theta}}) = \frac{1}{3\sigma_0^2} + \sqrt[3]{\frac{1}{27\sigma_0^6} + \frac{1}{2}\Bigg[\sum_{m \in \mathcal{A}} \frac{1}{\hat{\theta}_m + \frac{F(\hat{\theta}_m)}{f(\hat{\theta}_m)}}\Bigg] + \sqrt{\lambda(\boldsymbol{\theta})}} + \sqrt[3]{\frac{1}{27\sigma_0^6} + \frac{1}{2}\Bigg[\sum_{m \in \mathcal{A}} \frac{1}{\hat{\theta}_m + \frac{F(\hat{\theta}_m)}{f(\hat{\theta}_m)}}\Bigg] - \sqrt{\lambda(\boldsymbol{\theta})}}, \tag{77}$$

where function $\lambda : [\underline{\theta}, \bar{\theta}]^N \to \mathbb{R}_+$ is given as

$$\lambda(\boldsymbol{\theta}) = \frac{1}{27\sigma_0^6}\Bigg[\sum_{m \in \mathcal{A}} \frac{1}{\hat{\theta}_m + \frac{F(\hat{\theta}_m)}{f(\hat{\theta}_m)}}\Bigg] + \frac{1}{4}\Bigg[\sum_{m \in \mathcal{A}} \frac{1}{\hat{\theta}_m + \frac{F(\hat{\theta}_m)}{f(\hat{\theta}_m)}}\Bigg]^2.$$

According to (64), we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[\pi(\theta_n, \boldsymbol{\theta}_{-n}) - \frac{1}{2} \cdot [Q^{\mathrm{P}}(\theta_n, \boldsymbol{\theta}_{-n})]^2 \cdot \theta_n\big] = \frac{1}{2}\mathbb{E}_{\boldsymbol{\theta}_{-n}}\big[\int_{\theta_n}^{\bar{\theta}} \big[Q^{\mathrm{P}}(x, \boldsymbol{\theta}_{-n})\big]^2 \mathrm{d}x\big].$$

From the above equation, we can derive the optimal payment function as given in (46).

□

B Details of Simulations

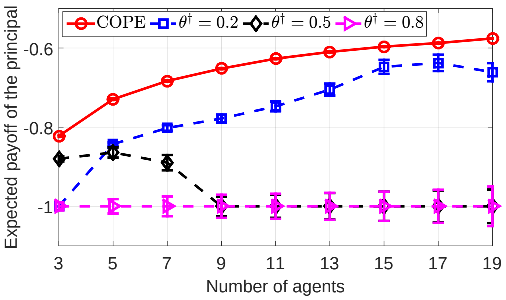
In this section, we first describe the details of the homogenous benchmark mechanism under both the linear and quadratic cost function. Then we compare the performance of COPE to the homogeneous benchmark in terms of expected prediction error and total payment made by the principal to the agents.

B.1 Homogenous Mechanism

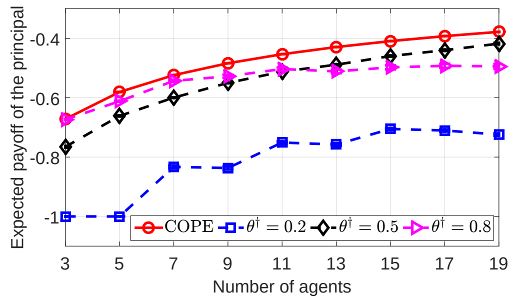
The homogenous mechanism assumes all agents to be identical (although in practice they are not), and hence does not elicit the cost parameters of individual agents. In the absence of this knowledge, the principal operates under the belief that every agent's cost parameter equals $\theta^\dagger \in [\underline{\theta}, \bar{\theta}]$. The principal thus chooses payment function $R_{\mathrm{hom}} := \alpha(\theta^\dagger) - \beta(\theta^\dagger) \cdot (x^* - \hat{y}_n)^2$, where the function $\alpha : [\underline{\theta}, \bar{\theta}] \to \mathbb{R}_+$ and the function $\beta : [\underline{\theta}, \bar{\theta}] \to \mathbb{R}_+$ are chosen to incentivize every agent n to exert optimal effort and report observations truthfully in a manner that maximizes the principal's payoff.

Recall that we assume the prior on $x^* \sim \mathcal{N}(\mu_0, \sigma_0^2)$. Let the function $q^\dagger : \mathbb{R}_+ \times [\underline{\theta}, \bar{\theta}] \to \mathbb{R}_+$ to be the effort that the principal requires every agent to exert, based on the principal's belief that every agent's cost parameter equals to $\theta^\dagger$. Then the principal makes the prediction as

$$\hat{x} = \frac{\mu_0/\sigma_0^2 + q^\dagger \sum_{n \in \mathcal{A}} g(\hat{y}_n)}{1/\sigma_0^2 + N \cdot q^\dagger}, \tag{78}$$



(a) Linear cost function.



(b) Quadratic cost function.

Figure 3: The principal's expected payoff under COPE and the homogeneous mechanism.

where $\hat{y}_n$ is the agent n's reported observation, and the function $g : \mathbb{R} \to \mathbb{R}$ is defined as

$$g(\hat{y}_n) = \hat{y}_n + \frac{\hat{y}_n - \mu_0}{q^\dagger \cdot \sigma_0^2}. \tag{79}$$

Linear Cost Function Under the linear cost function, the choice of function $q^\dagger$, α, and β are :

$$q^\dagger(N, \theta^\dagger) = \frac{1}{N}\left(\frac{1}{\sqrt{\theta^\dagger}} - \frac{1}{\sigma_0^2}\right),$$

$$\alpha(\theta^\dagger) = \left(1/\sigma_0^2 + q^\dagger\right) \cdot \theta^\dagger q^\dagger + \theta^\dagger q^\dagger, \quad \beta(\theta^\dagger) = \left(1/\sigma_0^2 + q^\dagger\right)^2 \cdot \theta^\dagger$$

The principal chooses the function α to make sure that the agent n with cost type $\theta^\dagger$ is willing to participate the prediction task, and chooses the function β to make sure that the agent n exerts the effort $q_n = q^\dagger(N, \theta^\dagger)$ as the principal desires.

Recall that the actual cost parameter of the agent $n \in \mathcal{A}$ is θ_n. Hence, the agent n will exert effort $q_n = \sqrt{\beta(\theta^\dagger)/\theta_n} - 1/\sigma_0^2$ and report $\hat{y}_n = \frac{\mu_0/\sigma_0^2 + y_n \cdot q_n}{1/\sigma_0^2 + q_n}$ to maximize her own expected payoff. Besides, if the expected payoff of the agent n is negative, he will not participate this prediction task.

Also recall that the principal knows the prior information of $x^* \sim \mathcal{N}(\mu_0, \sigma_0^2)$. Hence, the principal can always achieve a payoff of $-1/\sigma_0^2$ by not making any payments, and simply choosing the prior mean has his prediction. Hence, the principal does not pay anything and simply sets $\hat{x} = \mu_0$ if his expected payoff is smaller than $-1/\sigma_0^2$.

Quadratic Cost Function Under the quadratic cost function,$q^\dagger$ is the solution of the below equation.

$$\frac{1}{\left(1/\sigma_0^2 + N \cdot q^\dagger\right)^2} - \cdot \theta^\dagger \cdot q^\dagger = 0$$

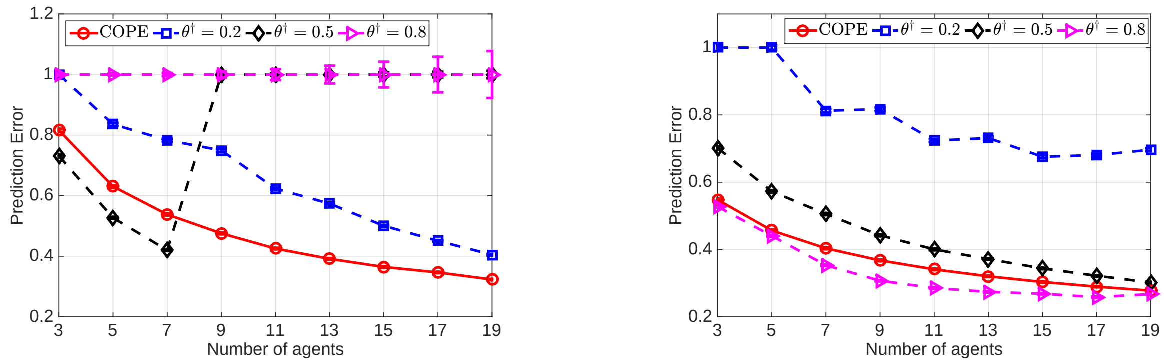
The function α, and β are :

$$\alpha(\theta^\dagger) = \left(1/\sigma_0^2 + q^\dagger\right) \cdot \theta^\dagger q^\dagger + \frac{1}{2}\theta^\dagger \left[q^\dagger\right]^2, \quad \beta(\theta^\dagger) = \left(1/\sigma_0^2 + q^\dagger\right)^2 \cdot \theta^\dagger q^\dagger.$$

Recall that the actually cost parameter of the agent $n \in \mathcal{A}$ is θ_n. Hence, the agent n will exert effort q_n to maximize her own expected payoff, where q_n is the solution of

$$\frac{\beta}{\left(1/\sigma_0^2 + q_n\right)^2} - \theta_n q_n = 0.$$

Besides, if the expected payoff of the agent n is negative, he will not participate this prediction task. Also, the principal does not pay anything and simply sets $\hat{x} = \mu_0$ if his expected payoff is smaller than $-1/\sigma_0^2$.



(a) Linear cost function.

(b) Quadratic cost function.

Figure 4: The principal's prediction error under COPE and the homogeneous mechanism.

B.2 Numerical Results

In the simulations, we draw $x^* \sim \mathcal{N}(0,1)$, and set $\underline{\theta} = 0$ and $\bar{\theta} = 1$. We vary the number of agents from $N = 3$ to $N = 19$. Without loss of generality, we normalize the principal's payoff so that it equals zero in the ideal (unachievable) case of zero prediction error and a zero payment. Note that the principal can always achieve a payoff of -1 by not making any payments, and simply choosing the prior mean has her prediction.

Figure 3 depicts the expected payoff of the principal under COPE and under the homogeneous mechanism for different values of $\theta^\dagger$. We use the red line with circle marker denotes COPE, the blue dash line with square marker denotes homogeneous mechanism with $\theta^\dagger = 0.2$, the dark dash line with diamond marker denotes homogeneous mechanism with $\theta^\dagger = 0.5$, and the magenta line with right-pointing triangle marker denotes homogeneous mechanism with $\theta^\dagger = 0.8$. One can draw the following insights from the figure. First, under COPE, the principal's expected payoff strictly increases with the number of agents. This is because COPE optimally exploits the presence of additional agents by making them exert different efforts based on their respective cost types. A second inference is that this feature of COPE allows COPE to outperform the homogeneous mechanism consistently, and the difference depends on the principal's belief of $\theta^\dagger$.

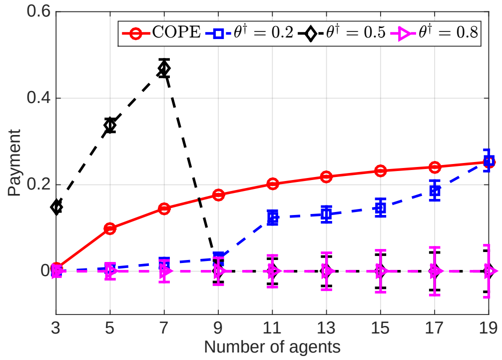
By comparing Figure 3a to Figure 3b, we can see that the belief of principal (i.e., $\theta^\dagger$) under the homogeneous mechanism leads to different results. Under the linear cost function, the lower value of $\theta^\dagger$ results in the good performance in terms of the principal's payoff. However, under the quadratic cost function, the high value of $\theta^\dagger$ results in the good performance. The reasons are as follows.

Under the homogeneous mechanism, setting different value of $\theta^\dagger$ may filtering different number and types of agents. Setting a high value of $\theta^\dagger$ would incentivize most of the agent to participate. This is because, for the agent $n \in \mathcal{A}$ whose cost parameter $\theta_n < \theta^\dagger$, she can put less effort to achieve the same performance as the agent with cost parameter $\theta^\dagger$ does.

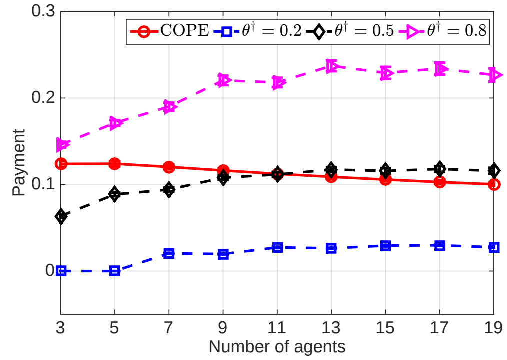
Under the linear cost function, similar as COPE, finding the most capable one would be optimal for the principal, as the marginal cost is nonnegative even the agent does not put any effort. Hence, setting the low value of $\theta^\dagger$ would eliminate the most of the agents, and have a high chance to find the agent with $\theta_n \leq \theta^\dagger$.

On the contrary, under the quadratic cost function, it would be optimal to recruit as many as the agents to improve the prediction accuracy. The benefit brought by the accuracy improvement would be higher than the payment paid to agents. Hence, setting the high value of $\theta^\dagger$ would help the principal recruit most of the agents.

Figure 4 depicts the expected prediction error (i.e., the expected value of $(x^* - \hat{x})^2$) made by the principal under COPE and under the homogeneous mechanism for different values of $\theta^\dagger$. We can see that COPE can achieve lowest prediction error at most cases. As there is a tradeoff between the



(a) Linear cost function.



(b) Quadratic cost function.

Figure 5: The principal's total payment under COPE and the homogeneous mechanism.

prediction accuracy and the payment made by the principal, the aim of COPE is to maximize the principal's expected payoff. Hence, the prediction error made by COPE is not the lowest at some cases.

Figure 5 shows the total payment made by COPE and under the homogeneous mechanism for different values of $\theta^\dagger$. In order to elicit the private information of the agents (i.e., their cost parameters), the payment made by the principal under COPE would be high, in return, the principal can explore the heterogeneous of the agents and improve his prediction accuracy.