Incentivizing Data Collaboration: A Mechanism Design Approach

Saeed Alaei[1], **Ali Daei Naby**[2], **Ali Makhdoumi**[3], **Azarakhsh Malekian**[2,1]

[1]Google
[2]Rotman School of Management, University of Toronto
[3]Fuqua School of Business, Duke University

Abstract

We study the problem of incentivizing strategic agents to truthfully contribute high-quality data in collaborative learning settings, where each agent benefits from improved estimation based on others' data. Each agent privately observes the quality of their data, and agents may misreport it if not incentivized properly. We cast this problem with a Bayesian mechanism design framework in which the platform aims to find the optimal data-sharing mechanism that jointly determines allocations and payments to maximize both estimation accuracy and platform revenue. We prove that the optimal mechanism that incentivizes truthful reporting takes the form of a *personalized threshold and pricing* mechanism, in which each agent is allocated the learned estimator if their reported quality exceeds a (personalized) threshold and is charged a price based on the relevance of other agents' data in the learning task. We analyze this mechanism in a canonical Gaussian mean estimation task, derive a closed-form solution to the optimal mechanism, and highlight how data correlation affects the mechanism. We further extend the model to allow agents to exert costly efforts to improve their data quality before collaboration. We show that "free-riding" is mitigated as the optimal data-sharing mechanism induces a supermodular game: each agent is incentivized to exert more effort when others exert more. Finally, we show that equilibrium efforts form a complete lattice, and in the highest-effort equilibrium, each agent increases effort as other agents' data becomes more relevant in the learning task.

1 Introduction

Collaborative learning is increasingly prevalent in settings where data and expertise are distributed across multiple stakeholders. A growing number of platforms now facilitate such collaboration—bringing together data contributors, coordinating computation, and monetizing the value created in the process. For example, cloud-based platforms like AWS and Microsoft Azure offer tools for federated or distributed learning across multiple agents. Snowflake, originally a data warehousing firm, now offers a data marketplace where firms can securely share and monetize their data, enabling cross-organizational analytics and model building. Similarly, Databricks supports collaborative machine learning by allowing data (and expertise) from different organizations to be used together. These platforms benefit not only from hosting and processing fees but also by charging for value-added services such as data access, model outputs, and performance improvements. In all these cases, effective collaboration hinges on aligning incentives among data holders, especially when the quality of shared data is private information of the agents, which is the topic of this paper.

Overview of Our Model and Main Results

In this paper, we study the optimal design of collaborative learning platforms involving multiple *agents*, each with its own dataset and its own learning objective. The data of each agent is useful for their personal estimation task but also for the learning tasks of others, which motivates collaborative learning. As a concrete example, consider firms estimating local market demand elasticities: a retailer in one region may benefit from demand estimates in nearby areas if consumer behavior is spatially correlated. The firms use a platform that facilitates collaboration among them but also charges them, and we are interested in finding the optimal mechanism that maximizes the platform's revenue but also incentivizes the firms to truthfully report their data quality.

We introduce a Bayesian mechanism design framework for this problem in which the platform designs a *data-sharing mechanism* that consists of an allocation probability function that, based on the reported qualities, determines the probability of giving each agent access to the learned model and a payment scheme, again based on the reported qualities, that determines how much to charge each agent. The goal of the platform in designing the allocation and payment is twofold: 1) incentivizing each agent to truthfully report their quality so that the overall performance of learning tasks is maximized and 2) maximizing the revenue generated from providing this service to the agents.

Our first main result characterizes the optimal mechanism by which a platform can elicit truthful reports of privately known data quality from agents and maximize its revenue. The optimal mechanism takes the form of a *personalized threshold and pricing* mechanism and, even though we allow for randomized allocations, deterministically decides which agents to allocate to. Here is the structure of the optimal mechanism:

- Collect the reported qualities of all agents and compute a personalized threshold and a personalized price for each agent $i \in \{1, \dots, n\}$ based on all the reported qualities and the relevance of other agents' data in the learning task of agent i.
- Allocate the learned model to each agent i, if her reported quality exceeds her personalized threshold and charge her the personalized price.

We carry out our analysis for general learning tasks in which the data of others are useful for the learning task of each agent. We then derive an explicit mechanism for our canonical example of estimating a parameter of interest from Gaussian signals. In the example mentioned above, the parameter of interest for each firm is the elasticity of demand that is correlated across the firms. Each firm collects some samples from its targeted population, resulting in a Gaussian estimation of its demand elasticity. We show that as the correlation between the parameter of agent i and those of others increases—or as the quality of others' data improves—the platform achieves higher model accuracy and charges agent i more. This finding confirms the intuition that if the "added value" of the platform for an agent is higher, she will be charged more. In fact, our result quantifies this added value and the corresponding payment.

In our second main result, we extend the analysis to a setting in which the agents can exert some costly effort to improve their data quality before collaborating on the platform. As already documented in the literature (e.g., [1]), such settings typically exhibit "free-riding" among agents—meaning each agent may be hesitant to exert the costly effort and rely on the others' efforts. However, in our setting of estimating parameters of interest from Gaussian signals, this free-riding effect is mitigated by the optimal data-sharing mechanism, as it induces a supermodular game: each agent is incentivized to exert more effort in response to greater effort from others. We then show that the set of all equilibria forms a complete lattice, meaning that all equilibria can be sorted by the total effort exerted. By focusing on the highest-effort equilibrium, we establish that the effort exerted by each agent increases as the parameters of interest become more correlated.

Related Work

Incentivizing Participation and Data Collaboration in Federated Learning. The issue of incentivizing participation in federated learning has been studied in a range of works [1–11]. See [4] for a survey. Our setting departs from this line of work in two key ways. First, data quality is private information that can be strategically misreported. Second, we account for heterogeneity in the relevance of agents' data to the learning task, which directly impacts incentives and model accuracy. [12–14] study incentives for truthful reporting in gradient updates and mean estimation settings. We differ from this line of work in two main aspects. First, we fully characterize the optimal mechanism for general collaborative learning tasks, including both gradient updates and mean estimation. Second, the platform's objective in our setting is to jointly optimize model accuracy and revenue, requiring a mechanism design approach that integrates both monetary and non-monetary incentives.

Mechanism Design. Our paper is related to the literature on mechanism design pioneered by [15]. However, in our setting, each agent's value depends not only on her type but also on the types of all other agents, which makes both the analysis and results different from the classical setting. This interdependence connects our work to the literature on mechanism design with externalities [16–18].

The rest of the paper proceeds as follows. In Section 2, we describe the setting and the problem. In Section 3, we characterize the optimal data-sharing mechanism. In Section 4, we extend our results to a setting in which the agents can exert effort in order to increase their data quality. Section 5 concludes, while the appendix includes all the omitted proofs from the text.

2 Environment

We consider n agents, represented by $[n] = \{1, \ldots, n\}$. Each agent $i \in [n]$ has access to a data set whose *quality* is privately known to the agent herself and not the others. We can think of the quality as the number of data points or any other data feature that impacts the performance of the learning task. The collaboration is happening on a platform that, from previous interactions with each agent, has some information about their data quality. We model this information by adopting a *Bayesian* setting, assuming each agent i's data quality, denoted by q_i, is independently drawn from a publicly known and absolutely continuous distribution $F_i(\cdot)$ with support $\mathcal{Q}_i$ over a bounded interval.

When sharing their data with the platform, agent $i \in [n]$ reports a quality $\hat{q}_i$, which may differ from the true data quality q_i. The designer then performs the learning tasks, using the shared data from all agents, and shares the results with the agents. The performance of the learning task depends on both the true and reported data qualities. For instance, when data quality is measured by size, the commonly used weighted average method in federated learning places more weight on data with higher reported quality ([19]). In such settings, having the reported quality equal to the true quality improves the performance of the learning task, e.g., deriving an unbiased estimator when using the weighted average method. Therefore, the goal of the platform is to encourage *truthful reporting* of the data qualities, with the goal of achieving the highest accuracy for the underlying learning tasks and also generating revenue from interacting with the agents.

A Data-sharing Mechanism. Based on the reported qualities, the designer determines, potentially randomizing, whether agent $i \in [n]$ receives the outcome of the learning task for task i and the corresponding payment. Specifically, a *Data-sharing Mechanism* is specified by the tuple $\{x_i(\cdot), p_i(\cdot)\}_{i=1}$ where $x_i : \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_n \to [0, 1]$ and $p_i : \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_n \to \mathbb{R}$ for all $i = 1, \cdots, n$. Given the reported qualities $\hat{\mathbf{q}} = [\hat{q}_1, \ldots, \hat{q}_n]$, agent $i \in [n]$ receives the estimator for task i with probability $x_i(\hat{\mathbf{q}})$, which we call *allocation probability* and pays $p_i(\hat{\mathbf{q}})$, which we call *payment*.

Accuracy and Values. For each task $i \in [n]$, given datasets with true qualities $\mathbf{q} = [q_1, \ldots, q_n]$ and reported qualities $\hat{\mathbf{q}} = [\hat{q}_1, \ldots, \hat{q}_n]$, the (expected) accuracy is $A_i(\mathbf{q}, \hat{\mathbf{q}})$.[1] Notice that, to keep the analysis general, we assume that accuracy is a function of both the reported qualities (that the platform may use in the learning tasks) and the true qualities of data. Each agent i's value for an estimator, $v_i(\mathbf{q}, \hat{\mathbf{q}})$, is a strictly increasing and smooth function of task i's accuracy.

We keep the analysis for general accuracy functions as long as they satisfy the following assumption.

Assumption 1 (Accuracy Function). *For $i \in [n]$ and $\mathbf{q} \in \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_n$, the following hold: (i) $A_i(\mathbf{q}, \mathbf{q})$ is non-decreasing in q_i; (ii) $A_i(\mathbf{q}, \hat{\mathbf{q}})$ is maximized at $\hat{\mathbf{q}} = \mathbf{q}$, i.e., $\mathbf{q} = \operatorname{argmax}_{\hat{\mathbf{q}}} A_i(\mathbf{q}, \hat{\mathbf{q}})$; (iii) $A_i(\cdot)$ is twice continuously differentiable, i.e., it belongs to $\mathcal{C}^2$.*

Assumption 1(i) formalizes the notion of data quality, stating that higher data quality results in higher accuracy. Assumption 1(ii) asserts that the platform uses the best use of shared information, e.g., assigning correct weight to each agent's data. Assumption 1(iii) is a technical condition that we need for the analysis.

[1] We can view the accuracy as a random variable whose mean is $A_i(\mathbf{q}, \hat{\mathbf{q}})$, where the randomness comes from data.

Platform's Problem and Truthful Mechanisms. In principle, the interaction between the platform and the agents and the strategy of the agents can be complex. However, in the appendix, we find conditions under which, without loss of generality, we can focus on the class of truthful direct mechanisms—the allocation and payment are such that the agents find it optimal to report truthfully. Our analysis is more involved compared to the *revelation principle* of classical mechanism design. We present the formal argument and details in the appendix, and in the rest of the paper focus on direct and incentive-compatible (IC) mechanisms.

2.1 Mean Estimation from Gaussian Signals

Consider a collaborative learning setting where agent $i \in [n]$ is interested in estimating a parameter θ_i. These parameters are interdependent, and we let their distribution be $\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, an n-dimensional Gaussian vector with known covariance matrix Σ. We let Σ_{ij} denote the (i, j)-th entry of Σ and σ_i^2 denote the (i, i)-th entry of Σ. Each agent i has access to a data set that essentially gives the agent a noisy signal of the underlying parameter, denoted by $s_i = \theta_i + z_i$, where $z_i \sim \mathcal{N}\left(0, \eta_i^2\right)$ is independently distributed across agents. This example falls into our setting as the data/signal of each agent can be used to obtain a better estimation for the others. Here, we define data quality of agent i as $q_i \coloneqq \sigma_i^2/\eta_i^2 \geq 0$. Note that noise variance η_i is strictly decreasing in quality q_i. All agents and the designer share the same prior F_i on the quality q_i, and only agent i knows the exact q_i.

We now define the accuracy functions. After agents report their qualities $\hat{\mathbf{q}}$ and contribute the vector of signals $\mathbf{s}$, the platform updates its posterior over θ. However, since the true qualities $\mathbf{q}$ are unobserved, the platform treats the reported qualities $\hat{\mathbf{q}}$ as correct. Thus, the platform's posterior $\pi_i(\mathbf{s}, \hat{\mathbf{q}})$ for θ_i is a function of signals and reported qualities. It then uses the posterior mean as the estimator $\hat{\theta}_i(\mathbf{s}, \hat{\mathbf{q}})$, defined as $\mathbb{E}_{\theta_i \sim \pi_i(\mathbf{s}, \hat{\mathbf{q}})}\left[\theta_i\right]$. We define the resulting accuracy for estimating parameter θ_i as its expected squared error under the true posterior distribution:

$$A_i\left(\mathbf{q}, \hat{\mathbf{q}}\right) = -\mathbb{E}_{\mathbf{s}}\left[\mathbb{E}_{\theta_i \sim \pi(\mathbf{s}, \mathbf{q})}\left(\hat{\theta}_i(\mathbf{s}, \hat{\mathbf{q}}) - \theta_i\right)^2\right]. \tag{1}$$

In particular, under truthful reports, $\hat{\theta}_i(\mathbf{s}, \mathbf{q})$ minimizes the MSE under the true posterior mean.

We further define the value functions as

$$v_i(\mathbf{q}, \hat{\mathbf{q}}) \coloneqq 1 + \frac{A_i(\mathbf{q}, \hat{\mathbf{q}})}{\sigma_i^2}. \tag{2}$$

3 Optimal Data-sharing Mechanism

3.1 Preliminary Definitions

Before providing the characterization of the optimal data-sharing mechanism, we introduce a definition and assumption that state how the notion of virtual values and regularity conditions extend from the classical setting to our setting. In what follows, we use the notation $\mathbf{q}_{-i}$ to refer to the vector of qualities excluding q_i, and interchangeably use $v_i\left(q_i, \mathbf{q}_{-i}\right)$ to represent $v_i\left(\left[q_i, \mathbf{q}_{-i}\right], \left[q_i, \mathbf{q}_{-i}\right]\right)$.

Assumption 2 (Virtual Values and Regularity)**.** *For any $i \in [n]$ and $\mathbf{q}_{-i} \in \prod_{j \neq i} \mathcal{Q}_j$, the virtual value function $\phi_{i,\mathbf{q}_{-i}}(\cdot)$, defined below, satisfies the single-crossing property at zero.*

$$\phi_{i,\mathbf{q}_{-i}}(q_i) \coloneqq v_i(q_i, \mathbf{q}_{-i}) - \frac{\partial v_i(q_i, \mathbf{q}_{-i})}{\partial q_i} \frac{1 - F_i(q_i)}{f_i(q_i)}. \tag{3}$$

In the appendix, we state simpler conditions that guarantee the above assumption holds. As we establish in the subsequent sections, Assumption 2 holds for our canonical example of mean estimation from Gaussian signals under mild conditions.

3.2 Our Characterization

Given the above definitions, we are now ready to present our first main result that fully characterizes the optimal data-sharing mechanism. For any vector of reported qualities $\hat{\mathbf{q}}$, a personalized threshold and pricing mechanism allocates the estimation for task $i \in [n]$ to agent i if her report $\hat{q}_i$ exceeds a threshold at a personalized price — both determined by the report of others $\hat{\mathbf{q}}_{-i}$.

Algorithm 1 Optimal Data-sharing Mechanism

Input: Accuracy and value functions $A_i(\cdot)$ and $v_i(\cdot)$, quality distributions F_i with support $\mathcal{Q}_i$, and reported quality vector $\hat{\mathbf{q}}$
for each agent $i \in [n]$ **do**
  Compute the function $\phi_{i,\hat{\mathbf{q}}_{-i}}(q)$ according to equation (3)
  Compute $\bar{q}_i(\hat{\mathbf{q}}_{-i}) \coloneqq \inf \left\{ q \in \mathcal{Q}_i \mid \phi_{i,\hat{\mathbf{q}}_{-i}}(q) \geq 0 \right\}$
  if $\hat{q}_i \geq \bar{q}_i(\hat{\mathbf{q}}_{-i})$ **then**
    Offer the estimation for task i to agent i at price $v_i\left(\bar{q}_i(\hat{\mathbf{q}}_{-i}), \hat{\mathbf{q}}_{-i}\right)$
  end if
end for

Theorem 1. *Under Assumptions 1 and 2, the personalized threshold and pricing mechanism described in Algorithm 1 is the optimal data-sharing mechanism.*

3.3 Our Characterization in the Case of Mean Estimation from Gaussian Signals

In this section, we only restrict quality distributions F_i to have a monotone hazard rate (MHR), i.e., non-decreasing $f_i(\cdot)/(1 - F_i(\cdot))$.[2]

Theorem 2. *For accuracy and value functions $v_i(\cdot)$ defined in equations* (1) *and* (2) *and quality distributions F_i with a monotone hazard rate (MHR) for $i \in [n]$, the personalized threshold and pricing mechanism described in Algorithm 1 is the optimal data-sharing mechanism for the mean estimation task from Gaussian signals.*

Using the characterization in Theorem 2, we explore a key trade-off associated with data centrality: when agent i's data becomes more correlated with that of other agents, does the platform charge agent i more or less? When agent i's data becomes more correlated, her data becomes more valuable for others' estimation tasks, motivating the platform to reduce the price offered to agent i. However, other agents' data already reveal much about agent i's signal, rendering her data. Lastly, despite less essential data, the platform achieves a more accurate estimation for her task, enabling it to charge her more. As we next show, the latter two forces outweigh the first: as agent i's data becomes more correlated with others, her payment increases.

Example 1. Consider a setting with three agents where data qualities are independently drawn from the same distribution, but the correlation structure is asymmetric: agent 1's data is correlated with that of agents 2 and 3 (with correlation factor ρ), while agents 2 and 3's data remain uncorrelated.[3] In Figure 1, we plot the expected payments in the optimal data-sharing mechanism. As we observe, when $|\rho|$ becomes larger, all agents' expected payments increase, with agent 1 paying the most.

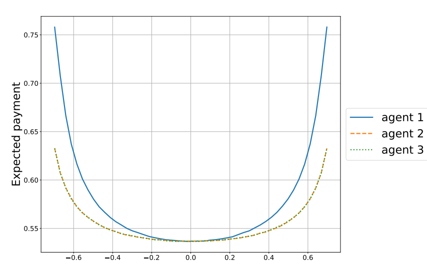

Figure 1: Expected payments of all agents in Example 1

In the appendix, we analytically capture the insights from the above example in the *small correlation regime*, defined as follows. For a sufficiently small $\varepsilon > 0$, we define the *ε-correlation regime* as when the correlation factors uniformly satisfy $|\sigma_{ij}| \leq \varepsilon$ for all $i \neq j$.

[2]This property is satisfied for most common distributions, including but not limited to uniform, exponential, and gamma, and is commonly used in mechanism design literature (see, e.g., [15]).

[3]The qualities are uniformly distributed on [0,10]. We let $\rho \in [\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$, and set $\eta_i = 1$ for $i \in [3]$.

4 What if Agents can Improve their Data Quality?

We extend our model to a setting in which agents can exert costly effort to improve the data qualities.

Effort-dependent Quality Distributions. We begin by presenting the model in general form, though our main focus will be on the canonical case of mean estimation from Gaussian signals. In the example above, firms incur an experimentation cost to improve the distribution of their data quality. For effort level $e_i \geq 0$, let F_{i,e_i} denote the resulting quality distribution incurred at a cost $c_i(e_i)$, where $c_i : \mathbb{R}_+ \to \mathbb{R}_+$ for all $i \in [n]$ with $c_i(0) = 0$ is a strictly increasing function that maps each effort level to its cost. Naturally, greater efforts improve the data quality that we capture by *first-order stochastic dominance*. The sequence of events unfolds as follows: (i) agent i publicly chooses an effort level e_i, incurs the corresponding cost $c_i(e_i)$, and improves her quality distribution to F_{i,e_i}; (ii) the platform then runs the optimal data-sharing mechanism based on the updated quality distributions.

Mean Estimation from Gaussian Signals. A well-known feature of effort-based models is that agents tend to free-ride by exerting less effort, as shown in [1] in the context of federated learning, where agents collect and share data to train a common model. In such settings, free-riding emerges because the utility function is submodular in the effort vector $\mathbf{e}$, implying that the best response of each agent decreases as the efforts of others increase. This directly follows from Topkis' Monotonicity Theorem [20], and we formally establish it in the appendix. A similar pattern arises in our setting if the mechanism does not account for data centrality: utilities remain submodular, leading to free-riding.

However, we establish a key result (under some assumptions): the optimal data-sharing mechanism reverses this tendency by inducing a supermodular game. That is, agents are incentivized to increase their effort in response to others' increased effort.[4] Moreover, we show that the set of equilibria forms a complete lattice and that under higher correlation, the highest-effort equilibrium features greater effort levels from the agents. Our findings are formally stated next.

Theorem 3. *For any $i \in [n]$, suppose $F_{i,e_i} = Unif[e_i, Q]$ for $e_i \in [0, Q]$ for some $Q > 4$, and the cost functions $c_i(\cdot)$ are such that $c_i\left(\sqrt{Q} - 2\right) \geq 1$. Under ε-correlation regime for sufficiently small $\varepsilon > 0$, (i) the mean estimation task with effort is a supermodular game; (ii) the set of all equilibria is non-empty and forms a complete lattice; and (iii) in the equilibrium with the highest collaboration, the optimal efforts are increasing in the absolute value of correlation factors.*

We now clarify the role of the assumptions. The assumption of uniform distributions and ε-correlation regime are made for tractability. The condition $c_i\left(\sqrt{Q} - 2\right) \geq 1$ is imposed to keep effort levels relatively small (in comparison to the upper bound Q on data quality), thereby avoiding a trivial scenario in which the mechanism allocates to all agents irrespective of their reports.

5 Conclusion

In this paper, we study the design of mechanisms that incentivize strategic agents to truthfully contribute data in collaborative learning environments, thereby improving model accuracy across all tasks. In particular, we develop a Bayesian mechanism design framework that optimally determines both allocation and payment rules to elicit truthful reports of data quality while maximizing the platform's revenue. Our key result characterizes the optimal mechanism as a personalized threshold and pricing mechanism, where each agent is allocated the learned estimator if their reported quality exceeds a threshold and is charged a price based on the relevance of others' data in the learning task. We provide a closed-form solution for a canonical Gaussian mean estimation problem and demonstrate how interdependencies in data affect both allocation and pricing. We also extend the model to allow agents to exert costly effort to improve their data quality, establishing that the typical "free-riding" effect in collaborative learning is mitigated by the structure of the proposed mechanism.

Our analysis relies on a few simplifying assumptions that demand further investigation in future works. One promising direction is to explore richer forms of data heterogeneity, including multidimensional qualities. Another avenue is to incorporate privacy constraints and study how the optimal mechanism changes when agents demand data protection. Finally, it would also be valuable to analyze dynamic environments where data quality evolves over time, or agents interact repeatedly.

[4]We detail the formal definition of supermodular games and their properties in the appendix. For further details, see [21].

References

[1] Sai Praneeth Karimireddy, Wenshuo Guo, and Michael I Jordan. Mechanisms that incentivize data sharing in federated learning. *arXiv preprint arXiv:2207.04557*, 2022.

[2] Mengmeng Chen, Xiaohu Wu, Xiaoli Tang, Tiantian He, Yew Soon Ong, Qiqi Liu, Qicheng Lao, and Han Yu. Free-rider and conflict aware collaboration formation for cross-silo federated learning. *Advances in Neural Information Processing Systems*, 37:54974–55004, 2024.

[3] Avrim Blum, Nika Haghtalab, Richard Lanas Phillips, and Han Shao. One for one, or all for all: Equilibria and optimality of collaboration in federated learning. In *International Conference on Machine Learning*, pages 1005–1014. PMLR, 2021.

[4] Xuezhen Tu, Kun Zhu, Nguyen Cong Luong, Dusit Niyato, Yang Zhang, and Juan Li. Incentive mechanisms for federated learning: From economic and game theoretic perspective. *IEEE transactions on cognitive communications and networking*, 8(3):1566–1593, 2022.

[5] Yuan Liu, Mengmeng Tian, Yuxin Chen, Zehui Xiong, Cyril Leung, and Chunyan Miao. A contract theory based incentive mechanism for federated learning. In *Federated and Transfer Learning*, pages 117–137. Springer, 2022.

[6] Chao Huang, Shuqi Ke, Charles Kamhoua, Prasant Mohapatra, and Xin Liu. Incentivizing data contribution in cross-silo federated learning. *arXiv preprint arXiv:2203.03885*, 2022.

[7] Zhepei Wei, Chuanhao Li, Haifeng Xu, and Hongning Wang. Incentivized communication for federated bandits. *Advances in Neural Information Processing Systems*, 36:54399–54420, 2023.

[8] Aniket Murhekar, Zhuowen Yuan, Bhaskar Ray Chaudhury, Bo Li, and Ruta Mehta. Incentives in federated learning: Equilibria, dynamics, and mechanisms for welfare maximization. *Advances in Neural Information Processing Systems*, 36:17811–17831, 2023.

[9] Yue Cui, Liuyi Yao, Yaliang Li, Ziqian Chen, Bolin Ding, and Xiaofang Zhou. An auction-based marketplace for model trading in federated learning. *arXiv preprint arXiv:2402.01802*, 2024.

[10] Enpei Zhang, Jingyi Chai, Rui Ye, Yanfeng Wang, and Siheng Chen. Incentivizing inclusive data contributions in personalized federated learning. In *ICLR 2024 Workshop: How Far Are We From AGI*, 2024.

[11] Lihui Yi, Xiaochun Niu, and Ermin Wei. Incentive analysis for agent participation in federated learning. *arXiv preprint arXiv:2503.09039*, 2025.

[12] Florian E Dorner, Nikola Konstantinov, Georgi Pashaliev, and Martin Vechev. Incentivizing honesty among competitors in collaborative learning and optimization. *Advances in Neural Information Processing Systems*, 36:7659–7696, 2023.

[13] Yiding Chen, Jerry Zhu, and Kirthevasan Kandasamy. Mechanism design for collaborative normal mean estimation. *Advances in Neural Information Processing Systems*, 36:49365–49402, 2023.

[14] Alex Clinton, Yiding Chen, Xiaojin Zhu, and Kirthevasan Kandasamy. Data sharing for mean estimation among heterogeneous strategic agents. *arXiv preprint arXiv:2407.15881*, 2024.

[15] Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

[16] Changrong Deng and Sasa Pekec. Money for nothing: exploiting negative externalities. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 361–370, 2011.

[17] Alexandre Belloni, Changrong Deng, and Saša Pekeč. Mechanism and network design with private negative externalities. *Operations Research*, 65(3):577–594, 2017.

[18] Anish Agarwal, Munther Dahleh, Thibaut Horel, and Maryann Rui. Towards data auctions with externalities. *Games and Economic Behavior*, 148:323–356, 2024.

[19] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[20] Donald M Topkis. Minimizing a submodular function on a lattice. *Operations research*, 26(2):305–321, 1978.

[21] Paul Milgrom and John Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica: Journal of the Econometric Society*, pages 1255–1277, 1990.

[22] Hannaneh Akrami, Bhaskar Ray Chaudhury, Jugal Garg, and Aniket Murhekar. On the theoretical foundations of data exchange economies. *arXiv preprint arXiv:2412.01968*, 2024.

[23] Jierui Lin, Min Du, and Jian Liu. Free-riders in federated learning: Attacks and defenses. *arXiv preprint arXiv:1911.12560*, 2019.

[24] Jiawen Kang, Zehui Xiong, Dusit Niyato, Shengli Xie, and Junshan Zhang. Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory. *IEEE Internet of Things Journal*, 6(6):10700–10714, 2019.

[25] Mengjing Chen, Yang Liu, Weiran Shen, Yiheng Shen, Pingzhong Tang, and Qiang Yang. A mechanism design approach for multi-party machine learning. In *International Workshop on Frontiers in Algorithmics*, pages 248–268. Springer, 2022.

[26] Marco Bornstein, Amrit Singh Bedi, Abdirisak Mohamed, and Furong Huang. Fact or fiction: Can truthful mechanisms eliminate federated free riding? *arXiv preprint arXiv:2405.13879*, 2024.

[27] Chenrui Wu, Yifei Zhu, Rongyu Zhang, Yun Chen, Fangxin Wang, and Shuguang Cui. Fedab: Truthful federated learning with auction-based combinatorial multi-armed bandit. *IEEE Internet of Things Journal*, 10(17):15159–15170, 2023.

[28] Han Yu, Zelei Liu, Yang Liu, Tianjian Chen, Mingshu Cong, Xi Weng, Dusit Niyato, and Qiang Yang. A fairness-aware incentive scheme for federated learning. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 393–399, 2020.

[29] Tra Huong Thi Le, Nguyen H Tran, Yan Kyaw Tun, Zhu Han, and Choong Seon Hong. Auction based incentive design for efficient federated learning in cellular wireless networks. In *2020 IEEE wireless communications and networking conference (WCNC)*, pages 1–6. IEEE, 2020.

[30] Ningning Ding, Zhixuan Fang, and Jianwei Huang. Optimal contract design for efficient federated learning with multi-dimensional private information. *IEEE Journal on Selected Areas in Communications*, 39(1):186–200, 2020.

[31] Adam Richardson, Aris Filos-Ratsikas, and Boi Faltings. Budget-bounded incentives for federated learning. In *Federated Learning: Privacy and Incentive*, pages 176–188. Springer, 2020.

[32] Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285–309, 1955.

[33] Jonathan Levin. Supermodular games. *Lectures Notes, Department of Economics, Stanford University*, 2003.

Supplementary Materials

The supplementary materials consist of three main parts:

- Appendix A: Additional results and details
- Appendix B: Proofs of the main results
- Appendix C: Auxiliary technical lemmas

A Additional Results and Details

A.1 Extended Related Work

Incentivizing Participation and Data Collaboration in Federated Learning. The issue of incentivizing participation in federated learning has been studied in a range of works [3–11]. See [4] for a survey. In the context of data exchange, [22] proposes fair outcomes based on data qualities. Our setting departs from this line of work in two key ways. First, data quality is private information that can be strategically fabricated (e.g., by sampling from fitted distributions). Second, we account for heterogeneity in the relevance of agents' data to the learning task, which directly impacts incentives and model accuracy. Our work is closely related to studies on free-riding behavior in data collaboration [1, 2], where agents exert less effort and benefit from others' data. In particular, [1] shows that when agents can exert costly effort to increase the number of data points (or their data quality), there will be free-riding, leading to inefficiently low total effort. Besides considering strategic data quality reporting and data heterogeneity, which are not present in this previous work, we fully identify the set of all possible equilibria. Specifically, we show that all equilibria form a complete lattice ordered by total effort, including both the lowest- and highest-effort outcomes.

Incentivizing Truthful Data Collaboration in Federated Learning. Untruthful data collaboration in our setting broadly relates to *free-rider attacks* in federated learning [23, 24], where low-quality local updates degrade the global model performance. [12] studies incentives for truthful reporting in both gradient updates and mean estimation settings. In the context of mean estimation, [13, 14] leverage non-monetary incentives to prevent data fabrication. We differ from this line of work in two main aspects. First, we fully characterize the optimal mechanism for general collaborative learning tasks, including both gradient updates and mean estimation. Second, the platform's objective in our setting is to jointly optimize model accuracy and revenue, requiring a mechanism design approach that integrates both monetary and non-monetary incentives. Similar to our work, [25] considers both monetary and non-monetary incentives for truthfulness in general learning tasks. However, their framework only allows misreporting to lower qualities, which does not capture the common scenario in free-rider attacks where agents overstate quality to avoid detection.

Mechanism Design. Finally, our paper is related to the literature on mechanism design pioneered by [15]. However, in our setting, the value that each agent obtains depends not only on her type but also on the types of all other agents, which makes both the analysis and results different from the classical setting. This interdependence connects our work to the literature on mechanism design with externalities [16–18], which typically models additive effects from others' allocations. In contrast, our setting features non-linear dependencies on others' reports, reflecting the role of truthfulness in determining model accuracy.

A.2 Clarifying Modeling Assumptions

We develop a general framework for collaborative learning in which data quality shapes agents' strategic incentives:

1. Collaboration takes place on a platform that allocates access to the trained model and sets payments based on reported qualities.
2. Model accuracy is a function of both true and reported qualities, as reported quality can be used to select better estimators.

Monetary Transfers. While payments are not the predominant practice in collaborative learning settings, recent studies in the literature suggest they can improve participation. In particular, [26] leverages monetary penalties to discourage free-riding in federated learning. In another line of work, [27–31] study mechanisms that incentivize participation through spending from a monetary budget. In contrast, our work investigates how monetary and non-monetary incentives can be used to jointly maximize accuracy and revenue in general learning tasks.

Accuracy Functions. In our model, agents report qualities based on *expected accuracies*, where the expectation is over data distributions parameterized by quality. This applies when agents report before observing realized data (e.g., reporting the precision in data collection such as staffing levels) or when accuracy cannot be computed from realized data—common in learning theory, where guarantees (e.g., generalization bounds) are stated in expectation over IID datasets.

Data Fabrication. We focus on a collaborative setting in which agents do not have any strategic incentive to harm others—a behavior that arises in competitive environments. Accordingly, each agent aims to maximize her own expected utility, which depends on her expected allocation, payment, and resulting model accuracy. Since the platform uses reported qualities to determine allocations and payments—while only the accuracy functions depend on both the true and reported qualities—agents might have incentives to misreport their data quality but have no incentive to fabricate the underlying data itself.

A.3 Mean Estimation from Gaussian Signals

To be more concrete, let us describe two applications of the mean estimation setting in cross-silo collaborative learning:

Example A.1 (Correlated Demand Elasticities Across Populations). Consider n firms, each operating in a different regional market and interested in estimating the demand elasticity θ_i for its own population. Each firm observes a local dataset (e.g., historical prices and quantities sold) that provides a noisy signal of its own elasticity, given by $s_i = \theta_i + z_i, \quad z_i \sim \mathcal{N}(0, \eta_i^2)$. This noise may result from sampling variability or modeling errors.[5] While each θ_i pertains to a different market, demand elasticities are often correlated across similar populations due to shared demographic or macroeconomic factors. Here, $\boldsymbol{\theta} \sim \mathcal{N}(0, \Sigma)$, where Σ encodes these correlations. Consequently, the signal s_i observed by firm i is informative about θ_i of firm i but also θ_j of the other firms, making collaborative learning beneficial.

Example A.2 (Estimating Match Quality from User Features in a Shared Market). Suppose there is a single population of users, each described by a latent feature vector $u \in \mathbb{R}^d$, distributed as $u \sim \mathcal{N}(0, \Sigma_0)$. Each of n firms offers a product or service, described by a known feature vector $x_i \in \mathbb{R}^d$, and seeks to estimate the quality of match with the user population. The match quality is quantified by the inner product $x_i^\top u$, and each firm observes a noisy signal of this quantity, given by $s_i = x_i^\top u + \epsilon_i$ for $\epsilon_i \sim \mathcal{N}(0, 1)$.[6] Since all signals depend on the shared latent vector u, the match qualities are jointly Gaussian and correlated as $\text{Cov}(x_i^\top u, x_j^\top u) = x_i^\top \Sigma_0 x_j$. Thus, s_j contains useful information for firm i's estimation of $x_i^\top u$. Sharing such signals across firms leads to improved learning and better alignment between user preferences and product offerings.

We now show that accuracy functions defined in equation (1) satisfy the general properties outlined in Assumption 1. We further derive the closed form of the accuracy and value functions under truthful reports.

Lemma A.1. *The accuracy and value functions defined in equations* (1) *and* (2) *satisfy the following:*

(i) The accuracy functions $A_i(\cdot)$ satisfy Assumption 1 for all $i \in [n]$

(ii) For any $i \in [n]$ and $\mathbf{q}$, the value function is given by

$$v_i(\mathbf{q}) = \frac{(\Sigma_{i1}, \dots, \Sigma_{in}) \left(\Sigma + \textit{diag}\left(\sigma_1^2/q_1, \cdots, \sigma_n^2/q_n\right)\right)^{-1} (\Sigma_{i1}, \dots, \Sigma_{in})^\top}{\sigma_i^2}. \quad \text{(A.1)}$$

[5] The assumption of Gaussian noise is justified by the Central Limit Theorem, as the signal typically aggregates many small random effects (e.g., consumer behaviors).

[6] The Gaussian noise reflects randomness in user feedback, measurement error, or behavioral variability.

A.4 Revelation Principle

Here, we provide conditions under which, without loss of generality, we can focus on direct mechanisms where agents truthfully report their private information. Unlike the classical mechanism design, the revelation principle does not necessarily hold, as agents' values may depend on both true and reported qualities. In particular, the reduction to direct mechanisms where agents truthfully report their data qualities depends on the designer's objective, which is defined over the tuple of expected accuracies and revenues. From the designer's perspective, the expected accuracy for task i does not directly depend on agent i's utility, which randomly receives the estimation. For simplicity, we assume the designer's objective depends only on the pair of aggregated expected accuracy and revenue.[7] Let $\preceq$ denote the designer's preference relation over pairs of aggregated expected accuracy and revenue.

We make use of the following assumption that says: the platform computes the *optimal* estimator given the reported qualities and the true qualities, but it is the case that the accuracy is the highest if the reported qualities that the platform uses are the true qualities.

Assumption A.1. *For any $\hat{\mathbf{q}}$ and $i \in [n]$, $A_i(\mathbf{q}, \hat{\mathbf{q}}) \leq A_i(\mathbf{q}, \mathbf{q})$. Additionally, if for some $\hat{\mathbf{q}}$, $A_i(\mathbf{q}, \hat{\mathbf{q}}) = A_i(\mathbf{q}, \mathbf{q})$ for all i, then either $\hat{\mathbf{q}} = \mathbf{q}$ or $A_i(\mathbf{q}, \hat{\mathbf{q}})$ is constant in $\hat{\mathbf{q}}$ for all i.*[8]

Proposition A.1. *Suppose Assumption A.1 holds, and the designer's preference relation $\preceq$ over the pairs of aggregated expected accuracy and revenue is lexicographic, with accuracy prioritized over revenue. Then, there exists an optimal truthful direct mechanism in which all agents choose to participate.*

Proposition A.1 guarantees that the reduction to truthful direct mechanisms is without loss of generality if the designer prioritizes accuracy over revenue. Furthermore, since any truthful direct mechanism maximizes the expected accuracy, we focus on characterizing the truthful direct mechanism that maximizes expected revenue, i.e., the optimal mechanism under the lexicographic order of preferences. This way, the platform's problem becomes finding the optimal data-sharing mechanism, which is the solution to the following functional optimization:

$$\max_{\{x_i(\cdot), p_i(\cdot)\}_{i=1}^n} \mathbb{E}_{\mathbf{q}} \left[\sum_{i=1}^n p_i(\mathbf{q}) \right] \tag{A.2}$$

$$\mathbb{E}_{\mathbf{q}_{-i}} \left[v_i(\mathbf{q}) x_i(\mathbf{q}) - p_i(\mathbf{q}) \right] \geq \mathbb{E}_{\mathbf{q}_{-i}} \left[v_i(\mathbf{q}, [q_i', \mathbf{q}_{-i}]) x_i(q_i', \mathbf{q}_{-i}) - p_i(q_i', \mathbf{q}_{-i}) \right], \quad \forall i, q_i' \neq q_i, \tag{A.3}$$

$$\mathbb{E}_{\mathbf{q}_{-i}} \left[v_i(\mathbf{q}) x_i(\mathbf{q}) - p_i(\mathbf{q}) \right] \geq 0, \quad \forall i, q_i, \tag{A.4}$$

where the objective is the sum of the payments received from agents, the constraints (A.3) are incentive compatibility constraints, ensuring that the agents report their qualities truthfully: the left-hand side is the expected utility agent i receives from truthful reporting, and the right-hand side is her expected utility from reporting quality q_i' while her true quality is q_i. The constraints (A.4) are individual rationality, ensuring that the agents receive a non-negative utility and, therefore, voluntarily participate and join the platform.

A.4.1 Mean Estimation from Gaussian Signals

To apply the revelation principle in our canonical example of mean estimation, it suffices to show that the accuracy functions $A_i(\mathbf{q}, \hat{\mathbf{q}})$ defined in equation (1) satisfy Assumption A.1.

Lemma A.2. *The accuracy functions $A_i(\cdot)$ defined in equation (1) satisfy Assumption A.1.*

A.5 Sufficient Conditions for Satisfying Assumption 2

Remark A.1. *Assumption 2 holds when the virtual value function $\phi_{i,\mathbf{q}_{-i}}(\cdot)$ is non-decreasing. Under Assumption 1, the virtual value function $\phi_{i,\mathbf{q}_{-i}}(\cdot)$ is non-decreasing if agent i's value $v_i(q_i, \mathbf{q}_{-i})$ is concave in q_i, for all $\mathbf{q}_{-i} \in \prod_{j \neq i} \mathcal{Q}_j$, and the quality distribution F_i has a monotone hazard rate (MHR), i.e., $f_i(\cdot)/(1 - F_i(\cdot))$ is non-decreasing.*

[7] Our results extend to any strictly increasing function of the accuracies and revenues.

[8] Notice that Assumption A.1 is a stronger variant of Assumption 1(ii).

A.6 Centrality Measures

In this subsection, we analytically derive the insights on data centrality effects in forming the optimal data-sharing mechanism in Example 1.

Proposition A.2. *Suppose the quality distributions have a monotone hazard rate for all $i \in [n]$. The following statements hold in the ε-correlation regime, for a sufficiently small $\varepsilon > 0$:*

(i) The mechanism in Algorithm 1 for $\hat{v}_i(\cdot)$ achieves the optimal expected revenue in the mean estimation task up to $\mathcal{O}\left(\varepsilon^3\right)$ error, where

$$\hat{v}_i(\mathbf{q}) = \frac{q_i}{1+q_i} + \left(\frac{1}{1+q_i}\right)^2 \sum_{j \neq i} \sigma_{ij}^2 \frac{q_j}{1+q_j}. \tag{A.5}$$

(ii) Suppose a sufficiently large $Q > 0$ exists such that F_i is the uniform distribution on $[0, Q]$ for all $i \in [n]$. There exist constants $\lambda_0, \lambda_1 > 0$ independent of n and Σ such that agent i's expected payment in the optimal mean estimation mechanism satisfies: $\mathbb{E}_{\mathbf{q}}[p_i(\mathbf{q})] = \lambda_0 + \lambda_1 \sum_{j \neq i} \sigma_{ij}^2 + \mathcal{O}\left(\varepsilon^3\right)$.

Proposition A.2(i) not only provides a tractable and approximately optimal mechanism that we leverage in the subsequent section, but also explains why increasing the absolute value of the correlation factors increases the payment. Moreover, Proposition A.2(ii) provides an intuitive *centrality measure* of agent i's data, captured by $\sum_{j \neq i} \sigma_{ij}^2$. We note that while more central agents pay more in the platform, we prove below that agents' utilities, in expectation, are increasing in their centrality measure due to receiving more accurate estimators.

We establish two additional properties for the expected payments and model accuracies under the approximate mechanism described in Proposition A.2(i).

Corollary A.1. *Suppose a sufficiently large $Q > 0$ exists such that F_i is the uniform distribution on $[0, Q]$ for all $i \in [n]$. Under the ε-correlation regime, for a sufficiently small $\varepsilon > 0$, for all i and quality vector of other agents $\mathbf{q}_{-i}$, there exist constants $\lambda_0', \lambda_0'', \lambda_1', \lambda_1'' > 0$ independent of n and Σ such that agent i's expected payment and utility, conditional on $\mathbf{q}_{-i}$, in the optimal mean estimation mechanism satisfy:*

$$\mathbb{E}_{q_i}[p_i(\mathbf{q})] = \lambda_0' + \lambda_1' \sum_{j \neq i} \sigma_{ij}^2 \frac{q_j}{1+q_j} + \mathcal{O}\left(\varepsilon^3\right), \tag{A.6}$$

$$\mathbb{E}_{q_i}[U_i(\mathbf{q})] = \lambda_0'' + \lambda_1'' \sum_{j \neq i} \sigma_{ij}^2 \frac{q_j}{1+q_j} + \mathcal{O}\left(\varepsilon^3\right). \tag{A.7}$$

Proposition A.2(i) and Corollary A.1 establish the following insight highlighted in the body of the paper:

As the correlation between the parameter of agent i and those of others increases—or as the quality of others' data improves—the platform achieves higher model accuracy and charges agent i more.

Specifically, the right-hand sides of equations (A.5) and (A.6), neglecting the error terms, are both increasing in the correlation factors and in the quality of others' data.[9] Additionally, taking expectation over $\mathbf{q}_{-i}$ in equation (A.7) shows that the utilities are increasing in centrality measure $\sum_{j \neq i} \sigma_{ij}^2$, confirming another insight discussed in the paper:

While more central agents pay more in the platform, agents' utilities, in expectation, are increasing in their centrality measure due to receiving more accurate estimators.

[9] Notice that each agent's value is strictly increasing in model accuracy.

B Proofs of Main Results

PROOF OF LEMMA A.1:

For $\mathbf{q} \in \mathbb{R}^n_{>0}$, define the matrix $\Sigma_{\mathbf{q}} \coloneqq \Sigma + \text{diag}\left(\sigma_1^2/q_1, \cdots, \sigma_n^2/q_n\right)$. We prove each part separately.

(i) We first establish that $A_i(\cdot)$ satisfies Assumption 1(i). Assume $i = 1$ without loss of generality. Since $\boldsymbol{\theta}$ is normally distributed with mean zero, the vector $[\theta_1, s_1, \ldots, s_n]$ is also normally distributed with mean zero and covariance matrix given by

$$\begin{pmatrix} \sigma_1^2 & \sigma_1^2 & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \sigma_1^2 & \frac{\sigma_1^2(q_1+1)}{q_1} & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{21} & \frac{\sigma_2^2(q_2+1)}{q_2} & \cdots & \Sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Sigma_{n1} & \Sigma_{n1} & \Sigma_{n2} & \cdots & \frac{\sigma_n^2(q_n+1)}{q_n} \end{pmatrix}.$$

Thus, $\theta_1 \mid s_1, \ldots, s_n$ is normally distributed with mean μ and variance σ^2, where

$$\mu = \hat{\theta}_1(\mathbf{q}), \quad \sigma^2 = \sigma_1^2 - \left(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n}\right) \Sigma_{\mathbf{q}}^{-1} \left(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n}\right)^\top.$$

Hence, the optimal estimator for $\theta_1 \mid s_1, \ldots, s_n$ is $\hat{\theta}_1(\mathbf{q})$, which achieves the least mean squared error equal to the variance σ^2. By equation (1), this leads to the following expression for the accuracy function:

$$A_1(\mathbf{q}) = \left(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n}\right) \Sigma_{\mathbf{q}}^{-1} \left(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n}\right)^\top - \sigma_1^2. \tag{B.1}$$

We next establish that $A_1(\mathbf{q})$ is non-decreasing in q_1 by showing that $\frac{\partial A_1}{\partial q_1}(\mathbf{q}) \geq 0$. Since $\Sigma_{\mathbf{q}}$ is symmetric, its inverse is also symmetric. Let $(c_{1i}, \ldots, c_{ni})^\top$ denote the i-th column of the inverse matrix $\Sigma_{\mathbf{q}}^{-1}$. We use the following identity from calculus:

$$\frac{\partial \Sigma_{\mathbf{q}}^{-1}}{\partial q_1} = -\Sigma_{\mathbf{q}}^{-1} \frac{\partial \Sigma_{\mathbf{q}}}{\partial q_1} \Sigma_{\mathbf{q}}^{-1}, \tag{B.2}$$

which is derived by taking the derivative of the equation $\Sigma_{\mathbf{q}} \Sigma_{\mathbf{q}}^{-1} = I$ with respect to q_1.

Additionally, since $\frac{\partial \Sigma_{\mathbf{q}}}{\partial q_1} = \text{diag}\left(-\sigma_1^2/q_1^2, 0, \cdots, 0\right)$, equation (B.2) implies that

$$\frac{\partial \Sigma_{\mathbf{q}}^{-1}}{\partial q_1} = \frac{\sigma_1^2}{q_1^2} (c_{11}, c_{12}, \ldots, c_{1n})^\top (c_{11}, c_{12}, \ldots, c_{1n}). \tag{B.3}$$

Finally, by equation (B.1), we obtain

$$\begin{aligned} \frac{\partial A_1}{\partial q_1}(\mathbf{q}) &= -(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n}) \frac{\partial \Sigma_{\mathbf{q}}^{-1}}{\partial q_1} (\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n})^\top. &\text{(B.4)} \\ &= \frac{\sigma_1^2}{q_1^2} (\Sigma_{11}, \Sigma_{12}, \ldots, \Sigma_{1n})(c_{11}, c_{12}, \ldots, c_{1n})^\top (c_{11}, c_{12}, \ldots, c_{1n})(\Sigma_{11}, \Sigma_{12}, \ldots, \Sigma_{1n})^\top \\ &= \frac{\sigma_1^2}{q_1^2} \left((\Sigma_{11}, \Sigma_{12}, \ldots, \Sigma_{1n})(c_{11}, c_{12}, \ldots, c_{1n})^\top\right)^2 \\ &= \frac{\sigma_1^2}{q_1^2} \left(-\frac{\Sigma_{11}}{q_1} c_{11} + \left(\Sigma_{11}\frac{1+q_1}{q_1}, \Sigma_{12}, \ldots, \Sigma_{1n}\right)(c_{11}, c_{12}, \ldots, c_{1n})^\top\right)^2 \\ &= \frac{\sigma_1^2}{q_1^2} \left(-\frac{\sigma_1^2}{q_1} c_{11} + 1\right)^2 &\text{(B.5)} \\ &\geq 0, \end{aligned}$$

where the second equality follows from equation (B.3), and the last equality follows from the facts that $\Sigma_{11} = \sigma_1^2$ and the inner product of the first row of $\Sigma_{\mathbf{q}}$ and the first column of $\Sigma_{\mathbf{q}}^1$ is equal to one.

We next prove that $A_i(\cdot)$ satisfies Assumption 1(ii). Assume $i = 1$ without loss of generality. As demonstrated earlier, the random variable $\theta_1 \mid s_1, \ldots, s_n$ follows a normal distribution with mean $\hat{\theta}_1(\mathbf{q})$. Thus, for any $\hat{\mathbf{q}} \in \prod_j \mathcal{Q}_j$, the estimator $\hat{\theta}_1(\hat{\mathbf{q}})$ achieves an accuracy at most as large as that of $\hat{\theta}_1(\mathbf{q})$. Taking the expectation over $s_1, \ldots, s_n$, we obtain $A_1(\mathbf{q}, \hat{\mathbf{q}}) \leq A_1(\mathbf{q}, \mathbf{q})$ as desired.

Finally, Assumption 1(iii) immediately follows by the definition of $A_i(\cdot)$.

(ii) The result in equation (A.1) directly follows from combining equations (2) and (B.1). □

PROOF OF PROPOSITION A.1:

For any mechanism $\mathcal{M}$, let $\beta = [\beta_1, \ldots, \beta_n]$ denote the strategy profile of agents $i = 1, \ldots, n$ in equilibrium. Let $\tilde{x}_i(\mathbf{q}) \coloneqq x_i(\beta(\mathbf{q}))$ and $\tilde{p}_i(\mathbf{q}) \coloneqq p_i(\beta(\mathbf{q}))$ denote agent i's expected allocation and payment. If $\beta(\cdot)$ is the identity function, then the outcome of mechanism $\mathcal{M}$ is equivalent to the outcome of mechanism $\tilde{\mathcal{M}}$ with allocations and payments $(\tilde{x}_i(\mathbf{q}), \tilde{p}_i(\mathbf{q}))$, where each agent i truthfully reports her quality q_i. Similarly, if the accuracy functions $A_i(\mathbf{q}, \hat{\mathbf{q}})$ are constant with respect to the report vector $\hat{\mathbf{q}}$, we can assume $\beta(\cdot)$ is the identity function without loss of generality.

Suppose the accuracy functions $A_i(\mathbf{q}, \hat{\mathbf{q}})$ are not constant with respect to the report vector $\hat{\mathbf{q}}$. Assume $\beta(\cdot)$ is not the identity function, i.e., there exists a subset $\mathcal{Q} \subseteq \prod_j \mathcal{Q}_j$ with non-zero measure such that $\beta(\mathbf{q}) \neq \mathbf{q}$ for $\mathbf{q} \in \mathcal{Q}$. In this case, the aggregated expected accuracy is given by $\mathbb{E}\left[\sum_i A_i(\mathbf{q}, \beta(\mathbf{q}))\right]$, which, by Assumption A.1, is strictly less than $\mathbb{E}\left[\sum_i A_i(\mathbf{q}, \mathbf{q})\right]$. Therefore, the designer strictly prefers the outcome of any direct mechanism over the outcome of mechanism $\mathcal{M}$ with strategy profile $\beta(\cdot)$. This concludes the proof since the space of truthful direct mechanisms is non-empty, as the designer can fully allocate to all agents for free. □

PROOF OF LEMMA A.2:

The first part of Assumption A.1 holds because it is the same as Assumption 1(ii), and $A_i(\cdot)$ satisfies Assumption 1 by Lemma A.1. To establish the second part, assume $i = 1$ without loss of generality. Let $\hat{\mathbf{q}}$ be such that

$$A_i(\mathbf{q}, \hat{\mathbf{q}}) = A_i(\mathbf{q}, \mathbf{q}), \text{ for all } i \in [n]. \tag{B.6}$$

We now show that $\hat{\mathbf{q}} = \mathbf{q}$, as desired. Given signals $s_1, \ldots, s_n$, we use the arguments in the proof of Lemma A.1(i) to show that the random variable $\theta_1 \mid s_1, \ldots, s_n$ follows a normal distribution with mean $\hat{\theta}_1(\mathbf{q})$. Thus, the estimator $\hat{\theta}_i(\hat{\mathbf{q}})$ achieves the least mean squared error for estimating $\theta_i \mid s_1, \ldots, s_n$ for almost every $s_1, \ldots, s_n$. Since for any random variable X, the best scalar that estimates X is unique and equal to $\mathbb{E}[X]$, it follows that $\hat{\theta}_i(\hat{\mathbf{q}}) = \hat{\theta}_i(\mathbf{q})$. Thus,

$$(\Sigma_{i1}, \ldots, \Sigma_{in}) \left(\Sigma_{\hat{\mathbf{q}}}^{-1} - \Sigma_{\mathbf{q}}^{-1}\right) (s_1, s_2, \ldots, s_n)^\top = 0, \quad \text{for almost every } s_1, \ldots, s_n.$$

This implies that $(\Sigma_{i1}, \ldots, \Sigma_{in}) \left(\Sigma_{\hat{\mathbf{q}}}^{-1} - \Sigma_{\mathbf{q}}^{-1}\right) = 0$. Repeating this argument for all $i \in [n]$ (see equation (B.6)), we obtain

$$\Sigma \left(\Sigma_{\hat{\mathbf{q}}}^{-1} - \Sigma_{\mathbf{q}}^{-1}\right) = 0. \tag{B.7}$$

Define the diagonal matrices

$$D_{\hat{\mathbf{q}}} \coloneqq \operatorname{diag}\left(\sigma_1^2/\hat{q}_1, \ldots, \sigma_n^2/\hat{q}_n\right), \quad D_{\mathbf{q}} \coloneqq \operatorname{diag}\left(\sigma_1^2/q_1, \ldots, \sigma_n^2/q_n\right).$$

We have $\Sigma_{\hat{\mathbf{q}}} = \Sigma + D_{\hat{\mathbf{q}}}$ and $\Sigma_{\mathbf{q}} = \Sigma + D_{\mathbf{q}}$. Substituting into equation (B.7), we obtain

$$\begin{aligned}
0 &= \Sigma \left(\Sigma + D_{\hat{\mathbf{q}}}\right)^{-1} - \Sigma \left(\Sigma + D_{\mathbf{q}}\right)^{-1} \\
&= \left(\Sigma + D_{\hat{\mathbf{q}}} - D_{\hat{\mathbf{q}}}\right) \left(\Sigma + D_{\mathbf{q}}\right)^{-1} - \left(\Sigma + D_{\mathbf{q}} - D_{\mathbf{q}}\right) \left(\Sigma + D_{\mathbf{q}}\right)^{-1} \\
&= \left(I - D_{\hat{\mathbf{q}}} \left(\Sigma + D_{\hat{\mathbf{q}}}\right)^{-1}\right) - \left(I - D_{\mathbf{q}} \left(\Sigma + D_{\mathbf{q}}\right)^{-1}\right) \\
&= -\left(\Sigma D_{\hat{\mathbf{q}}}^{-1} + I\right)^{-1} + \left(\Sigma D_{\mathbf{q}}^{-1} + I\right)^{-1}.
\end{aligned}$$

Hence, $\Sigma D_{\hat{\mathbf{q}}}^{-1} = \Sigma D_{\mathbf{q}}^{-1}$. Comparing the diagonals of both sides implies that $\hat{q}_i = q_i$ for all i. □

PROOF OF THEOREM 1:

Let $\mathcal{M}^*$ denote the personalized threshold and pricing mechanism described in Algorithm 1. We begin by showing that, under mechanism $\mathcal{M}^*$, truthful reporting constitutes an ex-post Nash equilibrium.[10] For any given quality vector $\mathbf{q}$, it suffices to verify ex-post individual rationality (IR) and ex-post incentive compatibility (IC).

Ex-post IR. We show that agent i receives non-negative utility from truthful reporting.

- Suppose $q_i \geq \bar{q}_i(\mathbf{q}_{-i})$. Then, agent i's utility from truthful reporting is
$$v_i(q_i, \mathbf{q}_{-i}) - v_i(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i}) \geq 0, \tag{B.8}$$
where the inequality follows from $q_i \geq \bar{q}_i(\mathbf{q}_{-i})$, using the monotonicity of $v_i(\cdot, \mathbf{q}_{-i})$ implied by Assumption 1(i).
- Suppose $q_i < \bar{q}_i(\mathbf{q}_{-i})$. Then, agent i's utility from truthful reporting is zero.

Ex-post IC. Suppose agent i reports a quality . By Assumption 1(ii), and since $v_i(\cdot) = g_i(A_i(\cdot))$ for some strictly increasing function g_i, we have
$$v_i(\mathbf{q}, [q_i, \mathbf{q}_{-i}]) \geq v_i(\mathbf{q}, [, \mathbf{q}_{-i}]). \tag{B.9}$$
We prove $U_i(\mathbf{q}, \mathbf{q}) \geq U_i(\mathbf{q}, [, \mathbf{q}_{-i}])$ by considering four cases:
(I-1) $q_i \geq \bar{q}_i(\mathbf{q}_{-i})$ and $\geq \bar{q}_i(\mathbf{q}_{-i})$,
(I-2) $q_i \geq \bar{q}_i(\mathbf{q}_{-i})$ and $< \bar{q}_i(\mathbf{q}_{-i})$,
(II-1) $q_i < \bar{q}_i(\mathbf{q}_{-i})$ and $\geq \bar{q}_i(\mathbf{q}_{-i})$,
(II-2) $q_i < \bar{q}_i(\mathbf{q}_{-i})$ and $< \bar{q}_i(\mathbf{q}_{-i})$.

(I-1) Suppose $q_i \geq \bar{q}_i(\mathbf{q}_{-i})$ and $\geq \bar{q}_i(\mathbf{q}_{-i})$. Hence,
$$U_i(\mathbf{q}, \mathbf{q}) = v_i(\mathbf{q}, \mathbf{q}) - v_i(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i}) \geq v_i(\mathbf{q}, [, \mathbf{q}_{-i}]) - v_i(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i}) = U_i(\mathbf{q}, [, \mathbf{q}_{-i}]),$$
where the inequality follows from equation (B.9).

(I-2) Suppose $q_i \geq \bar{q}_i(\mathbf{q}_{-i})$ and $< \bar{q}_i(\mathbf{q}_{-i})$. Hence,
$$U_i(\mathbf{q}, \mathbf{q}) = v_i(\mathbf{q}, \mathbf{q}) - v_i(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i}) \geq 0 = U_i(\mathbf{q}, [, \mathbf{q}_{-i}]),$$
where the inequality follows from equation (B.8).

(II-1) Suppose $q_i < \bar{q}_i(\mathbf{q}_{-i})$ and $\geq \bar{q}_i(\mathbf{q}_{-i})$. Hence,
$$U_i(\mathbf{q}, \mathbf{q}) = 0 \geq v_i(\mathbf{q}, \mathbf{q}) - v_i\left(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i}\right) \geq v_i(\mathbf{q}, [, \mathbf{q}_{-i}]) - v_i\left(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i}\right) = U_i\left(\mathbf{q}, [, \mathbf{q}_{-i}]\right),$$
where the first inequality follows from Assumption 1(i), and the second inequality follows from Assumption 1(ii).

(II-2) Suppose $q_i < \bar{q}_i(\mathbf{q}_{-i})$ and $< \bar{q}_i(\mathbf{q}_{-i})$. Hence,
$$U_i(\mathbf{q}, \mathbf{q}) = 0 = U_i(\mathbf{q}, [, \mathbf{q}_{-i}]).$$

Therefore, truthful reporting constitutes an ex-post Nash equilibrium. We next show that $\mathcal{M}^*$ is the optimal truthful data-sharing mechanism. Let $(x_i^*(\cdot), p_i^*(\cdot))$ denote agent i's expected allocation and payment under mechanism $\mathcal{M}^*$. Thus,
$$x_i^*(q_i, \mathbf{q}_{-i}) = \mathbf{1}\{q_i \geq \bar{q}_i(\mathbf{q}_{-i})\}, \quad p_i^*(q_i, \mathbf{q}_{-i}) = v_i(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i})\mathbf{1}\{q_i \geq \bar{q}_i(\mathbf{q}_{-i})\}. \tag{B.10}$$
We have
$$\begin{aligned}
\mathbb{E}_{\mathbf{q}}[p_i^*(\mathbf{q})] &= \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[p_i^*(q_i, \mathbf{q}_{-i})\right]\right] \\
&\overset{(a)}{=} \mathbb{E}_{\mathbf{q}_{-i}}\left[v_i(\bar{q}_i(\mathbf{q}_{-i}), \mathbf{q}_{-i})\left(1 - F_i(\bar{q}_i(\mathbf{q}_{-i}))\right)\right] \\
&\overset{(b)}{=} \mathbb{E}_{\mathbf{q}_{-i}}\left[\int_{\bar{q}_i(\mathbf{q}_{-i})}\left(v_i(q_i, \mathbf{q}_{-i}) - \frac{\partial v_i(q_i, \mathbf{q}_{-i})}{\partial q_i}\frac{1 - F_i(q_i)}{f_i(q_i)}\right) f_i(q_i)\right] \\
&= \mathbb{E}_{\mathbf{q}_{-i}}\left[\int_{\bar{q}_i(\mathbf{q}_{-i})} \phi_{i,\mathbf{q}_{-i}}(q_i) f_i(q_i)\right] \\
&\overset{(c)}{=} \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\phi^+_{i,\mathbf{q}_{-i}}(q_i)\right]\right], \quad \text{for all } i \in [n],
\end{aligned} \tag{B.11}$$

[10] This implies that the resulting equilibrium is prior-free, in contrast to a Bayesian Nash equilibrium.

where (a) follows from equation (B.10), (b) follows from integration by parts, and (c) follows from the fact that $\bar{q}_i(\mathbf{q}_{-i}) = \inf\{q_i \in \mathcal{Q}_i \mid \phi_{i,\mathbf{q}_{-i}}(q_i) \geq 0\}$ and that $\phi_{i,\mathbf{q}_{-i}}(q_i)$ satisfies the single-crossing property at zero by Assumption 2.[11]

Let $(x_i(\cdot), p_i(\cdot))$ denote a feasible solution of problem (A.2). We show $\mathbb{E}_{\mathbf{q}}[p_i(\mathbf{q})] \leq \mathbb{E}_{\mathbf{q}}[p_i^*(\mathbf{q})]$ for all $i \in [n]$, establishing the optimality of $\mathcal{M}^*$. Suppose the support $\mathcal{Q}_i$ starts at 0 for all $i \in [n]$; all arguments extend to the general case. We first use an auxiliary result with the proof deferred to Appendix B.

Lemma B.1 (Payment Identity)**.** *Suppose the support $\mathcal{Q}_i$ starts at 0. Let $(x_i(\cdot), p_i(\cdot))$ be a feasible solution of problem* (A.2)*. Under Assumption 1, the following identity holds:*

$$\mathbb{E}_{\mathbf{q}}[p_i(\mathbf{q})] = \mathbb{E}_{\mathbf{q}_{-i}}\left[p_i(0, \mathbf{q}_{-i}) - v_i(0, \mathbf{q}_{-i})\, x_i(0, \mathbf{q}_{-i})\right] + \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\phi_{i,\mathbf{q}_{-i}}(q_i) x_i(q_i, \mathbf{q}_{-i})\right]\right].$$

Moreover, by constraint (A.4) in problem (A.2) evaluated at $q_i = 0$, we have

$$\mathbb{E}_{\mathbf{q}_{-i}}\left[p_i(0, \mathbf{q}_{-i}) - v_i(0, \mathbf{q}_{-i})\, x_i(0, \mathbf{q}_{-i})\right] \leq 0.$$

Combining this with the payment identity in Lemma B.1, we have

$$\mathbb{E}_{\mathbf{q}}[p_i(\mathbf{q})] \leq \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\phi_{i,\mathbf{q}_{-i}}(q_i) x_i(q_i, \mathbf{q}_{-i})\right]\right] \leq \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\phi_{i,\mathbf{q}_{-i}}(q_i)^+\right]\right] = \mathbb{E}_{\mathbf{q}}[p_i^*(\mathbf{q})],$$

where the last equality follows from equation (B.11). □

PROOF OF THEOREM 2:

We apply Theorem 1 to establish the optimality of the mechanism in Algorithm 1 for the value functions $v_i(\cdot)$ defined in equation (2). To apply Theorem 1, it suffices to verify that Assumptions 1 and 2 are satisfied. By Lemma A.1(i), Assumption 1 holds. To verify Assumption 2, we use the following auxiliary result, with the proof provided in Appendix B.

Lemma B.2. *For the value function $v_i(\cdot)$ defined in equation (2) and quality distribution F_i with a monotone hazard rate (MHR), the virtual value function $\phi_{i,\mathbf{q}_{-i}}(\cdot)$ defined in equation (3) is non-decreasing for all $i \in [n]$ and $\mathbf{q}_{-i} \in \prod_{j \neq i} \mathcal{Q}_j$.*

Combining Lemma B.2 with Remark A.1, we conclude that Assumption 2 holds. □

PROOF OF PROPOSITION A.2:

(i) For $\hat{v}_i(\cdot)$ defined in equation (A.5), we define the approximate virtual value function as

$$\hat{\phi}_{i,\mathbf{q}_{-i}}(q_i) \coloneqq \hat{v}_i(\mathbf{q}) - \frac{\partial \hat{v}_i(\mathbf{q})}{\partial q_i} \frac{1 - F_i(q_i)}{f_i(q_i)}. \tag{B.12}$$

We start by using an auxiliary result, with the proof provided in Appendix B.

Lemma B.3. *Under ε-correlation regime for an $0 < \varepsilon < 1/\sqrt{3(n-1)}$, for any quality distribution F_i with a monotone hazard rate (MHR), the approximate virtual value function $\hat{\phi}_{i,\mathbf{q}_{-i}}(\cdot)$ defined in equation (B.12) is strictly increasing for all $i \in [n]$ and $\mathbf{q}_{-i} \in \prod_{j \neq i} \mathcal{Q}_j$.*

We first establish the following claim: it suffices to show that $\hat{\phi}_{i,\mathbf{q}_{-i}}(q_i)$ uniformly approximates $\phi_{i,\mathbf{q}_{-i}}(q_i)$ up to an error of $\mathcal{O}(\varepsilon^3)$ for all $i \in [n]$ and $\mathbf{q} \in \prod_j \mathcal{Q}_j$. By Lemmas B.2 and B.3, the virtual value functions $\phi_{i,\mathbf{q}_{-i}}(q_i)$ and $\hat{\phi}_{i,\mathbf{q}_{-i}}(q_i)$ are non-decreasing in q_i. Thus, the optimal mechanism described in part (i) yields an expected revenue of $r^* = \sum_{i=1}^n \mathbb{E}\left[\phi^+_{i,\mathbf{q}_{-i}}(q_i)\right]$ (see equation (B.11) in the proof of Theorem 1). Similarly, Lemma B.1 implies that the approximate mechanism described

[11] Note that $\phi^+_{i,\mathbf{q}_{-i}}$ takes a non-negative value at the upper end of support $\mathcal{Q}_i$.

in part (ii) yields an expected revenue given by

$$\begin{aligned}\hat{r} &= \sum_{i=1}^{n} \mathbb{E}\left[\phi_{i,\mathbf{q}_{-i}}(q_i)\mathbf{1}\left\{q_i \geq \bar{q}_i(\mathbf{q}_{-i})\right\}\right] \\ &= \sum_{i=1}^{n} \mathbb{E}\left[\left(\phi_{i,\mathbf{q}_{-i}}(q_i) - \hat{\phi}_{i,\mathbf{q}_{-i}}(q_i) + \hat{\phi}_{i,\mathbf{q}_{-i}}(q_i)\right)\mathbf{1}\left\{q_i \geq \bar{q}_i(\mathbf{q}_{-i})\right\}\right] \\ &= \sum_{i=1}^{n} \mathbb{E}\left[\left(\phi_{i,\mathbf{q}_{-i}}(q_i) - \hat{\phi}_{i,\mathbf{q}_{-i}}(q_i)\right)\mathbf{1}\left\{q_i \geq \bar{q}_i(\mathbf{q}_{-i})\right\} + \hat{\phi}^{+}_{i,\mathbf{q}_{-i}}(q_i)\right],\end{aligned}$$

where the last equality follows from the monotonicity of $\hat{\phi}_{i,\mathbf{q}_{-i}}(\cdot)$. Thus,

$$\begin{aligned}|r^* - \hat{r}| &\leq \sum_{i=1}^{n} \mathbb{E}\left[\left|\phi_{i,\mathbf{q}_{-i}}(q_i) - \hat{\phi}_{i,\mathbf{q}_{-i}}(q_i)\right| + \left|\phi^{+}_{i,\mathbf{q}_{-i}}(q_i) - \hat{\phi}^{+}_{i,\mathbf{q}_{-i}}(q_i)\right|\right] \\ &\leq 2\sum_{i=1}^{n} \mathbb{E}\left[\left|\phi_{i,\mathbf{q}_{-i}}(q_i) - \hat{\phi}_{i,\mathbf{q}_{-i}}(q_i)\right|\right].\end{aligned}$$

Therefore, to establish the desired result that $|r^* - \hat{r}| = \mathcal{O}(\varepsilon^3)$, it suffices to show

$$\left|\phi_{i,\mathbf{q}_{-i}}(q_i) - \hat{\phi}_{i,\mathbf{q}_{-i}}(q_i)\right| \leq C\varepsilon^3, \tag{B.13}$$

for some constant $C > 0$ independent of $i \in [n]$ and $\mathbf{q} \in \prod_j \mathcal{Q}_j$.

Assume $i = 1$ without loss of generality. By equation (A.1) in Lemma A.1(ii), we have

$$v_1(\mathbf{q}) = \frac{(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n})\Sigma_{\mathbf{q}}^{-1}(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n})^\top}{\sigma_1^2}.$$

We next establish the following:

$$v_1(\mathbf{q}) = \hat{v}_1(\mathbf{q}) + \mathcal{O}\left(\varepsilon^3\right), \tag{B.14}$$

$$\frac{\partial v_1(\mathbf{q})}{\partial q_1} = \frac{\partial \hat{v}_1(\mathbf{q})}{\partial q_1} + \mathcal{O}\left(\varepsilon^3\right). \tag{B.15}$$

Note that to obtain equation (B.15), one cannot directly differentiate equation (B.14) with respect to q_1, as the approximation error may depend on q_1 in a non-smooth way. We then conclude the proof by combining equations (B.14) and (B.15), which yields equation (B.13).

We first approximate the inverse matrix $\Sigma_{\mathbf{q}}$. Given $\mathbf{q} \in \prod \mathcal{Q}_i$, we express the covariance matrix $\Sigma_{\mathbf{q}}$ as $D + \varepsilon M$, where

$$D := \begin{pmatrix} \frac{\sigma_1^2}{\alpha_1} & 0 & \cdots & 0 \\ 0 & \frac{\sigma_2^2}{\alpha_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\sigma_n^2}{\alpha_n} \end{pmatrix}, \quad \alpha_j := \frac{q_j}{1+q_j}, \text{ for all } j \in [n], \tag{B.16}$$

$$M := \frac{1}{\varepsilon}\begin{pmatrix} 0 & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & 0 & \cdots & \Sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \cdots & 0 \end{pmatrix}.$$

Notice that α_j is a function of $\mathbf{q}$. Under the ε-correlation regime, we have $M_{jk} = \mathcal{O}(1)$. Therefore, since $\alpha_j \in [0,1]$ and $D_{jj}^{-1} = \alpha_j/\sigma_i^2 = \mathcal{O}(1)$ for all $j \in [n]$, all entries of M and MD^{-1} are uniformly bounded for all $i \in [n]$ and $\mathbf{q} \in \prod_j \mathcal{Q}_j$. Hence, there exists a constant $L > 0$ independent of $i \in [n]$ and $\mathbf{q} \in \prod_j \mathcal{Q}_j$ such that $\max\left\{\|M\|, \|MD^{-1}\|\right\} \leq L$. Thus, for sufficiently small

$\varepsilon < 1/(2L)$, we have $\|\varepsilon MD^{-1}\| < 1$, which implies that

$$
\begin{aligned}
\Sigma_{\mathbf{q}}^{-1} &= (D + \varepsilon M)^{-1} \\
&= \left(\left(I + \varepsilon MD^{-1}\right) D\right)^{-1} \\
&= D^{-1}\left(I + \varepsilon MD^{-1}\right)^{-1} \\
&= D^{-1}\left(I - \varepsilon MD^{-1} + \varepsilon^2 MD^{-1}MD^{-1} + \left(\sum_{k=3}^{\infty}(-1)^k\varepsilon^k\left(MD^{-1}\right)^{k-1} M\right) D^{-1}\right) \\
&= D^{-1}\left(I - \varepsilon MD^{-1} + \varepsilon^2 MD^{-1}MD^{-1} + RD^{-1}\right), \qquad \text{(B.17)}
\end{aligned}
$$

where all entries of matrix R are $\mathcal{O}\left(\varepsilon^3\right)$ uniformly for $i \in [n]$ and $\mathbf{q} \in \prod_j \mathcal{Q}_j$. This is because

$$
\|R\| \leq \sum_{k=3}^{\infty} \varepsilon^k \left\|MD^{-1}\right\|^{k-1} \|M\| \overset{(a)}{\leq} \sum_{k=3}^{\infty} \varepsilon^k L^k = \frac{\varepsilon^3 L^3}{1 - \varepsilon L} \overset{(b)}{\leq} 2\varepsilon^3 L^3 = \mathcal{O}\left(\varepsilon^3\right), \qquad \text{(B.18)}
$$

where (a) follows from $\max\left\{\|M\|, \|MD^{-1}\|\right\} \leq L$, and (b) follows from $\varepsilon < 1/(2L)$.

Therefore, we have

$$
\begin{aligned}
v_1(\mathbf{q}) &= \frac{(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n})\Sigma_{\mathbf{q}}^{-1}(\sigma_1^2, \Sigma_{12}, \ldots, \Sigma_{1n})^\top}{\sigma_1^2} \\
&\overset{(a)}{=} \frac{\left(\sigma_1^2, \varepsilon M_{12}, \cdots, \varepsilon M_{1n}\right) D^{-1}\left(I - \varepsilon MD^{-1} + \varepsilon^2 (MD^{-1})^2 + \mathcal{O}\left(\varepsilon^3\right)\right)\left(\sigma_1^2, \varepsilon M_{12}, \cdots, \varepsilon M_{1n}\right)^\top}{\sigma_1^2} \\
&= \sigma_1^2 D_{11}^{-1} + \varepsilon^2\left(\sum_{j\neq 1}\left(M_{1j}^2 D_{jj}^{-1} - 2M_{1j}\left(D^{-1}MD^{-1}\right)_{1j}\right) + \sigma_1^2\left(D^{-1}\left(MD^{-1}\right)^2\right)_{11}\right) + \mathcal{O}\left(\varepsilon^3\right) \\
&\overset{(b)}{=} \alpha_1 + \varepsilon^2\left(1 - 2\alpha_1 + \alpha_1^2\right)\sum_{j\neq 1}\frac{M_{1j}^2\alpha_j}{\sigma_1^2\sigma_j^2} + \mathcal{O}\left(\varepsilon^3\right) \\
&= \alpha_1 + (1 - \alpha_1)^2\sum_{j\neq 1}\frac{\left(\varepsilon M_{1j}\right)^2\alpha_j}{\sigma_1^2\sigma_j^2} + \mathcal{O}\left(\varepsilon^3\right) \\
&= \alpha_1 + (1 - \alpha_1)^2\sum_{j\neq 1}\sigma_{1j}^2\alpha_j + \mathcal{O}\left(\varepsilon^3\right) \\
&\overset{(c)}{=} \frac{q_1}{1 + q_1} + \left(1 - \frac{q_1}{1 + q_1}\right)^2\sum_{j\neq 1}\sigma_{1j}^2\frac{q_j}{1 + q_j} + \mathcal{O}\left(\varepsilon^3\right) \\
&\overset{(d)}{=} \hat{v}_1(\mathbf{q}) + \mathcal{O}\left(\varepsilon^3\right),
\end{aligned}
$$

where (a) follows from equations (B.17) and (B.18), (b) is derived using the following algebraic identities:

$$
\begin{aligned}
&D_{11}^{-1} = \frac{\alpha_1}{\sigma_1^2}, \\
&D_{jj}^{-1} = \frac{\alpha_j}{\sigma_j^2}, \quad \left(D^{-1}MD^{-1}\right)_{1j} = \alpha_1\frac{M_{1j}\alpha_j}{\sigma_1^2\sigma_j^2} \text{ for all } j \neq 1, \\
&\left(D^{-1}\left(MD^{-1}\right)^2\right)_{11} = \frac{\alpha_1^2}{\sigma_1^2}\sum_{j\neq 1}\frac{M_{1j}^2\alpha_j}{\sigma_1^2\sigma_j^2},
\end{aligned}
$$

(c) follows from substituting $\alpha_j = q_j/(1 + q_j)$ for all $j \in [n]$, as defined in equation (B.16), and (d) follows from equation (A.5). Hence, we proved the approximation error in equation (B.14).

We now establish equation (B.15), which completes the proof. Let c_{11} denote the first diagonal entry of the inverse matrix $\Sigma_{\mathbf{q}}^{-1}$. Using the approximation in equation (B.17), we have

$$\begin{aligned} c_{11} &= \frac{\alpha_1}{\sigma_1^2}\left(1+\varepsilon^2\sum_{j\neq 1}\frac{M_{1j}^2\alpha_1\alpha_j}{\sigma_1^2\sigma_j^2}\right)+\left(\frac{\alpha_1}{\sigma_1^2}\right)^2 R_{11} \\ &= \frac{\alpha_1}{\sigma_1^2}\left(1+\varepsilon^2\sum_{j\neq 1}\frac{M_{1j}^2\alpha_1\alpha_j}{\sigma_1^2\sigma_j^2}\right)+\alpha_1^2\mathcal{O}\left(\varepsilon^3\right). \end{aligned} \tag{B.19}$$

We have

$$\begin{aligned} \frac{\partial v_1(\mathbf{q})}{\partial q_1} &\overset{(a)}{=} \frac{1}{\sigma_1^2}\frac{\partial A_1(\mathbf{q})}{\partial q_1} \overset{(b)}{=} \frac{1}{q_1^2}\left(-\frac{\sigma_1^2}{q_1}c_{11}+1\right)^2 \\ &\overset{(c)}{=} (1-\alpha_1)^2\left(-\frac{\sigma_1^2(1-\alpha_1)}{\alpha_1^2}c_{11}+\frac{1}{\alpha_1}\right)^2 \\ &\overset{(d)}{=} (1-\alpha_1)^2\left(1-\varepsilon^2(1-\alpha_1)\sum_{j\neq 1}\frac{M_{1j}^2\alpha_j}{\sigma_1^2\sigma_j^2}-\sigma_1^2(1-\alpha_1)\mathcal{O}\left(\varepsilon^3\right)\right)^2 \\ &\overset{(e)}{=} (1-\alpha_1)^2\left(1-\varepsilon^2(1-\alpha_1)\sum_{j\neq 1}\frac{M_{1j}^2\alpha_j}{\sigma_1^2\sigma_j^2}+\mathcal{O}\left(\varepsilon^3\right)\right)^2 \\ &= (1-\alpha_1)^2\left(1-2\varepsilon^2(1-\alpha_1)\sum_{j\neq 1}\frac{M_{1j}^2\alpha_j}{\sigma_1^2\sigma_j^2}+\mathcal{O}\left(\varepsilon^3\right)\right) \\ &\overset{(f)}{=} (1-\alpha_1)^2\left(1-2(1-\alpha_1)\sum_{j\neq 1}\sigma_{1j}^2\alpha_j+\mathcal{O}\left(\varepsilon^3\right)\right) \\ &\overset{(g)}{=} \frac{\partial\alpha_1}{\partial q_1}\frac{\partial\hat{v}_1(\mathbf{q})}{\partial\alpha_1}+\mathcal{O}\left(\varepsilon^3\right) = \frac{\partial\hat{v}_1(\mathbf{q})}{\partial q_1}+\mathcal{O}\left(\varepsilon^3\right), \end{aligned}$$

where (a) follows from equation (2), (b) follows from equation (B.5), as derived in the proof of Lemma A.1(i), (c) follows from substituting $q_1 = \alpha_1/(1-\alpha_1)$, (d) follows from equation (B.19), (e) follows from $1-\alpha_1 \in [0,1]$, (f) follows from $\varepsilon M_{1j} = \Sigma_{1j}$ and $\sigma_{1j} = \Sigma_{1j}/(\sigma_1\sigma_j)$ for all $j \neq 1$, and (g) follows from $\partial\alpha_1/\partial q_1 = 1/(1+q_1)^2 = (1-\alpha_1)^2$ and $\hat{v}(\mathbf{q}) = \alpha_1 + (1-\alpha_1)^2\sum_{j\neq 1}\sigma_{1j}^2\alpha_j$.[12]

(ii) Assume that Q is sufficiently large such that

$$Q > 16. \tag{B.20}$$

As shown in part (i), the expected payment of each agent $i \in [n]$ under the mechanism with approximate value functions $\hat{v}_i(\cdot)$ matches her payment in the optimal mechanism up to an $\mathcal{O}\left(\varepsilon^3\right)$ error. Thus, since the approximation in part (ii) is of the same order, it suffices to approximate the expected payments under the approximate mechanism.

Assume that $\varepsilon > 0$ is sufficiently small such that

$$0 < \varepsilon < \frac{1}{\sqrt{3(n-1)}}. \tag{B.21}$$

For any given $\mathbf{q}_{-i} \in \prod_{j\neq i}\mathcal{Q}_j$, define $S_{\mathbf{q}_{-i}} \coloneqq \sum_{j\neq i}\sigma_{ij}^2 q_j/(1+q_j)$. Under the ε-correlation regime, we have $|\sigma_{ij}| \leq \varepsilon$ for all $j \neq i$. Thus,

$$S_{\mathbf{q}_{-i}} = \sum_{j\neq i}\sigma_{ij}^2\frac{q_j}{1+q_j} \leq \varepsilon^2\sum_{j\neq i}\frac{q_j}{1+q_j} \leq (n-1)\varepsilon^2 < 1/3, \tag{B.22}$$

[12] Note that we treat $\alpha_j = q_j/(1+q_j)$ as a function of q_j.

where the last inequality follows from equation (B.21). Moreover, from equations (A.5) and (B.12), we have

$$\begin{aligned}
\hat{\phi}_{i,\mathbf{q}_{-i}}(q_i) &= \hat{v}_i(q_i,\mathbf{q}_{-i}) - \frac{\partial \hat{v}_i(q_i,\mathbf{q}_{-i})}{\partial q_i}\frac{\bar{F}_{i,e_i}(q_i)}{f_{i,e_i}(q_i)} \\
&= \left(\frac{q_i}{1+q_i} + \left(\frac{1}{1+q_i}\right)^2 S_{\mathbf{q}_{-i}}\right) - \left(\frac{1}{(1+q_i)^2} - \frac{2}{(1+q_i)^3}S_{\mathbf{q}_{-i}}\right)(Q-q_i) \\
&= \left(\frac{q_i}{1+q_i} - \frac{Q-q_i}{(1+q_i)^2}\right) + \left(\frac{1}{(1+q_i)^2} + \frac{2(Q-q_i)}{(1+q_i)^3}\right)S_{\mathbf{q}_{-i}}. \qquad \text{(B.23)}
\end{aligned}$$

Therefore,

$$\hat{\phi}_{i,\mathbf{q}_{-i}}(0) = -Q + (1+2Q)S_{\mathbf{q}_{-i}} < -Q + \frac{(1+2Q)}{3} = \frac{1-Q}{3} < 0, \qquad \text{(B.24)}$$

where the first inequality follows from equation (B.22), and the last inequality follows from equation (B.20). Moreover, by Lemma B.3, the function $\hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(\cdot)$ is strictly increasing. Therefore, equation (B.24) implies that

$$\bar{q}_{\mathbf{q}_{-i}} = \inf\left\{q \in [0,Q] \mid \hat{\phi}_{i,\mathbf{q}_{-i}}(q) \geq 0\right\} > 0.$$

Combining this with the continuity of $\hat{\phi}_{i,\mathbf{q}_{-i}}(q)$ (see equation (B.23)), $\bar{q}_{\mathbf{q}_{-i}}$ is the solution to $\hat{\phi}_{i,\mathbf{q}_{-i}}(\bar{q}_{\mathbf{q}_{-i}}) = 0$. Thus, equation (B.23) implies that

$$\left(\frac{\bar{q}_{\mathbf{q}_{-i}}}{1+\bar{q}_{\mathbf{q}_{-i}}} - \frac{Q-\bar{q}_{\mathbf{q}_{-i}}}{(1+\bar{q}_{\mathbf{q}_{-i}})^2}\right) + \left(\frac{1}{(1+\bar{q}_{\mathbf{q}_{-i}})^2} + \frac{2(Q-\bar{q}_{\mathbf{q}_{-i}})}{(1+\bar{q}_{\mathbf{q}_{-i}})^3}\right)S_{\mathbf{q}_{-i}} = 0. \qquad \text{(B.25)}$$

Multiplying both sides by $\left(1+\bar{q}_{\mathbf{q}_{-i}}\right)^3$, we derive

$$S_{\mathbf{q}_{-i}} = g(\bar{q}_{\mathbf{q}_{-i}}), \qquad \text{(B.26)}$$

where we define

$$g(q) := \frac{-q(1+q)^2 + (Q-q)(1+q)}{(1+q) + 2(Q-q)}. \qquad \text{(B.27)}$$

By equation (B.26) and the fact that $S_{\mathbf{q}_{-i}} \in [0,1/3]$ from equation (B.22), Lemma C.2(i) implies that $\bar{q}_{\mathbf{q}_{-i}} = g^{-1}(S_{\mathbf{q}_{-i}})$. Hence,

$$\begin{aligned}
\mathbb{E}_{q_i}[p_i(\mathbf{q})] &= \frac{Q-\bar{q}_{\mathbf{q}_{-i}}}{Q}\hat{v}_i(\bar{q}_{\mathbf{q}_{-i}},\mathbf{q}_{-i}) \\
&= \frac{Q-\bar{q}_{\mathbf{q}_{-i}}}{Q}\left(\frac{\bar{q}_{\mathbf{q}_{-i}}}{1+\bar{q}_{\mathbf{q}_{-i}}} + \left(\frac{1}{1+\bar{q}_{\mathbf{q}_{-i}}}\right)^2 S_{\mathbf{q}_{-i}}\right) \\
&= Y\left(S_{\mathbf{q}_{-i}}\right),
\end{aligned}$$

where we define

$$Y(S) := \frac{Q-g^{-1}(S)}{Q}\left(\frac{g^{-1}(S)}{1+g^{-1}(S)} + \left(\frac{1}{1+g^{-1}(S)}\right)^2 S\right), \text{ for all } S \in [0,1/3]. \qquad \text{(B.28)}$$

By Lemma C.2(ii), given that $Q > 16$ from equation (B.20), we have that $Y(S)$ is strictly increasing in $S \in [0,1/3]$. Therefore, constants $\lambda_0, \hat{\lambda}_1 > 0$ independent of n and Σ exist such that

$$Y(S) = \lambda_0 + \hat{\lambda}_1 S + \mathcal{O}(S^2). \qquad \text{(B.29)}$$

Combining this with $S_{\mathbf{q}_{-i}} = \mathcal{O}(\varepsilon^2)$, we obtain

$$\begin{aligned}
\mathbb{E}_{\mathbf{q}}[p_i(\mathbf{q})] = \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[p_i(\mathbf{q})\right]\right] &= \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\lambda_0 + \hat{\lambda}_1 S_{\mathbf{q}_{-i}} + \mathcal{O}(\varepsilon^4)\right]\right] \\
&= \lambda_0 + \lambda_1 \sum_{j\neq i}\sigma_{ij}^2 + \mathcal{O}(\varepsilon^4), \qquad \text{(B.30)}
\end{aligned}$$

where we define

$$\lambda_1 := \hat{\lambda}_1\left(1 - \frac{1}{Q}\log(1+Q)\right)$$

This concludes the proof as the error in equation (B.30) is $\mathcal{O}\left(\varepsilon^4\right)$. □

PROOF OF COROLLARY A.1:

The result in equation (A.6) follows directly from the proof of Proposition (A.2)(ii). Specifically, the approximation in equation (B.29) establishes the claim. Similarly, the result in equation (A.7) follows from the proof of Theorem 3(i). Specifically, when the effort level $e_i = 0$ and ε is sufficiently small—specifically, $\varepsilon < 1/\sqrt{3(n-1)}$—equation (B.36) yields

$$\mathbb{E}_{q_i}[U_i(\mathbf{q})] = \frac{1}{Q} H\left(S_{\mathbf{q}_{-i}}\right),$$

where $H(\cdot)$ is defined as in equation (B.37) and $S_{\mathbf{q}_{-i}} \coloneqq \sum_{j \neq i} \sigma_{ij}^2 q_j/(1+q_j)$. Since by Lemma C.2(iii), function $H(\cdot)$ is strictly increasing, its first-order Taylor approximation leads to the result in equation (A.7). □

PROOF OF THEOREM 3:

In what follows, we use $\mathbf{e}$ and $[e_i, \mathbf{e}_{-i}]$ interchangeably to refer to effort profiles. We begin by showing that each agent participates in the game and exerts an effort level in the interval $\left[0, \sqrt{Q}-2\right]$. Since $c_i(0) = 0$ for all $i \in [n]$, each agent can obtain non-negative utility from collaboration without exerting any effort.[13] Thus, all agents participate. Moreover, since $c_i(\cdot)$ is strictly increasing, we have $c_i(e_i) \geq 1$ for $e_i \geq \sqrt{Q}-2$. Additionally, agent i's value for any quality vector $\mathbf{q}$ satisfies

$$\begin{aligned}\hat{v}_i(\mathbf{q}) = \frac{q_i}{1+q_i} + \left(\frac{1}{1+q_i}\right)^2 \sum_{j \neq i} \sigma_{ij}^2 \frac{q_j}{1+q_j} &\leq \frac{q_i}{1+q_i} + \frac{1}{3}\left(\frac{1}{1+q_i}\right)^2 \\ &< \frac{q_i}{1+q_i} + \frac{1}{1+q_i} = 1.\end{aligned}$$

Therefore, agent i's expected value is strictly less than one for all effort profiles $\mathbf{e}$. Combining this with $c_i(e_i) \geq 1$ for $e_i \geq \sqrt{Q}-2$, we conclude that all agents choose effort levels in $\left[0, \sqrt{Q}-2\right]$. In what follows, we always assume

$$\mathbf{e} \in \left[0, \sqrt{Q}-2\right]^n. \tag{B.31}$$

(i) For each agent $i \in [n]$, we show that her expected utility from participating in the mechanism has increasing differences in $(e_i, \mathbf{e}_{-i})$. This implies that the game is supermodular, since the cost functions depend only on each agent's own effort level. Let $F_{\mathbf{e}_{-i}}$ denote the joint distribution over other agents' data qualities. For any given $\mathbf{q}_{-i}$, define $S_{\mathbf{q}_{-i}} \coloneqq \sum_{j \neq i} \sigma_{ij}^2 q_j/(1+q_j)$. We have

$$S_{\mathbf{q}_{-i}} = \sum_{j \neq i} \sigma_{ij}^2 \frac{q_j}{1+q_j} \leq \sum_{j \neq i} \sigma_{ij}^2 < 1/3, \tag{B.32}$$

where the first inequality follows from $q_j \geq 0$ for all $j \in [n]$. Moreover,

$$\begin{aligned}\hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(q_i) &= \hat{v}_i(q_i, \mathbf{q}_{-i}) - \frac{\partial \hat{v}_i(q_i, \mathbf{q}_{-i})}{\partial q_i} \frac{\bar{F}_{i,e_i}(q_i)}{f_{i,e_i}(q_i)} \\ &= \left(\frac{q_i}{1+q_i} + \left(\frac{1}{1+q_i}\right)^2 S_{\mathbf{q}_{-i}}\right) - \left(\frac{1}{(1+q_i)^2} - \frac{2}{(1+q_i)^3} S_{\mathbf{q}_{-i}}\right)(Q-q_i) \\ &= \left(\frac{q_i}{1+q_i} - \frac{Q-q_i}{(1+q_i)^2}\right) + \left(\frac{1}{(1+q_i)^2} + \frac{2(Q-q_i)}{(1+q_i)^3}\right) S_{\mathbf{q}_{-i}}.\end{aligned} \tag{B.33}$$

Notice that the virtual value function $\hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(q_i)$ is independent of e_i, since the hazard rates do not depend on efforts.

We first derive some useful properties. From Lemma C.2(i), the inverse function $g^{-1}(S)$ exists for $S \in [0, 1/3]$, where function $g(\cdot)$ is defined in equation (B.27). Define $\bar{q}_0 \coloneqq g^{-1}(1/3)$. Since $g(\bar{q}_0) = 1/3$, Lemma C.1(ii) implies that

$$\left(\frac{\bar{q}_0}{1+\bar{q}_0} - \frac{Q-\bar{q}_0}{(1+\bar{q}_0)^2}\right) + \left(\frac{1}{(1+\bar{q}_0)^2} + \frac{2(Q-\bar{q}_0)}{(1+\bar{q}_0)^3}\right)\frac{1}{3} = 0. \tag{B.34}$$

[13] See the description of the setting in Section 4, where we assume $c_i(0) = 0$.

Additionally, Lemma C.2(i) implies that

$$\bar{q}_0 \geq \sqrt{Q} - 2 \geq e_i, \tag{B.35}$$

where the last inequality follows from equation (B.31). Furthermore, by equation (B.33), Lemma C.1(i) implies that $\hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(q_i)$ is strictly increasing. Combining this with equation (B.35), we have

$$\begin{aligned}\hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(e_i) &\leq \hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(\bar{q}_0) \\ &= \left(\frac{1}{(1+\bar{q}_0)^2} + \frac{2(Q-\bar{q}_0)}{(1+\bar{q}_0)^3}\right)\left(S_{\mathbf{q}_{-i}} - \frac{1}{3}\right) < 0,\end{aligned}$$

where the equality follows from equation (B.34). As a result, since

$$\bar{q}_{i,\mathbf{q}_{-i},e_i} = \inf\left\{q \in [e_i, Q] \mid \hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(q) \geq 0\right\},$$

we have $\hat{\phi}_{e_i,i,\mathbf{q}_{-i}}(\bar{q}_{i,\mathbf{q}_{-i},e_i}) = 0$. Combining this with equation (B.33), Lemma C.1(ii) implies that

$$\bar{q}_{i,\mathbf{q}_{-i},e_i} = g^{-1}(S_{\mathbf{q}_{-i}}).$$

This implies that the threshold $\bar{q}_{i,\mathbf{q}_{-i},e_i}$ is independent of e_i for any $e_i \leq \sqrt{Q} - 2$. To simplify the notation, let $\bar{q} \coloneqq \bar{q}_{i,\mathbf{q}_{-i},e_i}$. We have

$$\begin{aligned}&\mathbb{E}_{q_i}[U_i(q_i,\mathbf{q}_{-i})] = \frac{1}{Q-e_i}\int_{\bar{q}}^{Q} v_i(q_i,\mathbf{q}_{-i}) - v_i(\bar{q},\mathbf{q}_{-i})\, dq_i \\ &= \frac{1}{Q-e_i}\left(\int_{\bar{q}}^{Q} \frac{q_i}{1+q_i} + \left(\frac{1}{1+q_i}\right)^2 S_{\mathbf{q}_{-i}}\, dq_i - (Q-\bar{q})\left(\frac{\bar{q}}{1+\bar{q}} + \left(\frac{1}{1+\bar{q}}\right)^2 S_{\mathbf{q}_{-i}}\right)\right) \\ &= \frac{1}{Q-e_i}\left(\log\left(\frac{1+\bar{q}}{1+Q}\right) + \frac{Q-\bar{q}}{1+\bar{q}} - \frac{(Q-\bar{q})^2}{(1+\bar{q})^2}\frac{S_{\mathbf{q}_{-i}}}{1+Q}\right) \\ &= \frac{1}{Q-e_i}H(S_{\mathbf{q}_{-i}}) \qquad\qquad (\text{B.36})\end{aligned}$$

where we define

$$H(S) \coloneqq \log\left(\frac{1+g^{-1}(S)}{1+Q}\right) + \frac{Q-g^{-1}(S)}{1+g^{-1}(S)} - \frac{\left(Q-g^{-1}(S)\right)^2}{\left(1+g^{-1}(S)\right)^2}\frac{S}{1+Q}, \text{ for all } S \in [0,1/3]. \tag{B.37}$$

Consequently,

$$\mathbb{E}_{\mathbf{q}\sim F_{\mathbf{e}}}[U_i(\mathbf{q})] = \frac{1}{Q-e_i}\mathbb{E}_{\mathbf{q}_{-i}\sim F_{\mathbf{e}_{-i}}}\left[H(S_{\mathbf{q}_{-i}})\right]. \tag{B.38}$$

Hence, for any $j \neq i$, we have

$$\frac{\partial \mathbb{E}_{\mathbf{q}\sim F_{\mathbf{e}}}[U_i(\mathbf{q})]}{\partial e_i \partial e_j} = \frac{1}{(Q-e_i)^2}\frac{\partial}{\partial e_j}\mathbb{E}_{\mathbf{q}_{-i}\sim F_{\mathbf{e}_{-i}}}\left[H(S_{\mathbf{q}_{-i}})\right]. \tag{B.39}$$

To establish supermodularity, it suffices to show that $\mathbb{E}_{\mathbf{q}_{-i}\sim F_{\mathbf{e}_{-i}}}\left[H(S_{\mathbf{q}_{-i}})\right]$ is non-decreasing in e_j. We prove a stronger, pointwise property: for any fixed $\mathbf{q}_{-i,j}$, the function $\mathbb{E}_{q_j\sim F_{j,e_j}}[H(S_{q_j,\mathbf{q}_{-i,j}})]$ is non-decreasing in e_j. Define the function $G(q_j) \coloneqq H(S_{q_j,\mathbf{q}_{-i,j}})$. By Lemma C.2(iii), $H(S)$ is non-decreasing in $S \in [0,1/3]$. Moreover,

$$S_{q_j,\mathbf{q}_{-i,j}} = \sigma_{ij}^2\frac{q_j}{1+q_j} + \sum_{k\neq i,j}\sigma_{ik}^2\frac{q_k}{1+q_k} \in [0,1/3]$$

is non-decreasing in q_j. Therefore, $G(q_j) = H(S_{q_j,\mathbf{q}_{-i,j}})$ is non-decreasing in q_j. Moreover, for any $e_j \geq e'_j$, F_{j,e_j} (uniform on $[e_j, Q]$) stochastically dominates F_{j,e'_j} (uniform on $[e'_j, Q]$). This implies that $\mathbb{E}_{q_j\sim F_{j,e_j}}[G(q_j)]$ is non-decreasing in e_j, as desired.

(ii) We invoke two classical results:

TOPKIS' MONOTONICITY THEOREM ([20]): *Let $X \subset \mathbb{R}$ be a compact set and T be some partially ordered set. Assume that the function $f : X \times T \to \mathbb{R}$ is continuous in x for all $t \in T$ and has increasing differences in (x, t). Define $x(t) \equiv \operatorname{argmax}_{x \in X} f(x, t)$. Then, we have:*

(i) For all $t \in T$, $x(t)$ is nonempty and has a greatest and least element, denoted by $\bar{x}(t)$ and $\underline{x}(t)$, respectively.
(ii) For all $t' \geq t$, we have $\bar{x}(t') \geq \bar{x}(t)$ and $\underline{x}(t') \geq \underline{x}(t)$.

TARSKI'S FIXED POINT THEOREM ([32]): *Let $(L, \leq)$ be a complete lattice, and let $T : L \to L$ be an order-preserving mapping. Then:*

(i) The set of fixed points of T, defined as $\{x \in L \mid T(x) = x\}$, is a non-empty complete lattice.
(ii) In particular, T has a least fixed point and a greatest fixed point in L.

For each agent $i \in [n]$, let $\overline{BR}_i(\mathbf{e}_{-i})$ denote her best response effort to others' efforts.[14] That is,

$$\overline{BR}_i(\mathbf{e}_{-i}) = \underset{e_i \in \left[0, \sqrt{Q}-2\right]}{\operatorname{argmax}} \; \mathbb{E}_{\mathbf{q} \sim F_{(e_i, \mathbf{e}_{-i})}} \left[U_i(\mathbf{q})\right] - c_i(e_i).$$

As shown in part (i), $\mathbb{E}_{\mathbf{q} \sim F_{(e_i, \mathbf{e}_{-i})}} \left[U_i(\mathbf{q})\right] - c_i(e_i)$ has increasing differences in (e_i, e_j) for any $j \neq i$. Thus, by Topkis' Monotonicity Theorem, $\overline{BR}_i(\mathbf{e}_{-i})$ is non-decreasing in $\mathbf{e}_{-i}$.

We now define the best response mapping $\overline{BR}(\mathbf{e}) : \left[0, \sqrt{Q} - 2\right]^n \to \left[0, \sqrt{Q} - 2\right]^n$ as

$$\overline{BR}(\mathbf{e}) \coloneqq \left(\overline{BR}_1(\mathbf{e}_{-1}), \ldots, \overline{BR}_n(\mathbf{e}_{-n})\right).$$

This mapping is order preserving: for any $\mathbf{e}' \geq \mathbf{e}$, $\overline{BR}(\mathbf{e}') \geq \overline{BR}(\mathbf{e})$. Since $\left[0, \sqrt{Q} - 2\right]^n$ is a complete lattice, Tarski's Fixed Point Theorem implies that the set of fixed points of $\overline{BR}$ is non-empty and forms a complete lattice, i.e., it has a greatest and least element. We conclude the proof by noting that the set of fixed points of the best response mapping is the set of all equilibria.

(iii) Let $\hat{\Sigma} = (\sigma_{ij}^2)_{i \neq j}$ denote the matrix of squared correlation factors. We show that the expected utility of agent i has increasing differences in $(e_i, \hat{\Sigma})$. This complements part (ii), where we established increasing differences in (e_i, e_j).

By equation (B.38), we have

$$\begin{aligned}
\frac{\partial \mathbb{E}_{\mathbf{q} \sim F_{\mathbf{e}}} \left[U_i(\mathbf{q})\right]}{\partial e_i \partial \hat{\Sigma}_{ij}} &= \frac{1}{(Q - e_i)^2} \frac{\partial}{\partial \hat{\Sigma}_{ij}} \mathbb{E}_{\mathbf{q}_{-i} \sim F_{\mathbf{e}_{-i}}} \left[H(S_{\mathbf{q}_{-i}})\right] \\
&= \frac{1}{(Q - e_i)^2} \mathbb{E}_{\mathbf{q}_{-i} \sim F_{\mathbf{e}_{-i}}} \left[\frac{\partial}{\partial \hat{\Sigma}_{ij}} H(S_{\mathbf{q}_{-i}})\right] \\
&= \frac{1}{(Q - e_i)^2} \mathbb{E}_{\mathbf{q}_{-i} \sim F_{\mathbf{e}_{-i}}} \left[\frac{q_j}{1 + q_j} H'(S_{\mathbf{q}_{-i}})\right] \geq 0,
\end{aligned}$$

where the last inequality follows from Lemma C.2(iii) and that $S_{\mathbf{q}_{-i}} \in [0, 1/3]$ for all $\mathbf{q}_{-i} \geq 0$ (see equation (B.32)). Finally, since $S_{\mathbf{q}_{-i}}$ is independent of $\hat{\Sigma}_{kj}$ for $k, j \neq i$, we conclude that agent i's expected utility has increasing differences in $(e_i, \hat{\Sigma})$.

We now apply standard arguments from supermodular game theory (e.g., [33]). As shown in part (ii), agent i's best response is non-decreasing in others' efforts. A similar argument shows it is also non-decreasing in σ_{ij}^2. Define the best response mapping $\overline{BR}(\mathbf{e}, \hat{\Sigma}) \coloneqq \left(\overline{BR}_1(\cdot), \ldots, \overline{BR}_n(\cdot)\right)$. For any Nash equilibrium $\mathbf{e}^{NE}$, we have $\overline{BR}(\mathbf{e}^{NE}, \hat{\Sigma}) = \mathbf{e}^{NE}$. Thus, the set $\mathcal{E}(\hat{\Sigma}) \coloneqq \{\overline{BR}(\mathbf{e}, \hat{\Sigma}) \geq \mathbf{e}\}$ includes all Nash equilibria.

We show that $\mathcal{E}(\hat{\Sigma})$ has the largest element $\mathbf{e}(\hat{\Sigma})$, and it is the largest Nash equilibrium in the lattice order. Let $\mathbf{e}_0 \in \mathcal{E}(\hat{\Sigma})$. Define the sequence of best response dynamics $\mathbf{e}_{i+1} \coloneqq \overline{BR}(\mathbf{e}_i)$ for

[14] If the maximizer is not unique, we select the greatest one.

$i \geq 0$. Because $\mathbf{e}_0 \in \mathcal{E}(\hat{\Sigma})$ and $\overline{BR}(\cdot, \hat{\Sigma})$ is non-decreasing, $\{\mathbf{e}_l\}_{l=0}^{\infty}$ is a non-decreasing sequence. Because $\mathbf{e}_l \in \left[0, \sqrt{Q} - 2\right]$ is upper bounded, the sequence converges to some $\mathbf{e}_\infty \in \left[0, \sqrt{Q} - 2\right]$. By continuity of the best response mapping, we obtain $\overline{BR}(\mathbf{e}_\infty) = \mathbf{e}_\infty$, and thus $\mathbf{e}_\infty$ is a Nash equilibrium. Let $\bar{\mathbf{e}}(\hat{\Sigma})$ denote the largest Nash equilibrium. Since $\mathbf{e}_\infty$ is a Nash equilibrium, we have $\bar{\mathbf{e}}(\hat{\Sigma}) \geq \mathbf{e}_\infty \geq \mathbf{e}_0$. Therefore, $\bar{\mathbf{e}}(\hat{\Sigma}) = \max \mathcal{E}(\hat{\Sigma})$. Finally, because $\overline{BR}(\mathbf{e}, \cdot)$ is non-decreasing, we obtain $\mathcal{E}(\hat{\Sigma}) \subseteq \mathcal{E}(\hat{\Sigma}')$ for any $\hat{\Sigma} \leq \hat{\Sigma}'$ (entry-wise). This implies that $\bar{\mathbf{e}}(\hat{\Sigma})$ is non-decreasing in $\hat{\Sigma}$ as desired. $\square$

C Auxiliary Technical Lemmas

PROOF OF LEMMA B.1:

For agent i, define her interim expected utility as:

$$U_i(q_i) \coloneqq \mathbb{E}_{\mathbf{q}_{-i}}\left[v_i\left(q_i, \mathbf{q}_{-i}\right) x_i\left(q_i, \mathbf{q}_{-i}\right) - p_i(q_i, \mathbf{q}_{-i})\right]. \tag{C.1}$$

We first show that

$$U_i(q_i) = U_i(0) + \mathbb{E}_{\mathbf{q}_{-i}}\left[\int_0^{q_i} \frac{\partial v_i(s, \mathbf{q}_{-i})}{\partial q_i} x_i(s, \mathbf{q}_{-i})\, ds\right], \quad \text{for all } q_i \in \mathcal{Q}_i.$$

We then derive the payment identity by taking the expectation over q_i, as we will elaborate on later in the proof.

Using equation (C.1), we express problem (A.2) as:

$$\max_{x_i(\cdot), p_i(\cdot): i\in[n]} \sum_{i=1}^{n} \mathbb{E}_{\mathbf{q}}\left[p_i(\mathbf{q})\right] \tag{main}$$

$$s.t.\ U_i(q_i) \geq U_i(q_i') + \mathbb{E}_{\mathbf{q}_{-i}}\left[\left(v_i\left(\left[q_i, \mathbf{q}_{-i}\right], \left[q_i', \mathbf{q}_{-i}\right]\right) - v_i\left(q_i', \mathbf{q}_{-i}\right)\right) x_i\left(q_i', \mathbf{q}_{-i}\right)\right], \forall i, q_i' \neq q_i, \tag{IC}$$

$$U_i(q_i) \geq 0, \forall i, q_i. \tag{IR}$$

Moreover, by Assumption 1(iii), the accuracy function $A_i(\cdot)$ belongs to $\mathcal{C}^2$. Hence, since $v_i(\cdot) = g_i(A_i(\cdot))$ for a strictly increasing function $g_i \in \mathcal{C}^2$, we have $v_i(\cdot) \in \mathcal{C}^2$. Thus, a uniform bound $D < \infty$ exists such that

$$\left|\frac{v_i([q_i', \mathbf{q}_{-i}], [q_i, \mathbf{q}_{-i}]) - v_i([q_i, \mathbf{q}_{-i}], [q_i, \mathbf{q}_{-i}])}{q_i' - q_i}\right| < D, \text{ for all } q_i, q_i' \in \mathcal{Q}_i, \mathbf{q}_{-i} \in \prod_{j\neq i} \mathcal{Q}_j. \tag{C.2}$$

Additionally, for any $\mathbf{q}_{-i} \in \prod_{j\neq i} \mathcal{Q}_j$, we have $q_i = \operatorname{argmax} A_i(\mathbf{q}, [, \mathbf{q}_{-i}])$ by Assumption 1(ii). Since $v_i(\cdot) = g_i(A_i(\cdot))$ for a strictly increasing function g_i, we obtain $q_i = \operatorname{argmax} v_i(\mathbf{q}, [, \mathbf{q}_{-i}])$. This implies that $\frac{\partial v_i(\mathbf{q}, \mathbf{q})}{\partial} = 0$. Combining this with $v_i(\cdot) \in \mathcal{C}^2$, we have

$$\begin{aligned}
\lim_{q_i' \to q_i} \frac{v_i([q_i', \mathbf{q}_{-i}], [q_i, \mathbf{q}_{-i}]) - v_i([q_i, \mathbf{q}_{-i}], [q_i, \mathbf{q}_{-i}])}{q_i' - q_i} &= \frac{\partial v_i}{\partial q_i}(\mathbf{q}, \mathbf{q}) \\
&= \frac{\partial v_i}{\partial q_i}(\mathbf{q}, \mathbf{q}) + \frac{\partial v_i}{\partial}(\mathbf{q}, \mathbf{q}) \\
&= \frac{\partial v_i}{\partial q_i}(q_i, \mathbf{q}_{-i}), \text{ for all } q_i \in \mathcal{Q}_i,
\end{aligned} \tag{C.3}$$

Moreover, for $q_i' > q_i$, constraint (IC) implies that

$$\begin{aligned}
\mathbb{E}_{\mathbf{q}_{-i}}\left[\frac{v_i([q_i', \mathbf{q}_{-i}], [q_i', \mathbf{q}_{-i}]) - v_i([q_i, \mathbf{q}_{-i}], [q_i', \mathbf{q}_{-i}])}{q_i' - q_i} x_i(q_i', \mathbf{q}_{-i})\right] &\geq \frac{U_i(q_i') - U_i(q_i)}{q_i' - q_i} \\
\geq \mathbb{E}_{\mathbf{q}_{-i}}\left[\frac{v_i([q_i', \mathbf{q}_{-i}], [q_i, \mathbf{q}_{-i}]) - v_i([q_i, \mathbf{q}_{-i}], [q_i, \mathbf{q}_{-i}])}{q_i' - q_i} x_i(q_i, \mathbf{q}_{-i})\right]&.
\end{aligned} \tag{C.4}$$

Combining this with $x_i(\cdot) \in [0, 1]$ and equation (C.2), we obtain

$$\left|U_i(q_i) - U_i(q_i')\right| \leq D\left|q_i - q_i'\right|, \quad \text{for all } q_i > q_i'.$$

Therefore, $U_i(q_i)$ is Lipschitz continuous, and thus, its derivative $U_i'(q_i)$ exists almost everywhere and satisfies

$$U_i(q_i) = U_i(0) + \int_0^{q_i} U_i'(s)\, ds, \quad \text{for all } q_i \in \mathcal{Q}_i. \tag{C.5}$$

Let $\mathcal{D}_{U_i} \coloneqq \{q_i \in \mathcal{Q}_i \mid U_i'(q_i) \text{ exists}\}$ and $q_i \in \mathcal{D}_{U_i}$. We next characterize $U_i'(q_i)$ by taking the limit $q_i' - q_i \to 0$ in the right-hand side of equation (C.4). Given the pointwise convergence by equation (C.3) and uniform bound by equation (C.2), the Dominated Convergence Theorem yields

$$U_i'(q_i) \geq \mathbb{E}_{\mathbf{q}_{-i}}\left[\frac{\partial v_i(q_i, \mathbf{q}_{-i})}{\partial q_i} x_i(q_i, \mathbf{q}_{-i})\right].$$

A similar argument, by using constraint (IC) for $q_i' < q_i$, implies that

$$U_i'(q_i) \le \mathbb{E}_{\mathbf{q}_{-i}}\left[\frac{\partial v_i(q_i,\mathbf{q}_{-i})}{\partial q_i} x_i(q_i,\mathbf{q}_{-i})\right].$$

As a result, we obtain

$$U_i'(q_i) = \mathbb{E}_{\mathbf{q}_{-i}}\left[\frac{\partial v_i(q_i,\mathbf{q}_{-i})}{\partial q_i} x_i(q_i,\mathbf{q}_{-i})\right], \quad \text{for all } q_i \in \mathcal{D}_{U_i}.$$

Since $\mathcal{Q}_i \setminus \mathcal{D}_{U_i}$ is of measure zero and U_i is Lipschitz continuous, equation (C.5) implies that

$$U_i(q_i) = U_i(0) + \mathbb{E}_{\mathbf{q}_{-i}}\left[\int_0^{q_i} \frac{\partial v_i(s,\mathbf{q}_{-i})}{\partial q_i} x_i(s,\mathbf{q}_{-i})\, ds\right], \quad \text{for all } q_i \in \mathcal{Q}_i. \tag{C.6}$$

Additionally, combining equations (C.1) and (C.6) yields

$$p_i(q_i) = -U_i(0) + \mathbb{E}_{\mathbf{q}_{-i}}\left[v_i(q_i,\mathbf{q}_{-i})x_i(q_i,\mathbf{q}_{-i}) - \int_0^{q_i} \frac{\partial v_i(s,\mathbf{q}_{-i})}{\partial s} x_i(s,\mathbf{q}_{-i})\, ds\right].$$

Thus, taking the expectation over q_i implies that

$$\mathbb{E}_{q_i}[p_i(q_i)] = -U_i(0) + \mathbb{E}_{\mathbf{q}}\left[v_i(\mathbf{q})x_i(\mathbf{q})\right] - \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\int_0^{q_i} \frac{\partial v_i(s,\mathbf{q}_{-i})}{\partial s} x_i(s,\mathbf{q}_{-i})\, ds\right]\right]. \tag{C.7}$$

Moreover, we have

$$\begin{aligned}
\mathbb{E}_{q_i}\left[\int_0^{q_i} \frac{\partial v_i(s,\mathbf{q}_{-i})}{\partial s} x_i(s,\mathbf{q}_{-i})\, ds\right] &= \int_{q_i\in\mathcal{Q}_i}\int_0^{q_i} \frac{\partial v_i(s,\mathbf{q}_{-i})}{\partial s} f_i(q_i) x_i(s,\mathbf{q}_{-i})\, ds\, dq_i \\
&= \int_{s\in\mathcal{Q}_i}\int_s \frac{\partial v_i(s,\mathbf{q}_{-i})}{\partial s} f_i(q_i) x_i(s,\mathbf{q}_{-i})\, dq_i\, ds_i \\
&= \int_{s\in\mathcal{Q}_i} \frac{\partial v_i(s,\mathbf{q}_{-i})}{\partial s} (1-F_i(s)) x_i(s,\mathbf{q}_{-i})\, ds_i \\
&= \mathbb{E}_{q_i}\left[\frac{\partial v_i(q_i,\mathbf{q}_{-i})}{\partial q_i}\frac{1-F_i(q_i)}{f_i(q_i)} x_i(q_i,\mathbf{q}_{-i})\right].
\end{aligned} \tag{C.8}$$

Combining equations (C.7) and (C.8), we have

$$\begin{aligned}
\mathbb{E}_{q_i}[p_i(q_i)] &= -U_i(0) + \mathbb{E}_{\mathbf{q}}[v_i(\mathbf{q})x_i(\mathbf{q})] - \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\frac{\partial v_i(q_i,\mathbf{q}_{-i})}{\partial q_i}\frac{1-F_i(q_i)}{f_i(q_i)} x_i(q_i,\mathbf{q}_{-i})\right]\right] \\
&= -U_i(0) + \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\left(v_i(q_i,\mathbf{q}_{-i}) - \frac{\partial v_i(q_i,\mathbf{q}_{-i})}{\partial q_i}\frac{1-F_i(q_i)}{f_i(q_i)}\right) x_i(q_i,\mathbf{q}_{-i})\right]\right] \\
&\overset{(a)}{=} -U_i(0) + \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\phi_{i,\mathbf{q}_{-i}}(q_i) x_i(q_i,\mathbf{q}_{-i})\right]\right] \\
&\overset{(b)}{=} \mathbb{E}_{\mathbf{q}_{-i}}\left[p_i(0,\mathbf{q}_{-i}) - v_i\left(0,\mathbf{q}_{-i}\right) x_i\left(0,\mathbf{q}_{-i}\right)\right] + \mathbb{E}_{\mathbf{q}_{-i}}\left[\mathbb{E}_{q_i}\left[\phi_{i,\mathbf{q}_{-i}}(q_i) x_i(q_i,\mathbf{q}_{-i})\right]\right],
\end{aligned}$$

where (a) follows from the definition of the the virtual value function $\phi_{i,\mathbf{q}_{-i}}(q_i)$ in equation (3), and (b) follows from the the definition of the interim utility $U_i(\cdot)$ in equation (C.1). □

PROOF OF LEMMA B.2:

We apply Remark A.1 to establish that the virtual value function $\phi_{i,\mathbf{q}_{-i}}(\cdot)$ is non-decreasing. Since F_i has a monotone hazard rate, it suffices to show that $v_i(\mathbf{q})$ is non-decreasing and concave in q_i.

Assume $i = 1$ without loss of generality. By equation (2), $v_1(\mathbf{q})$ is non-decreasing in q_i since the accuracy function $A_i(\mathbf{q})$ is non-decreasing in q_i using Lemma A.1(i). To establish the concavity of $v_1(\mathbf{q})$, we first derive two properties of the inverse matrix $\Sigma_{\mathbf{q}}^{-1}$, where

$$\Sigma_{\mathbf{q}} = \begin{pmatrix} \frac{\sigma_1^2(1+q_1)}{q_1} & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \frac{\sigma_2^2(1+q_2)}{q_2} & \cdots & \Sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \cdots & \frac{\sigma_n^2(1+q_n)}{q_n}. \end{pmatrix}.$$

Fixing $\mathbf{q}_{-1}$, let c_{q_1} denote the first diagonal entry of $\Sigma_{\mathbf{q}}^{-1}$, and m_{1j} denote the determinant of the matrix obtained by removing row 1 and column j from $\Sigma_{\mathbf{q}}$. We have

$$c_{q_1} = \frac{m_{11}}{\frac{\sigma_1^2(1+q_1)}{q_1} m_{11} + \sum_{j\geq 2}(-1)^{1+j}\Sigma_{1j} m_{1j}} = \frac{1}{\frac{\sigma_1^2(1+q_1)}{q_1} + \gamma}, \tag{C.9}$$

where $\gamma := \sum_{j\geq 2}(-1)^{1+j}\Sigma_{1j}\frac{m_{1j}}{m_{11}}$. Note that γ does not depend on q_1, since only the first diagonal entry of Σ_q is a function of q_1. We next show

$$\sigma_1^2 + \gamma \geq 0. \tag{C.10}$$

Since $\Sigma_{\mathbf{q}}$ is a covariance matrix for all $\mathbf{q}$, it follows that $\Sigma_{\mathbf{q}}$ is positive definite.[15] Hence, $\Sigma_{\mathbf{q}}^{-1}$ is positive definite as well, which implies that $c_{q_1} > 0$ for all values of q_1. Combining this with equation (C.9), we conclude that

$$0 \leq \lim_{q_1\to\infty} \frac{1}{c_{q_1}} = \lim_{q_1\to\infty} \frac{\sigma_1^2(1+q_1)}{q_1} + \gamma = \sigma_1^2 + \gamma.$$

We now establish the desired result that $v_i(\mathbf{q})$ is concave in q_i. We show the partial derivative $\frac{\partial v_1(\mathbf{q})}{\partial q_1}$ is non-increasing in q_i. Using equation (B.5), we derive

$$\begin{aligned}\frac{\partial v_1(\mathbf{q})}{\partial q_1} = \frac{1}{q_1^2}\left(-c_{11}\frac{\sigma_1^2}{q_1} + 1\right)^2 &= \frac{1}{q_1^2}\left(-\frac{\frac{\sigma_1^2}{q_1}}{\frac{\sigma_1^2(q_1+1)}{q_1} + \gamma} + 1\right)^2 \\ &= \frac{1}{q_1^2}\left(\frac{\sigma_1^2 + \gamma}{\frac{\sigma_1^2(q_1+1)}{q_1} + \gamma}\right)^2 \\ &= \left(\frac{\sigma_1^2 + \gamma}{q_1(\sigma_1^2+\gamma) + \sigma_1^2}\right)^2,\end{aligned} \tag{C.11}$$

where the second equality follows from equation (C.9). We conclude the proof by noting that γ does not depend on q_1 and $\sigma_1^2 + \gamma \geq 0$ by equation (C.10). □

PROOF OF LEMMA B.3:

We use a similar argument as in the proof of Lemma B.2. Specifically, we apply Remark A.1 to establish that the approximate virtual value function $\hat{\phi}_{i,\mathbf{q}_{-i}}(\cdot)$ is non-decreasing. Since F_i has a monotone hazard rate, it suffices to show that $\hat{v}_i(\mathbf{q})$ is non-decreasing and concave in q_i.

Assume $i = 1$ without loss of generality. For any $0 < \varepsilon < 1/\sqrt{3(n-1)}$, we have

$$1/3 > (n-1)\varepsilon^2 \geq \sum_{j\neq i}\sigma_{ij}^2 \geq \sum_{j\neq i}\sigma_{ij}^2\frac{q_j}{1+q_j}, \quad \text{for all } i \in [n], \tag{C.12}$$

where the second inequality follows from $\varepsilon \geq |\sigma_{ij}|$ for all $j \neq i$ under the ε-correlation regime, and the last inequality follows from $q_j \geq 0$. Moreover,

$$\begin{aligned}\frac{\partial \hat{v}^{(i)}(\mathbf{q})}{\partial q_i} &= \frac{1}{(1+q_i)^3}\left((1+q_i) - 2\sum_{j\neq i}\sigma_{ij}^2\frac{q_j}{1+q_j}\right) > 0, \\ \frac{\partial^2 \hat{v}^{(i)}(\mathbf{q})}{\partial q_i^2} &= \frac{-2}{(1+q_i)^4}\left((1+q_i) - 3\sum_{j\neq i}\sigma_{ij}^2\frac{q_j}{1+q_j}\right) < 0,\end{aligned}$$

where both inequalities follow from $q_i \geq 0$ and equation (C.12). Combining these two inequalities completes the proof. □

[15] As mentioned in the paper, we always implicitly assume that $\Sigma_{\mathbf{q}}$ is invertible. Thus, $\Sigma_{\mathbf{q}}$ is positive semi-definite and invertible, and therefore positive definite.

Additional Technical Lemmas

Lemma C.1. *For $S \in [0, 1/3]$ and $Q > 0$ the following statements hold:*

(i) The function

$$B(q) \coloneqq \left(\frac{q}{1+q} - \frac{Q-q}{(1+q)^2}\right) + \left(\frac{1}{(1+q)^2} + \frac{2(Q-q)}{(1+q)^3}\right) S_{\mathbf{q}_{-i}} \tag{C.13}$$

is strictly increasing in $q \in [0, Q]$.

(ii) $B(q) = 0$ if and only if $g(q) = S$, where function $g(\cdot)$ is defined as in equation (B.27).

PROOF OF LEMMA C.1: We prove each part separately.

(i) First, we rewrite $B(q)$ as:

$$B(q) = \left(\frac{q}{1+q} + \frac{1}{(1+q)^2}S\right) - \left(\frac{1}{(1+q)^2} - \frac{2}{(1+q)^3}S\right)(Q-q).$$

For all $q \in [0, Q]$, we compute the derivatives:

$$\frac{\partial}{\partial q}\left(\frac{q}{1+q} + \frac{1}{(1+q)^2}S\right) = \frac{1}{(1+q)^3}\left((1+q) - 2S\right) > 0,$$
$$\frac{\partial}{\partial q}\left(\frac{1}{(1+q)^2} - \frac{2}{(1+q)^3}S\right) = \frac{-2}{(1+q)^4}\left((1+q) - 3S\right) < 0,$$

where both inequalities follow from $S \leq 1/3$. Since $Q - q$ is strictly decreasing in $q \in [0, Q]$, we conclude that $B(q)$ is strictly increasing in $q \in [0, Q]$.

(ii) Multiplying both sides of the equation $B(q) = 0$ by $(1+q)^3$ and algebraic simplification implies the desired result. □

Lemma C.2. *Suppose $Q > 4$. Let functions g, Y, and H be defined as equations* (B.27), (B.28)*, and* (B.37)*. For $S \in [0, 1/3]$, the following statements hold:*

(i) The inverse function $g^{-1}(S)$ exists and is strictly decreasing in S. Additionally,

$$g^{-1}(S) \in \left[\sqrt{Q} - 2, \sqrt{1+Q} - 1\right]. \tag{C.14}$$

(ii) If $Q > 16$, then the function $Y(S)$ exists and is strictly increasing in S.

(iii) The function $H(S)$ exists and is strictly increasing in S.

PROOF OF LEMMA C.2: We prove each part separately.

(i) We first show that $g^{-1}(S)$ exists for $S \in [0, 1/3]$. It suffices to show that for any $S \in [0, 1/3]$, a unique $q \in [0, Q]$ exists that satisfies $g(q) = S$. By Lemma C.1(ii), this is equivalent to $B(q) = 0$, where $B(\cdot)$ is defined in equation (C.13). To show that a unique $q \in [0, Q]$ exists that satisfies $B(q) = 0$, it suffices to show that $B(q)$ is strictly increasing in $q \in [0, Q]$, $B(0) \leq 0$, and $B(Q) \geq 0$. The monotonicity follows from Lemma C.1(i). Moreover,

$$B(0) = S - (1-2S)Q \overset{(a)}{\leq} 9S - 4 \overset{(b)}{<} 0,$$
$$B(Q) = \frac{Q}{1+Q} + \frac{1}{(1+Q)^2}S > 0,$$

where (a) follows from $Q > 4$, and (b) follows from $S \in [0, 1/3]$. Therefore, by monotonicity of $B(\cdot)$, there exists a unique $q \in (0, Q)$ such that $B(q) = 0$. Hence, $g^{-1}(S) \in (0, Q)$ exists for all $S \in [0, 1/3]$.

We next show that $g^{-1}(S)$ is strictly decreasing in $S \in [0, 1/3]$. We first show $g(q)$ is strictly decreasing at $g^{-1}(0)$. Define $\tilde{q} := g^{-1}(0)$. Thus,

$$0 = g(\tilde{q}) = \frac{-\tilde{q}(1+\tilde{q})^2 + (Q-\tilde{q})(1+\tilde{q})}{(1+\tilde{q}) + 2(Q-\tilde{q})} = \frac{(1+\tilde{q})\left(-\tilde{q}(1+\tilde{q}) + Q - \tilde{q}\right)}{(1+\tilde{q}) + 2(Q-\tilde{q})},$$

where the second equality follows from the definition of $g(\cdot)$ in equation (B.27). Therefore,

$$Q = \tilde{q}^2 + 2\tilde{q}. \tag{C.15}$$

As a result, we derive

$$\begin{aligned} g'(\tilde{q}) &= \frac{(-2\tilde{q}(1+\tilde{q}) - (1+\tilde{q}) + Q - \tilde{q})(1 - \tilde{q} + 2Q)}{(1-\tilde{q}+Q)^2} \\ &= \frac{(-1 - 2\tilde{q}(1+\tilde{q}) - (1+\tilde{q}) - \tilde{q})(1 - \tilde{q} + 2Q)}{(1-\tilde{q}+Q)^2} < 0. \end{aligned}$$

Thus, since $g^{-1}(S)$ exists for all $S \in [0, 1/3]$, we have that $g(q)$ is strictly monotone for q in between $g^{-1}(0)$ and $g^{-1}(1/3)$. Combining this with the fact that $g(\cdot)$ is strictly decreasing at $g^{-1}(0)$, we conclude that $g^{-1}(\cdot)$ is strictly decreasing. This implies that for all $S \in [0, 1/3]$,

$$g^{-1}(S) \in \left[g^{-1}(1/3), g^{-1}(0)\right].$$

We next show the result in equation (C.14). By equation (C.15), we have $g^{-1}(0) = \sqrt{1+Q} - 1$. Moreover, we show $g^{-1}(1/3) \geq \sqrt{Q} - 2$, completing the proof. Define $\bar{q}_0 := g^{-1}(1/3)$. We have

$$\frac{1}{3} = g(\bar{q}_0) = \frac{-\bar{q}_0(1+\bar{q}_0)^2 + (Q-\bar{q}_0)(1+\bar{q}_0)}{1 - \bar{q}_0 + 2Q},$$

which implies that

$$Q = \frac{1 + 2\bar{q}_0 + 3\bar{q}_0^2 + 3\bar{q}_0(1+\bar{q}_0)^2}{1 + 3\bar{q}_0} = (1+\bar{q}_0)^2 + \frac{2\bar{q}_0^2}{1+3\bar{q}_0} \leq (1+\bar{q}_0)^2 + 2\bar{q}_0 < (2+\bar{q}_0)^2.$$

Thus, $\bar{q}_0 > \sqrt{Q} - 2$ as desired.

(ii) Define the interval

$$:= \left[\sqrt{Q} - 2, \sqrt{1+Q} - 1\right]. \tag{C.16}$$

Since $g^{-1}(S)$ exists for all $S \in [0, 1/3]$, equation (B.28) implies that $Y(S)$ exists for all these values. We have $Y(S) = K(g^{-1}(S))$, where

$$K(q) := \frac{Q-q}{Q}\left(\frac{q}{1+q} + \left(\frac{1}{1+q}\right)^2 g(q)\right).$$

To establish that $Y(S)$ is strictly increasing for $S \in [0, 1/3]$, it suffices to show $K(q)$ is strictly decreasing for $q \in$. Note that $g^{-1}(S) \in$ is strictly decreasing in S by part (i). By definition of $g(\cdot)$ in equation (B.27), we have

$$\begin{aligned} K(q) &= \frac{Q-q}{Q}\left(\frac{q}{1+q} + \left(\frac{1}{1+q}\right)^2 \frac{-q(1+q)^2 + (Q-q)(1+q)}{1+q+2(Q-q)}\right) \\ &= \frac{Q-q}{Q}\left(\frac{q\left(1+q+2(Q-q)\right) - q(1+q) + (Q-q)}{(1+q)\left(1+q+2(Q-q)\right)}\right) \\ &= \frac{Q-q}{Q}\left(\frac{(Q-q)(1+2q)}{(1+q)\left(1+q+2(Q-q)\right)}\right) \\ &= \frac{(Q-q)^2(1+2q)}{Q(1+q)\left(1+q+2(Q-q)\right)}. \end{aligned}$$

Thus, we derive

$$K'(q) = \frac{2Q^3 + Q^2\left(-6q^2 - 14q - 2\right) + Q\left(8q^3 + 12q^2 - 4q - 2\right) + -2q^4 + 6q^2 + 2q}{Q\left((1+q)\left(1+q+2(Q-q)\right)\right)^2}.$$

We next show that the numerator is negative for all $q \in$, which shows the monotonicity of $K(\cdot)$. It suffices to establish the following for all $q \in$:

$$2Q^3 + \left(-2q^2 - 14q - 2\right)Q^2 \leq 0, \tag{C.17}$$

$$-4q^2Q^2 + \left(8q^3 + 12q^2\right)Q \leq 0, \tag{C.18}$$

$$Q(-4q-2) - 2q^4 + 6q^2 + 2q < 0. \tag{C.19}$$

Note that summing these three inequalities implies the desired result. We show each separately:

Equation (C.17)**.** Since $q \in$, we have $q \geq \sqrt{Q} - 2$. Hence, given that $Q > 16$, we have

$$(q+2)^2 \geq Q > 16. \tag{C.20}$$

Thus,

$$2Q^3 + \left(-2q^2 - 14q - 2\right)Q^2 \leq 2Q^3 + \left(-2q^2 - 8q - 8\right)Q^2 = 2Q^3 - 2\left(q+2\right)^2 Q^2 \leq 0,$$

where both inequalities follow from equation (C.20).

Equation (C.18)**.** Since $q \in$, equation (C.16) implies that $(1+q)^2 \leq 1 + Q$. Thus,

$$q^2 + 2q \leq Q, \tag{C.21}$$

which leads to

$$-4q^2Q^2 + \left(8q^3 + 12q^2\right)Q \leq -4q^2\left(q^2 + 2q\right)Q + \left(8q^3 + 12q^2\right)Q = \left(-4q^4 + 12q^2\right)Q \leq 0,$$

where the last inequality follows from $q > 2$ by equation (C.20).

Equation (C.19)**.** By $q > 2$, we have

$$Q(-4q-2) - 2q^4 + 6q^2 + 2q \leq Q(-4q-2) \leq 0.$$

(iii) Since $g^{-1}(S)$ exists for all $S \in [0, 1/3]$, equation (B.37) implies that $H(S)$ exists for all $S \in [0, 1/3]$. We have $H(S) = L(g^{-1}(S))$, where

$$L(q) := \log\left(\frac{1+q}{1+Q}\right) + \frac{Q-q}{1+q} - \frac{(Q-q)^2}{(1+q)^2}\frac{g(q)}{1+Q}.$$

To establish that $H(S)$ is strictly increasing for $S \in [0, 1/3]$, it suffices to show that $L(q)$ is strictly decreasing for $q \in$. Note that $g^{-1}(S) \in$ is strictly decreasing in S by part (i). By definition of $g(\cdot)$ in equation (B.27), we derive

$$L'(q) = \frac{(Q-q)N(q)}{(1+q)^2(1+Q)(-q+2Q+1)^2},$$

where we define

$$N(q) := -2Q^3 + (2q^2 + 8q)Q^2 + \left(-6q^3 - 15q^2 - 4q - 1\right)Q + 2q^4 + 2q^3 - 5q^2 - 4q - 1.$$

We next show that $N(q) \leq 0$ for all $q \in$, which shows the monotonicity of $L(\cdot)$ using $Q - q > 0$. It suffices to establish the following:

$$-2Q^3 + \left(2q^2 + 8q\right)Q^2 + \left(-4q^3 - 8q^2\right)Q \leq 0, \tag{C.22}$$

$$\left(-2q^3 - 7q^2 - 4q - 1\right)Q + 2q^4 + 2q^3 \leq 0, \tag{C.23}$$

$$-5q^2 - 4q - 1 < 0. \tag{C.24}$$

Note that summing these three inequalities implies the desired result. We show each separately:

Equation (C.22)**.** Dividing both sides by $2Q > 0$, it suffices to show that $T(Q) \leq 0$, where

$$T(Q) := -Q^2 + (q^2 + 4q)Q + \left(-2q^3 - 4q^2\right).$$

First, we derive $T\left(q^2 + 2q\right) = 0$. Moreover, by equation (C.21), we have $q^2 + 2q \leq Q$. Combining this with $T'\left(q^2 + 2q\right) = -q^2 \leq 0$, we obtain $T(Q) \leq T\left(q^2 + 2q\right) = 0$.

Equation (C.23)**.** Since $Q > q$, we have

$$\left(-2q^3 - 7q^2 - 4q - 1\right)Q + 2q^4 + 2q^3 \leq -5q^3 - 4q^2 - q \leq 0.$$

Equation (C.24)**.** The result directly follows from $q \geq 0$. □

Submodularity Under Full Allocations and No Payments

As established in Theorem 3(i), the optimal data-sharing mechanism induces a supermodular game. Here, we show that if the mechanism allocates fully to agents with no payments, the game becomes submodular. This confirms the intuition that data collaboration games are submodular in the absence of payment incentives. We show this result in the general setting described in Section 4.

Proposition C.1. *Under the full allocation mechanism with no payments, the mean estimation task with effort is a submodular game. Additionally, each agent's best response is decreasing in the other agents' effort.*

PROOF OF PROPOSITION C.1: To prove submodularity, since the cost function $c_i(e_i)$ for agent i solely depends on her effort e_i, it suffices to show that the expected utility from participating in the mechanism is submodular in $(e_i, \mathbf{e}_{-i})$. Under full allocation, it suffices to show that the expected value of agent i is submodular in (e_i, e_j) for all $j \neq i$.

We first show that the value function $v_i(\mathbf{q})$ has decreasing differences in (q_i, q_j). Let c_{ii} denote the i-th diagonal entry of the inverse matrix $\Sigma_{\mathbf{q}}^{-1}$. Using the same argument as in equation (C.11), we have

$$\frac{\partial v_i(\mathbf{q})}{\partial q_i} = \frac{1}{q_i^2}\left(-c_{ii}\frac{\sigma_i^2}{q_i} + 1\right)^2. \tag{C.25}$$

Thus,

$$\frac{\partial^2 v_i(\mathbf{q})}{\partial q_i \partial q_j} = \frac{-2\sigma_i^2}{q_i^3} \cdot \frac{\partial c_{ii}}{\partial q_j}\left(-c_{ii}\frac{\sigma_i^2}{q_i} + 1\right) \leq 0, \tag{C.26}$$

where the inequality follows from $\frac{\partial c_{ii}}{\partial q_j} \geq 0$ and $-c_{ii}\frac{\sigma_i^2}{q_i} + 1 \geq 0$, as we establish below.

First, by equation (A.1), we derive

$$v_i(\mathbf{q}) = 1 - \frac{1}{q_i}\left(-c_{ii}\frac{\sigma_i^2}{q_i} + 1\right).$$

By definition, $v_i(\mathbf{q})$ is the relative decrease in MSE compared to the variance. Thus, $v_i(\mathbf{q}) \in [0,1]$ for all $\mathbf{q}$. Combining this with equation (C), we conclude that $\left(-c_{ii}\frac{\sigma_i^2}{q_i} + 1\right) \geq 0$. Moreover, using equation (B.2), we derive

$$\frac{\partial \Sigma_{\mathbf{q}}^{-1}}{\partial q_j} = \frac{\sigma_j^2}{q_j^2}(c_{j1}, c_{j2}, \ldots, c_{jn})^\top (c_{j1}, c_{j2}, \ldots, c_{jn}).$$

Therefore,

$$\frac{\partial c_{ii}}{\partial q_j} = \frac{\sigma_j^2}{q_j^2} c_{ji}^2 \geq 0.$$

As a result, $v_i(\mathbf{q})$ has decreasing differences in (q_i, q_j) for all $j \neq i$.

We next show that $\mathbb{E}_{\mathbf{q}\sim F_{\mathbf{e}}}[v_i(\mathbf{q})]$ has decreasing differences in (e_1, e_j). Conditional on q_l for $l \neq i, j$, the expected utility of agent i is

$$\begin{aligned}
&\mathbb{E}_{q_i \sim F_{i,e_i}, q_j \sim F_{j,e_j}}\left[v_i(\mathbf{q})\right] = \int v_i(q_i, \mathbf{q}_{-i}) f_{i,e_i}(q_i) f_{j,e_j}(q_j)\, dq_i\, dq_j \\
&= v_i(0,0,\mathbf{q}_{-i,j}) + \int \frac{\partial v_i(0, q_j, \mathbf{q}_{-i,j})}{\partial q_j} \bar{F}_{j,e_j}(q_j)\, dq_j + \int \frac{\partial v_i(q_i, 0, \mathbf{q}_{-i,j})}{\partial q_i} \bar{F}_{i,e_i}(q_i)\, dq_i + \\
&\quad + \int \frac{\partial^2 v_i(q_i, q_j, \mathbf{q}_{-i,j})}{\partial q_i \partial q_j} \bar{F}_{i,e_i}(q_i) \bar{F}_{j,e_j}(q_j)\, dq_i\, dq_j,
\end{aligned}$$

where we define the tail distributions $\bar{F}_{l,e_l}(\cdot) \coloneqq 1 - F_{l,e_l}(\cdot)$ for $l \in [n]$. Hence,

$$\frac{\partial^2}{\partial e_i \partial e_j}\left(\mathbb{E}_{q_i \sim F_{i,e_i}, q_j \sim F_{j,e_j}}\left[v_i(\mathbf{q})\right]\right) = \int \frac{\partial^2 v_i(q_i, q_j, \mathbf{q}_{-i,j})}{\partial q_i \partial q_j} \frac{\partial \bar{F}_{i,e_i}(q_i)}{\partial e_i} \frac{\partial \bar{F}_{j,e_j}(q_j)}{\partial e_j}\, dq_i\, dq_j \leq 0,$$

where the last inequality follows from equation (C.26) and the fact that F_{l,e_l} stochastically dominates $F_{l,e_l'}$ for $l \in [n]$ and $e_l \geq e_l'$. Taking the expectation over $\mathbf{q}_{-i,j}$ implies that the game is submodular. Thus, by Topkis' Monotonicity Theorem, each agent's best response is non-increasing in the effort levels of other agents. $\square$