Online Supplement: "Pay With Your Data: Optimal Data-Sharing Mechanisms for AI Services"

Sameer Mehta* Chandrasekhar Manchiraju† Milind Dawande‡ Ganesh Janakiraman‡
*Rotterdam School of Management, Erasmus University, mehta@rsm.nl
†Eli Broad College of Business, Michigan State University, manchiraju@broad.msu.edu
‡Naveen Jindal School of Management, The University of Texas at Dallas, {milind, ganesh}@utdallas.edu

Note: Due to the M&SOM 16-page limit for the online supplement, some extensions on a) hybrid mechanism, b) diminishing improvement in quality, and c) procuring data from external sources, and the proofs of some technical results are presented in a separate file. This separate file is available at https://dx.doi.org/10.2139/ssrn.4552550

Appendix A Proof of Theorem 1 (Optimal manual data-sharing mechanism)

From (2), under mechanism μ, the utility to a consumer of type c when she reports her type as $\hat{c}$ is:

$$U^{\mu}_{\mathrm{M}}(\hat{c};c) = 1 - c\cdot k\, q^{\mu}_{\mathrm{M}}(\hat{c}) - p^{\mu}_{\mathrm{M}}(\hat{c})$$

Let $U^{\mu}_{\mathrm{M}}(c) := U^{\mu}_{\mathrm{M}}(c;c)$ denote the equilibrium utility to the consumer of type c under an incentive compatible mechanism μ. From the incentive-compatibility constraints (IC-M), we have

$$U^{\mu}_{\mathrm{M}}(c) = \max_{\hat{c}\in[0,\bar{c}]} 1 - c\cdot k\, q^{\mu}_{\mathrm{M}}(\hat{c}) - p^{\mu}_{\mathrm{M}}(\hat{c}). \tag{32}$$

Note that $U^{\mu}_{\mathrm{M}}(c)$ is a maximum of a family of affine functions. Therefore $U^{\mu}_{\mathrm{M}}(\cdot)$ is a convex function of its argument. Also, for any $\hat{c}, c \in [0,\bar{c}]$, we have

$$\begin{aligned}1 - \hat{c}\cdot k\, q^{\mu}_{\mathrm{M}}(c) - p^{\mu}_{\mathrm{M}}(c) &= \left(1 - c\cdot k\, q^{\mu}_{\mathrm{M}}(c) - p^{\mu}_{\mathrm{M}}(c)\right) + k\, q^{\mu}_{\mathrm{M}}(c)\cdot(c-\hat{c})\\ &= U^{\mu}_{\mathrm{M}}(c) + (-k\, q^{\mu}_{\mathrm{M}}(c))\cdot(\hat{c}-c).\end{aligned}$$

Thus, the incentive-compatibility constraints are equivalent to the following condition:

$$U^{\mu}_{\mathrm{M}}(\hat{c}) \ge U^{\mu}_{\mathrm{M}}(c) + (-k\, q^{\mu}_{\mathrm{M}}(c))\cdot(\hat{c}-c)\ \forall \hat{c}, c\in[0,\bar{c}]. \tag{33}$$

Since $U^{\mu}_{\mathrm{M}}(\cdot)$ is a convex function, it is absolutely continuous and hence, is differentiable almost everywhere in the interior of its domain. Thus, we have $\frac{dU^{\mu}_{\mathrm{M}}(c)}{dc} = -k\, q^{\mu}_{\mathrm{M}}(c)$. Further, since $U^{\mu}_{\mathrm{M}}(\cdot)$ is convex, $q^{\mu}_{\mathrm{M}}(\cdot)$ is a non-increasing function. Moreover, since every absolutely continuous function is the definite integral of its derivative, we have

$$U^{\mu}_{\mathrm{M}}(\bar{c}) = U^{\mu}_{\mathrm{M}}(c) - k\int_{c}^{\bar{c}} q^{\mu}_{\mathrm{M}}(x)\ dx. \tag{34}$$

The above equation implies that $U^{\mu}_{\mathrm{M}}(c)$ is decreasing in c. Also, from the individual rationality (IR-M) constraints, we have $U^{\mu}_{\mathrm{M}}(\bar{c}) \ge 0$. Using contradiction, it is straightforward to see that in an optimal mechanism,

$$U^{\mu}_{\mathrm{M}}(\bar{c}) = 0. \tag{35}$$

Combining (34) and (35) yields $U^{\mu}_{\mathrm{M}}(c) = k\int_{c}^{\bar{c}} q^{\mu}_{\mathrm{M}}(x)\ dx\ \forall c\in[0,\bar{c}]$. Then, using the definition of $U^{\mu}_{\mathrm{M}}(c)$, we have

$$p^{\mu}_{\mathrm{M}}(c) = 1 - c\cdot k\, q^{\mu}_{\mathrm{M}}(c) - k\int_{c}^{\bar{c}} q^{\mu}_{\mathrm{M}}(x)\ dx\ \ \forall c\in[0,\bar{c}]. \tag{36}$$

The problem $\mathcal{P}_{\text{M}}$ of the firm is to maximize Π^{μ}_{M} in (3) subject to the (IC-M) and (IR-M) constraints. To this end, we first maximize Π^{μ}_{M} by dropping the (IC-M) and (IR-M) constraints and later show that the optimal solution does satisfy these constraints. Using (36) in (3) and from the assumption that the distribution of cost type c is $\mathrm{U}[0, \bar{c}]$, the objective of the firm can be simplified to

$$\eta + 1 + \frac{1}{\bar{c}} \int_0^{\bar{c}} (\eta\,\tau - k\,c) q^{\mu}_{\text{M}}(c)\ dc - \frac{k}{\bar{c}} \int_0^{\bar{c}} \int_c^{\bar{c}} q^{\mu}_{\text{M}}(x)\ dx\, dc. \tag{37}$$

Interchanging the order of integration in the last term, we obtain

$$\int_0^{\bar{c}} \int_c^{\bar{c}} q^{\mu}_{\text{M}}(x)\ dx\, dc = \int_0^{\bar{c}} \int_0^{x} q^{\mu}_{\text{M}}(x)\ dc\, dx = \int_0^{\bar{c}} x\, q^{\mu}_{\text{M}}(x)\ dx. \tag{38}$$

Finally, using (37) and (38), and dropping the constraints, the firm's objective is

$$\max_{\mu} \frac{1}{\bar{c}} \int_0^{\bar{c}} (\eta\,\tau - 2k\,c) q^{\mu}_{\text{M}}(c)\ dc.$$

Let MDS denote an optimal mechanism for the above problem. The optimization problem can be solved through point-wise maximization. The optimal data-sharing function can then be obtained as follows:

Case 1: If $\bar{c} > \frac{\eta\tau}{2k}$, then

$$q^{\text{MDS}}_{\text{M}}(c) = \begin{cases} 1, & 0 \le c \le \dfrac{\eta\tau}{2k}, \\ 0, & \dfrac{\eta\tau}{2k} < c \le \bar{c}, \end{cases}$$

Case 2: If $0 < \bar{c} \le \frac{\eta\tau}{2k}$, then

$$q^{\text{MDS}}_{\text{M}}(c) = 1,\ \forall\, 0 \le c \le \bar{c}.$$

The optimal pricing function $p^{\text{MDS}}_{\text{M}}(c)$ can be obtained by substituting $q^{\text{MDS}}_{\text{M}}(c)$ in (36). Using (33) and (34) and the fact that $q^{\text{MDS}}_{\text{M}}(c)$ is a non-increasing function, it is straightforward to verify that MDS is an incentive compatible mechanism. Furthermore, the MDS mechanism also satisfies $U^{\text{MDS}}_{\text{M}}(c) \ge 0$ for $c \in [0, \bar{c}]$ and is therefore individually rational for the consumers. Thus, the MDS mechanism satisfies the (IC-M) and (IR-M) constraints and hence, is optimal for $\mathcal{P}_{\text{M}}$. ■

Appendix B Proof of Theorem 2 (Optimal algorithmic data-sharing mechanism)

From the expression for the revenue (Π_{A}) in Problem $\mathcal{P}_{\text{A}}$, we see that any value of p_{A} such that $p_{\text{A}} \le 1 - \bar{c}k\beta$ and $p_{\text{A}} \ge 1$ is suboptimal. When $1 - \bar{c}k\beta \le p_{\text{A}} \le 1$, the revenue under the algorithmic data-sharing mechanism Π_{A} is a concave function of p_{A}. The interior optimum of this concave function is $\frac{1-\eta\tau\gamma}{2}$. The result follows. ■

Appendix C Proof of Proposition 1

Using Theorem 2, the optimal revenue of the ADS mechanism (Π^{ADS}) can be written as follows:

$$\Pi^{\text{ADS}} = \begin{cases} \eta + 1 + \eta\tau\gamma - \bar{c}k\beta, & 0 < \bar{c} \le \dfrac{1+\eta\tau\gamma}{2k\beta}, \\ \eta + \dfrac{(1+\eta\tau\gamma)^2}{4\bar{c}k\beta}, & \bar{c} \ge \dfrac{1+\eta\tau\gamma}{2k\beta}. \end{cases} \tag{39}$$

Using the above expression, we show how Π^{ADS} changes with false-negative and false-positive error rates.

(i) **False-positive error rate:** From (39), it is easy to verify that Π^{ADS} decreases with an increase in α.

(ii) **False-negative error rate:** We first consider the case where $\bar{c} \leq \frac{\eta\tau}{2}$. In this case, the optimal revenue is

$$\Pi^{\text{ADS}} = \eta + 1 + \eta\tau(1-\alpha)(1-k) + (\eta\tau - \bar{c}) \cdot k \cdot \beta. \tag{40}$$

Thus, Π^{ADS} is increasing in β.

Now consider the case where $\bar{c} \geq \frac{\eta\tau}{2}$. Let $\tilde{\beta} = \frac{1+\eta\tau(1-\alpha)(1-k)}{k\cdot(2\bar{c}-\eta\tau)}$. In this case, the optimal revenue can be written as follows:

$$\Pi^{\text{ADS}} = \begin{cases} \eta + 1 + \eta\tau(1-\alpha)(1-k) + (\eta\tau - \bar{c}) \cdot k \cdot \beta, & 0 \leq \beta \leq \tilde{\beta}, \\ \eta + \dfrac{1}{4\bar{c}k} \cdot \dfrac{(1+\eta\tau\gamma)^2}{\beta}, & \tilde{\beta} < \beta \leq 1, \end{cases} \tag{41}$$

Let us focus on the subcase where $\frac{\eta\tau}{2} \leq \bar{c} \leq \eta\tau$. In this subcase, if $0 \leq \beta \leq \tilde{\beta}$, from (41), we have $\frac{\partial \Pi^{\text{ADS}}}{\partial \beta} > 0$. In the same subcase where $\frac{\eta\tau}{2} \leq \bar{c} \leq \eta\tau$, consider the region where $\tilde{\beta} \leq \beta \leq 1$. Let us define $\beta' = \frac{1+\eta\tau(1-\alpha)(1-k)}{k\cdot\eta\tau}$. Here, since $\frac{\eta\tau}{2} \leq \bar{c} \leq \eta\tau$, we have $\beta' \leq \tilde{\beta}$. Further, it can be verified from (41) that $\frac{\partial \Pi_{\text{A}}}{\partial \beta} \geq 0$ if $\beta \geq \beta'$. Thus, if $\tilde{\beta} \leq \beta \leq 1$, then Π^{ADS} is increasing in β. Therefore, if $\frac{\eta\tau}{2} \leq \bar{c} \leq \eta\tau$, then Π^{ADS} is increasing in β.

We now consider the other subcase where $\bar{c} \geq \eta\tau$. Here, if $0 \leq \beta \leq \tilde{\beta}$, from (41), we have $\frac{\partial \Pi^{\text{ADS}}}{\partial \beta} \leq 0$. Since $\bar{c} \geq \eta\tau$, we also have $\beta' \geq \tilde{\beta}$. Further, from (41), it can be verified that $\frac{\partial \Pi^{\text{ADS}}}{\partial \beta} \leq 0$ if $\tilde{\beta} \leq \beta \leq \beta'$ and $\frac{\partial \Pi^{\text{ADS}}}{\partial \beta} \geq 0$ if $\beta' \leq \beta \leq 1$. Therefore, for the subcase where $\bar{c} \leq \eta\tau$, we have $\frac{\partial \Pi^{\text{ADS}}}{\partial \beta} \leq 0$ for $0 \leq \beta \leq \beta'$ and $\frac{\partial \Pi^{\text{ADS}}}{\partial \beta} \geq 0$ for $\beta' \leq \beta \leq 1$. ■

Appendix D Proof of Proposition 2

(i) **Consumer surplus under MDS mechanism:** From Theorem 1, if $\bar{c} \leq \frac{\eta\tau}{2k}$, then $q^{\mu}_{\text{M}}(c) = 1$ and $p^{\text{MDS}}_{\text{M}}(c) = 1 - \bar{c}k$ for $c \in [0, \bar{c}]$. Then, the consumer surplus under the MDS mechanism can be written as

$$\text{CS}^{\text{MDS}} = \int_{c=0}^{\bar{c}} \frac{1}{\bar{c}} \cdot (1 - c \cdot k \cdot 1 - (1 - \bar{c} \cdot k)) \; dc \; = \; \frac{\bar{c}k}{2}.$$

Thus, CS^{MDS} is increasing in $\bar{c}$ for the case where $\bar{c} \leq \frac{\eta\tau}{2k}$. We now consider the case where $\bar{c} \geq \frac{\eta\tau}{2k}$. From Theorem 1, using the optimal quantity of data shared and the price of the service, the consumer surplus is

$$\text{CS}^{\text{MDS}} = \int_{c=0}^{\frac{\eta\tau}{2k}} \frac{1}{\bar{c}} \cdot \left(1 - c \cdot k \cdot 1 - \left(1 - \frac{\eta\tau}{2}\right)\right) \; dc \; + \int_{c=\frac{\eta\tau}{2k}}^{\bar{c}} \frac{1}{\bar{c}} \cdot (1 - c \cdot k \cdot 0 - 1) \; dc \; = \; \frac{\eta^2\tau^2}{8\bar{c}k}.$$

Therefore, CS^{MDS} is decreasing in $\bar{c}$ for the case where $\bar{c} \geq \frac{\eta\tau}{2k}$.

(ii) **Consumer surplus under ADS mechanism:** From Theorem 2, if $\bar{c} \leq \frac{1+\eta\tau\gamma}{2k\beta}$, the optimal price of the service is $p^{\text{ADS}}_{\text{A}} = 1 - \bar{c}k\beta$. Recall that under ADS mechanism, $\beta \cdot k$ amount of the sensitive data of a customer gets shared with the firm. Thus, the consumer surplus under the ADS mechanism is

$$\text{CS}^{\text{ADS}} = \int_{c=0}^{\bar{c}} \frac{1}{\bar{c}} \cdot (1 - ck\beta - (1 - \bar{c}k\beta)) \; dc \; = \; \frac{\bar{c}k\beta}{2}.$$

Clearly, CS^{ADS} is increasing in $\bar{c}$ under the case where $\bar{c} \leq \frac{1+\eta\tau\gamma}{2k\beta}$. Let us consider the case where $\bar{c} \geq \frac{1+\eta\tau\gamma}{2k\beta}$. In this case, the optimal price of the service is $p^{\text{ADS}}_{\text{A}} = \frac{1-\eta\tau\gamma}{2}$. Then, the consumer surplus is

$$\text{CS}^{\text{ADS}} = \int_{c=0}^{\frac{1+\eta\tau\gamma}{2k\beta}} \frac{1}{\bar{c}} \cdot \left(1 - ck\beta - \left(\frac{1-\eta\tau\gamma}{2}\right)\right) \; dc \; = \; \frac{(1+\eta\tau\gamma)^2}{8\bar{c}k\beta}.$$

Therefore, in this case, CS^{ADS} is decreasing in $\bar{c}$. ■

Appendix E Proof of Theorem 3

The objective in (21) is $\sum_{t=1}^{T} N_t\big(Q_t^{\text{M}} - \theta_t^{\text{M}} T\, d_t^{*\text{M}}\big) + \eta\, N_T\, Q_{T+1}^{\text{M}}$. We know that $Q_1^{\text{M}} = 1$ and $Q_t^{\text{M}} = \Big(1 + \tau \cdot \sum_{s=1}^{t-1} d_s^{*\text{M}}\Big)$ for $t \geq 2$. Thus, the objective is now:

$$\sum_{t=1}^{T} N_t \cdot \left(1 + \tau \cdot \sum_{s=1}^{t-1} d_s^{*\text{M}}\right) - \sum_{t=1}^{T} \theta_t^{\text{M}} \cdot N_t \cdot T d_t^{*\text{M}} + \eta N_T \cdot \left(1 + \tau \cdot \sum_{s=1}^{T} d_s^{*\text{M}}\right)$$

Let $\pi^{\text{M}}(\boldsymbol{\theta}^{\text{M}})$ be the revenue of the firm as function of the discount vector $\boldsymbol{\theta}^{\text{M}}$. The optimal amount of data shared in any period $1 \leq s \leq T$ is a function of θ_s^{M} and λ_s. For any $1 \leq s \leq T$, let $d_s^{*\text{M}}(\theta_s^{\text{M}}; \lambda_s)$ represent the optimal data shared function. Then, the optimization problem of the firm can be simplified as follows:

$$\begin{aligned}
\max_{\boldsymbol{\theta}^{\text{M}}} \pi^{\text{M}}(\boldsymbol{\theta}^{\text{M}}) = & -\theta_1^{\text{M}} \cdot N_1 \cdot T \cdot d_1^*(\theta_1^{\text{M}}; \lambda_1) + \tau \left(\sum_{t=2}^{T} N_t + \eta N_T\right) \cdot d_1^*(\theta_1^{\text{M}}; \lambda_1) \\
& -\theta_2^{\text{M}} \cdot N_2 \cdot T \cdot d_2^*(\theta_2^{\text{M}}; \lambda_2) + \tau \left(\sum_{t=3}^{T} N_t + \eta N_T\right) \cdot d_2^*(\theta_2^{\text{M}}; \lambda_2) + \ldots. + \\
& -\theta_{T-1}^{\text{M}} \cdot N_{T-1} \cdot T \cdot d_{T-1}^*(\theta_{T-1}^{\text{M}}; \lambda_{T-1}) + \tau(N_T + \eta N_T) \cdot d_{T-1}^*(\theta_{T-1}^{\text{M}}; \lambda_{T-1}) \\
& -\theta_T^{\text{M}} \cdot N_T \cdot T \cdot d_T^*(\theta_T^{\text{M}}; \lambda_T) + \tau\eta N_T \cdot d_T^*(\theta_T^{\text{M}}; \lambda_T).
\end{aligned}$$

Observe that this problem decouples over periods. Therefore, for any period $1 \leq t \leq T$, the problem of the firm can be written as follows:

$$\max_{\theta_t^{\text{M}}} d_t^*(\theta_t^{\text{M}}; \lambda_t) \cdot \left(\tau \cdot \left(\sum_{s=t+1}^{T} N_s + \eta N_T\right) - \theta_t^{\text{M}} \cdot N_t \cdot T\right) \tag{42}$$

From (19), it can be observed that $d_t^*(\theta_t^{\text{M}}; \lambda_t)$ is concave and increasing in $(\theta_t^{\text{M}} + \lambda_t)$ for $0 \leq \theta_t^{\text{M}} + \lambda_t \leq 2\bar{c}$ and is constant after $\theta_t^{\text{M}} + \lambda_t \geq 2\bar{c}$. We know that λ_t is increasing in t from (14). Therefore, as t increases, the value of θ_t^{M} that maximizes $d_t^*(\theta_t^{\text{M}}; \lambda_t)$ decreases. Moreover, $\tau \cdot \big(\sum_{s=t+1}^{T} N_s + \eta N_T\big)$ is decreasing in t and $\theta_t^{\text{M}} \cdot N_t \cdot T$ is increasing in t. The result follows. ∎

Appendix F Proof of Proposition 3

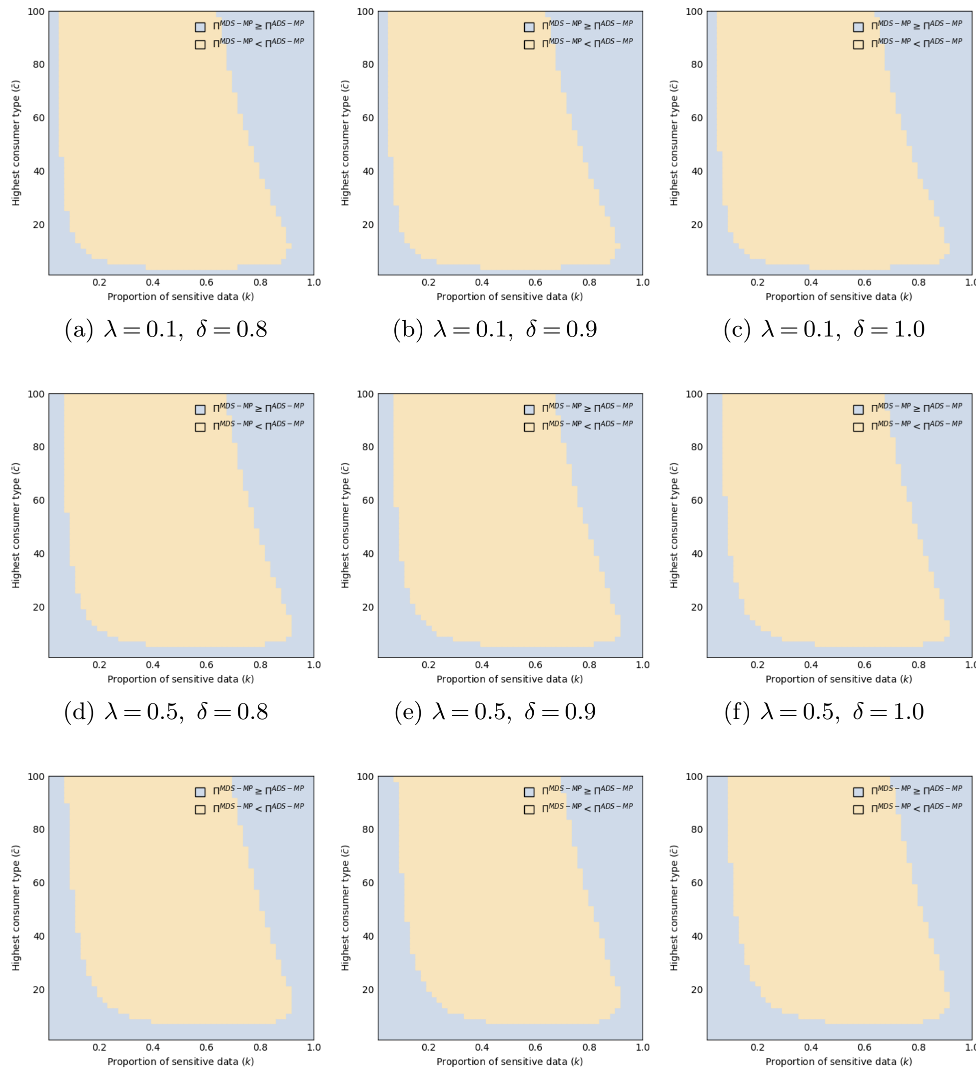
From (42) in the proof of Theorem 3, the problem of the firm in any period $1 \leq t \leq T$ is

$$\max_{\theta_t^{\text{M}}} d_t^*(\theta_t^{\text{M}}; \lambda_t) \cdot \left(\tau \cdot \left(\sum_{s=t+1}^{T} N_s + \eta N_T\right) - \theta_t^{\text{M}} \cdot N_t \cdot T\right).$$

We know from (19) that the function $d_t^*(\theta_t^{\text{M}}; \lambda_t)$ is concave and increasing in $(\theta_t^{\text{M}} + \lambda_t)$. Let $\hat{\theta}_t^{\text{M}}$ be the value of θ_t^{M} at which the function $d_t^*(\theta_t^{\text{M}}; \lambda_t)$ reaches its maximum. From (14), it is easy to see that for any $1 \leq t \leq T$, the quantity λ_t is increasing in λ and δ. Therefore, with an increase in λ or δ, the function $d_t^*(\theta_t^{\text{M}}; \lambda_t)$ reaches is its maximum at some $\theta_t^{\text{M}} \leq \hat{\theta}_t^{\text{M}}$. Moreover, the quantity $\big(\tau \cdot \big(\sum_{s=t+1}^{T} N_s + \eta N_T\big) - \theta_t^{\text{M}} \cdot N_t \cdot T\big)$ is decreasing in θ_t^{M}. The result follows. ∎

Appendix G Numerical Analysis of Mechanisms under Multi-Period Setup

Figure 9 shows that the revenue comparison between the optimal MDS-MP and ADS-MP mechanisms is qualitatively similar to that in the baseline model. Across all values of λ and δ, the ADS-MP mechanism dominates in a broad intermediate region of the $(k, \bar{c})$ space, whereas the MDS-MP mechanism performs better when $\bar{c} \cdot k$ is either sufficiently low or sufficiently high. The variation in λ and δ mainly shifts the boundary between the two mechanisms at the margin rather than altering the shape of the regions where one mechanism dominates the other.



(a) $\lambda = 0.1,\ \delta = 0.8$ (b) $\lambda = 0.1,\ \delta = 0.9$ (c) $\lambda = 0.1,\ \delta = 1.0$

(d) $\lambda = 0.5,\ \delta = 0.8$ (e) $\lambda = 0.5,\ \delta = 0.9$ (f) $\lambda = 0.5,\ \delta = 1.0$

(g) $\lambda = 0.9,\ \delta = 0.8$ (h) $\lambda = 0.9,\ \delta = 0.9$ (i) $\lambda = 0.9,\ \delta = 1.0$

Figure 9 **Comparison of revenue under the optimal manual data-sharing mechanism and the optimal algorithmic data-sharing mechanism across different values of λ and δ. Parameter values: $\tau = 0.5$, $\eta = 4$, $\beta = 0.1$, $\alpha = 0.1$.**

Appendix H Endogenous Market Size in the Operational Phase

Recall that in our base model, if N is the market size (number of consumers) in the development phase, then the market size in the operational phase is $\eta \cdot N$. Consequently, the revenue obtained by the firm in the operational phase is $\eta \cdot N \cdot (1 + \tau \cdot d)$, where d is the expectation (over the consumer types) of the average amount of data collected in the development phase and τ is the effectiveness in using the data collected. We now endogenize the market size in the operational phase as follows: Given the market size of N in the development phase, the market size in the operational phase is $\big(\eta(1+\phi \cdot d)\big) \cdot N$, where $\phi \geq 0$. Here, η captures the increase in the market size due to the relatively long duration of the operational phase and $(1 + \phi \cdot d)$ captures the increase in the market size resulting from the improvement in the quality of the AI service.

Below, we characterize an optimal manual data-sharing mechanism under this setting.

THEOREM 4. *The optimal mechanism, which we refer to as* MDS-END*, for the setting where the market size in the operational phase is endogenous on the amount of data collected in the development phase, is as follows:*

- *If the highest type $\bar{c}$ is sufficiently low, i.e., $0 \le \bar{c} \le \frac{\eta(\phi+\tau+\phi\tau)}{2k}$, then*

$$(q_{\mathrm{H}}^{\text{MDS-END}}(c), p_{\mathrm{H}}^{\text{MDS-END}}(c)) = (1, 1-\bar{c}k) \;\; \forall c \in [0, \bar{c}].$$

- *If the highest type $\bar{c}$ is sufficiently high, i.e., $\bar{c} > \frac{\eta(\phi+\tau+\phi\tau)}{2k}$, then*

$$(q_{\mathrm{H}}^{\text{MDS-END}}(c), p_{\mathrm{H}}^{\text{MDS-END}}(c)) = \begin{cases} \left(1, \dfrac{\eta(\phi+\tau+\phi\tau)}{2}\right), & 0 \le c \le \dfrac{\eta(\phi+\tau+\phi\tau)}{2k}, \\ (0,1), & \dfrac{\eta(\phi+\tau+\phi\tau)}{2k} < c \le \bar{c}. \end{cases}$$

The proof of this result is similar to the proof of Theorem 1 and is omitted for brevity.

We note that the structure of the MDS-END mechanism is similar to that of the MDS mechanism (see Theorem 1) with a subtle difference. In the MDS-END mechanism, a larger proportion of the consumers share their data with the firm compared to the MDS mechanism. Moreover, under the MDS-END mechanism, the consumers who share their data with the firm avail the service at a lower price compared to that under the MDS mechanism. This is because in the MDS-END mechanism, due to the endogenous market size in the operational phase, the amount of data collected in the development phase has a greater contribution to the total revenue compared to that in the MDS mechanism.

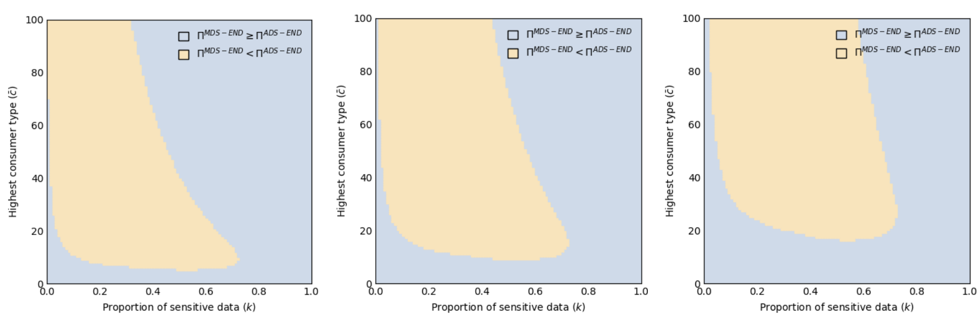


Figure 10 Comparison of the revenues under optimal manual and algorithmic data-sharing mechanisms for endogenous market size in the operational phase. Parameter values: $\tau = 0.5$, $\eta = 4$, $\beta = 0.1$, $\alpha = 0.1$. **Left panel:** $\phi = 0.5$, **middle panel:** $\phi = 1$, **right panel:** $\phi = 2$.

Next, we compare the firm's revenue under the manual and algorithmic data-sharing mechanisms. Figure 10 shows the revenue comparison across various values for the parameter ϕ. We observe that when the parameter ϕ is low (left panel), the parametric regions where the manual and algorithmic data-sharing mechanisms outperform each other are, respectively, similar to those in our base model (see Figure 5 in Section 6). As the parameter ϕ increases (middle and right panels), we observe two phenomena:

- When the consumer population is less privacy sensitive (i.e., low values of $\bar{c}$), the parametric region where the manual data-sharing mechanism outperforms the algorithm data-sharing mechanism expands. First, ceteris paribus, the manual data-sharing mechanism collects more data in the development phase compared to the algorithmic data-sharing mechanism below a certain threshold on $\bar{c}$. Next, as the value of ϕ increases, this threshold increases, which leads to MDS-END yielding a higher revenue over a larger parametric region.

- When the consumer population is more privacy sensitive (i.e., high values of $\bar{c}$), the parametric region where the algorithmic data-sharing mechanism outperforms the manual data-sharing mechanism expands. Here, note that ceteris paribus, the algorithmic data-sharing mechanism collects more data in the development phase compared to the manual data-sharing mechanism. However, above a certain threshold (right-side boundary of the yellow region) on k (proportion of sensitive data), the total revenue under the algorithmic data-sharing mechanism is lower than that under the manual data-sharing mechanism. The reason is as follows: As k increases, the amount of data collected under both mechanisms decreases. However, above that threshold on k, the amount of data collected by the algorithmic data-sharing mechanism does not sufficiently contribute to the revenue under the operational phase and consequently, the total revenue. Finally, as ϕ increases, the impact of the amount of data collected on the operational phase revenue increases because of which the above-mentioned threshold on k also increases. This leads to the algorithmic data-sharing mechanism yielding a higher revenue over a larger parametric region.

Appendix I Private Value and Private Privacy Cost

Thus far in our analysis, we assumed that all consumers obtain a unit value from consuming the service and they are heterogeneous only in the privacy cost incurred from sharing sensitive data. In this section, we relax this assumption and generalize our model by considering the case where consumers are heterogeneous in both value from the service as well as privacy cost. In particular, we assume that a consumer is characterized by the tuple (θ, c) of private information, where θ is the value she obtains from consuming the service and c is her privacy cost. We assume that $\theta \sim U(\underline{\theta}, \bar{\theta})$, and as before, $c \sim U(0, \bar{c})$, with both the types independently distributed.

Notice that endowing consumers with a private valuation for the service as well as private privacy costs results in a two-dimensional mechanism-design problem. It is well known that a complete characterization of such problems in full generality is intractable (see, e.g., Daskalakis et al. 2014). In view of this difficulty, we obtain a numerical solution to the mechanism-design problem by discretizing the type space. The main message from this section is that the key insights from the analysis of our base model continue to hold for this extension. Before presenting the results of our numerical analysis, we illustrate a few structural properties of an optimal two-dimensional manual data-sharing mechanism.

A two-dimensional manual data-sharing mechanism ω is characterized by the functions $(z^{\omega}_{2\text{M}}, q^{\omega}_{2\text{M}}, p^{\omega}_{2\text{M}})$, where $z^{\omega}_{2\text{M}} : \left[\underline{\theta}, \bar{\theta}\right] \times [0, \bar{c}] \to \{0, 1\}$ is an indicator function that denotes the purchase decision and $(q^{\omega}_{2\text{M}}, p^{\omega}_{2\text{M}})$ denote, respectively, the quantity of data shared and the price paid by the consumers. Under mechanism ω, the utility to a consumer of type (θ, c) when she reveals her type as $(\hat{\theta}, \hat{c})$ is given by:

$$U^{\omega}_{2\text{M}}(\hat{\theta}, \hat{c}; \theta, c) = \theta \cdot z^{\omega}_{2\text{M}}(\hat{\theta}, \hat{c}) - ck \cdot q^{\omega}_{2\text{M}}(\hat{\theta}, \hat{c}) - p^{\omega}_{2\text{M}}(\hat{\theta}, \hat{c}). \tag{43}$$

THEOREM 5. *An optimal solution to the two-dimensional manual data-sharing mechanism has the following structure:*

(i) **(Threshold-based participation):** *A consumer purchases the service only if the value she receives from consuming it exceeds a threshold that is dependent on her privacy cost. That is, for a consumer of type (θ, c), there exists a threshold $\Theta^{\omega}(c) \in \left[\underline{\theta}, \bar{\theta}\right]$, such that $z^{\omega}_{2\text{M}}(\theta, c) = 1$ if and only if $\theta > \Theta^{\omega}(c)$.*

(ii) **(Structure of threshold):** *The threshold $\Theta^\omega(c)$ is monotonically non-decreasing and concave in c.*

(iii) **(Structure of data shared and price paid):** *As long as the value a consumer receives from consuming the service exceeds the threshold specified in (i) above, the quantity of data shared and the price paid by that consumer depends only on her privacy cost and is independent of her value from the service. Thus, for any two consumers of type (θ_1, c) and (θ_2, c) with $\theta_1, \theta_2 \geq \Theta^\omega(c)$, we have $q^\omega_{2\mathrm{M}}(\theta_1, c) = q^\omega_{2\mathrm{M}}(\theta_2, c)$ and $p^\omega_{2\mathrm{M}}(\theta_1, c) = p^\omega_{2\mathrm{M}}(\theta_2, c)$.*

The proof of this result is provided in the supplementary appendix. In Figure 11, we illustrate the structure of an optimal solution to the two-dimensional manual data-sharing mechanism obtained by discretizing the type space. Not surprisingly, consumers with high privacy costs and low value for the service do not participate in the mechanism (yellow region in the left panel). The firm collects data only from consumers who incur less privacy costs from sharing their data (blue and red regions in the middle panel). We note here that consumers with intermediate value and cost types share partial data with the firm (red region). Finally, the right panel shows that consumers who share more data with the firm pay a lower price for the service.

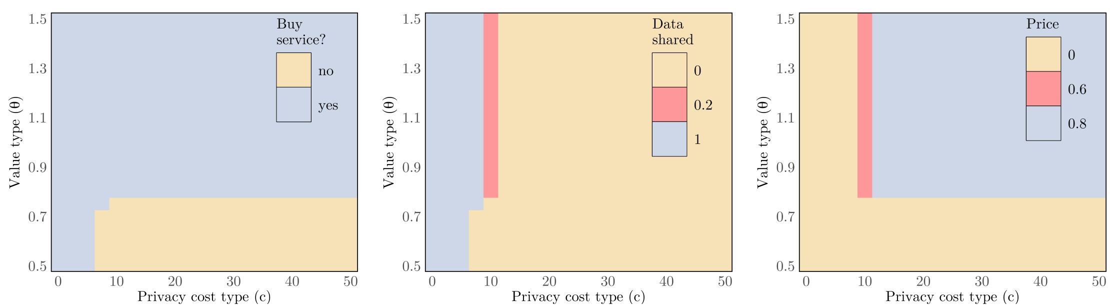


Figure 11 **Structure of optimal two-dimensional manual data sharing mechanism. Parameter values:** $\eta = 4, \tau = 0.5, k = 0.1, \bar{c} = 50$. **Discretization level:** $\theta \in [0.5, 1.5]$ **in uniform steps of 0.05 and** $c \in [0, 50]$ **in uniform steps of 2.5.**

Recall from Section 6 that a key factor that impacts the performance of the manual and algorithmic data-sharing mechanisms is the amount of data collected by the mechanisms in the development phase. Figure 12 (left and middle panels) shows that the optimal two-dimensional manual and algorithmic mechanisms exhibit similar data-collection effects as their one-dimensional counterparts. The right panel of that figure also confirms that the parametric region where the manual data-sharing mechanism dominates the algorithmic data-sharing mechanism (and vice versa) is similar.

Appendix J Selective Data Sharing

In this section, we consider the setting where in the manual data-sharing mechanism, the consumers have the ability to select portions of their data to share with the firm. Recall from Section 3 that a fraction k $(\in [0,1])$ of the one unit of data owned by the consumers is sensitive data and the remaining fraction $(1-k)$ is non-sensitive. Our base model assumes that the sensitive segments in consumers' data are uniformly spread across the data. The rationale behind this assumption is that the sensitive data possessed by the consumers is typically scattered across a large number of files due to which they may not be able to selectively share the non-sensitive portions of their data with the firm. Thus, under the manual data-sharing mechanism, if a

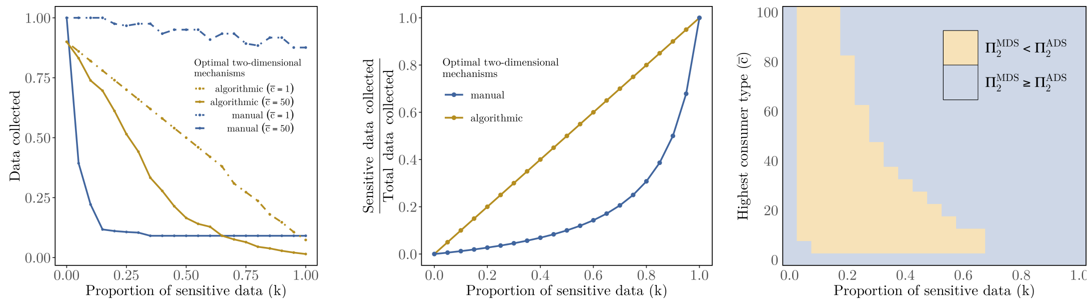


Figure 12 **Left panel: Comparison of revenue under optimal two-dimensional manual and algorithmic data-sharing mechanisms. Middle panel: Amount of data collected under the two mechanisms. Right panel: Ratio of sensitive data collected to total data collected under the two mechanisms. Parameter values:** $\tau = 0.5$**,** $\eta = 4$**,** $\beta = 0.1$**,** $\alpha = 0.1$**.**

consumer shares an amount q of their data, then the sensitive data that gets shared with the firm is simply $k \cdot q$.

Here, we consider the following generalization that endows consumers with 'cherry-picking' power to limit the amount of sensitive data they share. Let $\xi \in [0,1]$ denote the cherry-picking power of consumers.

- **No cherry-picking power**: When consumers have no cherry-picking power and they share q amount of data with the firm, then the amount of sensitive data that gets shared with the firm is kq. This situation corresponds to our base model and we represent this case by $\xi = 0$.
- **Perfect cherry-picking power**: In this case, when a consumer shares q amount of data with the firm, the sensitive data that gets shared with the firm is $\big(q-(1-k)\big)^+$. The rationale is as follows: when consumers can perfectly cherry-pick their data – (i) if the amount of data shared (q) by the consumer is lower than the total amount of her non-sensitive data $(1-k)$, she can completely avoid sharing any of her sensitive data with the firm, and (ii) if the amount of data shared (q) by the consumer is higher than the total amount of her non-sensitive data $(1-k)$, she cannot avoid sharing $(q-(1-k))$ amount of sensitive data with the firm. We represent this case by $\xi = 1$.
- **Intermediate cherry-picking power**: We model an intermediate level of cherry-picking power as a convex combination of the above two cases. In particular, when a consumer has cherry-picking power $\xi \in [0,1]$ and she shares q amount of data, then the amount of sensitive data shared with the firm is $(1-\xi)\cdot \underbrace{kq}_{\text{no cherry-picking}} + \xi \cdot \underbrace{\big(q-(1-k)\big)^+}_{\text{perfect cherry-picking}}$.

To characterize an optimal manual data-sharing mechanism in this setting, it is convenient to define two thresholds: $\hat{c}_1 := \frac{\eta\tau}{2[(1-\xi)k+\xi]}$ and $\hat{c}_2 := \frac{\eta\tau}{2(1-\xi)k}$.

Theorem 6. *The optimal mechanism, which we refer to as* MDS-SDS*, for the above-defined generalized model with consumers having the ability to share their data selectively, is as follows:*

(i) If the highest type $\bar{c}$ *is low, i.e.,* $0 \leq \bar{c} \leq \hat{c}_1$*, then*

$$(q_{\mathrm{M}}^{\text{MDS-SDS}}(c), p_{\mathrm{M}}^{\text{MDS-SDS}}(c)) = (1, 1-\bar{c}k) \;\; \forall c \in [0,\bar{c}].$$

(ii) If the highest type $\bar{c}$ is intermediate, i.e., $\hat{c}_1 < \bar{c} \leq \hat{c}_2$, then

$$(q_{\text{M}}^{\text{MDS-SDS}}(c), p_{\text{M}}^{\text{MDS-SDS}}(c)) = \begin{cases} \big(1, 1 - \bar{c}k(1-k)(1-\xi) - \hat{c}_1 k[1-(1-k)(1-\xi)]\big), & 0 \leq c \leq \hat{c}_1, \\ \big(1-k, 1 - \bar{c}k(1-k)(1-\xi)\big), & \hat{c}_1 < c \leq \bar{c}. \end{cases}$$

(iii) If the highest type $\bar{c}$ is high, i.e., $\bar{c} > \hat{c}_2$, then

$$(q_{\text{M}}^{\text{MDS-SDS}}(c), p_{\text{M}}^{\text{MDS-SDS}}(c)) = \begin{cases} \big(1, 1 - \frac{\eta\tau}{2}\big), & 0 \leq c \leq \hat{c}_1, \\ \big(1-k, 1 - \frac{\eta\tau}{2}(1-k)\big), & \hat{c}_1 < c \leq \hat{c}_2, \\ (0, 1), & \hat{c}_2 < c \leq \bar{c}. \end{cases}$$

The proof of this result is provided in the supplementary appendix. We now compare the revenue of the manual data-sharing mechanism with that of the algorithmic data-sharing mechanism when consumers can selectively share their data. Figure 13 shows the revenue comparison across various values for the cherry-picking parameter ξ (left panel: $\xi = 0.25$, middle panel: $\xi = 0.5$, right panel: $\xi = 0.75$). We observe that when the cherry-picking power is not very high (left and middle panel), the parametric regions where the algorithmic and the manual data-sharing mechanisms outperform each other, respectively, are similar to those in our base model (see Figure 5 in Section 6.1). As the cherry-picking power becomes sufficiently high (right panel), the parametric region where the manual data-sharing mechanism yields a better revenue relative to the algorithmic data-sharing mechanism expands. There are two key reasons behind this: (i) Note that a higher value of cherry-picking power (ξ) helps consumers avoid sharing sensitive data with the firm. In this sense, the cherry-picking power in the manual data-sharing mechanism corresponds to the true-positive rate $(1-\beta)$ of the redaction algorithm in the algorithmic data-sharing mechanism. Since $(1-\beta) = 0.9$ for the algorithmic data-sharing mechanism, when $\xi = 0.25$ (left panel) and $\xi = 0.5$ (middle panel), the manual data-sharing mechanism under the selective data-sharing case performs similar to our base case. (ii) When $\xi = 0.75$ (right panel), the cherry-picking power under the manual data-sharing mechanism is comparable to the true-positive rate ($(1-\beta) = 0.9$) of the algorithmic data-sharing mechanism. Furthermore, when $\bar{c} > \hat{c}_2$ (i.e., the consumer population is highly privacy sensitive), the manual data-sharing mechanism is offered at three price-points to the consumers (see Theorem 6 (iii)). Thus, it can better price-discriminate compared to the manual data-sharing mechanism without cherry-picking (i.e., our base case), consequently resulting in a higher revenue compared to the algorithmic data-sharing mechanism.

Appendix K Endogenous Data-Effectiveness

Recall that in our base model, the value of the service improves from 1 in the development phase to $(1 + \tau d)$ in the operational phase, where d is the average amount of data collected in the development phase and τ is the effectiveness in using the data collected. In this section, we endogenize the parameter τ using a two-stage model. In the first stage, the firm chooses τ by incurring an investment cost of $K(\tau)$, where $K(\cdot)$ is a convex increasing function. In the second stage, the firm designs its data-sharing policy (the focus of the base model).

Recall from Section 4 that in the MDS setting, the firm chooses a mechanism μ to maximize Π_{M}^{μ}, where Π_{M}^{μ} is as given in (3). Similarly, in the ADS setting the firm chooses the price p_{A} to maximize Π_{A}, where Π_{A} is

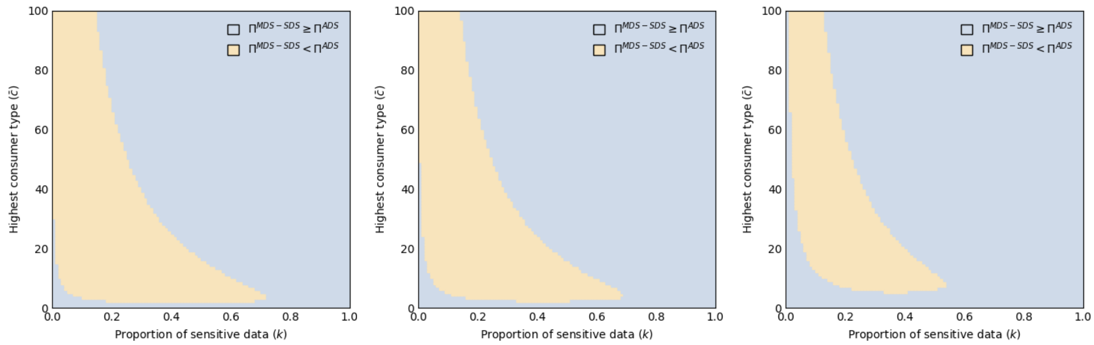


Figure 13 Comparison of revenue under optimal manual data-sharing mechanism with selective data sharing and the optimal algorithmic data-sharing mechanism. Parameter values: $\tau = 0.5$, $\eta = 4$, $\beta = 0.1$, $\alpha = 0.1$. **Left panel:** $\xi = 0.25$, **middle panel:** $\xi = 0.5$, **right panel:** $\xi = 0.75$.

defined in (4). In the two-stage model, these problems constitute the second-stage optimization conditional on τ. Let

$$\Pi_{\mathrm{M}}^*(\tau) := \max_{\mu} \ \Pi_{\mathrm{M}}^{\mu}(\tau),$$
$$\Pi_{\mathrm{A}}^*(\tau) := \max_{p_{\mathrm{A}}} \ \Pi_{\mathrm{A}}(\tau),$$

denote the firm's optimal revenue as a function of τ under the optimal MDS mechanism and the optimal ADS mechanism, respectively. In the first stage, the firm chooses $\tau \in [0,1]$ by incurring investment cost $K(\tau)$ and solves

$$\max_{\tau \in [0,1]} \ \Pi_{\mathrm{M}}^*(\tau) - K(\tau) \quad \text{in the MDS setting, and} \tag{44}$$
$$\max_{\tau \in [0,1]} \ \Pi_{\mathrm{A}}^*(\tau) - K(\tau) \quad \text{in the ADS setting.} \tag{45}$$

Using the envelope theorem, we have

$$\frac{d\Pi_{\mathrm{M}}^*(\tau)}{d\tau} = \frac{\partial \Pi_{\mathrm{M}}^{\mu}}{\partial \tau}\bigg|_{\mu = \mathrm{MDS}} = \eta N \, d_{\mathrm{M}}^{\mathrm{MDS}} = \begin{cases} \eta N, & \bar{c} \leq \dfrac{\eta\tau}{2k}, \\ \eta N \cdot \dfrac{\eta\tau}{2k\bar{c}}, & \bar{c} > \dfrac{\eta\tau}{2k}, \end{cases} \tag{46}$$

$$\frac{d\Pi_{\mathrm{A}}^*(\tau)}{d\tau} = \frac{\partial \Pi_{\mathrm{A}}}{\partial \tau}\bigg|_{p_{\mathrm{A}} = p_{\mathrm{A}}^{\mathrm{ADS}}} = \eta N \, d_{\mathrm{A}} = \begin{cases} \eta N \gamma, & \bar{c} \leq \dfrac{1+\eta\tau\gamma}{2k\beta}, \\ \eta N \gamma \cdot \dfrac{1+\eta\tau\gamma}{2\bar{c}k\beta}, & \bar{c} > \dfrac{1+\eta\tau\gamma}{2k\beta}. \end{cases} \tag{47}$$

Based on this, we have the following insights:

- **Optimal data-sharing mechanisms**: Framing endogenous-τ as a two-stage problem makes it immediate that, conditional on any given τ, the second-stage optimization (which is our base model) remains unchanged. Hence, the optimal MDS and ADS mechanisms retain the same structure as characterized in Theorems 1 and 2.

- **Complementarity between τ and data collection**: Recall from our baseline analysis (Sections 4 and 5) that the equilibrium amount of data collected by the firm under both the MDS and ADS mechanisms is increasing in τ (until the upper bound is reached). This is because a higher effectiveness in leveraging data (τ) increases the operational-phase revenue, which in turn incentivizes the firm to offer steeper discounts in the development phase to acquire more data. Conversely, the expressions in (46) and (47) show that a larger volume of collected data increases the marginal return to investing in τ (until all consumers participate and share their data). Together, these two directions imply that τ and the amount of data collected by the firm are strategic complements and create a virtuous cycle: investments to improve τ incentivizes greater data collection, which in turn increases the marginal payoff from making those investments.

Appendix L Strategic Consumers

In this section, we consider a setting where consumers can delay their purchase of the AI service in anticipation of an improved quality of the service in the future. That is, we allow the consumers to be strategic in timing the purchase of the service. To incorporate strategic consumers, we modify our base model as follows: Recall from Section 3 that our base model consists of an initial *development* phase in which the firm collects data from consumers to improve the service and sells this improved-quality service in the subsequent *operational* phase.

To account for strategic decision-making by the consumers, we partition the development phase into two sub-phases (i.e., into two periods), namely $\mathcal{D}_1$ and $\mathcal{D}_2$. As in our base model, we consider a total of N consumers across these two periods. These consumers are strategic in that they choose whether or not to buy the AI service and, if they choose to buy, then whether to buy in period $\mathcal{D}_1$ or to wait and buy the improved-quality service in period $\mathcal{D}_2$. Moreover, as in the base model, upon purchasing the service, the consumers also decide whether to share their data with the firm or not.

For ease of exposition and analytical convenience, we make the following two simplifying assumptions:

- **Two-price menus**: We focus on mechanisms that have the same structure as the MDS mechanism derived in Corollary 1 of Theorem 1 – that is, a two-price menu with (i) a higher price for consumers who purchase the service but do not share any data and (ii) a relatively low price for consumers who purchase the service and share their entire data. Thus, the firm needs to decide on the following four prices:
  1. Price $p_{s,1}$ for consumers who purchase the service in period $\mathcal{D}_1$ and share their entire data,
  2. Price $p_{ns,1}$ for consumers who purchase the service in period $\mathcal{D}_1$ and do not share any data,
  3. Price $p_{s,2}$ for consumers who purchase the service in period $\mathcal{D}_2$ and share their entire data, and
  4. Price $p_{ns,2}$ for consumers who purchase the service in period $\mathcal{D}_2$ and do not share any data.
- **Binary consumer types**: In contrast to the continuous consumer types in our base model, we assume in this section that a consumer can be either of *high* type with probability h or *low* type with probability $(1-h)$. For a unit amount of sensitive data shared, high-type consumers incur a privacy cost of c_h whereas low-type consumers incur a privacy cost of c_ℓ, where $c_h > c_\ell$. As before, consumers do not incur a privacy cost when they do not share their data. In addition to the privacy cost, consumers incur a waiting cost of $w > 0$ if they wait (and do not purchase the service) in period $\mathcal{D}_1$ and purchase the service in $\mathcal{D}_2$. The parameters h, c_h, c_ℓ, and w are common knowledge.

Let $m_{s,i}$ and $m_{ns,i}$ denote, respectively, the number of consumers who share and do not share their data with the firm when they purchase the service in period $\mathcal{D}_i$ $(i = 1, 2)$. In line with our base model, (i) we have

$m_{s,1}+m_{ns,1}+m_{s,2}+m_{ns,2}=N$, and (ii) the valuation of the AI service in period $\mathcal{D}_1$ is 1 and increases to $\left(1+\tau\cdot\frac{m_{s,1}}{N}\right)$ in period $\mathcal{D}_2$, and further increases to $\left(1+\tau\cdot\frac{m_{s,1}+m_{s,2}}{N}\right)$ in the operational phase.

The sequence of events is as follows: (i) First, the firm announces the prices – namely, $p_{s,1}, p_{ns,1}, p_{s,2}$, and $p_{ns,2}$. (ii) Based on these prices, consumers decide whether or not to buy the service and, if yes, then whether to buy it in period $\mathcal{D}_1$ or in period $\mathcal{D}_2$. In addition, they also need to decide whether or not to share their data with the firm upon purchasing the service. Of course, while making these decisions, consumers anticipate the decisions of other consumers. (iii) Next, after obtaining the data from consumers who purchase the service and share their data in $\mathcal{D}_1$, the firm improves the quality of the AI service and offers it to the waiting consumers in period $\mathcal{D}_2$ who then make their purchasing and data-sharing decisions. (iv) Finally, the firm further improves the quality of the AI service based on the data collected in period $\mathcal{D}_2$ and offers it in the operational phase.

Based on the above discussion, when consumers are strategic, the firm's revenue maximization problem is as follows:

$$\max_{p_{ns,1},p_{s,1},p_{ns,2},p_{s,2}} \underbrace{p_{ns,1}\cdot m_{ns,1}+p_{s,1}\cdot m_{s,1}}_{\text{Revenue from } \mathcal{D}_1}+\underbrace{p_{ns,2}\cdot m_{ns,2}+p_{s,2}\cdot m_{s,2}}_{\text{Revenue from } \mathcal{D}_2}+\underbrace{\eta N\left(1+\tau\cdot\frac{m_{s,1}+m_{s,2}}{N}\right)}_{\text{Revenue from operational phase}}. \qquad (\mathcal{P}_{\mathrm{s}})$$

In the result below, we present the most interesting equilibria that occur when consumers are strategic:

THEOREM 7 **(Equilibria for strategic consumers)**. *When consumers are strategic, we have:*

(i) *If $c_\ell k \le \eta\tau \le c_h k$ and $\tau h \le w \le \tau(1-h)$, then low-type consumers purchase the service in $\mathcal{D}_1$ and share their entire data with the firm, whereas high-type consumers purchase the service in $\mathcal{D}_2$ and do not share their data with the firm.*

(ii) *If $\eta\tau \ge c_h k$ and $\frac{c_h k - c_\ell k}{\eta\tau - c_\ell k} \le h \le \min\left\{\frac{1}{2}, \frac{\tau}{\tau+w}, \frac{\tau-(c_h-c_\ell)k)}{\tau}\right\}$, then low-type consumers purchase the service in $\mathcal{D}_1$ and share their entire data with the firm, whereas high-type consumers purchase the service in $\mathcal{D}_2$ and share their entire data with the firm.*

(iii) *If $\eta\tau \ge c_h k$ and $h \ge \max\left\{\frac{1}{2}, \frac{w}{\tau}, \frac{c_h k - c_\ell k}{\eta\tau - c_\ell k}, \frac{c_h k - c_\ell k + w}{c_h k - c_\ell k + 2w}\right\}$, then high-type consumers purchase the service in $\mathcal{D}_1$ and share their entire data with the firm, whereas low-type consumers purchase the service in $\mathcal{D}_2$ and share their entire data with the firm.*

The proof of this result is provided in the supplementary appendix. We now explain the insights from Theorem 7. Note that for each of the statements in the theorem, the equilibrium is determined by two conditions: (i) a *participation* condition that determines the period ($\mathcal{D}_1$ or $\mathcal{D}_2$) in which the consumers purchase the service, and (ii) a *data-sharing* condition that determines whether the participating consumers share their data with the firm or not.

Recall that the key tradeoff faced by the firm in our context is that a higher amount of data collected in the development phase results in a higher revenue in the operational phase; however, to collect more data in the development phase, the firm needs to set lower prices to incentivize consumers to share their data.

- **Theorem 7-(i)**: Note that the revenue to the firm in the operational phase is proportional to $\eta\tau$. Further, $c_\ell k$ and $c_h k$ denote the privacy cost incurred per unit amount of data shared with the firm by low- and high-type consumers, respectively. The higher these privacy costs, the higher the incentives that the firm needs to provide (by setting lower prices) to the consumers to enable data sharing. In Theorem 7-(i), the data-sharing condition ($c_\ell k \le \eta\tau \le c_h k$) broadly states that the firm is able to offer sufficient incentives to offset the privacy cost incurred by low-type consumers, but not enough to offset the privacy cost incurred by high-type consumers. This explains why low-type consumers share their data with the firm, whereas high-type consumers do not share their data. On the other hand, for the participation condition ($\tau h \le w \le \tau(1-h)$), note that τh (resp., $\tau(1-h)$) denotes the increase in consumers' valuation of the service in period $\mathcal{D}_2$ (compared to that in period $\mathcal{D}_1$), when only high-type (resp., low-type) consumers share their data in period $\mathcal{D}_1$. Thus, the participation condition $\tau h \le w \le \tau(1-h)$ implies that the waiting cost w is too high ($\ge \tau h$) for the low-type consumers – thus, they purchase the service in $\mathcal{D}_1$. Consequently, the valuation of the service increases to $\tau(1-h)$ in period $\mathcal{D}_2$. Since $w \le \tau(1-h)$, the high-type consumers find it optimal to incur the waiting cost and purchase the improved service in $\mathcal{D}_2$.

- **Theorem 7-(ii)**: Since $\eta\tau \ge c_h k \ge c_\ell k$, the firm has sufficient incentives to offset the privacy costs of both low- and high-type consumers. Thus, both type of consumers share their entire data with the firm. Further, the participation condition $\frac{c_h k - c_\ell k}{\eta\tau - c_\ell k} \le h \le \min\left\{\frac{1}{2}, \frac{\tau}{\tau+w}, \frac{\tau-(c_h-c_\ell)k)}{\tau}\right\}$ states that the proportion of low-type consumers is sufficiently high (even higher than the high-type consumers) due to which the firm provides them enough incentives to purchase the service in period $\mathcal{D}_1$. On the other hand, high-type consumers purchase the service in period $\mathcal{D}_2$ since the improvement in the quality of the AI service in that period outweighs the waiting cost they incur.

- **Theorem 7-(iii)**: Similar to that in Theorem 7-(ii), the data-sharing condition implies that both consumer types share their entire data with the firm. However, in contrast, the participation condition $h \ge \max\left\{\frac{1}{2}, \frac{w}{\tau}, \frac{c_h k - c_\ell k}{\eta\tau - c_\ell k}, \frac{c_h k - c_\ell k + w}{c_h k - c_\ell k + 2w}\right\}$ states that the proportion of high-type consumers is sufficiently high (even higher than that of low-type consumers) due to which the firm incentivizes them to participate in period $\mathcal{D}_1$. On the other hand, low-type consumers purchase the service in period $\mathcal{D}_2$ since the improvement in the quality of the AI service in that period outweighs their waiting cost.

Appendix M Socially Optimal Data-Sharing Mechanisms

In this section, we study data-sharing mechanisms that maximize social welfare. The social welfare under a mechanism is the total surplus (i.e., the sum of the firm's revenue and consumers' surplus) generated across the development and operational phases under that mechanism. Since price is a transfer between the consumer and the firm, it cancels in the welfare calculation.

Social welfare under manual data-sharing mechanism: Recall that each consumer receives utility 1 from consuming the service. Under the manual data-sharing mechanism, if a type-c consumer shares $q(c)$ amount of data, then she incurs a privacy cost of $ck \cdot q(c)$. Integrating over all consumer types, the social welfare in the development phase is:

$$\frac{N}{\bar{c}} \int_0^{\bar{c}} \big(1 - ck \cdot q(c)\big)\, dc.$$

The total data collected in the development phase is $1/\bar{c}\int_0^{\bar{c}} q(c)dc$. This improves the service quality in the operational phase to $1+\tau\cdot\left(1/\bar{c}\int_0^{\bar{c}} q(c)dc\right)$. With η consumers in the operational phase, the social welfare in this phase is:

$$\eta N\left(1+\tau\cdot\frac{1}{\bar{c}}\int_0^{\bar{c}} q(c)dc\right).$$

Combining the two phases, the total social welfare is:

$$\frac{N}{\bar{c}}\int_0^{\bar{c}}\big(1-ck\cdot q(c)\big)\,dc+\eta N\left(1+\tau\cdot\frac{1}{\bar{c}}\int_0^{\bar{c}} q(c)dc\right)=N(1+\eta)+\frac{N}{\bar{c}}\int_0^{\bar{c}}(\eta\tau-ck)q(c)dc.$$

The socially optimal level of data-sharing under the manual data-sharing mechanism can be obtained by solving the point-wise optimization problem:

$$q_{\text{M}}^{\text{so}}(c)\in\arg\max_{q\in[0,1]}(\eta\tau-ck)\,q,$$

which yields the following result. If the highest type $\bar{c}$ is sufficiently low, i.e., if $0<\bar{c}\leq\frac{\eta\tau}{k}$, then

$$q_{\text{M}}^{\text{so}}(c)=1\ \forall c\in[0,\bar{c}],$$

and if the highest type $\bar{c}$ is sufficiently high, i.e., if $\bar{c}>\frac{\eta\tau}{k}$, then

$$q_{\text{M}}^{\text{so}}(c)=\begin{cases}1, & \text{if } c\leq\dfrac{\eta\tau}{k},\\ 0, & \text{if } c>\dfrac{\eta\tau}{k}.\end{cases}$$

This shows that if the consumer population is sufficiently insensitive to privacy, i.e., if $\bar{c}\leq\frac{\eta\tau}{2k}$, then the revenue-maximizing and socially optimal MDS mechanisms coincide, and all consumers share their data. If $\frac{\eta\tau}{2k}<\bar{c}\leq\frac{\eta\tau}{k}$, then full data-sharing remains socially optimal, but the revenue-maximizing MDS mechanism already excludes the intermediate types $c\in\left(\frac{\eta\tau}{2k},\bar{c}\right]$ from sharing. Finally, if $\bar{c}>\frac{\eta\tau}{k}$, then the socially optimal mechanism itself is threshold-based: full data-sharing is socially desirable whenever the marginal social benefit, $\eta\tau$, exceeds the marginal cost, ck, and no sharing is optimal otherwise. In contrast, Theorem 1 shows that the revenue-maximizing MDS mechanism uses the lower threshold $\frac{\eta\tau}{2k}$, which generates a downward distortion in data-sharing relative to the social optimum.

Social welfare under algorithmic data-sharing mechanism: Under algorithmic data-sharing mechanism, a type-c consumer either purchases the service and shares her data, or does not participate. Let $y(c)\in\{0,1\}$ denote whether a type-c consumer purchases the service. If she purchases, then the redaction algorithm collects $k\beta$ amount of sensitive data and $(1-k)(1-\alpha)$ amount of non-sensitive data from her. Hence, the total amount of data collected from a participating consumer is $\gamma:=k\beta+(1-k)(1-\alpha)$. Since only the sensitive data retained by the redaction algorithm leads to privacy loss, a participating type-c consumer incurs a privacy cost of $ck\beta$. Integrating over all consumer types, the social welfare in the development phase is:

$$\frac{N}{\bar{c}}\int_0^{\bar{c}}\big(1-ck\beta\big)y(c)\,dc.$$

The total data collected in the development phase is $\frac{\gamma}{\bar{c}}\int_0^{\bar{c}} y(c)\,dc$. This improves the service quality in the operational phase to $1+\tau\cdot\left(\frac{\gamma}{\bar{c}}\int_0^{\bar{c}} y(c)\,dc\right)$. With η consumers in the operational phase, the social welfare in this phase is:

$$\eta N\left(1+\tau\cdot\frac{\gamma}{\bar{c}}\int_0^{\bar{c}} y(c)\,dc\right).$$

Combining the two phases, the total social welfare is:

$$\frac{N}{\bar{c}}\int_0^{\bar{c}}(1-ck\beta)y(c)\,dc+\eta N\left(1+\tau\cdot\frac{\gamma}{\bar{c}}\int_0^{\bar{c}}y(c)\,dc\right)=\eta N+\frac{N}{\bar{c}}\int_0^{\bar{c}}(1+\eta\tau\gamma-ck\beta)y(c)\,dc.$$

The socially optimal participation rule under the algorithmic data-sharing mechanism can be obtained by solving the point-wise optimization problem:

$$y_{\text{A}}^{\text{SO}}(c)\in\arg\max_{y\in\{0,1\}}(1+\eta\tau\gamma-ck\beta)y,$$

which yields the following result. If the highest type $\bar{c}$ is sufficiently low, i.e., if $0<\bar{c}\le\frac{1+\eta\tau\gamma}{k\beta}$, then

$$y_{\text{A}}^{\text{SO}}(c)=1\ \forall c\in[0,\bar{c}],$$

and if the highest type $\bar{c}$ is sufficiently high, i.e., if $\bar{c}>\frac{1+\eta\tau\gamma}{k\beta}$, then

$$y_{\text{A}}^{\text{SO}}(c)=\begin{cases}1, & \text{if } c\le\dfrac{1+\eta\tau\gamma}{k\beta},\\ 0, & \text{if } c>\dfrac{1+\eta\tau\gamma}{k\beta}.\end{cases}$$

Thus, if the consumer population is sufficiently insensitive to privacy, i.e., if $\bar{c}\le\frac{1+\eta\tau\gamma}{2k\beta}$, then the revenue-maximizing and socially optimal ADS mechanisms coincide, and all consumers participate. In this case, the redaction algorithm processes the uploaded data of every participating consumer and retains γ amount of data from each consumer. If $\frac{1+\eta\tau\gamma}{2k\beta}<\bar{c}\le\frac{1+\eta\tau\gamma}{k\beta}$, then full participation remains socially optimal, but the revenue-maximizing ADS mechanism already excludes the intermediate types $c\in\left(\frac{1+\eta\tau\gamma}{2k\beta},\bar{c}\right]$. Finally, if $\bar{c}>\frac{1+\eta\tau\gamma}{k\beta}$, then participation is socially desirable whenever the direct consumption benefit, 1, together with the marginal social benefit from the retained data, $\eta\tau\gamma$, exceeds the marginal privacy cost, $ck\beta$, and non-participation is optimal otherwise. In contrast, Theorem 2 shows that the revenue-maximizing ADS mechanism uses the lower threshold $\frac{1+\eta\tau\gamma}{2k\beta}$, which again yields a downward distortion in participation relative to the social optimum.

The analysis above reveals a common structural pattern across both mechanisms: whenever the threshold structure is relevant, the socially optimal threshold is exactly twice the revenue-maximizing threshold. This factor-of-two distortion is a consequence of the standard monopoly pricing under uniformly distributed types: to attract a marginal type c into sharing data (or participating), the firm must reduce the price paid by all less privacy-sensitive types, which exactly doubles the effective marginal cost of data collection (from ck to $2ck$) relative to the social planner's problem. The social planner does not face such an information-rent consideration, and therefore internalizes only the true privacy cost. Consequently, the planner can set the threshold twice as high.

This also delivers a natural "win-win" insight. Under MDS, if $\bar{c}\le\frac{\eta\tau}{2k}$, and under ADS, if $\bar{c}\le\frac{1+\eta\tau\gamma}{2k\beta}$, the decentralized revenue-maximizing mechanism attains the first best. In these regions, privacy costs are low enough that both the firm and the social planner prefer full coverage, so the monopoly distortion disappears. Once $\bar{c}$ exceeds the relevant threshold, however, the firm under-collects data relative to the social optimum because it internalizes only the revenue consequences of data collection, not the full welfare gains created in the operational phase. In that case, implementing the first best would require an external intervention such as a subsidy.