Supplemental Material to
Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences

By

Kai Hao Yang

October 2020

COWLES FOUNDATION DISCUSSION PAPER NO. 2258S

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Supplemental Material for "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences"

Kai Hao Yang*

October 9, 2020

S.1 Constructing an Implementable Quasi-Perfect Scheme: An Algorithmic Approach

The proof of Lemma 3 provides a direct construction of a ψ-upper separation scheme satisfying (8), provided that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$. However, this construction might be difficult to compute for some demands. This section introduces an iteration process that can be directly applied to construct a ψ-quasi-perfect scheme satisfying (8). To begin with, given any $m \in \Delta^f(V)$, for any $V^* \subseteq \text{supp}(m)$ and for any $v^* \in V^*$, define an operation "project m to v^* along V^*" as the following: Given $m \in \Delta^f(V)$, construct $\hat{m}$ by letting

$$\hat{m}(v) := \begin{cases} \lambda m(v), & \text{if } v \notin V^* \\ m(v), & \text{if } v = v^* \\ 0, & \text{if } v \in V^*,\, v \neq v^* \end{cases},$$

where

$$\lambda := \frac{m(v)}{\sum_{v' \in V^*} m(v')}$$

Furthermore, call $\hat{m}$ "the projection of m to v^* along V^*" and call $x - \hat{m}$ "the residual". As a convention, if $V^* = \emptyset$, define the projection of any m along V^* as m itself. In addition, for any finite set $V^* \subseteq V$, for any $\hat{m} \in \mathbb{R}_+^{V^*}$, define the operation "normalize $\hat{m}$" by constructing

$$m^*(v) := \frac{\hat{m}(v)}{\sum_{v' \in V^*} \hat{m}(v')}.$$

and call $m^* \in \Delta^f(V)$ the "normalization" of $\hat{m}$. Now, take and fix any $\psi : C \to \mathbb{R}_+$ and any $c \in C$, for any step function $D \in \mathcal{D}$, let $v^{(n)}$ denote the n-th largest element in $\text{supp}(D)$ (as a convention, for any $n > |\text{supp}(D)|$, $n \in \mathbb{N}$, write $v^{(n)} = v^{(n-1)}$) and define the following four cases

- **Case 1:** $\psi(c) \leq \overline{\boldsymbol{p}}_D(c) < v^{(1)}$.
- **Case 2:** $\overline{\boldsymbol{p}}_D(c) = v^{(1)} > v^{(2)} \geq \psi(c)$ and $(v^{(2)} - c)D(v^{(2)}) \geq (v' - c)D(v')$ for all $v' < \psi(c)$.

*Cowles Foundation for Economic Research, Yale University, kaihao.yang@yale.edu.

- **Case 3:** $\overline{\boldsymbol{p}}_D(c) = v^{(1)} > v^{(2)} \geq \psi(c)$ and there exists $v' < \psi(c)$ such that $(v^{(2)} - c)D(v^{(2)}) < (v' - c)D(v')$.

- **Case 4:** $\overline{\boldsymbol{p}}_D(c) = v^{(1)} \geq \psi(c) > v^{(2)}$.

Given any step function $D \in \mathcal{D}$, consider the following iteration process. Let $\mathcal{D}_0 := \{D\}$ and let $s^0 := \delta_{\{D\}}$. For each $k \in \mathbb{N}$, given a collection of step functions $\mathcal{D}_{k-1} \subseteq \mathcal{D}$ with $|\mathcal{D}_{k-1}| < \infty$ and $s^{k-1} \in \Delta(\mathcal{D}_{k-1})$. Enumerate $\mathcal{D}_{k-1}$ as $\{D_1, \ldots, D_{n_{k-1}}\}$. For each $D_j \in \mathcal{D}_{k-1}$, construct D_j^1, D_j^2 and λ_j^1, λ_j^2 as follows: Let $\hat{m}_j^1$ be the projection of m^{D_j} to v^j along V^j and let $\hat{m}_j^2$ be the residual. Then normalize $\hat{m}_j^1$ to m_j^1 and $\hat{m}_j^2$ to m_j^2. Moreover, let

$$\lambda_j^1 := \begin{cases} \frac{m^{D_j}(v^j)}{\sum_{v \in V^j} m^{D_j}(v)}, & \text{if } V^j \neq \emptyset \\ 1, & \text{if } V^j = \emptyset \end{cases}$$

and let

$$\lambda_j^2 := 1 - \lambda_j^1,$$

where $V^j \subseteq \text{supp}(D_j)$ and $v^j \in V^j$ are defined respectively according to the cases into which D_j falls as the following.

- In **Case 1**, $v^j := \overline{\boldsymbol{p}}_{D_j}(c)$ and $V^j := \{v \in \text{supp}(D_j) : v \geq v^j\}$.

- In **Case 2**, $v^j := v^{(2)}$ and $V^j := \{v^{(1)}, v^{(2)}\}$.

- In **Case 3**, $v^j := v^{(2)}$ and $V^j := \{v^{(2)}, v^{(5)}\}$.

- In **Case 4**, $v^j := v^{(1)}$ and $V^j := \{v^{(1)}\}$.

- Otherwise, $V^j := \emptyset$.

Then, for each $j \in \{1, \ldots, n_{k-1}\}$ and for each $i \in \{1, 2\}$, let

$$D_j^i(p) := m_j^i([p, \overline{v}]),$$

for all $p \in V$ and let $\mathcal{D}_k$ be $\{D_j^1, D_j^2\}_{j=1}^{n_{k-1}}$ and define $s^k \in \Delta(\mathcal{D}_k)$ by letting

$$s^k(D_j^i) := \lambda_j^i s^{k-1}(D_j), \, \forall j \in \{1, \ldots, n_{k-1}\}, \, \forall i \in \{1, 2\}.$$

The following lemma shows that such iteration process leads to a desirable $\psi(c)$-quasi-perfect segmentation for c when the market demand is D_0.

Lemma S2. *Consider any step function $D \in \mathcal{D}$ and any $\psi \in \mathbb{R}_+^C$ such that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_D(c)$, for all $c \in C$. For any $c \in C$, there exists $K \in \mathbb{N}$ such that s^K is a $\psi(c)$-quasi-perfect segmentation for c with*

$$\psi(z) \leq \overline{\boldsymbol{p}}_{D'}(z), \, \forall z \in [\underline{c}, c], \, \forall D' \in \text{supp}(s^K).$$

Proof. Consider any step function $D \in \mathcal{D}$ and any nondecreasing function $\psi \in \mathbb{R}_+^C$ such that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_D(c)$. Take and fix any $c \in C$, I first show that whenever $\psi \leq \overline{\boldsymbol{p}}_D$ on C, for any $k \in \mathbb{N} \cup \{0\}$, $s^k \in \mathcal{S}_D$ and

$$\psi(z) \leq \overline{\boldsymbol{p}}_{D'}(z), \, \forall z \in [\underline{c}, c], \, \forall D' \in \text{supp}(s^k) \tag{S1}$$

Clearly, when $k = 0$, $s^0 \in \mathcal{S}_D$ and (S1) holds. Now suppose that for $k \in \mathbb{N} \cup \{0\}$, $s^k \in \mathcal{S}_D$ and (S1) holds. Consider any $D_j \in \text{supp}(s^k) = \mathcal{D}_k = \{D_1, \dots, D_{n_k}\}$. Notice that since $\psi(c) \leq \overline{\boldsymbol{p}}_{D_j}(c)$, D_j must belongs to one of **Case1**, **Case 2**, **Case 3** or **Case 4**. In **Case 1**, by the iteration process,

$$m^{D_j} = \hat{m}_j^1 + \hat{m}_j^2 = \lambda_j^1 m_j^1 + \lambda_j^2 m_j^2$$

where $\hat{m}_j^1$ is the projection of m^{D_j} on $\overline{\boldsymbol{p}}_{D_j}(c)$ along $\{v \in \text{supp}(D_j) : v \geq \overline{\boldsymbol{p}}_{D_j}(c)\}$, $\hat{m}_j^2 = m^{D_j} - \hat{m}_j^1$ and m_j^1 and m_j^2 are normalizations of $\hat{m}_j^1$ and $\hat{m}_j^2$, respectively. As a result, for any $v' \leq \overline{\boldsymbol{p}}_{D_j}(c)$ and for any $i \in \{1, 2\}$,

$$\hat{m}_j^i([v', \overline{v}]) = \lambda_j^i D_j(v').$$

Thus, as $\overline{\boldsymbol{p}}_{D_j}(c) \geq \psi(c)$, for any $v' \leq \psi(c)$,

$$(\overline{\boldsymbol{p}}_{D_j}(c) - c)\hat{m}_j^i([\overline{\boldsymbol{p}}_{D_j}(c), \overline{v}]) = \lambda_j^1(\overline{\boldsymbol{p}}_{D_j}(c) - c)D_j(\overline{\boldsymbol{p}}_{D_j}(c)) \geq \lambda_j^1(v' - c)D_j(v') = (v' - c)\hat{m}_j^i([v', \overline{v}]),$$

which implies that $\overline{\boldsymbol{p}}_{D_j}(c) = \overline{p}_{D_j^i}(c) \geq \psi(c)$. Furthermore, for any $z \in [\underline{c}, c)$ and for any $i \in \{1, 2\}$, whenever $\overline{\boldsymbol{p}}_{D_j^i}(z) < \psi(c)$, it must be that for any $v' \leq \overline{\boldsymbol{p}}_{D_j}(z) \leq \overline{\boldsymbol{p}}_{D_j}(c)$,

$$\lambda_i^j(v' - z)D_j(v') = (v' - z)\hat{m}_j^i([v', \overline{v}]) \leq (\overline{\boldsymbol{p}}_{D_j^i}(z) - z)\hat{m}_j^i([\overline{\boldsymbol{p}}_{D_j^i}(z), \overline{v}]) = \lambda_j^i(\overline{\boldsymbol{p}}_{D_j^i}(z) - z)D_j(\overline{\boldsymbol{p}}_{D_j^i}(z)),$$

which implies that $\overline{\boldsymbol{p}}_{D_j^i}(z) \geq \overline{\boldsymbol{p}}_{D_j}(z)$. Together with $\overline{\boldsymbol{p}}_{D_j}(z) \geq \psi(z)$ for all $z \in [\underline{c}, c]$ and $\psi(c) \geq \psi(z)$, we have

$$\overline{\boldsymbol{p}}_{D_j^i}(z) \geq \min\{\psi(c), \overline{\boldsymbol{p}}_{D_j}(z)\} \geq \psi(z), \forall z \in [\underline{c}, c], \forall i \in \{1, 2\}.$$

In **Case 2**. By the iteration process,

$$m^{D_j} = \hat{m}_j^1 + \hat{m}_j^2 = \lambda_j^1 m_j^1 + \lambda_j^2 m_j^2$$

where $\hat{m}_j^1$ is the projection of D^{m_j} to $v^{(2)}$ along $\{v^{(1)}, v^{(2)}\}$, $\hat{m}_j^2 = m^{D_j} - \hat{m}_j^1$ and m_j^1 and m_j^2 are normalizations of $\hat{m}_j^1$ and $\hat{m}_j^2$, respectively. Moreover,

$$(v^{(2)} - c)D_j(v^{(2)}) \geq (v' - c)D_j(v'), \forall v' < \psi(c), v' \in \text{supp}(D_j)$$

and $\psi(c) \leq v^j = v^{(2)} < v^{(1)}$. If $\text{supp}(D_j) = \{v^{(1)}, v^{(2)}\}$, then clearly

$$\psi(z) \leq \overline{\boldsymbol{p}}_{D_j^i}(z), \forall z \in [\underline{c}, c], \forall i \in \{1, 2\}$$

since ψ is nondecreasing. On the other hand, if $\{v^{(1)}, v^{(2)}\} \subsetneq \text{supp}(D_j)$, then for any $v' \in \text{supp}(D_j)$ such that $v' < v^{(2)}$, and for any $i \in \{1, 2\}$

$$\hat{m}_j^i([v', \overline{v}]) = \lambda_j^i D_j(v').$$

This implies that, for all $v' < \psi(c)$, since $v^{(2)} \geq \psi(c)$,

$$(v^{(2)} - c)\hat{m}_j^1([v^{(2)}, \overline{v}]) = \lambda_j^1(v^{(2)} - c)D_j(v^{(2)}) \geq \lambda_j^1(v' - c)D_j(v') = (v' - c)\hat{m}_j^1([v', \overline{v}])$$

and that

$$(v^{(1)} - c)\hat{m}_j^2([v^j, \overline{v}]) = \lambda_j^2(v^{(1)} - c)D_j(v^j) > \lambda_j^2(v' - c)D_j(v') = (v' - c)\hat{m}_j^2([v', \overline{v}]).$$

Together, this implies that $\boldsymbol{P}_{D_j^i}(c) \cap \{v \in \text{supp}(D_i^j) : v \geq \psi(c)\} \neq \emptyset$ and hence $\overline{\boldsymbol{p}}_{D_j^i}(c) \geq \psi(c)$ for each $i \in \{1,2\}$. Moreover, for any $z \in [\underline{c}, c)$ and for any $i \in \{1,2\}$, if $\overline{\boldsymbol{p}}_{D_j^i}(z) < \psi(c)$, then

$$\lambda_j^i(\overline{\boldsymbol{p}}_{D_j^i}(z) - z)D_j(\overline{\boldsymbol{p}}_{D_j^i}(z)) = (\overline{\boldsymbol{p}}_{D_j^i}(z) - z)\hat{m}_j^i([\overline{\boldsymbol{p}}_{D_j^i}(z), \overline{v}]) > (v - z)\hat{m}_j^i([v, \overline{v}]) = \lambda_j^i(v - z)D_j(v),$$

for all $v \in \{v^{(1)}, v^{(2)}\}$ and

$$\lambda_j^i(v' - z)D_j(v') = (v' - z)\hat{m}_j^i([v', \overline{v}]) \leq (\overline{\boldsymbol{p}}_{D_j^i}(z) - z)\hat{m}_j^i([\overline{\boldsymbol{p}}_{D_j^i}(z), \overline{v}]) = \lambda_j^i(\overline{\boldsymbol{p}}_{D_j^i}(z) - z)D_j(\overline{\boldsymbol{p}}_{D_j^i}(z)),$$

for all $v' < v^{(2)}$. Thus, $\overline{\boldsymbol{p}}_{D_j}(z) = \overline{\boldsymbol{p}}_{D_j^i}(z)$ for each $i \in \{1,2\}$ whenever $\overline{\boldsymbol{p}}_{D_j}(z) \leq \psi(c)$. Together, we have

$$\overline{\boldsymbol{p}}_{D_j^i}(z) \geq \min\{\psi(c), \overline{\boldsymbol{p}}_{D_j}(z)\} \geq \psi(z), \ \forall z \in [\underline{c}, c], \ \forall i \in \{1,2\}.$$

In **Case 3**, notice that the condition implies $\overline{\boldsymbol{p}}_{D_j}(c) = v^{(1)} > v^{(2)} > v^{(5)}$. Again by the iteration process

$$m^{D_j} = \hat{m}_j^1 + \hat{m}_j^2 = \lambda_j^1 m_1^j + \lambda_j^2 m_j^2,$$

where $\hat{m}_j^1$ is the projection of m^{D_j} to $v^{(2)}$ along $\{v^{(2)}, v^{(5)}\}$, $\hat{m}_j^2 = m^{D_j} - \hat{m}_j^1$ and m_j^1 and m_j^2 are normalizations of $\hat{m}_j^1$ and $\hat{m}_j^2$, respectively. Then for all $v' \leq v^{(5)}$,

$$\hat{m}_j^i([v', \overline{v}]) = \lambda_j^i D_j(v').$$

Therefore, for any $v' < v^{(5)}$

$$(v' - c)\hat{m}_j^1([v', \overline{v}]) = \lambda_j^1(v' - c)D_j(v') \leq \lambda_j^1(\overline{\boldsymbol{p}}_{D_j}(c) - c)D_j(\overline{\boldsymbol{p}}_{D_j}(c)) = (v^{(1)} - c)\hat{m}_j^1([v^{(1)}, \overline{v}]).$$

Also,

$$(v^{(5)} - c)\hat{m}_j^1([v^{(5)}, \overline{v}]) = (v^{(5)} - c)\hat{m}_j^1([v^{(2)}, \overline{v}]) < (v^{(2)} - c)\hat{m}_j^1([v^{(2)}, \overline{v}]).$$

Similarly, for any $v' \in \text{supp}(D_j)$ with $v' < v^{(2)}$,

$$(v' - c)\hat{m}_j^2([v', \overline{v}]) = \lambda_j^2(v' - c)D_j(v') \leq \lambda_j^2(\overline{\boldsymbol{p}}_{D_j}(c) - c)D_j(\overline{\boldsymbol{p}}_{D_j}(c)) = (v^{(1)} - c)\hat{m}_j^2([v^{(1)}, \overline{v}]).$$

Together, since $\psi(c) \leq v^{(2)} < v^{(1)}$, $\overline{\boldsymbol{p}}_{D_j^i}(c) \geq \psi(c)$ for all $i \in \{1,2\}$. Furthermore, for any $z \in [\underline{c}, c)$ and for any $i \in \{1,2\}$, if $\overline{\boldsymbol{p}}_{D_j^i}(z) < v^{(2)}$, then for any $v' \in \text{supp}(D_j)$ such that either $v' < v^{(2)}$ or $v' = v^{(1)}$,

$$\lambda_j^i(v' - z)D_j(v') = (v' - z)\hat{m}_j^i([v', \overline{v}]) \leq (\overline{\boldsymbol{p}}_{D_j^i}(z) - z)\hat{m}_j^i([\overline{\boldsymbol{p}}_{D_j^i}(z), \overline{v}]) = \lambda_j^i(\overline{\boldsymbol{p}}_{\hat{m}_j^i}(z) - z)D_j(\overline{\boldsymbol{p}}_{\hat{m}_j^i}(z)).$$

Moreover, since $(v^{(2)} - c)D_j(v^{(2)}) < (v' - c)D_j(v')$ for some $v' < \psi(c) \leq v^{(2)}$, it must be that $(v^{(2)} - z)D_j(v^{(2)}) < (v' - z)D_j(v')$. Together, these imply that $\overline{\boldsymbol{p}}_{D_j^i}(z) \geq \overline{\boldsymbol{p}}_{D_j}(z)$ whenever $\overline{\boldsymbol{p}}_{D_j^i}(z) < v^{(2)}$. Therefore, since $\overline{\boldsymbol{p}}_{D_j}(z) \geq \psi(z)$ for all $z \in [\underline{c}, c]$, and since $\psi(z) \leq \psi(c) \leq v^{(2)}$ for all $z \in [\underline{c}, c]$

$$\overline{\boldsymbol{p}}_{D_j^i}(z) \geq \min\{v^{(2)}, \overline{\boldsymbol{p}}_{D_j}(z)\} \geq \psi(z), \ \forall z \in [\underline{c}, c], \ \forall i \in \{1,2\}.$$

In **Case 4**, clearly $\hat{m}_j^1 = m^{D_j}$ and $\lambda_j^1 = 1$, $\lambda_j^2 = 0$,

$$m^{D_j} = \hat{m}_j^1 = m_j^1.$$

It then follows immediately that

$$\overline{\boldsymbol{p}}_{D_j^1}(z) = \overline{\boldsymbol{p}}_{D_j}(z) \geq \psi(z), \, \forall z \in [\underline{c}, c].$$

Together, in all the possible cases,

$$m^{D_j} = \lambda_j^1 m_j^1 + \lambda_j^2 m_j^2$$

and hence

$$D_j = \lambda_j^1 D_j^1 + \lambda_j^2 D_j^2,$$

for some $\lambda_j^1, \lambda_j^2 \in [0,1]$ such that $\lambda_j^1 + \lambda_j^2 = 1$ and

$$\boldsymbol{p}_{D_j^i}(z) \geq \psi(z), \, \forall z \in [\underline{c}, c]$$

and for all $i \in \{1,2\}$ such that $\lambda_j^i > 0$. By the iteration process, for any $D_j \in \mathrm{supp}(s^k) = \mathcal{D}^k$,

$$s^{k+1}(D_j^i) = \lambda_i^j s^k(D_j).$$

This implies that

$$\sum_{i\in\{1,2\},\, D_j \in \mathcal{M}^{k+1}} D_j^i s^k(D_j^i) = \sum_{D_j \in \mathcal{M}^k} (\lambda_j^1 D_j^1 + \lambda_j^2 D_j^2) s^k(D_j) = \sum_{D_j \in \mathcal{M}^k} D_j s^k(D_j) = D$$

and that

$$\sum_{D \in \mathcal{D}^{k+1}} s^{k+1}(D) = \sum_{i\in\{1,2\},\, D_j \in \mathcal{D}^k} s^{k+1}(D_j^i) = \sum_{D_j \in \mathcal{D}^k} (\lambda_j^1 + \lambda_j^2) s^k(D_j) = \sum_{D_j \in \mathcal{D}^k} s^k(D_j) = 1.$$

Therefore, $s^{k+1} \in \mathcal{S}_D$. Furthermore, for any $D' \in \mathrm{supp}(s^{k+1})$, $D' = D_j^i$ for some $j \in \{1, \ldots, n_k\}$ and some $i \in \{1,2\}$ such that $\lambda_j^i > 0$. Therefore,

$$\overline{\boldsymbol{p}}_{D'}(z) \geq \psi(z), \, \forall z \in [\underline{c}, c].$$

By induction, $s^k \in \mathcal{S}_D$ and satisfies (S1).

Now notice that for any $k \in \mathbb{N}$ and for any $D_j \in \mathcal{D}^k$, if either D_j is in **Case 1** or **Case 3**, or if D_j is in **Case 2** and $v^{(1)} > v^{(2)}$, then

$$D_j = \lambda_j^1 D_j^1 + \lambda_j^2 D_j^2$$

with $\lambda_j^1 > 0$ and $\lambda_j^2 > 0$. This implies that $|\mathrm{supp}(D_j^i)| = |\mathrm{supp}(D_j)| - 1$ for each $i \in \{1,2\}$. Since $|\mathrm{supp}(D)| < \infty$ and since $s^k \in \mathcal{S}_D$ satisfies (S1) for all $k \in \mathbb{N}$, there exists $\bar{K} \in \mathbb{N}$ such that all of $D_j \in \mathcal{D}^K$ are either in **Case 4** or with $|\mathrm{supp}(D_j)| = 1$. Let $K \in \mathbb{N}$ be the smallest number such that m_j is either in **Case 4** or with $|\mathrm{supp}(D_j)| = 1$ for all $D_j \in \mathcal{D}^K$. Then s^K is a $\psi(c)$-quasi-perfect segmentation for c with

$$\overline{\boldsymbol{p}}_{D'}(z) \geq \psi(z), \, \forall z \in [\underline{c}, c], \, \forall D' \in \mathrm{supp}(s^K),$$

as desired.

Since c is arbitrary, the construction above leads to a ψ-quasi-perfect scheme $\sigma \in \mathcal{S}_D^C$ that satisfies (8). ■

S.2 Surplus Extraction in General Environments

In this section, I generalize the surplus extraction result (Theorem 3) in the main text. As shown in Theorem S1 below, Theorem 3 in fact does not depend on any assumptions about D_0 and G (except that G has a density $g > 0$).

Theorem S1 (Generalized Surplus Extraction)**.** *For any D_0 and G, there exists an incentive feasible mechanism that maximizes the data broker's revenue. Furthermore, under any revenue-maximizing mechanism for the data broker, the consumers retain zero surplus.*

Proof. I first show that the optimal solution for the data broker's revenue-maximization problem exists. To see this, notice that since the data broker's revenue maximization problem is to choose $\sigma \in \mathcal{S}^C$ and $\tau \in \mathbb{R}^C$ that satisfies the incentive compatibility and the individual rationality constraints to maximize $\mathbb{E}_G[\tau(c)]$ and since the data broker's revenue is bounded by R^* for all incentive feasible mechanisms and since there is only one agent with one-dimensional private information, it is without loss to restrict the space of feasible transfers to $[-\bar{R}, \bar{R}]$ for some $\bar{R} \in [R^*, \infty)$. Then, by the Lebesgue dominated convergence theorem, the data broker's objective function is continuous in the choice variable (σ, τ) under the product topology on $\mathcal{S}^C \times [\bar{R}, \bar{R}]^C$. Moreover, since $\Delta(\mathcal{D})$ is compact and since $\mathcal{S} \subset \Delta(\mathcal{D})$ is closed and hence also compact, Tychonoff's theorem implies that $\mathcal{S}^C \times [-\bar{R}, \bar{R}]^C$ is compact under the product topology. Also, by Lemma 7, the feasible set given by the incentive compatibility and individual rationality constraints is a closed set in $\mathcal{S}^C \times [-\bar{R}, \bar{R}]^C$, which ensures that there exists a solution to the data broker's problem.

To see that any incentive compatible and individually rational mechanism that maximizes the data broker's revenue must give the consumers zero surplus, use Lemma 1 and by (5) and rewrite the data broker's revenue maximization into

$$\begin{aligned}
&\sup_{\sigma \in \mathcal{S}^C} \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi} \\
\text{s.t. } &\int_c^{c'} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z))(\sigma(\mathrm{d}D|z) - \sigma(\mathrm{d}D|c')) \right) \mathrm{d}z \geq 0,\ \forall c, c' \in C \\
&\bar{\pi} + \int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z)) \sigma(\mathrm{d}D|z) \right) \mathrm{d}z \geq \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\, \mathrm{d}z,\ \forall c \in C.
\end{aligned} \tag{S2}$$

Consider any incentive compatible and individually rational mechanism (σ, τ). Suppose that the consumer surplus is positive. That is,

$$\int_C \int_{\mathcal{D}} \int_{\{v \geq \boldsymbol{p}_D(c)\}} (v - \boldsymbol{p}_D(c)) D(\mathrm{d}v) \sigma(\mathrm{d}D|c) G(\mathrm{d}c) > 0$$

for some $\boldsymbol{p} \in \boldsymbol{P}$. First notice that since by Lemma 1, the data broker's expected revenue is

$$\mathbb{E}_G[\tau(c)] = \int_C \int_{\mathcal{D}} \left[\pi_D(c) - \frac{G(c)}{g(c)} D(\boldsymbol{p}_D(c)) \right] \sigma(\mathrm{d}D|c) G(\mathrm{d}c) - \bar{\pi},$$

and hence if

$$\int_{\mathcal{D}} D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c) < \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c)$$

for a set of $c \in C$ that has G positive measure, then the mechanism (σ, τ) cannot be optimal. As such, it is without loss of generality to suppose that

$$\int_C \int_{\mathcal{D}} \int_{\{v \geq \boldsymbol{p}_D(c)\}} (v - \overline{\boldsymbol{p}}_D(c)) D(\mathrm{d}v) \sigma(\mathrm{d}D|c) G(\mathrm{d}c) > 0 \tag{S3}$$

I will now construct a mechanism $(\hat{\sigma}, \hat{\tau})$ that strictly improves the data broker's revenue. Indeed, since $\Delta^f(\mathcal{D})^C$ is dense in $\Delta(\mathcal{D})^C$ (under the product topology) and since $\sigma \in \mathcal{S}^C$ is implementable, there exists a sequence of step functions $\{D_n\} \subset \mathcal{D}$ and a sequence of implementable $\{\sigma_n\} \subset \mathcal{S}^C_{D_n}$ such that $\{\sigma_n\} \to \sigma$ pointwise. For any $c \in C$, any $n \in \mathbb{N}$, and any $D \in \mathrm{supp}(\sigma_n(c))$, since $\overline{\boldsymbol{p}}_D \in \mathbb{R}^C$ is nondecreasing and clearly $c \leq \overline{\boldsymbol{p}}_D(c) \leq \overline{\boldsymbol{p}}_D(c)$ for all $c \in C$, by Lemma 3, there exists a $\overline{\boldsymbol{p}}_D$-quasi-perfect scheme $K^D_n \in \mathcal{S}^C_D$ that satisfies (8). By Lemma 11, K^D_n has the following properties:

$$\sum_{D' \in \mathrm{supp}(K^D_n(c))} D' K^D_n(D'|c) = D, \qquad \sum_{D' \in \mathrm{supp}(K^D_n(c))} D'(\overline{\boldsymbol{p}}_{D'}(c)) K^D_n(D'|c) = D(\overline{\boldsymbol{p}}_D(c)),$$

and

$$\sum_{D' \in \mathrm{supp}(K^D_n(c))} (\overline{\boldsymbol{p}}_{D'}(c) - \overline{\boldsymbol{p}}_D(c)) D'(\overline{\boldsymbol{p}}_{D'}(c)) K^D_n(D'|c) = \sum_{v \geq \overline{\boldsymbol{p}}_D(c)} (v - \overline{\boldsymbol{p}}_D(c)) m^D(v).$$

Furthermore, by Lemma 3 and Lemma 11,

$$\sum_{D' \in \mathrm{supp}(K^D_n(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c)) K^D_n(D'|c') \leq \sum_{D' \in \mathrm{supp}(K^D_n(c'))} D'(\overline{\boldsymbol{p}}_D(c)) K^D_n(D'|c') = D(\overline{\boldsymbol{p}}_D(c)),$$

for any $c, c' \in C$ with $c < c'$. Also, by Lemma 11 and by the same argument as in the proof of Lemma 12,

$$\sum_{D' \in \mathrm{supp}(K^D_n(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c)) K^D_n(D'|c') \geq D(\overline{\boldsymbol{p}}_D(c)) = \sum_{D' \in \mathrm{supp}(K^D_n(c'))} D'(\overline{\boldsymbol{p}}_D(c')) K^D_n(D'|c'),$$

for any $c, c' \in C$ with $c > c'$. Moreover, by Lemma 11, the function $K^D_n : C \to \mathcal{D}$ is measurable for all $n \in \mathbb{N}$ and all $D \in \mathrm{supp}(\sigma_n(c))$.

Now define $\hat{\sigma}_n \in \Delta^f(\mathcal{D})^C$ as

$$\hat{\sigma}_n(D'|c) := \sum_{D \in \mathrm{supp}(\sigma_n(c))} K^D_n(D'|c) \sigma_n(D|c), \forall D' \in \bigcup_{D \in \mathrm{supp}(\sigma_n(c))} \mathrm{supp}(K^D_n(c)),$$

for all $c \in C$ and all $n \in \mathbb{N}$. As both functions σ_n and K^D_n are measurable, $\hat{\sigma}_n$ is also measurable. Moreover, by definition, for any $n \in \mathbb{N}$ and any $c \in C$,

$$
\begin{aligned}
\int_{\mathcal{D}} D'\hat{\sigma}_n(\mathrm{d}D'|c) &= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} D'\hat{\sigma}_n(D'|c) \\
&= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} D' \left(\sum_{D \in \mathrm{supp}(\sigma_n(c))} K_n^D(D'|c)\sigma_n(D|c) \right) \\
&= \sum_{D \in \mathrm{supp}(\sigma_n(c))} \left(\sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} D' K_n^D(D'|c) \right) \sigma_n(D|c) \\
&= \sum_{D \in \mathrm{supp}(\sigma_n(c))} D\sigma_n(D|c) \\
&= \int_{\mathcal{D}} D\sigma_n(\mathrm{d}D|c) \\
&= D_n
\end{aligned} \tag{S4}
$$

and hence $\hat{\sigma}_n(c) \in \mathcal{S}_{D_n}$ for all $n \in \mathbb{N}$ and for all $c \in C$. Also, for any $n \in \mathbb{N}$ and any $c \in C$,

$$
\begin{aligned}
&\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_{D'}(c) - \phi_G(c))D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(\mathrm{d}D'|c) \\
=& \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} (\overline{\boldsymbol{p}}_{D'}(c) - \phi_G(c))D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(D'|c) \\
=& \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} (\overline{\boldsymbol{p}}_{D'}(c) - \phi_G(c))D'(\overline{\boldsymbol{p}}_{D'}(c)) \left(\sum_{D \in \mathrm{supp}(\sigma_n(c))} K_n^D(D'|c)\sigma_n(D|c) \right) \\
=& \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} \sum_{D \in \mathrm{supp}(\sigma_n(c))} \left[(\overline{\boldsymbol{p}}_{D'}(c) - \overline{\boldsymbol{p}}_D(c))D'(\overline{\boldsymbol{p}}_{D'}(c)) - (\phi_G(c) - \overline{\boldsymbol{p}}_D(c))D'(\overline{\boldsymbol{p}}_{D'}(c)) \right] K_n^D(D'|c)\sigma_n(D|c) \\
=& \sum_{D \in \mathrm{supp}(\sigma_n(c))} \left(\sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} (\overline{\boldsymbol{p}}_{D'}(c) - \overline{\boldsymbol{p}}_D(c))D'(\overline{\boldsymbol{p}}_{D'}(c))K_n^D(D'|c) \right) \sigma_n(D|c) \\
&+ \sum_{D \in \mathrm{supp}(\sigma_n(c))} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) \left(\sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} D'(\overline{\boldsymbol{p}}_{D'}(c))K_n^D(D'|c) \right) \sigma_n(D|c) \\
=& \sum_{D \in \mathrm{supp}(\sigma_n(c))} \sum_{v \geq \overline{\boldsymbol{p}}_D(c)} (v - \overline{\boldsymbol{p}}_D(c))m^D(v)\sigma_n(D|c) + \sum_{D \in \mathrm{supp}(\sigma_n(c))} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c))D(\overline{\boldsymbol{p}}_D(c))\sigma_n(D|c).
\end{aligned} \tag{S5}
$$

Furthermore, for any $n \in \mathbb{N}$ and any $c \in C$,

$$\begin{aligned}
\int_{\mathcal{D}} D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(\mathrm{d}D'|c) &= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(D'|c) \\
&= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c))} D'(\overline{\boldsymbol{p}}_{D'}(c)) \left(\sum_{D \in \mathrm{supp}(\sigma_n(c))} K_n^D(D'|c)\sigma_n(D|c) \right) \\
&= \sum_{D \in \mathrm{supp}(\sigma_n(c))} \left(\sum_{D' \in \mathrm{supp}(K_n^D(c))} D'(\overline{\boldsymbol{p}}_{D'}(c))K_n^D(D'|c) \right) \sigma_n(D|c) \qquad \text{(S6)} \\
&= \sum_{D \in \mathrm{supp}(\sigma_n(c))} D(\overline{\boldsymbol{p}}_D(c))\sigma_n(D|c) \\
&= \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma_n(\mathrm{d}D|c)
\end{aligned}$$

and for any $c, c' \in C$ such that $c' > c$,

$$\begin{aligned}
\int_{\mathcal{D}} D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(\mathrm{d}D'|c') &= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(D'|c') \\
&= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c)) \left(\sum_{D \in \mathrm{supp}(\sigma_n(c'))} K_n^D(D'|c')\sigma_n(D|c') \right) \\
&= \sum_{D \in \mathrm{supp}(\sigma_n(c'))} \left(\sum_{D' \in \mathrm{supp}(K_n^D(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c))K_n^D(D'|c') \right) \sigma_n(D|c') \qquad \text{(S7)} \\
&\leq \sum_{D \in \mathrm{supp}(\sigma_n(c'))} D(\overline{\boldsymbol{p}}_D(c))\sigma_n(D|c') \\
&= \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma_n(\mathrm{d}D|c'),
\end{aligned}$$

whereas for any $c, c' \in C$ such that $c' < c$,

$$\begin{aligned}
\int_{\mathcal{D}} D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(\mathrm{d}D'|c') &= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c))\hat{\sigma}_n(D'|c') \\
&= \sum_{D' \in \mathrm{supp}(\hat{\sigma}_n(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c)) \left(\sum_{D \in \mathrm{supp}(\sigma_n(c'))} K_n^D(D'|c')\sigma_n(D|c') \right) \\
&= \sum_{D \in \mathrm{supp}(\sigma_n(c'))} \left(\sum_{D' \in \mathrm{supp}(K_n^D(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c))K_n^D(D'|c') \right) \sigma_n(D|c') \qquad \text{(S8)} \\
&\geq \sum_{D \in \mathrm{supp}(\sigma_n(c'))} \left(\sum_{D' \in \mathrm{supp}(K_n^D(c'))} D'(\overline{\boldsymbol{p}}_{D'}(c'))K_n^D(D'|c') \right) \sigma_n(D|c') \\
&= \sum_{D \in \mathrm{supp}(\sigma_n(c'))} D(\overline{\boldsymbol{p}}_D(c'))\sigma_n(D|c') \\
&= \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c'))\sigma_n(\mathrm{d}D|c').
\end{aligned}$$

As such, since $\sigma_n \in \mathcal{S}_{D_n}^C$ is implementable, by Lemma 1, and by (S6), (S7) and (S8),

$$\begin{aligned}&\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\hat{\sigma}_n(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\hat{\sigma}_n(\mathrm{d}D|c')\right)\mathrm{d}z\\ \geq&\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma_n(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma_n(\mathrm{d}D|c')\right)\mathrm{d}z\\ \geq&0,\end{aligned}$$

for all $c, c' \in C$ and for all $n \in \mathbb{N}$, whereas (S6) implies that

$$\int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\hat{\sigma}_n(\mathrm{d}D|z)\right)\mathrm{d}z = \int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma_n(\mathrm{d}D|z)\right)\mathrm{d}z$$

for all $c \in C$ and for all $n \in \mathbb{N}$. Thus, by Lemma 1, there exists $\hat{\tau}_n \in \mathbb{R}^C$ such that $(\hat{\sigma}_n, \hat{\tau}_n)$ is incentive feasible. For each $n \in \mathbb{N}$, let

$$\hat{V}_n(c) := \int_{\mathcal{D}} \pi_D(c)\hat{\sigma}_n(\mathrm{d}D|c) - \hat{\tau}_n(c),\ \forall c \in C,$$

Lemma 1 then implies that $\hat{V}_n$ is nonincreasing and convex for all $n \in \mathbb{N}$ and that $\{\hat{V}_n\}$ is uniformly bounded. Therefore, by Helly's selection theorem, after possibly taking a subsequence,[1] $\{\hat{\sigma}_n\} \to \hat{\sigma}$ pointwise for some $\hat{\sigma} \in \Delta(\mathcal{D})^C$ and $\hat{V}_n \to \hat{V}$ for some nonincreasing and convex function $\hat{V}$. Notice that since the convergence is pointwise and since $\hat{\sigma}_n$ is measurable for all $n \in \mathbb{N}$, the mapping $\hat{\sigma} \in \Delta(\mathcal{D})^C$ is measurable. Moreover, for any $c \in C$, since $\sigma(c) \in \mathcal{S}$ and since $\{\sigma_n\} \to \sigma$, by (S4), for any bounded continuous function $f : V \to \mathbb{R}$,

$$\begin{aligned}\int_V f(v)\left(\int_{\mathcal{D}} D(\mathrm{d}v)\hat{\sigma}(\mathrm{d}D|c)\right) &= \int_V\int_{\mathcal{D}} f(v)D(\mathrm{d}v)\hat{\sigma}(\mathrm{d}D|c)\\ &= \lim_{n\to\infty}\int_V\int_{\mathcal{D}} f(v)D(\mathrm{d}v)\hat{\sigma}_n(\mathrm{d}D|c)\\ &= \lim_{n\to\infty}\int_V f(v)\left(\int_{\mathcal{D}} D(\mathrm{d}v)\hat{\sigma}_n(\mathrm{d}D|c)\right)\\ &= \lim_{n\to\infty}\int_V f(v)\left(\int_{\mathcal{D}} D(\mathrm{d}v)\sigma_n(\mathrm{d}D|c)\right)\\ &= \lim_{n\to\infty}\int_V\int_{\mathcal{D}} f(v)D(\mathrm{d}v)\sigma_n(\mathrm{d}D|c)\\ &= \int_V\int_{\mathcal{D}} f(v)D(\mathrm{d}v)\sigma(\mathrm{d}D|c)\\ &= \int_V f(v)\left(\int_{\mathcal{D}} D(\mathrm{d}v)\sigma(\mathrm{d}D|c)\right)\\ &= \int_V f(v)D_0(\mathrm{d}v),\end{aligned} \tag{S9}$$

where the second equality follows from the fact that $\{\hat{\sigma}_n(c)\} \to \hat{\sigma}(c)$ under the weak-*topology, the fourth equality follows from (S4), the sixth equality follows from the fact that $\{\sigma_n(c)\} \to \sigma(c)$ under the weak-*

[1]Similar to footnote 29, see, for instance, Porter (2005) for a generalized version of Helly's selection theorem. To apply this theorem, notice that the family of functions $\{\hat{\sigma}_n\}$ is of bounded p-variation due to incentive compatibility of the mechanism (σ, τ). Furthermore, for any $c \in C$, the set $\mathrm{cl}(\{\hat{\sigma}_n(c)\})$ is closed in a compact metric space $\Delta(\mathcal{D})$ and hence is itself compact. As such, there exists a pointwise convergent subsequence of $\{\hat{\sigma}_n\}$.

topology, the last equality follows from the fact that $\sigma(c) \in \mathcal{S}$, and the rest equalities are from interchanging the order of integrals. Therefore, by the Riesz representation theorem,

$$\int_{\mathcal{D}} D(p)\hat{\sigma}(\mathrm{d}D|c) = D_0(p),\, \forall p \in V,\, c \in C,$$

which implies that $\hat{\sigma} \in \mathcal{S}^C$. Notice that (S9) also implies $\{D_n\} \to D_0$.

Now, define $\hat{\tau} \in \mathbb{R}^C$ as

$$\hat{\tau}(c) := \int_{\mathcal{D}} \pi_D(c)\hat{\sigma}(\mathrm{d}D|c) - \hat{V}(c),\, \forall c \in C.$$

Since $(\hat{\sigma}_n, \hat{\tau}_n)$ is incentive compatible for all $n \in \mathbb{N}$, for any $c, c' \in C$

$$\begin{aligned}
&\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}(\mathrm{d}D|c) - \hat{\tau}(c)\\
=&\hat{V}(c)\\
=&\lim_{n\to\infty}\left(\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}_n(\mathrm{d}D|c) - \hat{\tau}_n(c)\right)\\
\geq&\lim_{n\to\infty}\left(\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}_n(\mathrm{d}D|c') - \hat{\tau}_n(c')\right)\\
=&\lim_{n\to\infty}\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}_n(\mathrm{d}D|c') - \lim_{n\to\infty}\int_{\mathcal{D}} \pi_D(c')\hat{\sigma}_n(\mathrm{d}D|c') + \lim_{n\to\infty}\hat{V}_n(c')\\
=&\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}(\mathrm{d}D|c') - \int_{\mathcal{D}} \pi_D(c')\hat{\sigma}(\mathrm{d}D|c') + \hat{V}(c')\\
=&\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}(\mathrm{d}D|c) - \hat{\tau}(c'),
\end{aligned}$$

where the second equality follows from the definition of $\hat{\tau}$ and $\hat{V}_n$ and from fact that $\{\hat{V}_n\} \to \hat{V}$ pointwise; the inequality follows from incentive compatibility of $(\hat{\sigma}_n, \hat{\tau}_n, \overline{\boldsymbol{p}})$; the third equality also follows from the definition of $\hat{V}_n$; the fourth equality follows from the fact that $\{\hat{V}_n\} \to \hat{V}$ and $\{\hat{\sigma}_n\} \to \hat{\sigma}$ pointwise and from Lemma 7; and the last equality again follows from the definition of $\hat{\tau}$. Similarly, for any $c \in C$,

$$\begin{aligned}
&\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}(\mathrm{d}D|c) - \hat{\tau}(c)\\
=&\hat{V}(c)\\
=&\lim_{n\to\infty}\left(\int_{\mathcal{D}} \pi_D(c)\hat{\sigma}_n(\mathrm{d}D|c) - \hat{\tau}_n(c)\right)\\
\geq&\lim_{n\to\infty}\pi_{D_n}(c)\\
=&\pi_0(c),
\end{aligned}$$

where the inequality follows from individual rationality of $(\hat{\sigma}_n, \hat{\tau}_n)$ and the last equality follows from Lemma 7 and that $\{D_n\} \to D_0$. Together, the mechanism $(\hat{\sigma}, \hat{\tau})$ is incentive feasible.

Finally, it remains to show that $(\hat{\sigma}, \hat{\tau})$ strictly improves the data broker's revenue. To see this, notice that by Lemma 1, since $(\hat{\sigma}, \hat{\tau})$ is incentive feasible, the data broker's revenue is

$$\int_C \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c))D(\overline{\boldsymbol{p}}_D(c))\hat{\sigma}(\mathrm{d}D|c)G(\mathrm{d}c) - \hat{V}(\bar{c}).$$

Moreover, by possibly adding a constant to the transfer, the expected revenue can be made into

$$\mathbb{E}_G[\hat{\tau}(c)] = \int_C \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \hat{\sigma}(\mathrm{d}D|c) G(\mathrm{d}c) - \bar{\pi}.$$

On the other hand, again by Lemma 1, since (σ, τ) is incentive feasible, the data broker's revenue is

$$\mathbb{E}_G[\tau(c)] = \int_C \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) G(\mathrm{d}c) - U(\bar{c}),$$

where $U \in \mathbb{R}^C$ is the producer's indirect utility under the mechanism (σ, τ). Furthermore, for all $n \in \mathbb{N}$, since $\sigma_n \in \mathcal{S}^C_{D_n}$ is implementable, there exists $\tau_n \in \mathbb{R}^C$ such that (σ_n, τ_n) is incentive feasible and hence, by Lemma 1,

$$\mathbb{E}_G[\tau_n(c)] = \int_C \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma_n(\mathrm{d}D|c) G(\mathrm{d}c) - U_n(\bar{c}), \tag{S10}$$

where U_n is the indirect utility of the producer under (σ_n, τ_n) and $\{U_n(\bar{c})\} \to U(\bar{c})$. By Lemma 1 again, U_n is nonincreasing and convex while $\{U_n\}$ is uniformly bounded. Therefore, by Helly's selection theorem, after possibly taking a subsequence, $\{U_n\} \to U$ pointwise. Together with $\{\sigma_n\} \to \sigma$, this implies that $\{\tau_n\} \to \tau$ pointwise. As a result,

$$\begin{aligned}
&\mathbb{E}_G[\hat{\tau}(c)] + \bar{\pi} \\
=& \int_C \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \hat{\sigma}(\mathrm{d}D|c) G(\mathrm{d}c) \\
\geq& \lim_{n\to\infty} \int_C \left(\int_{\mathcal{D}} \pi_D(c) \hat{\sigma}_n(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \liminf_{n\to\infty} \int_C (\phi_G(c) - c) \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \hat{\sigma}_n(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\
=& \limsup_{n\to\infty} \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \hat{\sigma}_n(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\
=& \limsup_{n\to\infty} \left[\int_C \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) \sigma_n(\mathrm{d}D|c) G(\mathrm{d}c) + \int_C \int_{\mathcal{D}} \int_{\{v \geq \overline{\boldsymbol{p}}_D(c)\}} (v - \overline{\boldsymbol{p}}_D(c)) D(\mathrm{d}v) \sigma_n(\mathrm{d}D|c) G(\mathrm{d}c) \right] \\
\geq& \lim_{n\to\infty} (\mathbb{E}_G[\tau_n(c)] + U_n(\bar{c})) + \liminf_{n\to\infty} \int_C \int_{\mathcal{D}} \int_{\{v \geq \overline{\boldsymbol{p}}_D(c)\}} (v - \overline{\boldsymbol{p}}_D(c)) D(\mathrm{d}v) \sigma_n(\mathrm{d}D|c) G(\mathrm{d}c) \\
\geq& \mathbb{E}_G[\tau(c)] + U(\bar{c}) + \int_C \int_{\mathcal{D}} \int_{\{v \geq \overline{\boldsymbol{p}}_D(c)\}} (v - \overline{\boldsymbol{p}}_D(c)) D(\mathrm{d}v) \sigma(\mathrm{d}D|c) G(\mathrm{d}c) \\
>& \mathbb{E}_G[\tau(c)] + U(\bar{c}),
\end{aligned}$$

where the first inequality follows from the fact that $\{\hat{\sigma}_n(c)\} \to \hat{\sigma}(c)$ for all $c \in C$, Lemma 7 and Lemma 9; the second equality follows from Lemma 7 and properties of the lim inf and lim sup operators;[2] the third equality follows from (S5); the second inequality follows (S10); the third inequality follows from the fact that $\{\tau_n\} \to \tau$ and $\{U_n(\bar{c})\} \to U(\bar{c})$, the Lebesgue dominated convergence theorem, as well as the fact that $\{\sigma_n\} \to \sigma$ pointwise and that the integrand of the second term is lower-semicontinuous in D, which can be shown by using exactly the same arguments as in the proof of Lemma 10; and the last inequality follows from (S3). Therefore, since $U(\bar{c}) \geq \bar{\pi}$ as (σ, τ) is individually rational,

$$\mathbb{E}_G[\hat{\tau}(c)] > \mathbb{E}_G[\tau(c)].$$

Together, $(\hat{\sigma}, \hat{\tau})$ is an incentive feasible mechanism that strictly improves the data broker's revenue, this completes the proof. ■

[2] See footnote 27.

S.3 Optimal Mechanism in General Environments

Theorem S2 (Generalized Optimal Mechanism)**.** *Suppose that $D_0 \in \mathcal{D}$ is continuous. Then the set of optimal mechanisms is exactly the set of φ^*-quasi-perfect mechanisms for some nondecreasing function $\varphi^* \in \mathbb{R}_+^C$.*

The proof of Theorem S2 requires three main steps. Let $\boldsymbol{q}^*$ denote the solution of the price-controlling data broker's (reduced) problem given by (28) in the main text. **Step 1** applies a strong duality argument to the price-controlling data broker's problem and back out the associated Lagrange multiplier $\bar{M}$ for the individual rationality constraint of the price-controlling data broker's problem. **Step 2** then constructs the candidate cutoff function φ^* by modifying $\boldsymbol{q}^*$ and shows that there is a φ^*-quasi-perfect mechanism for the data broker that is incentive feasible. Finally, **Step 3** constructs the Lagrange multiplier for the incentive compatibility constraints. Together with $\bar{M}$, these two Lagrange multipliers warrants the candidate mechanism constructed in **Step 2**, confirming optimality. Finally, by repeating the same arguments as in the proof of Theorem 1, it then follows that any optimal mechanism must be a φ^*-quasi-perfect mechanism.

Step 1: A Strong Duality Result for the Price-Controlling Data Broker's Problem

Lemma S3. *Given any continuous $D_0 \in \mathcal{D}$. There exists a nondecreasing function $\psi^* \in \mathbb{R}_+^C$ with $\psi^*(c) \geq c$ for all $c \in C$ and a Borel measure $\bar{\mu}$ on C with associated CDF $\bar{M}$ such that:*

1. *$D_0 \circ \psi^*$ is a solution of the for the price-controlling data broker's problem (28).*

2. *ψ^* is a solution of the transformed primal problem*

$$\begin{aligned} &\max_{\psi\in\Psi} \int_C \left(\int_{\{v \geq \psi(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) - \bar{\pi} \\ &\text{s.t. } \bar{\pi} + \int_c^{\bar{c}} D_0(\psi(z))\,\mathrm{d}z \geq \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in C \end{aligned} \tag{S11}$$

   where $\Psi \subseteq \mathbb{R}_+^C$ is the collection of all nonnegative and nondecreasing functions on C.

3. *ψ^* solves*

$$\begin{aligned} \max_{\psi\in\Psi} \Bigg[&\int_C \left(\int_{\{v \geq \psi(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) \\ &+ \int_C \left(\int_c^{\bar{c}} D_0(\psi(z))\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}(\mathrm{d}c) \Bigg], \end{aligned} \tag{S12}$$

4. *The complementary slackness condition holds. That is,*

$$\int_C \bar{M}(c)(D_0(\psi^*(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c = 0. \tag{S13}$$

Proof. First recall that by Lemma 16 and (5), the price-controlling data broker's problem (28) can be written as

$$\max_{\boldsymbol{q}\in\mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y\right) G(\mathrm{d}c) - \bar{\pi} \tag{S14}$$
$$\text{s.t. } \bar{\pi} + \int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z \geq \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in C,$$

where $\mathcal{Q}$ is the collection of nonincreasing functions in $[0,1]^C$. By Lemma 16, this problem has a solution.

For any $\varepsilon > 0$, consider the following ε-relaxation of the price-controlling data broker's problem (S14).

$$\max_{\boldsymbol{q}\in\mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y\right) G(\mathrm{d}c) \tag{S15}$$
$$\text{s.t. } \int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z \geq \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in [\underline{c}, \bar{c} - \varepsilon].$$

Let $\boldsymbol{q}_\varepsilon$ be the solution to (S15). By Helly's selection theorem, after possibly taking a subsequence, $\{\boldsymbol{q}_\varepsilon\} \to \boldsymbol{q}^*$ pointwise for some nonincreasing function $\boldsymbol{q}^* \in [0,1]^C$ as $\varepsilon \to 0$. On the other hand, notice that the objective function of (S15) is a concave functional of $\boldsymbol{q}$ and the choice set of (S15) is a compact and convex set in the normed linear space $L^1(C)$ and there exists some $\boldsymbol{q}_0$ in the choice set $\mathcal{Q}$ such that

$$\int_c^{\bar{c}} \boldsymbol{q}_0(z)\,\mathrm{d}z > \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in [\underline{c}, \bar{c} - \varepsilon]$$

(for instance, $\boldsymbol{q}_0(c) = 1$ for all $c \in C$). As a result, by the strong duality theorem (c.f., Luenberger (1969), p.224, Theorem 1), there exists a Borel measure $\bar{\mu}_\varepsilon$ defined on $[\underline{c}, \bar{c} - \varepsilon]$ such that

$$\boldsymbol{q}_\varepsilon \in \operatorname*{argmax}_{\boldsymbol{q}\in\mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y\right) G(\mathrm{d}c) + \int_C \left(\int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z\right) \bar{\mu}_\varepsilon(\mathrm{d}c), \tag{S16}$$

and that

$$\int_C \left(\int_c^{\bar{c}} q_\varepsilon(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z\right) \bar{\mu}_\varepsilon(\mathrm{d}c) = 0. \tag{S17}$$

For any $\varepsilon > 0$, let $\bar{M}_\varepsilon$ be the CDF associated with $\bar{\mu}_\varepsilon$. Since $\bar{M}_\varepsilon$ is nondecreasing for all $\varepsilon > 0$, by Helly's selection theorem, after possibly taking a subsequence, $\{\bar{M}_\varepsilon\}$ converges pointwise to some nondecreasing function M. Let $\bar{M}$ be the right-limit of M and let $\bar{\mu}$ be the Borel measure associated with $\bar{M}$.

Now I show that that $\boldsymbol{q}^*$ is a solution of (S14), that $\boldsymbol{q}^*$ solves

$$\max_{\boldsymbol{q}\in\mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y\right) + \int_C \left(\int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z)\,\mathrm{d}z)\right) \bar{\mu}(\mathrm{d}c), \tag{S18}$$

and that

$$\int_C \bar{M}(c)(\boldsymbol{q}(c) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c = 0 \tag{S19}$$

To see (S19), since $\{\boldsymbol{q}_\varepsilon\} \to \boldsymbol{q}^*$ and $\{\bar{M}_\varepsilon\} \to \bar{M}$ Lenesgue-almost everywhere, by the Lebesgue dominated convergence theorem,

$$\begin{aligned}
&\int_C \left(\int_c^{\bar{c}} \boldsymbol{q}^*(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}(\mathrm{d}c) \\
=& \int_C \bar{M}(c)(\boldsymbol{q}^*(c) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c \\
=& \lim_{\varepsilon \to 0} \int_C \bar{M}_\varepsilon(c)(\boldsymbol{q}_\varepsilon(c) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c \\
=& \lim_{\varepsilon \to 0} \int_C \left(\int_c^{\bar{c}} \boldsymbol{q}_\varepsilon(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}_\varepsilon(\mathrm{d}c) \\
=& 0,
\end{aligned}$$

where the first and the third equalities follows from interchanging the order of integrals, the second equality results from the Lebesgue dominated convergence theorem and the last equality follows from (S17).

On the other hand, to show that $\boldsymbol{q}^*$ solves (S18), notice that by interchanging the order of integrals, for any nonincreasing function $\boldsymbol{q} \in \mathcal{Q}$,

$$\begin{aligned}
&\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y \right) G(\mathrm{d}c) + \int_C \left(\int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}_\varepsilon(\mathrm{d}c) \\
=& \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_\varepsilon^{\bar{M}}(c))\,\mathrm{d}y \right) - \int_C \left(\int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}_\varepsilon(\mathrm{d}c)
\end{aligned}$$

for any $\varepsilon > 0$, where $\phi_\varepsilon^{\bar{M}}(c) := \phi_G(c) - \bar{M}_\varepsilon(c)/g(c)$ for all $c \in C$. Similarly, for any nonincreasing function $\boldsymbol{q} \in [0,1]^C$,

$$\begin{aligned}
&\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y \right) G(\mathrm{d}c) + \int_C \left(\int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}(\mathrm{d}c) \\
=& \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \psi^{\bar{M}}(c))\,\mathrm{d}y \right) - \int_C \left(\int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}(\mathrm{d}c),
\end{aligned}$$

where $\phi^{\bar{M}}(c) := \phi_G(c) - \bar{M}(c)/g(c)$ for all $c \in C$. As such,

$$\begin{aligned}
&\operatorname*{argmax}_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y \right) G(\mathrm{d}c) + \int_C \left(\int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}_\varepsilon(\mathrm{d}c) \\
=& \operatorname*{argmax}_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_\varepsilon^{\bar{M}}(c))\,\mathrm{d}y \right)
\end{aligned}$$

and

$$\begin{aligned}
&\operatorname*{argmax}_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y \right) G(\mathrm{d}c) + \int_C \left(\int_c^{\bar{c}} q(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}(\mathrm{d}c) \\
=& \operatorname*{argmax}_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi^{\bar{M}}(c))\,\mathrm{d}y \right).
\end{aligned} \tag{S20}$$

Moreover, since $\{\bar{M}_\varepsilon\} \to \bar{M}$ pointwise almost everywhere, $\{\phi_\varepsilon^{\bar{M}}\} \to \psi^{\bar{M}}$ pointwise almost everywhere as well. Now suppose that

$$\boldsymbol{q}^* \notin \operatorname*{argmax}_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi^{\bar{M}}(c))\,\mathrm{d}y \right).$$

That is, there exists a nonincreasing function $\boldsymbol{q} \in \mathcal{Q}$ such that

$$\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}c) > \int_C \left(\int_0^{\boldsymbol{q}^*(c)} (D_0^{-1}(y) - \phi^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}c).$$

Then since, by the Lebesgue dominated convergence theorem,

$$\lim_{\varepsilon \to 0} \int_C \left(\int_0^{\boldsymbol{q}_\varepsilon(c)} (D_0^{-1} - \phi_\varepsilon^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}c) = \int_C \left(\int_0^{\boldsymbol{q}^*(c)} (D_0^{-1}(y) - \phi^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}v)$$

and

$$\lim_{\varepsilon \to 0} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_\varepsilon^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}c) = \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}c),$$

for $\varepsilon > 0$ small enough,

$$\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi_\varepsilon^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}c) > \int_C \left(\int_0^{\boldsymbol{q}_\varepsilon(c)} (D_0^{-1}(y) - \phi_\varepsilon^{\bar{M}}(c))\,\mathrm{d}y \right) G(\mathrm{d}c),$$

a contradiction. As such, it must be that

$$\boldsymbol{q}^* \in \operatorname*{argmax}_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(y) - \phi^{\bar{M}}(c))\,\mathrm{d}y \right).$$

Together with (S20), it must be that $\boldsymbol{q}^*$ solves (S12).

Notice that by the Lebesgue dominated convergence theorem and by the fact that $\boldsymbol{q}_\varepsilon$ is feasible in the primal problem (S15), for all $c \in C$

$$\begin{aligned} &\int_c^{\bar{c}} \boldsymbol{q}^*(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \\ =&\lim_{\varepsilon \to 0} \int_c^{\bar{c}} \boldsymbol{q}_\varepsilon(z)\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \\ \geq& 0. \end{aligned}$$

Together with (S12) and (S13), the weak duality argument implies that $\boldsymbol{q}^*$ is a solution to the maximization problem (S14).

Now define $\psi^* \in \mathbb{R}_+^C$ as

$$\psi^*(c) := D_0^{-1}(\boldsymbol{q}^*(c)), \forall c \in C.$$

Then, since D_0 is continuou, $\boldsymbol{q}^* \equiv D_0 \circ \psi^*$, which proves assertion 1. Moreover, for any nondecreasing function $\psi \in \mathbb{R}_+^C$, $D_0 \circ \psi \in \mathcal{Q}$ and

$$\int_{\{v \geq \psi(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) = \int_0^{D_0 \circ \psi(c)} (D_0^{-1}(y) - \phi_G(c))\,\mathrm{d}y, \forall c \in C.$$

Therefore, by (S18), ψ^* solves

$$\max_{\psi \in \Psi} \int_C \left(\int_{\{v \geq \psi(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) + \int_C \left(\int_c^{\bar{c}} D_0(\psi(z))\,\mathrm{d}z - \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \right) \bar{\mu}(\mathrm{d}c),$$

which proves assertion 3. Also, by(S19),

$$\int_C \bar{M}(c)(D_0(\psi^*(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c = 0,$$

which proves assertion 4. Finally, since D_0 is continuous, the price-controlling data broker's problem (S14) can be written as

$$\begin{aligned}&\max_{\psi\in\Psi} \int_C \left(\int_{\{v\geq\psi(c)\}} (v-\phi_G(c))D_0(\mathrm{d}v)\right) G(\mathrm{d}c) - \bar{\pi}\\ &\text{s.t. } \bar{\pi} + \int_c^{\bar{c}} D_0(\psi(z))\,\mathrm{d}z \geq \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\end{aligned}$$

which then implies that ψ^* also solves (S11), proving assertion 2, which in turn implies that $\psi^*(c) \geq c$ for all $c \in C$. This completes the proof. ■

Step 2: Constructing a Candidate Cutoff Function φ^*

To construct an optimal mechanism for the data broker from the price-controlling data broker's optimal mechanism, the pricing constraints (i.e., $\boldsymbol{p} \in \boldsymbol{P}$) and the double-deviation constraints (i.e., condition 2 in Lemma 1) need to be accommodated. To this end, Lemma S4 below extends Lemma 3 and constructions a $\psi(c)$-quasi-perfect segmentation for c, even when $\psi(c) > \overline{\boldsymbol{p}}_0(c)$. The decomposition is a modification of the output-minimizing decomposition in Bergemann et al. (2013). The modification is made to accommodate double deviation constraints as well. To state this lemma, first define a crucial function $\psi_0 \in V^C$ as follows.

$$\psi_0(c) := \sup\left\{\hat{v} \in V : \int_{\{v\geq\hat{v}\}} (v-c)D_0(\mathrm{d}v) \geq \pi_0(c)\right\}, \forall c \in C. \tag{S21}$$

Clearly, $\psi_0(c) \geq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$.

Lemma S4. *For any step function $D_0 \in \mathcal{D}$, any $c \in C$ and any ψ such that $\overline{\boldsymbol{p}}_0(c) \leq \psi \leq \psi_0(c)$, there exists a segmentation $s \in \mathcal{S}$ such that:*

1. $\sum_{D\in\mathrm{supp}(s)} Ds(D) = D_0$.
2. *For each $D \in \mathrm{supp}(s)$, the set*
$$\{v \in \mathrm{supp}(D) : v \geq \psi\}$$
is nonempty and is a singleton.
3. *For each $D \in \mathrm{supp}(s)$, $\max(\mathrm{supp}(D)) \in \boldsymbol{P}_D(c)$.*
4. *For any $z \in [\underline{c}, c]$ and for any $D \in \mathrm{supp}(s)$,*
$$\overline{\boldsymbol{p}}_D(z) \geq \overline{\boldsymbol{p}}_0(z)$$

Proof. Given any step function $D_0 \in \mathcal{D}$, any $c \in C$ and any ψ such that $\overline{\boldsymbol{p}}_0(c) \le \psi \le \psi_0(c)$. Let $m^0 := m^{D_0}$ and define $\{\hat{m}^w\}_{w\ge\psi}$ recursively as follows

$$\hat{m}^w(v) := \begin{cases} 0, & \text{if } v \ge \psi \text{ and } v \ne w \\ m^0(v), & \text{if } v = w \\ \beta(w,v)\hat{m}^w(v), & \text{if } \overline{\boldsymbol{p}}_0(c) \le v \le \psi \\ \alpha(w)\hat{m}^w(v), & \text{if } v < \overline{\boldsymbol{p}}_0(c) \end{cases}, \forall v \in \text{supp}(D_0),\ \forall w \in \text{supp}(D_0) \text{ s.t. } w \ge \psi,$$

where

$$\beta(w,v) := \frac{(w-c)m^0(w) - (v-c)\sum_{v'>v}\hat{m}^w(v')}{\sum_{w\ge\psi}\left[(w-c)m^0(w) - (v-c)\sum_{v'>v}\hat{m}^w(v')\right]}, \forall w, v \in \text{supp}(s) \text{ s.t. } x \ge \psi,\ \overline{\boldsymbol{p}}_0(c) \le v \le \psi.$$

and

$$\alpha(w) := \frac{\sum_{v'\ge\overline{\boldsymbol{p}}_0(c)}\hat{m}^w(v')}{\sum_{v'\ge\overline{\boldsymbol{p}}_0(c)} m^0(v')}, \forall w \in \text{supp}(D_0) \text{ s.t. } w \ge \psi.$$

By construction,

$$\sum_{w\ge\psi}\alpha(w) = \sum_{w\ge\psi}\beta(w,v) = 1$$

for all $v \in \text{supp}(D_0)$ with $\overline{\boldsymbol{p}}_0(c) \le v \le \psi$. As such,

$$\sum_{w\ge\psi}\hat{m}^w(v) = m^0(v),\ \forall v \in \text{supp}(D_0). \tag{S22}$$

Notice that for any $v \in \text{supp}(D_0)$ with $\overline{\boldsymbol{p}}_0(c) \le v \le \psi$,

$$\begin{aligned} &\sum_{w\ge\psi}\left[(w-c)m^0(w) - (v-c)\sum_{v'>v}\hat{m}^w(v')\right] \\ =&\sum_{v'\ge\psi}(v'-c)m^0(v') - (v-c)\sum_{v'>v}m^0(v') \\ \ge&(\overline{\boldsymbol{p}}_0(c)-c)\sum_{v'\ge\overline{\boldsymbol{p}}_0(c)}m^0(v') - (v-c)\sum_{v'>v}m^0(v') \\ \ge&(v-c)\sum_{v'\ge v}m^0(v') - (v-c)\sum_{v'>v}m^0(v') \\ =&(v-c)m^0(v), \end{aligned}$$

where the first equality follows from (S22), the first inequality follows from the definition of $\psi_0(c)$ and that $\psi \le \psi_0(c)$, the second inequality follows from the fact that $\overline{\boldsymbol{p}}_0(c) \in \boldsymbol{P}_0(c)$. As such, for any $v \in \text{supp}(D_0)$ with $\overline{\boldsymbol{p}}_0(c) \le v \le \psi$ and for any $w \in \text{supp}(D_0)$ with $w \ge \psi$,

$$\begin{aligned} &\hat{m}^w(v) \le \frac{(w-c)m^0(w) - (v-c)\sum_{v'>v}\hat{m}^w(v')}{(v-c)m^0(v)}m^0(v) \\ \Longleftrightarrow\ &(v-c)\hat{m}^w(v) + (v-c)\sum_{v'>v}\hat{m}^w(v') \le (w-c)m^0(w) \\ \Longleftrightarrow\ &(v-c)\sum_{v'\ge v}\hat{m}^w(v') \le (w-c)\hat{m}^w(w), \end{aligned} \tag{S23}$$

which also, inductively, implies that $\beta(w,v) \geq 0$ for all $w \geq \psi$ and for all v such that $\overline{\boldsymbol{p}}_0(c) \leq v \leq \psi$ (see details from the proof of Lemma 3). On the other hand, for any $v \in \text{supp}(D_0)$ with $v < \overline{\boldsymbol{p}}_0(c)$ and any $w \in \text{supp}(D_0)$ with $w \geq \psi$, Notice that by the definition of $\alpha(w)$,

$$\sum_{v' \geq v} \hat{m}^w(v') = \alpha(w) \sum_{v \leq v' \overline{\boldsymbol{p}}_0(c)} m^0(v) + \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} \hat{m}^w(c') = \alpha(w) \sum_{v' \geq v} m^0(v'). \tag{S24}$$

Also, for any $v \in \text{supp}(D_0)$ with $v < \overline{\boldsymbol{p}}_0(c)$ and any $w \in \text{supp}(D_0)$ with $w \geq \psi$,

$$\begin{aligned}
&(\overline{\boldsymbol{p}}_0(c) - c) \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} m^0(v') \geq (v - c) \sum_{v' \geq v} m^0(v') \\
\iff &(v - c) \sum_{v \leq v' \leq \overline{\boldsymbol{p}}_0(c)} m^0(v') \leq (\overline{\boldsymbol{p}}_0(c) - v) \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} m^0(v') \\
\iff &(v - c) \sum_{v \leq v' \leq \overline{\boldsymbol{p}}_0(c)} m^0(v') \frac{\sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} \hat{m}^w(v')}{\sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} m^0(v')} \leq (\overline{\boldsymbol{p}}_0(c) - v) \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} \hat{m}^w(v') \qquad \text{(S25)} \\
\iff &\alpha(w)(v - c) \sum_{v \leq v' \leq \overline{\boldsymbol{p}}_0(c)} m^0(v') \leq (\overline{\boldsymbol{p}}_0(c) - v) \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} \hat{m}^w(v').
\end{aligned}$$

Together, these imply that

$$\begin{aligned}
&(w - c)\hat{m}^w(v) \qquad \text{(S26)} \\
\geq &\alpha(w)(v - c) \sum_{v \leq v' \leq \overline{\boldsymbol{p}}_0(c)} m^0(v') - (\overline{\boldsymbol{p}}_0(c) - v) \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} \hat{m}^w(v') + (\overline{\boldsymbol{p}}_0(c) - c) \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} \hat{m}^w(v') \\
= &\alpha(w)(v - c) \sum_{v \leq v' \leq \overline{\boldsymbol{p}}_0(c)} m^0(v') + (v - c) \sum_{v' \geq \overline{\boldsymbol{p}}_0(c)} \hat{m}^w(v') \\
= &\alpha(w)(v - c) \sum_{v' \geq v} m^0(v') \qquad \text{(S27)} \\
= &(v - c) \sum_{v' \geq v} \hat{m}^w(v'),
\end{aligned}$$

where the inequality follows from (S23) and (S25), the second and the third equality follows from (S24). Moreover, by (S24) for any $v \in \text{supp}(D_0)$ with $v \leq \overline{\boldsymbol{p}}_0(c)$, for any $z \in [\underline{c}, c]$, and any $w \in \text{supp}(D_0)$ with $w \geq \psi$,

$$\begin{aligned}
&(w - z)\hat{m}^w(w) < (v - z) \sum_{v' \geq v} \hat{m}^w(v') \\
\Rightarrow &(\overline{\boldsymbol{p}}_0(z) - z) \sum_{v' \geq \overline{\boldsymbol{p}}_0(z)} \hat{m}^w(v') \geq (v - z) \sum_{v' \geq v} \hat{m}^w(v'), \forall v \in \text{supp}(\hat{m}^w). \qquad \text{(S28)}
\end{aligned}$$

Now for each $w \in \text{supp}(D_0)$ with $w \geq \psi$, let

$$m^w(v) := \frac{\hat{m}^w(v)}{\sum_{v \in V} \hat{m}^w(v)}$$

and let

$$D_w(p) := m^w([p, \overline{v}])$$

and define $s \in \Delta^f(\mathcal{D})$ by

$$s(D_w) := \sum_{v \in V} \hat{m}^w(v), \, \forall w \in \text{supp}(D_0), \, w \geq \psi.$$

By (S22), and by the facts that $\alpha \geq 0$ and $\beta \geq 0$, such s satisfies assertion 1 and therefore is indeed a segmentation. Furthermore, by construction of $\hat{m}^w$, s satisfies assertion 2. In addition, since m^w is proportional to $\hat{m}^w$ for all $w \in \text{supp}(D_0)$ with $w \geq \psi$, (S23) and (S26) ensures that s satisfies assertion 3. Finally, (S28) and the same proportionality imply that s satisfies assertion 4. This completes the proof. ■

Lemma S5. *Given any $D_0 \in \mathcal{D}$ and G, for any nondecreasing function $\psi \in \mathbb{R}_+^C$ such that $c \leq \psi(c) \leq \psi_0(c)$ for all $c \in C$, there exists a ψ-quasi-perfect scheme $\sigma \in \mathcal{S}^C$ such that*

$$\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c) \leq D_0(\min\{\overline{\boldsymbol{p}}_0(z), \psi(z)\}) \right) \mathrm{d}z \geq 0, \tag{S29}$$

for all $c, c' \in C$ with $c < c'$. In particular,

$$\int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) = \int_{\{v \geq \psi(c)\}} v D_0(\mathrm{d}v), \tag{S30}$$

for Lebesgue-almost all $c \in C$ and

$$\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) = D_0(\psi(c)), \tag{S31}$$

for Lebesgue-almost all $c \in C$.

Proof. Consider first the case where D_0 is a step function. Combining Lemma 11 and Lemma S4, for each $c \in C$, there exists a finite $\psi(c)$-quasi-perfect segmentation for c, $\sigma(c) \in \mathcal{S}$. As such, for any $D \in \text{supp}(s)$,

$$D(\overline{\boldsymbol{p}}_D(c)) = m^D(\max(\text{supp}(D)))$$

and

$$\overline{\boldsymbol{p}}_D(c) D(\overline{\boldsymbol{p}}_D(c)) = \max(\text{supp}(D)) m^D(\max(\text{supp}(D))).$$

Since for any $D \in \text{supp}(\sigma(c))$, m^D has positive probability on one and only one value that is above ψ, by Bayes' plausibility of $\sigma(c)$,

$$\sum_{\{D \in \text{supp}(\sigma(c)):\max(\text{supp}(D))=v\}} m^D(v)\sigma(D|c) = m^0(v), \, \forall v \in \text{supp}(D_0), \, v \geq \psi(c).$$

Therefore,

$$\begin{aligned}
\sum_{D \in \text{supp}(\sigma(c))} D(\overline{\boldsymbol{p}}_D(c)) &= \sum_{v \geq \psi(c)} \left(\sum_{\{D \in \text{supp}(\sigma(c)):\max(\text{supp}(D))=v\}} D(\overline{\boldsymbol{p}}_D(c))\sigma(D|c) \right) \\
&= \sum_{v \geq \psi(c)} \left(\sum_{\{D \in \text{supp}(\sigma(c)):\max(\text{supp}(D))=v\}} m^D(v)\sigma(D|c) \right) \\
&= \sum_{v \geq \psi(c)} m^0(v) \\
&= D_0(\psi(c))
\end{aligned} \tag{S32}$$

and

$$\sum_{D\in\text{supp}(\sigma(c))}\overline{\boldsymbol{p}}_D(c)D(\overline{\boldsymbol{p}}_D(c))\sigma(D|c)=\sum_{v\geq\psi(c)}\left(\sum_{\{D\in\text{supp}(\sigma(c)):\max(\text{supp}(D))=v\}}\overline{\boldsymbol{p}}_D(c)D(\overline{\boldsymbol{p}}_D(c))\sigma(D|c)\right)$$
$$=\sum_{v\geq\psi(c)}\left(v\sum_{\{D\in\text{supp}(\sigma(c)):\max(\text{supp}(D))=v\}}m^D(v)\sigma(D|c)\right) \tag{S33}$$
$$=\sum_{v\geq\psi(c)}vm^0(v).$$

Furthermore, for each $c\in C$ if $\psi(c)\geq\overline{\boldsymbol{p}}_0(c)$, assertion 4 of Lemma S4 implies that

$$\sum_{D\in\text{supp}(\sigma(c))}D(\overline{\boldsymbol{p}}_D(z))\sigma(D|c)\leq\sum_{D\in\text{supp}(\sigma(c))}D(\overline{\boldsymbol{p}}_0(z))\sigma(D|c)=D_0(\overline{\boldsymbol{p}}_0(z)),\,\forall z\in[\underline{c},c].$$

On the other hand, if $\psi(c)\leq\overline{\boldsymbol{p}}_0(c)$, Lemma 3 ensures that

$$\sum_{D\in\text{supp}(\sigma(c))}D(\overline{\boldsymbol{p}}_D(z))\sigma(D|c)\leq\sum_{D\in\text{supp}(\sigma(c))}D(\overline{\boldsymbol{p}}_0(z))\sigma(D|c)=D_0(\psi(z)),\,\forall z\in[\underline{c},c].$$

Together, since ψ is nondecreasing,

$$\sum_{D\in\text{supp}(\sigma(c))}D(\overline{\boldsymbol{p}}_D(z))\sigma(D|c)\leq D_0(\min\{\overline{\boldsymbol{p}}_0(z),\psi(z)\}),\,\forall z\in[\underline{c},c]. \tag{S34}$$

Now consider any $D_0\in\mathcal{D}$. Since the set of step functions in $\mathcal{D}$ is dense in $\mathcal{D}$, there exists sequence of step functions $\{D_n\}\subset\mathcal{D}$ such that $\{D_n\}\to D_0$. For each $n\in\mathbb{N}$ and for each $c\in C$, let $\sigma_n(c)\in\mathcal{S}$ be the ψ-quasi-perfect segmentation for c defined above. By Helly's selection theorem, after possibly taking a subsequence, $\{\sigma_n\}\to\sigma$ pointwise for some $\sigma:C\to\Delta(\mathcal{D})$.[3] By Lemma 13, it follows that $\sigma\in\mathcal{S}^C$ and that σ is a ψ-quasi-perfect scheme. In particular, for Lebesgue-almost all $c\in C$,

$$\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c)=D_0(\psi(c)),$$

Similarly, by Lemma 11, for all $c\in C$,

$$\int_{\mathcal{D}}\overline{\boldsymbol{p}}_D(c)D(\overline{\boldsymbol{p}}_D(c))=\int_{\{v\geq\psi(c)\}}vD_0(\mathrm{d}v)$$

Finally, by (S34), Lemma 9, Lemma 10 and the reversed Fatou's lemma, for any $c,c'\in C$ with $c<c'$,

$$\int_c^{c'}\left(\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c)\right)\mathrm{d}z=\liminf_{n\to\infty}\int_c^{c'}\left(\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D(z))\sigma_n(\mathrm{d}D|c)\right)\mathrm{d}z$$
$$\leq\limsup_{n\to\infty}\int_c^{c'}D_n(\min\{\overline{\boldsymbol{p}}_{D_n}(z),\psi(z)\})\,\mathrm{d}z$$
$$\leq\int_c^{c'}D_0(\min\{\overline{\boldsymbol{p}}_{D_0}(z),\psi(z)\})\,\mathrm{d}z.$$

This completes the proof. ■

[3]Notice that the function $\sigma_n:C\to\Delta^f(\Delta(\mathcal{D}))$ is of p-bounded variation. See footnote 29 for details.

With Lemma S4 and Lemma S5, the candidate cutoff function φ^* can then be constructed. To begin with. Notice that since $\boldsymbol{q}^* \equiv D_0 \circ \psi^*$ is a solution of the price-controlling data broker's problem (28), the individual rationality constraint implies that

$$\int_c^{\bar{c}} D_0(\psi^*(z))\,\mathrm{d}z \geq \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in C. \tag{S35}$$

On the other hand, since $\bar{M}$ is nondecreasing,

$$c^* := \inf\{c \in C : \bar{M}(c) > 0\}$$

is well-defined and $\bar{M}(c) > 0$ for all $c > c^*$ and $\bar{M}(c) = 0$ for all $c < c^*$. As such, by (S12) and by the proof of Theorem 1, for any $c \in [\underline{c}, c^*]$, $\psi^*(c) = \varphi_G(c)$.

Furthermore, as the functions

$$c \mapsto \int_c^{\bar{c}} D_0(\psi^*(z))\,\mathrm{d}z$$

and

$$c \mapsto \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z$$

are continuous on C, the (possibly empty) set

$$C_s := \left\{c \in (c^*, \bar{c}] : \int_c^{\bar{c}} D_0(\psi^*(z))\,\mathrm{d}z > \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z\right\}$$

can be written as a countable union of disjoint open intervals.

$$C_s = \bigcup_{j=1}^{\infty} (\underline{c}_j, \bar{c}_j).$$

Finally, by (S13), for any $j \in \mathbb{N}$, $\bar{M}(c) = \bar{M}_j$ for all $c \in (\underline{c}_j, \bar{c}_j)$, for some $\bar{M}_j \geq 0$ and

$$\int_{\underline{c}_j}^{\bar{c}_j} D_0(\psi^*(z))\,\mathrm{d}z = \int_{\underline{c}_j}^{\bar{c}_j} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z.$$

Together, since $\bar{M}$ is right-continuous, there exists $\tilde{c} \in C_s$ such that

$$(c^*, \tilde{c}) \cap C_s = \emptyset.$$

As such, it is without loss of generality to assume that $\underline{c}_{j+1} \geq \bar{c}_j$ for all $j \in \mathbb{N}$.

Now define a sequence of functions $\{\varphi_j\}$ recursively according to the following procedure. Let $\varphi_0 := \psi^*$. For each $j \in \mathbb{N}$, given a function φ_{j-1}, with the properties that

$$\varphi_{j-1}(c) = \psi^*(c),\ \forall c \notin C_s \cap [\cup_{l=1}^{j-1} (\underline{c}_l, \bar{c}_l)] \tag{S36}$$

and that

$$\int_c^{\bar{c}} D_0(\varphi_{j-1}(z))\,\mathrm{d}z \geq \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in [\underline{c}, \bar{c}_j], \tag{S37}$$

with equality at $\bar{c}_l$, $\underline{c}_l$, for all $l \in \{1, \dots, j-1\}$, let

$$\hat{\varphi}_j(c) := \begin{cases} \varphi_{j-1}(c), & \text{if } c \in [\underline{c}, \underline{c}_j) \\ \psi^*(c), & \text{if } c \in [\underline{c}_j, \bar{c}] \end{cases}$$

and let

$$\delta_j(c) := \sup_s \left[\min_{c' \in [\underline{c}, c]} \left(\int_{c'}^{c} D_0(\hat{\varphi}_j(z))\,\mathrm{d}z + \int_{\{v \geq \psi^*(z)\}} (v-c) D_0(\mathrm{d}v) - \int_{\mathcal{D}} \pi_D(c') s(\mathrm{d}D) \right) \right], \tag{S38}$$

for all $c \in [\underline{c}_j, \bar{c}_j]$,where the supremum is taken over all $\psi^*(c)$-quasi-perfect segmentations for c. Also, let $C_j^+ := \{c \in [\underline{c}_j, \bar{c}_j] : \delta_j(c) \geq 0\}$ and $C_j^- := \{c \in [\underline{c}_j, \bar{c}_j] : \delta_j(c) < 0\}$. Then, if $V_j^- = \emptyset$, let $c_j^* := \bar{c}_j$. Otherwise, if $C_j^- \neq \emptyset$, let c_j^* be $\inf C_j^-$. Clearly, $c_j^* \in [\underline{c}_j, \bar{c}_j]$. If $c_j^* = \bar{c}_j$ then define φ_j as

$$\varphi_j := \varphi_{j-1}.$$

On the other hand, if $c_j^* < \bar{c}_j$, then by definition of C_s, and by (S35) and (S37), it must be that $D_0(\psi^*(c)) < D_0(\overline{\boldsymbol{p}}_0(c))$ for all $c \in (\underline{c}_j, c_j^*]$. Therefore, as $\overline{\boldsymbol{p}}_0$ is nondecreasing and D_0 is nonincreasing, for any $c \in (\underline{c}_j, c_j^*]$, there exists a unique $r_j(c) > c$ such that

$$(r_j(c) - c) D_0(\psi^*(c)) + \int_{\underline{c}_j}^{c} D_0(\psi^*(z))\,\mathrm{d}z = \int_{\underline{c_j}}^{r_j(c)} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z. \tag{S39}$$

Thus, since $\psi_0(c) \geq \boldsymbol{p}_0(c) > c$ for all $c \in C$, $\psi^*(c) < \overline{\boldsymbol{p}}_0(c)$ for all $c \in (\hat{c}_j, \bar{c}_j)$ for some $\hat{c}_j \in (c_j^*, \bar{c}_j)$, there exists unique $c_j^1 \in (\underline{c}_j, c_j^*]$ and $c_j^2 \in (c_j^*, \bar{c}_j]$ such that

$$(c_j^2 - r_j(c_j^1)) D_0(\psi^*(c_j^2)) + \int_{c_j^2}^{\bar{c}_j} D_0(\psi^*(z))\,\mathrm{d}z = \int_{r_j(c_j^1)}^{c_j^2} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \tag{S40}$$

and that $\psi^*(c_j^2) \leq \psi_0(r_j((c_j^1))$.

Then, define

$$\varphi_j(c) := \begin{cases} \varphi_{j-1}(c), & \text{if } c \in [\underline{c}, \underline{c}_j) \\ \psi^*(c), & \text{if } c \in [\underline{c}_j, c_j^1) \\ \psi^*(c_j^1), & \text{if } c \in [c_j^1, r_j(c_j^1)) \\ \psi^*(c_j^2), & \text{if } c \in [r_j(c_j^1), c_j^2) \\ \psi^*(c), & \text{if } c \in [c_j^2, \bar{c}] \end{cases}.$$

It can then be verified that the constructed φ_j (in both cases) satisfies (S36) and (S37). As such, the sequence of functions $\{\varphi_j\}$ is well-defined.

Finally, since φ_j is nondecreasing for all $j \in \mathbb{N}$, Helly'e selection theorem then ensures that, by possibly taking a subsequence, $\{\varphi_j\} \to \varphi^*$ for some nondecreasing function $\varphi^* \in \mathbb{R}_+^C$.

Lemma S7. *There exists an incentive compatible and individually rational φ^*-quasi-perfect mechanism.*

Proof. By construction, $c \leq \varphi^*(c) \leq \psi_0(c)$ for all $c \in C$. As such, by Lemma S5, there exists a φ^*-quasi-perfect scheme $\tilde{\sigma} \in \mathcal{S}^C$ such that

$$\int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \tilde{\sigma}(\mathrm{d}D|c) \right) \mathrm{d}z \leq \int_{c'}^{c} D_0(\min\{\varphi^*(z), \overline{\boldsymbol{p}}_0(z)\})\,\mathrm{d}z, \tag{S41}$$

for all $c, c' \in C$ with $c' < c$. On the other hand, by Lemma 13 , the choice set of the maximization problem in (S38) is closed and hence is compact as it is a subset of the compact set $\mathcal{D}$. Also, the objective function is continuous in s as it is a pointwise infimum of a family of affine functions. Thus, for any $c \in C_s$, the solution of (S38), $\hat{\sigma}(c)$, exists. Now define a segmentation scheme σ^* as

$$\sigma^*(c) := \begin{cases} \tilde{\sigma}(c), & \text{if } c \notin \cup_{j=1}^{\infty}[\underline{c}_j, r(c_j^1)] \\ \hat{\sigma}(c), & \text{if } c \in \cup_{j=1}^{\infty}[\underline{c}_j, r(c_j^1)] \end{cases}.$$

Then, define τ^* as the associated transfer given by assertion 1 in Lemma 1 with the constant being chosen such that $U(\bar{c}) = \bar{\pi}$.

By construction, (σ^*, τ^*) is a φ^*-quasi-perfect mechanism. I will then show that the mechanism (σ^*, τ^*) is incentive feasible. To verify individual rationality, notice that since σ^* is a φ^*-quasi-perfect scheme, and since $\varphi^*(c) \geq c$ for all $c \in C$, by Lemma 11, for all $c \in C$,

$$\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma^*(\mathrm{d}D|c) = D_0(\varphi^*(c)),\ \forall c \in C. \tag{S42}$$

Moreover, by construction of $\{\varphi_j\}$, for any $j \in \mathbb{N}$, for all $c \in C$,

$$\int_c^{\bar{c}} D_0(\varphi_j(z))\,\mathrm{d}z \geq \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z.$$

Together with the Lebesgue dominated convergence theorem, the fact that D_0 is continuous, and the fact that $\{\varphi_j\} \to \varphi^*$,

$$\int_c^{\bar{c}} D_0(\varphi^*(z))\,\mathrm{d}z = \lim_{j\to\infty} \int_c^{\bar{c}} D_0(\varphi_j(z))\,\mathrm{d}z \geq \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z.$$

Together with Lemma 1, the mechanism (σ^*, τ^*) is indeed individually rational.

To see incentive compatibility, by Lemma 1, it suffices to check whether assertion 2 of Lemma 1 is satisfied. In fact, by Lemma 12, since by construction, $\varphi^*(c) \geq c$ for all $c \in C$, it suffices to check the upward deviations. As such, consider any $c', c \in C$ with $c' < c$. Clearly, if $c \leq c^*$, since $\varphi^* \equiv \varphi_G$ on $[\underline{c}, c^*]$, as shown in the proof of Theorem 1, assertion 2 holds. If, on the other hand, $c > c^*$ and $c \notin C_s$, let $j(c) \in \mathbb{N}$ be the largest $j \in \mathbb{N}$ such that $\bar{c}_j < c$. By construction of $\varphi_{j(c)}$,

$$\int_{\underline{c}_j}^{\bar{c}_j} D_0(\varphi_{j(c)}(z))\,\mathrm{d}z = \int_{\underline{c}_j}^{\bar{c}_j} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall j \in \{1, \dots, j(c)\}$$

and $\varphi_{j(c)} \equiv \varphi_{j(c)-1}$ on $[\underline{c}, \underline{c}_{j(c)})$. As such,

$$\int_{c'}^{c} D_0(\varphi_{j(c)}(z))\,\mathrm{d}z \geq \int_{c'}^{c} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z.$$

Together with (S41), this implies that

$$\begin{aligned}
&\int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
\geq &\int_{c'}^{c} D_0(\varphi^*(z))\,\mathrm{d}z - \int_{c'}^{c} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \\
= &\int_{c'}^{c} D_0(\varphi_{j(c)}(z))\,\mathrm{d}z - \int_{c'}^{c} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \\
\geq &0,
\end{aligned}$$

confirming assertion 2.

Finally, if $c \in C_s$, let $j(c)$ be the unique $j \in \mathbb{N}$ for which $c \in (\underline{c}_j, \bar{c}_j)$. If $c \leq c^*_{j(c)}$, then by definition of $\delta_{j(c)}$ and by (S42),

$$\begin{aligned}
&\int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
=&\int_{c'}^{c} D_0(\varphi^*(z))\,\mathrm{d}z - \int_{\mathcal{D}} (\pi_D(c') - \pi_D(c))\sigma^*(\mathrm{d}D|c) \\
=&\int_{c'}^{c} D_0(\hat{\varphi}_{j(c)}(z))\,\mathrm{d}z + \int_{\{v \geq \psi^*(c)\}} (v-c) D_0(\mathrm{d}v) - \int_{\mathcal{D}} \pi_D(c')\sigma^*(\mathrm{d}D|c) \\
\geq&\delta_{j(c)}(c) \\
\geq&0,
\end{aligned}$$

where the first equality follows from (S42), (5), and the fundamental theorem of calculus; the second equality follows from the construction of φ^* and the property of a φ^*-quasi-perfect scheme; the first inequality follows from the definition of $\delta_{j(c)}$ and the last equality follows from the fact that $c \leq c^*_{j(c)}$. If, on the other hand, $c \in (c^*_{j(c)}, r_{j(c)}(c^1_{j(c)}))$, then

$$\begin{aligned}
&\int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
=&\int_{c'}^{c^*_{j(c)}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
&+\int_{c^*_{j(c)}}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
\geq&\delta_{j(c)}(c^*_{j(c)}) \\
\geq&0,
\end{aligned} \tag{S43}$$

where the first inequality follows from the fact that $\varphi_{j(c)}(z) = \varphi_{j(c)}(c^1_{j(c)})$ for all $z \in [c^1_{j(c)}, r_{j(c)}(c^1_{j(c)}))$, (5), the fundamental theorem of calculus and $c^1_j \leq c^*_j$. Lastly, if $c \in [r_{j(c)}(c^1_{j(c)}), \bar{c}_{j(c)}]$,

$$\begin{aligned}
&\int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
=&\int_{c'}^{r_{j(c)}(c^1_{j(c)})} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
&+\int_{r_{j(c)}(c^1_{j(c)})}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z \\
\geq&\int_{c'}^{r_{j(c)}(c^1_{j(c)})} D_0(\varphi_{j(c)}(z))\,\mathrm{d}z - \int_{c'}^{r_{j(c)}(c^1_{j(c)})} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \\
&+\int_{r_{j(c)}(c^1_{j(c)})}^{c} D_0(\varphi_{j(c)}(z))\,\mathrm{d}z - \int_{r_{j(c)}(c^1_{j(c)})}^{c} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z \\
\geq&0,
\end{aligned}$$

where the first inequality follows from (S41) and the second inequality follows from (S39), (S40) and from the construction of $\varphi_{j(c)}$ (i.e., the fact that $\psi^*(c^2_{j(c)}) > \overline{\boldsymbol{p}}_0(c^2_{j(c)})$). Together, assertion 2 of Lemma 1 is always satisfied. ■

Step 3: Verifying Optimality and Uniqueness

In what follows, I construct a Lagrange multiplier Λ for the incentive compatibility constraints. Together with $\bar{M}$, a weak duality argument can then be used to conclude that the mechanism (σ^*, τ^*) given by Lemma S7 is optimal for the data broker. To this end, notice that by construction of φ^* and by Lemma S3, $\varphi^* \equiv \varphi_G$ on $[\underline{c}, c^*]$; $\varphi^* \equiv \overline{\boldsymbol{p}}_0$ almost everywhere on $[c^*, \bar{c}] \backslash C_s$. Moreover, for any $j \in \mathbb{N}$, the incentive compatibility constraints are binding on both $[c_j^1, r_j(c_j^1)]$ and $(r_j(c_j^1), c_j^2]$. That is,

$$\int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) \right) \mathrm{d}z = 0$$

whenever $c', c \in [c_j^1, r_j(c_j^1)]$ or $c', c \in (r_j(c_j^1), c_j^2]$.

For each $j \in \mathbb{N}$, define $\underline{\Lambda}_j : [c_j^1, r_j(c_j^1)] \to \mathbb{R}_+$ and $\bar{\Lambda}_j : [r_j(c_j^1), c_j^2] \to \mathbb{R}_+$ as the solution of the following two differential equations:

$$(r_j(c_j^1) - c)\Lambda'(c) = (\psi^*(c) - \psi^*(c_j^1))g(c) + \Lambda(c),\ \Lambda(c_j^1) = 0. \tag{S44}$$

and

$$(c - r_j(c_j^1))\Lambda'(c) = \bar{\Lambda} - \Lambda(c) + (\psi^*(c_j^2) - \psi^*(c))g(c),\ \Lambda(c_j^2) = \bar{\Lambda}. \tag{S45}$$

Clearly, for each $j \in \mathbb{N}$, $\underline{\Lambda}_j$ and $\bar{\Lambda}_j$ are nondecreasing. It can then be verified that the mechanism (σ^*, τ^*) is optimal for the data broker.

Proof of Theorem S1. Consider the dual of a relaxed program of the data broker's problem

$$\begin{aligned}
\bar{d} := \sup_{\sigma} & \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\
& + \int_C \left[\int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right] \bar{M}(\mathrm{d}c) \\
& + \sum_{j=1}^{\infty} \int_{c_j^1}^{r_j(c_j^1)} \left[\int_c^{r_j(c_j^1)} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) \right) \mathrm{d}z \right] \underline{\Lambda}_j(\mathrm{d}c) \\
& + \sum_{j=1}^{\infty} \int_{r_j(c_j^1)}^{c_j^2} \left[\int_{r_j(c_j^1)}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c) \right) \right] \bar{\Lambda}_j(\mathrm{d}c),
\end{aligned} \tag{S46}$$

where the supremum is taken over all segmentation schemes $\sigma \in \mathcal{S}^C$ such that

$$c \mapsto \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c)$$

is nonincreasing. Notice that for the segmentation scheme σ^* and the recommended price $\overline{\boldsymbol{p}} \in \boldsymbol{P}$,

$$c \mapsto \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma^*(\mathrm{d}D|c)$$

is indeed nonincreasing and the complementary slackness conditions with the multipliers $\bar{M}$, $\{\underline{\Lambda}_j, \bar{\Lambda}_j\}$ hold. That is,

$$\sum_{j=1}^{\infty} \int_{c_j^1}^{r_j(c_j^1)} \left[\int_c^{r_j(c_j^1)} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|c) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) \right) \mathrm{d}z \right] \underline{\Lambda}_j(\mathrm{d}c) = 0, \tag{S47}$$

$$\sum_{j=1}^{\infty} \int_{r_j(c_j^1)}^{c_j^2} \left[\int_{r_j(c_j^1)}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*(\mathrm{d}D|c) \right) \right] \bar{\Lambda}_j(\mathrm{d}c) = 0 \tag{S48}$$

and

$$\int_C \left[\int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*(\mathrm{d}D|z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right] \bar{M}(\mathrm{d}c) = 0, \tag{S49}$$

where the last condition holds due to (S13) and the construction of φ^* (in particular, (S37)). As such, together with the implementability of σ^* implied by Lemma S7, it suffices to show that σ^* indeed solves the maximization problem given by the dual (S46). To this end, first notice that, after interchanging the order of integral, for any feasible σ,

$$\begin{aligned} &\int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\ &\quad + \int_C \left[\int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma(\mathrm{d}D|z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right] \bar{M}(\mathrm{d}c) \\ &= \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi^{\bar{M}}(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{K}, \end{aligned}$$

where

$$\phi^{\bar{M}}(c) := \phi_G(c) - \frac{\bar{M}(c)}{g(c)},$$

and

$$\bar{K} := \int_C \left(\int_c^{\bar{c}} D(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z \right) \bar{M}(\mathrm{d}c)$$

is a constant that does not depend on the choice variable σ. By Lemma S3, since ψ^* solves the dual the price-controlling data broker, (S12) implies that ψ^* must be the ironing of $\phi^{\bar{M}}$. In particular, for any feasible σ,

$$\begin{aligned} &\int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi^{\bar{M}}(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\ &\leq \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi^*(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c). \end{aligned} \tag{S50}$$

Also, since σ^* is a φ^*-quasi-perfect scheme and since φ^* is constant on an interval $[c_1, c_2]$ whenever ψ^* is a constant on $[c_1, c_2]$.

$$\begin{aligned} &\int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi^{\bar{M}}(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^*(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\ &= \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi^*(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^*(\mathrm{d}D|c) \right) G(\mathrm{d}c). \end{aligned} \tag{S51}$$

Moreover, by interchanging the orders of integrals again and by the use of the fundamental theorem of calculus, the objective in the dual program (S46) can be written as

$$\begin{aligned} &\int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi^{\bar{M}}(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{K} \\ &+ \sum_{j=1}^{\infty} \left[\int_{c_1^j}^{r_j(c_1^j)} \left(\int_{\mathcal{D}} (\pi_D(c) - \pi_D(r_j(c_j^1)))) \sigma(\mathrm{d}D|c) \right) \underline{\Lambda}_j'(c) \, \mathrm{d}c - \int_{c_j^1}^{r_j(c_j^1)} \underline{\Lambda}_j(c) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \, \mathrm{d}c \right] \\ &+ \sum_{j=1}^{\infty} \left[\int_{r_j(c_j^1)}^{c_j^2} (\bar{\Lambda}_j(c_j^2) - \bar{\Lambda}_j(c)) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \, \mathrm{d}c - \int_{r_j c_j^1)}^{c_j^2} \left((\pi_D(r_j(c_j^1)) - \pi_D(c)) \sigma(\mathrm{d}D|c) \right) \bar{\Lambda}_j'(c) \, \mathrm{d}c \right] \end{aligned}$$

and hence, by (S50),

$$\begin{aligned}
\bar{d} \leq \sup_{\sigma} \int_C & \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi^*(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) g(c)\,\mathrm{d}c \\
& + \sum_{j=1}^{\infty} \Bigg[\int_{c_1^j}^{r_j(c_1^j)} \left(\int_{\mathcal{D}} (\pi_D(c) - \pi_D(r_j(c_j^1))) \sigma(\mathrm{d}D|c) \right) \underline{\Lambda}_j'(c)\,\mathrm{d}c \\
& \qquad - \int_{c_j^1}^{r_j(c_j^1)} \underline{\Lambda}_j(c) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c)\,\mathrm{d}c \Bigg] \\
& + \sum_{j=1}^{\infty} \Bigg[\int_{r_j(c_j^1)}^{c_j^2} (\bar{\Lambda}_j(c_j^2) - \bar{\Lambda}_j(c)) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c)\,\mathrm{d}c \\
& \qquad - \int_{r_j c_j^1)}^{c_j^2} \left((\pi_D(r_j(c_j^1)) - \pi_D(c)) \sigma(\mathrm{d}D|c) \right) \bar{\Lambda}_j'(c)\,\mathrm{d}c \Bigg] - \bar{K} \\
=: \widetilde{d} &
\end{aligned} \tag{S52}$$

Now notice that for any feasible σ, for any $j \in \mathbb{N}$, for any $c \in [c_j^1, r_j(c_j^1)]$,

$$\begin{aligned}
& \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi^*(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) g(c) \\
& - \underline{\Lambda}_j(c) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) + \left(\int_{\mathcal{D}} (\pi_D(c) - \pi_D(r_j(c_j^1))) \sigma(\mathrm{d}D|c) \right) \underline{\Lambda}_j'(c) \\
= & \int_{\mathcal{D}} \pi_D(c) \sigma(\mathrm{d}D|c) (g(c) + \underline{\Lambda}_j'(c)) \\
& - \left(\psi^*(c) + \frac{\underline{\Lambda}_j(c)}{g(c)} - c \right) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) g(c) - \int_{\mathcal{D}} \pi_D(r_j(c_j^1))) \sigma(\mathrm{d}D|c) \underline{\Lambda}_j'(c) \\
\leq & \int_{\mathcal{D}} \pi_D(c) \sigma(\mathrm{d}D|c) (g(c) + \underline{\Lambda}_j'(c)) - \left(\psi^*(c) + \frac{\underline{\Lambda}_j(c)}{g(c)} - c \right) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) g(c) \\
& - \int_{\mathcal{D}} \pi_D(c) \sigma(\mathrm{d}D|c) \underline{\Lambda}_j'(c) + \underline{\Lambda}_j'(c)(r_j(c_j^1) - c) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \\
= & \int_{\mathcal{D}} \pi_D(c) \sigma(\mathrm{d}D|c) g(c) - \left[\left(\psi^*(c) + \frac{\underline{\Lambda}_j(c)}{g(c)} - c \right) g(c) - \underline{\Lambda}_j'(c)(r_j(c_j^1 - c)) \right] \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \\
= & \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi^*(c_j^1)) D(\overline{\boldsymbol{p}}_D(c)) g(c),
\end{aligned}$$

where the first inequality follows from the fundamental theorem of calculus and the fact that

$$c \mapsto \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c)$$

is nonincreasing and the last equality follows from the definition of $\underline{\Lambda}_j$ (i.e., (S44)). Similarly, by the definition of $\bar{\Lambda}_j$, for any $j \in \mathbb{N}$, any $c \in (r_j(c_j^1), c_j^2]$, we have

$$\begin{aligned}
& \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi^*(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) g(c) \\
& + (\bar{\Lambda}_j(c_j^2) - \bar{\Lambda}_j(c)) \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) - \left(\int_{\mathcal{D}} (\pi_D(r_j(c_j^1))) - \pi_D(c)) \sigma(\mathrm{d}D|c) \right) \bar{\Lambda}_j'(c) \\
\leq & \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi^*(c_j^2)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) g(c).
\end{aligned}$$

Together, since the intervals $\{(c_j^1, r_j(c_j^1)), (r_j(c_j^1), c_j^2)\}$ are disjoint, and since $\varphi^*(c) = \psi^*(c_j^1)$ for all $c \in [c_j^1, r_j(c_j^1)]$ and $\varphi^*(c) = \psi^*(c_j^2)$ for all $c \in (r_j(c_j^1), c_j^2]$

$$\begin{aligned}\widetilde{d} \leq& \sup_{\sigma} \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \varphi^*(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{K} \\ \leq& \int_C \left(\int_{\{v \geq \varphi^*(c)\}} (v - \varphi^*(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{K} \\ =& \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \varphi^*(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^*(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{K},\end{aligned}$$

where the last equality follows from the fact that σ^* is a φ^*-quasi-perfect scheme. Finally, combining with (S51), and the fact that σ^* is feasible in the dual program (S46), σ^* is indeed a solution to the dual program (S46). Together with the complementary slackness conditions (S47), (S48) and (S49). the mechanism (σ^*, τ^*) is indeed optimal for the data broker.

Finally, since (σ^*, τ^*) is optimal, for any other optimal mechanism (σ, τ) of the data broker, it must be that

$$\int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) = \int_C \left(\int_{\{v \geq \varphi^*(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c).$$

Then, by the same arguments as in the proof of Theorem 1, together with (S47), (S48), (S50), (S51) as well as the construction of φ^*, it must be that σ is a φ^*-quasi-perfect scheme. Moreover, by optimality, the indirect utility of the producer with marginal cost $\overline{c}$ must be $\overline{\pi}$. Therefore, (σ, τ) must be a φ^*-quasi-perfect mechanism. This completes the proof. ■

S.4 Gains from Trade and Assumption 1

In the main text, it is shown that a sufficient condition for Assumption 1 is that $\phi_G \leq \overline{\boldsymbol{p}}_0$ on C. In this section, I provide an economic interpretation of this condition. Specifically, under a notion of gains from trade, this condition is equivalent to gains from trade being large enough. To this end, take and fix $D_0 \in \mathcal{D}$ for some $\overline{v} > 0$ and define the *location family of* D_0, $\{D_0^k\}_{k \geq 0}$ by

$$D_0^k(p) := D_0(p - k).$$

As such, for each $k \geq 0$, $D_0^k \in \mathcal{D}([\underline{v} + k, \overline{v} + k])$. Given such location family $\{D_0^k\}_{k \geq 0}$ and any distribution of production cost G, a natural way to rank the gains from trade in these economies is by k. As such, an economy (D_0^k, G) has larger gains from trade than an economy $(D_0^{k'}, G)$ if and only if $k > k'$. With this definition, the following lemma provides an interpretation of the sufficient condition $\phi_G \leq \overline{\boldsymbol{p}}_0$.

Proposition S1. *Given any $D_0 \in \mathcal{D}$ and any distribution of marginal cost G, consider the induced location family $\{D_0^k\}_{k \geq 0}$. There exists $\bar{k} \geq 0$ such that $\phi_G \leq \overline{\boldsymbol{p}}_{D_0^k}$ if and only if $k \geq \bar{k}$.*

Proof. For any $k \geq 0$, notice that for any $v \in \mathrm{supp}(D_0^k)$ and any $c \in C$,

$$\begin{aligned}(v - c) D_0^k(v) = (v - c) D_0(v - k) =& ((v - k) - (c - k)) D_0(v - k) \\ \leq& (\overline{\boldsymbol{p}}_0(c - k) - (c - k)) D_0(\overline{\boldsymbol{p}}_0(c - k)) \\ =& (\overline{\boldsymbol{p}}_0(c - k) + k - c) D_0^k(\overline{\boldsymbol{p}}_0(c - k) + k).\end{aligned}$$

As such, for $\boldsymbol{p}_{D_0^k}(c) := \boldsymbol{p}_0(c-k)+k$,

$$\boldsymbol{p}_{D_0^k}(c) \in \boldsymbol{P}_{D_0^k}(c),\ \forall c \in C,\ \forall k \geq 0.$$

Furthermore, since

$$\lim_{k\to\infty} \boldsymbol{p}_0(c-k) = 0,\ \forall c \in C,$$

there exists some $\bar{k} \geq 0$ such that

$$\inf_{c\in C}[\boldsymbol{p}_0(c-k)+k-\phi_G(c)] \geq 0 \iff k \geq \bar{k}.$$

As such, $\boldsymbol{p}_{D_0^k}(c) \geq \phi_G(c)$ for all $c \in C$ if and only if $k \geq \bar{k}$, as desired. ■

With Proposition S1, together with an observation that Assumption 1 is implied by $\phi_G \leq \overline{\boldsymbol{p}}_0$, Theorem 7 can be interpreted as the following: Data brokership is outcome-equivalent to price-controlling data brokership if gains from trade are large enough.

S.5 Comparisons across Market Regimes in General Environments

Based on the characterization of Theorem S2, it is noteworthy that by construction of φ^*, since $\psi^*(c) \geq c$ for all $c \in C$, $\varphi^*(c) \geq c$ for all $c \in C$, with strict inequality for positive measure of $c \in C$. Thus, the arguments that lead to Theorem 6 are still valid even without Assumption 1. That is, since the consumer surplus is always zero by Theorem S1 and since the optimal mechanism does not induce perfect-price-discrimination in general since $\varphi^*(c) > c$, vertical integration between the data broker and the producer is always Pareto-improving and is strictly Pareto-improving if $\bar{c} > \underline{v}$. As such, Theorem 6 can be generalized even without Assumption 1, as long as D_0 is continuous.

On the other hand, notice that by construction of φ^*, for any $c \in C$,

$$\int_c^{\bar{c}} D_0(\varphi^*(z))\,\mathrm{d}z \leq \int_c^{\bar{c}} D_0(\psi^*(z))\,\mathrm{d}z.$$

This implies that the producer's profit under data brokership is always smaller than that under price-controlling data brokership. As a result, even if Theorem 7 nay not hold when Assumption 1 fails, as long as D_0 is continuous, data brokership and price-controlling data brokership can be Pareto-ranked. Furthermore, since $\psi^*(c) \geq c$ for all $c \in C$, price-controlling data brokership is further Pareto-dominated by vertical integration. Finally, notice that the proof of Lemma 4 does not rely on any assumptions about (D_0, G) Together, the next theorem summarizes these generalizations.

Theorem S3 (Ranking Regimes)**.** *Suppose that $D_0 \in \mathcal{D}$ is continuous. Then vertical integration Pareto-dominates price-controlling data brokership and exclusive retail, which in turn Pareto-dominates data brokership.*

S.6 The Purchase of Consumer Data

Section 5.2 of the main text discussed an alternative model that requires the data broker to purchase the data from the consumers and concludes that the data broker would pay the consumers their ex-ante surplus and sell the data to the producers using the optimal mechanisms characterized by Theorem 1, which leads to a Pareto-improving outcome comparing to uniform pricing. This result hinges on two crucial assumptions: (i) The data broker can only buy all the consumer data or buy nothing. (ii) The consumers have to decide whether to accept the offered compensation before learning their value. While relaxing (ii) could also be an interesting and relevant question, this section considers a relaxation of (i).

For simplicity, throughout this section, assume that $\phi_G \leq \overline{\boldsymbol{p}}_0$ on C (so that Assumption 1 is satisfied) and that G is regular. Suppose that, prior to interacting with the producer, the data broker can make a take-it-or-leave-it offer $(\bar{s}, x)$ to the consumers, where $\bar{s} \in \mathcal{S}$ stands for the amount of data that will be obtained by the data broker and $x \in \mathbb{R}$ stands for the monetary transfer from the data broker to the consumers. Given data $\bar{s}$, the data broker then designs an incentive feasible mechanism (σ, τ) to sell the data to the producer. The distinction from the main model is that for each report $c \in C$, it is required that

$$\delta_{\{m^0\}} \leq_{\text{MPS}} \sigma(c) \leq_{\text{MPS}} \bar{s},$$

where $\leq_{\text{MPS}}$ is the mean-preserving spread order. That is, given data $\bar{s}$, the data broker can only provide information about the consumers' values that is Blackwell-less informative than that can be provided by $\bar{s}$. Under this setting, the following proposition shows that it is optimal for the data broker to buy all the data, which implies that assumption (i) imposed in the main text is in fact without loss.

Proposition S2. *Suppose that $\phi_G \leq \overline{\boldsymbol{p}}_0$. If the consumers own their data and if the data broker can provide monetary transfers to the consumers, then the optimal mechanisms for the data broker are to purchase all the data by paying the consumers their ex-ante surplus under uniform pricing, and then use any optimal selling mechanism (σ, τ) characterized by Theorem 1 to sell to the producer. In particular, the outcome is Pareto improving comparing to uniform pricing in the ex-ante sense.*

Proof. Consider any take-it-or-leave-it offer $(\bar{s}, x)$ made to the consumers. Since the consumers are ex-ante homogeneous, they make the same decisions. If the consumers reject the offer, then the data broker cannot provide any further information to the producer and hence consumer surplus would be that given by optimal uniform pricing. On the other hand, if the consumers accept the offer $(\bar{s}, x)$, by the same arguments as the existence proof if Theorem S1, there exists an optimal mechanism $(\sigma^{\bar{s}}, \tau^{\bar{s}})$ for the data broker when selling to the producer and the consumer surplus would be

$$\int_C \left(\int_V \left(\int_{\{v \geq \boldsymbol{p}_D^{\bar{s}}(c)\}} (v - \overline{\boldsymbol{p}}_D(c)) \right) D(\mathrm{d}v) \sigma^{\bar{s}}(\mathrm{d}D|c) \right) G(\mathrm{d}c).$$

Therefore, the consumers will accept (s, x) if and only if

$$\begin{aligned} x \geq & \int_C \left(\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) \\ & - \int_C \left(\int_{\mathcal{D}} \left(\int_{\{v \geq \boldsymbol{p}_D^{\bar{s}}(c)\}} (v - \overline{\boldsymbol{p}}_D(c)) \right) D(\mathrm{d}v) \sigma^{\bar{s}}(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\ =: & x^*(s) \end{aligned}$$

As a result, by Lemma 1, the data broker's expected revenue from purchasing $s \in \mathcal{S}$ is

$$\int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^{\bar{s}}(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi} - x^*(\bar{s}).$$

Notice that for any $c \in C$ and for any $s \in \mathcal{S}$

$$\begin{aligned}
&\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^{\bar{s}}(\mathrm{d}D|c) + \int_{\mathcal{D}} \left(\int_{\{v \geq \overline{\boldsymbol{p}}_D(c)\}} (v - \overline{\boldsymbol{p}}_D(c)) D(\mathrm{d}v) \right) \sigma^{\bar{s}}(\mathrm{d}D|c) \\
&= \int_{\mathcal{D}} \left(\int_{\{v \geq \overline{\boldsymbol{p}}_D(c)\}} (v - \phi_G(c))^+ D(\mathrm{d}v) \right) \sigma^{\bar{s}}(\mathrm{d}D|c) \\
&\leq \int_{\mathcal{D}} \left(\int_V (v - \phi_G(c))^+ D(\mathrm{d}v) \right) \sigma^{\bar{s}}(\mathrm{d}D|c) \qquad \text{(S53)} \\
&= \int_V (v - \phi_G(c))^+ D_0(\mathrm{d}v),
\end{aligned}$$

where the last equality follows from the fact that $\sigma^{\bar{s}}(c) \in \mathcal{S}$. As such, for any $\bar{s} \in \mathcal{S}$

$$\begin{aligned}
R^{**} &:= \int_C \left(\int_V (v - \phi_G(c))^+ D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \int_C \left(\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) \\
&\geq \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^{\bar{s}}(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi} - x^*(\bar{s}),
\end{aligned}$$

where the inequality follows from the definition of $x^*(\bar{s})$ and from (S53). Notice that by Proposition 1, $R^{**} \geq 0$ and thus R^{**} is an upper bound of the data broker's profit across all mechanisms.

Finally, since $\phi_G \leq \overline{\boldsymbol{p}}_0$ and since G is regular, $\bar{\varphi}_G \equiv \phi_G$ and hence Theorem 1 implies that the set of incentive feasible ϕ_G- quasi-perfect mechanisms is nonempty and is exactly the set of optimal mechanisms. Furthermore, any such mechanism gives revenue

$$\int_C \left(\int_V (v - \phi_G(c))^+ D_0(\mathrm{d}v) \right) G(\mathrm{d}c).$$

Thus, it is optimal for the data broker to purchase the value-revealing segmentation by paying the consumers their ex-ante surplus and then use any optimal mechanism characterized by Theorem 1 to sell to the producer, as this attains the upper bound R^{**}. This completes the proof. ■

S.7 Implementing Segmentations by Revealing Characteristics

Throughout the paper, I define a market segmentation s as a probability distribution over $\mathcal{D}$, which can be interpreted as splitting the market demand D_0 into several market segments described by a demand function. Although this is a common way to formalize market segmentations in the price discrimination literature, a more intuitive way to describe a market segmentation is through consumer characteristics. After all, in many real-world practices, a data broker does not create market segmentations directly. Instead, he sells consumer data to the producers and the information contained in consumer data generates market segmentations.

If the data broker is able to introduce additional noises to the existing consumer data (which is in fact commonly seen in practice, where the purchased consumer data often contain information such as purchasing propensity and consumer score etc., see Federal Trade Commission (2014) for more details), then it is straightforward to create any market segmentation $s \in \mathcal{S}$. To do so, notice that any $s \in \mathcal{D}$ induces a joint distribution of values and market segments

$$\mu(\mathrm{d}v, \mathrm{d}D) := m^D(\mathrm{d}v) s(\mathrm{d}D)$$

with $\text{marg}_V \mu = m^0$. Since both $\mathcal{D}$ and V (endowed with their Borel σ-algebras) are regular probability spaces, there exists a transition probability $\pi : V \to \Delta(\mathcal{D})$ such that

$$\mu(\mathrm{d}v, \mathrm{d}D) = \pi(\mathrm{d}D|v)m^0(\mathrm{d}v).$$

As a result, since the data broker is assumed to have enough of consumer data to perfectly estimate each consumer's value, he can first find out the consumers' values and then for each $v \in V$, create a variable D according to $\pi(\cdot|v)$. This would then implement the market segmentation s.

However, there are also situations in which the data broker only sells existing consumer data. In this case, market segmentations can only be created by partitioning the consumer characteristics. It is less clear whether a data broker can create any market segmentation $s \in \mathcal{S}$ by only selling existing consumer data and partitioning the consumer characteristics. In what follows, I provide a formal foundation for this ability.

Formally, let $(\Theta, \mathcal{F}, \mathbb{P})$ be a probability space. The set Θ describes the collection of consumer characteristics. For instance, if $\Theta = \Theta_1 \times \Theta_2 \times \dots \Theta_K$ with $\Theta_k \subset \mathbb{R}$ for all k, then Θ can be thought of as the realization of K different consumer characteristics that are described numerically (e.g., income, tax record, age, time spent on a certain website etc.). Alternatively, if $\Theta = \{f : [0,1] \to \{0,1\} | f \text{ measurable}\}$, then Θ can be interpreted as a large collection (a continuum) of consumer characteristics that take binary values (e.g., gender, purchase history of a certain product etc.). The measure $\mathbb{P}$ describes the distribution of consumers. The consumers' values of a product is determined by their characteristics $\theta \in \Theta$. Formally let $\mathbf{V} : \Theta \to V$ be a measurable function. For any characteristic $\theta \in \Theta$, the consumers with characteristic θ would have value $\mathbf{V}(\theta)$. Since the distribution of value is given by m^0, it must be that $\mathbb{P} \circ \mathbf{V}^{-1} = m^0$.

Suppose that the data broker has data about every consumer's characteristics, and that the data broker can only provide data by partitioning the set Θ (e.g., by revealing to the producer whether a particular consumer purchased a related product, their gender, their income level and education etc.). The results below ensure that as long as the characteristic space $(\Theta, \mathcal{F}, \mathbb{P})$ is "rich" enough, the data broker can still induce any market segmentation $s \in \mathcal{S}$ even if he cannot generate additional data by himself.

To begin with, I first introduce an useful lemma in probability theory.

Lemma S8. *Let $\mathcal{B}(V)$ denote the Borel σ-algebra on V. Then there exists a countable partition $\mathcal{A}$ of V such that for any $A \in \mathcal{B}(V)$, there exists $A' \in \sigma(\mathcal{A})$ such that*

$$m^0(A \Delta A') := m^0((A \backslash A') \cup (A' \backslash A)) = 0$$

and vice versa.

With this lemma, I now state another useful lemma for the proof of the main theorems. To this end, let $\mathcal{A}$ be the countable partition given by Lemma S8.

Lemma S9. *For any Borel measure $m \in \Delta(V)$, suppose that there exists $F \in \mathcal{F}$ such that*

$$\mathbb{P}(F \cap \mathbf{V}^{-1}(A)) = m(A) \tag{S54}$$

for all $A \in \mathcal{A}$. Then (S54) holds for any $A \in \mathcal{B}(V)$

Proof. Let $F \in \mathcal{F}$ be the set such that (S54) holds for all $A \in \mathcal{A}$. Consider the σ-algebra generated by $\mathcal{A}$, $\sigma(\mathcal{A})$. Notice that $\mathcal{A} \cup \{\emptyset\}$ is a π-system since $\mathcal{A}$ is a partition. Furthermore, let $\mathcal{L}$ be the collection of subsets A of V such that (S54) holds. It is then straightforward to verify that, by using the fact that every $A \in \mathcal{A}$ satisfies (S54), $\mathcal{L}$ is a λ-system. By Dynkin's π-λ theorem, since $\mathcal{A} \cup \{\emptyset\} \subseteq \mathcal{L}$, it must be that $\sigma(\mathcal{A}) \subseteq \mathcal{L}$. That is, for any $A \in \sigma(\mathcal{A})$,

$$\mathbb{P}(F \cap \mathbf{V}^{-1}(A)) = m(A).$$

Finally, for any $A \in \mathcal{B}(V)$, by Lemma S8, there exists $A' \in \sigma(\mathcal{A})$ such that $m^0(A \Delta A') = 0$, which implies that $m^0(A \backslash A') = 0$. As a result,

$$\begin{aligned}\mathbb{P}(F \cap \mathbf{V}^{-1}(A)) =& \mathbb{P}([F \cap \mathbf{V}^{-1}(A')] \cup [\mathbf{V}^{-1}(A) \backslash \mathbf{V}^{-1}(A')]) \\ =& \mathbb{P}(F \cap \mathbf{V}^{-1}(A')) + \mathbb{P}(\mathbf{V}^{-1}(A) \backslash \mathbf{V}^{-1}(A')) \\ =& \mathbb{P}(F \cap \mathbf{V}^{-1}(A')) + \mathbb{P}(\mathbf{V}^{-1}(A \backslash A')) \\ =& m(A) + m^0(A \backslash A') \\ =& m(A),\end{aligned}$$

where the fourth equality holds because $A' \in \sigma(\mathcal{A})$ and $\mathbb{P} \circ \mathbf{V}^{-1} = m^0$ while the last equality is due to $m^0(A \backslash A') = 0$ as given by Lemma S8. This completes the proof. ■

I now state the main theorem for $s \in \mathcal{S}$ that has countable support.

Theorem S4 (Generating Countable Segmentations)**.** *Suppose that* $(\Theta, \mathcal{F}, \mathbb{P})$ *is nonatomic. Then for any* $s \in \mathcal{S}$ *with* $\text{supp}(s)$ *being countable, there exists a countable partition* $\mathcal{P}$ *of* Θ *such that for any* $D \in \text{supp}(s)$, *there exists* $F \in \mathcal{P}$ *such that*

$$\mathbb{P}(F \cap \mathbf{V}^{-1}(A)) = m^D(A)s(D),$$

for all $A \in \mathcal{B}(V)$.

Proof. Since $\text{supp}(s)$ is a countable set, there exists $\{D_n\} \subseteq \mathcal{D}$ such that

$$D_0 = \sum_{n=1}^{\infty} D_n s(D_n).$$

Equivalently, for any $A \in \mathcal{B}(V)$,

$$m^0(A) = \sum_{n=1}^{\infty} m^{D_n}(A)s(D_n).$$

Notice that for $m := m^{D_1}s(D_1)$,

$$m(A) \leq m^0(A)$$

for all $A \in \mathcal{B}(V)$.

I now show that there exists $F_1 \in \mathcal{F}$ such that

$$\mathbb{P}(F_1 \cap \mathbf{V}^{-1}(A)) = m^D(A),$$

for all $A \in \mathcal{B}(V)$. To see this, take the countable partition $\mathcal{A}$ of V given by Lemma S8. For any $A \in \mathcal{A}$, since

$$\mathbb{P}(\mathbf{V}^{-1}(A)) = m^0(A) \geq m^D(A),$$

by Sierpinski's intermediate value theorem, there exists $F_A \subseteq \mathbf{V}^{-1}(A)$, $F_A \in \mathcal{F}$ such that

$$\mathbb{P}(F_A) = m^D(A).$$

Now let

$$F_1 := \bigcup_{A \in \mathcal{A}} F_A.$$

Since $\mathcal{A}$ is countable and $F_A \in \mathcal{F}$ for all $A \in \mathcal{A}$, $F_1 \in \mathcal{F}$. Moreover, since $\mathcal{A}$ is a partition of V, $\mathbf{V}^{-1}(A) \cap \mathbf{V}^{-1}(A') = \emptyset$, and hence $F_A \cap F_{A'} = \emptyset$ for all $A, A' \in \mathcal{A}$, $A \neq A'$. As a result, for any $A \in \mathcal{A}$,

$$\mathbb{P}(F_1 \cap \mathbf{V}^{-1}(A)) = \mathbb{P}(F_A \cap \mathbf{V}^{-1}(A)) = \mathbf{P}(F_A) = m^D(A).$$

By Lemma S9, this then implies that for any $A \in \mathcal{B}(V)$,

$$\mathbb{P}(F_1 \cap \mathbf{V}^{-1}(A)) = m^D(A).$$

Now let

$$\tilde{m} := \sum_{n=2}^{\infty} m^{D_n} s(D_n).$$

and let

$$\Theta_1 := \Theta \backslash F_1.$$

It then follows from $m^0 = m + \tilde{m}$ that

$$\begin{aligned}
\mathbb{P}(\Theta_1 \cap \mathbf{V}^{-1}(A)) =& \mathbf{P}(\mathbf{V}^{-1}(A)) - \mathbf{P}(F_1 \cap \mathbf{V}^{-1}(A)) \\
=& m^0(A) - m^D(A) \\
=& \tilde{m}(A),
\end{aligned}$$

for all $A \in \mathcal{B}(V)$.

Inductively, there exists a disjoint sequence of sets $\{F_n\} \subseteq \mathcal{F}$ such that

$$\mathbb{P}(F_n \cap \mathbf{V}^{-1}(A)) = m^{D_n}(A) s(D_n),$$

for all $A \in \mathcal{B}(V)$. In particular,

$$\begin{aligned}
\mathbb{P}(\cup_{n=1}^{\infty} F_n) =& \sum_{n=1}^{\infty} \mathbb{P}(F_n) \\
=& \sum_{n=1}^{\infty} \mathbb{P}(F_n \cap \mathbf{V}^{-1}(V)) \\
=& \sum_{n=1}^{\infty} m^{D_n}(V) s(D_n) \\
=& \sum_{n=1}^{\infty} s(D_n) \\
=& 1
\end{aligned}$$

and hence $\cup_{n=1}^{\infty} F_n = \Theta$. Together, the partition $\mathcal{P} := \{F_n\}_{n=1}^{\infty}$ is as desired. ■

Essentially, Theorem S4 means that, any segmentation $s \in \mathcal{S}$ with countable support and be generated by finding a partition of the characteristic space Θ as long as $(\Theta, \mathcal{F}, \mathbb{P})$ is rich enough (in the sense that it is nonatomic). This implies that the data broker does not need to create any additional variable when selling consumer data. Instead, he can always generate an arbitrary segmentation with countable support by partially revealing the consumer characteristics so that each segment corresponds to a certain (range) of characteristic.

Although Theorem S4 enables the data broker to create any countable market segmentation by partitioning the characteristics, it is not always enough to only be able to generate countable segmentation. After all, as the main characterization of this paper, Theorem 1, suggests, optimal mechanisms involves uncountably many market segments when $|\text{supp}(D_0)| = \infty$. Nevertheless, Theorem S5 below ensures that any arbitrary segmentation can be generated by a partition on Θ, as long as the characteristic space $(\Theta, \mathcal{F}, \mathbb{P})$ is "even richer".

Definition S1. *Given any probability space* $(\Theta, \mathcal{F}, \mathbb{P})$ *and any measurable* $\mathbf{V} : \Theta \to V$*. Say that* $(\Theta, \mathcal{F}, \mathbb{P})$ *is rich relative to* $\mathbf{V}$ *if for any* $A \in \mathcal{B}(V)$*,* $(\mathbf{V}^{-1}(A), \mathcal{F}|_{\mathbf{V}^{-1}(A)}, \tilde{\mathbb{P}}_{\mathbb{A}})$ *is isomorphic to* $(I, \mathcal{B}([0,1]), L)$ *modulo zero for some interval* $I \subseteq [0,1]$*, where*

$$\mathcal{F}|_{\mathbf{V}^{-1}(A)} := \{F \in \mathcal{F} : F \subseteq \mathbf{V}^{-1}(A)\},$$

$$\tilde{\mathbb{P}}(F) := \mathbb{P}(F \cap \mathbf{V}^{-1}(A)),$$

for any $F \in \mathcal{F}|_{\mathbf{V}^{-1}(A)}$ *and* L *is the Lebesgue measure.*

For example, suppose that $\Theta \subseteq \mathbb{R}^n$ for some $n \geq 2$, $\mathcal{F} = \mathcal{B}(\Theta)$, $\mathbb{P}$ is absolutely continuous with respect to the Lebesgue measure on Θ, and that $\mathbf{V}$ is such that

$$\{\theta \in \Theta | \mathbf{V}(\theta) = v\}$$

has Hausdorff dimension greater than 2 for all $v \in V$. Then $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative to $\mathbf{V}$. Intuitively, a characteristic space $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative to $\mathbf{V}$ if for any fixed $v \in V$, there are enough of variations in characteristics $\theta \in \mathbf{V}^{-1}(v)$ for the data broker to generate additional noises given v.

Theorem S5 (Generating Arbitrary Segmentation)**.** *Suppose that* $(\Theta, \mathcal{F}, \mathbb{P})$ *is rich relative to* $\mathbf{V}$*. Then for any* $s \in \mathcal{S}$*, there exists a random variable* $\mathbf{D} : \Theta \to \mathcal{D}$ *such that*

$$\mathbb{P}(\mathbf{D}^{-1}(B) \cap \mathbf{V}^{-1}(A)) = \int_B m^D(A) s(\mathrm{d}D),$$

for all $A \in \mathcal{B}(V)$ *and for any measurable* $B \subseteq \mathcal{D}$*.*

Proof. Consider the countable partition $\mathcal{A}$ given by Lemma S8. For any $A \in \mathcal{A}$, define a Borel measure $\nu_A \in \Delta(\mathcal{D})$ as

$$\nu_A(B) := \int_B m^D(A) s(\mathrm{d}D)$$

and let $\mathbb{P}_A \in \Delta(\mathbf{V}^{-1}(A))$ be defined as

$$\mathbb{P}_A(F) := \mathbb{P}(F),$$

for all $F \in \mathcal{F}$ such that $F \subseteq \mathbf{V}^{-1}(A)$. Then since $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative $\mathbf{V}$, and since $\mathcal{D}$, endowed with the Borel σ-algebra is a standard Borel space, there exists a measurable function $\mathbf{D}_A : \mathbf{V}^{-1}(A) \to \mathcal{D}$ such that

$$\mathbb{P}_A(\mathbf{D}_A^{-1}(B)) = \nu_A(B),$$

for all measurable $B \subseteq \mathcal{D}$. Since $\mathcal{A}$ is a partition of V, $\{\mathbf{V}^{-1}(A)\}_{A \in \mathcal{A}}$ is a partition of Θ. Now define $\mathbf{D} : \Theta \to \mathcal{D}$ as

$$\mathbf{D}(\theta) := \mathbf{D}_A(\theta),$$

for all $\theta \in \mathbf{V}^{-1}(A)$.Since $\mathbf{D}_A$ is measurable for all $A \in \mathcal{A}$, $\mathbf{D}$ is measurable. Moreover, notice that for any $A \in \mathcal{A}$ and for any measurable $B \subseteq \mathcal{D}$,

$$\begin{aligned}
\mathbb{P}(\mathbf{D}^{-1}(B) \cap \mathbf{V}^{-1}(A)) =& \mathbb{P}(\mathbf{D}_A^{-1}(B) \cap \mathbf{V}^{-1}(A)) \\
=& \mathbb{P}(\mathbf{D}_A^{-1}(B)) \\
=& \nu_A(B) \\
=& \int_B m^D(A) s(\mathrm{d}D).
\end{aligned}$$

where the second equality follows from the fact that $\mathbf{D}_A^{-1}(V) \subseteq \mathbf{V}^{-1}(A)$. Finally, by Lemma S9, it then follows that for any $A \in \mathcal{B}(V)$ and for any measurable $B \subseteq \mathcal{D}$,

$$\mathbb{P}(\mathbf{D}^{-1}(B) \cap \mathbf{V}^{-1}(A)) = \int_B m^D(A) s(\mathrm{d}D),$$

as desired. ■

As an example, consider the case when $\Theta = [0,1]^2$, $\mathcal{F} = \mathcal{B}([0,1]^2)$, $\mathbb{P}$ is the Lebesgue measure on $[0,1]^2$ and $\mathbf{V}(\theta_1, \theta_2) = \theta_1 + \theta_2$. That is, the characteristics θ_1, θ_2 jointly (and linearly) determines a consumer's value. Then, the partition $\{\mathcal{P}_c\}_{c \in C}$ defined below implements the caonnical $\overline{\varphi}_G$-quasi perfect mechanism. To define $\mathcal{P}_c$, for any $v \geq \overline{\varphi}_G(c)$, let F_v be defined as

$$F_v := \{\theta \in \Theta | \theta_1 + \theta_2 = v\} \bigcup \left\{ \theta \in \Theta | \theta_1 + \theta_2 \leq \overline{\varphi}_G(c),\, \theta_2 = \Xi_c^{-1}\left(\frac{D_0(\overline{\varphi}_G(c)) - D_0(v)}{D_0(\overline{\varphi}_G(c))} \right) \right\},$$

where Ξ_c is the CDF of the marginal of θ_2 conditional on $\theta_1 + \theta_2 \leq \overline{\varphi}_G(c)$. Then, let

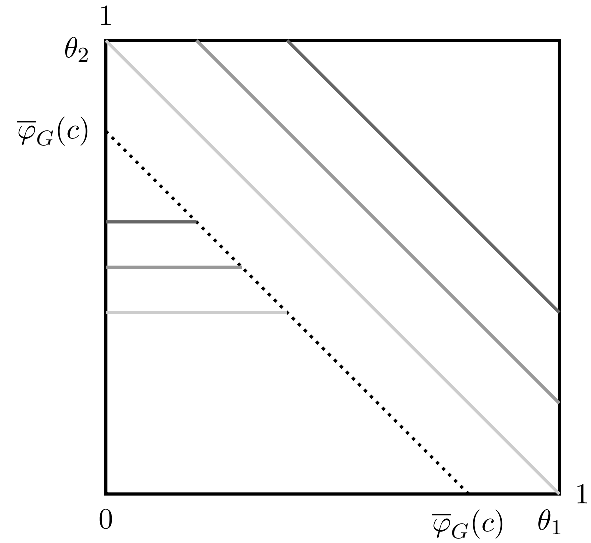
$$\mathcal{P}_c : \bigcup_{v \in [\overline{\varphi}_G(c), \overline{v}]} F_v,$$

as illustrated in Figure S1, where different colors indicate different elements in the partition. Notice that every $v \geq \overline{\varphi}_G(c)$, among the consumers with characteristics $\theta \in F_v$, the distribution of their values is exactly $D_v^{\overline{\varphi}_G(c)}$.

References

Bergemann, D., B. Brooks, and S. Morris (2013): "The Limits of Price Discrimination," Discussion Paper, Cowles Foundation for Research in Economics, Yale University.

Figure S1: Partitioning Characteristics



Luenberger, D. G. (1969): *Optimization by Vector Space Methods*, New York, NY: John Wiley & Sons.

Federal Trade Commission (2014): "Data Brokers: A Call for Transparency and Accountability," https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (accessed June 20, 2019).

Porter, J. E. (2005): "Helly's Selection Principle for Functions with Bounded p-Variation," *Rocky Mountain Journal of Mathematics*, 35, 675–679.