Voluntary Disclosure, Location and Price Choice in a Hotelling Model

Nenad Kos

Goonj Mohan

PRELIMINARY AND INCOMPLETE. NOT FOR CIRCULATION.

Abstract

We study voluntary disclosure of consumer data in a market where a seller uses the information to personalize both product design and pricing. Disclosure benefits the consumers through improved product matching but but also enables the seller to extract a larger share of the surplus through more personalized pricing. We examine how voluntary disclosure and welfare outcomes differ by consumer type.We show that the greedy partition maximizes ex-ante consumer surplus. The seller sequentially segments consumers, in each next step selecting a segment with maximal per segment profit. In equilibrium all consumers purchase, but niche consumers disclose more data, leading to more tailored product designs yet also higher prices. As the price effect dominates the design benefit, niche consumers are ultimately worse off than mass consumers.

1 Introduction

A consumer's online activity generates a steady stream of data that reveals her preferences and her willingness to pay. The products she searches for, the filters she applies, the time she spends browsing particular items, and even actions such as adding products to wishlists or reading reviews all help firms infer her underlying tastes. Contextual information—such as device type, shopping time, or location—further sharpens these estimates. Over time, the accumulation of such digital traces enables companies to form increasingly precise predictions about what she wants and how much she is prepared to spend.

With regulations such as the GDPR giving consumers greater control over their data, individuals can decide how much information to reveal online. Disclosure is a double-edged sword. Sharing more information can work in consumer's favor by allowing the platform to suggest products that more closely align with her tastes. Yet, the same disclosure also equips the platform with a clearer picture of her willingness to pay, enabling it to extract more surplus. Against this backdrop, we examine disclosure behavior across consumer types and find that niche consumers share more data than mass consumers and, as a result, face higher prices.

We study a setting with a monopolist seller and a continuum of consumers distributed along the interval $[0, 1]$ in the spirit of Hotelling, with an increasing density function f.[1] The interval captures horizontal product differentiation, while a consumer's location represents her preferred product design. Naturally, each consumer prefers products closer to her ideal design. Consumers choose how much data to disclose by specifying the set their type belongs to, and the seller responds with a personalized product design and price tailored to the information revealed.

We study a setting with a monopolist seller and a continuum of consumers distributed along the Hotelling interval $[0, 1]$.[2] The interval captures horizontal product differentiation, while a consumer's location represents her preferred product design. Each consumer prefers products closer to her ideal design. Consumers choose how much data to disclose by specifying the set their type belongs to, and the seller responds

[1] Alternatively, this can be interpreted as a single consumer whose preferred location is unknown.
[2] Alternatively, a single consumer whose preferred location is unknown.

with a personalized product design and price tailored to the information revealed.

The analysis, so far, pertains to environments where the density over the interval is log-concave and monotonic. To fix ideas, this discussion focuses on increasing densities. The analysis starts by examining the monopolist's optimal product design and pricing when consumers disclose that they are below some threshold. First, we show that in such a case the seller offers a product that is purchased by an interval of consumers with the upper bound of the interval being the threshold. Namely, the increasing density implies that the closer one gets to threshold the more consumers prefer that product. In addition, the lower the threshold the higher is the price the seller charges after learning that the consumers are below the threshold.

The analysis, thus far, considers environments where the density over the interval is log-concave and monotonic. To be specific, this discussion focuses on increasing densities. We first examine the monopolist's optimal product design and pricing when consumers reveal that their type lies below a certain threshold. In this case, the seller offers a product purchased by an interval of consumers, with the upper bound of the interval equal to the disclosure threshold. The reason is that the density being increasing implies the demand is higher for consumers closer to the threshold. Additionally, the lower the threshold, the higher the price the seller can charge once he knows that the consumers are below it.

Using these results we construct an equilibrium disclosure called the "greedy partition". In the first segment, consumers at the highest point are served at a price derived from the seller's indifference condition: the point where lowering the price to increase quantity sold yields the same profit as raising the price to increase price per consumer. Subsequent segments are formed by iteratively removing the covered segment and applying the same condition until all consumers are covered. Theorem 1 shows that the greedy partition maximizes the ex-ante consumer surplus.

Niche consumers are covered by a smaller segment and are worse off relative to mass consumers. This result follows from the seller indifference condition: at any fixed price, a marginal decrease in price causes a larger increase in sales of a mass product than in sales of a niche product. The greedy partition may have either finitely or infinitely many segments: if there is zero mass at the lowest point, the partition is

infinite; otherwise, it is finite.

Related literature: Our work is closest to Ali et al. (2023) and Hidir and Vellodi (2021). Ali et al. (2023) analyze optimal consumer segmentation, where the seller personalizes price but keeps product design fixed. We extend their framework to include both price and design personalization. Hidir and Vellodi (2021) study optimal consumer segmentation with personalization of both price and design, assuming uniform consumer distribution. By generalizing the analysis to distributions beyond the uniform, we are able to characterize how each consumer's conformity to preferences in the larger society (the density at that type) affects their willingness to disclose information and through that affects the offered products and prices.

2 Model

A monopolist, who chooses the design of a product, is facing a consumer. A consumer's preferred design (taste/location), x, is commonly known to be drawn from a distribution F over $[0,1]$ that allows for an increasing density f with $f(x) > 0$ for all $x \in (0,1]$. A consumer likes better a design that is closer to her preferred design. In particular, consumer x values the product with design (position at) ℓ at $1 - |x - \ell|$. A consumer's willingness-to-pay for the preferred design is normalized to 1, while $|x - \ell|$ captures the disutility from preference misalignment (or "transportation cost").

A consumer's preferred design is her private information, the seller only knows the distribution of tastes F. The consumer can, however, communicate her preferences by sending a message $m \in \mathcal{M}$, where $\mathcal{M} = 2^{[0,1]}$ with the interpretation that the consumer's taste belongs to the subset m of $[0,1]$. Given m, the seller offers a product with design ℓ and price $p \geq 0$. The monopolist can be thought of as either producing the product or as a platform that can offer a product of different designs.

Payoff of a consumer x (the consumer surplus) who is offered a product ℓ at price p is $1 - p - |x - \ell|$ if she buys the product and 0 otherwise. We assume that a consumer always breaks her indifference in favor of purchasing the product. The seller's payoff is p if the consumer buys the product and 0 otherwise; the seller has no production costs.

To summarize, the consumer and the seller play the following three stage game.

- $t = 1$: the consumer sends a message m about her preferred design;
- $t = 2$: the seller offers a personalized product to the consumer, with design ℓ and price p.
- $t = 3$: the consumer decides whether to buy the offered product or not.

The solution concept we adopt is the perfect Bayesian equilibrium (equilibrium, for short): the seller's belief system is updated via Bayes' Rule whenever possible, and both the consumer and the seller behave sequentially rationally. The goal of this paper is to characterize the equilibria that maximize the ex-ante consumer surplus.

Given (ℓ, p), a consumer x purchases the product if and only if $1 - |x - \ell| \geq p$. Let $\Delta(p) := 1 - p$ denote the reach at price p; that is, $\Delta(p)$ denotes the largest distance from the seller at which the consumer is willing to buy the good. $\Delta(.)$ is continuous and strictly decreasing in p, so the seller can be interpreted as choosing reach Δ (around ℓ) instead of p.

Henceforth, if a consumer x sends a message m, the seller will respond by offering a product with design ℓ and reach Δ. Payoff of the consumer is $1 - (1 - \Delta) - |x - \ell| = \Delta - |x - \ell|$ if she buys the product and 0 otherwise, payoff of the seller is $1 - \Delta$ if the consumer buys the product and 0 otherwise.

Discussion of the Model. *Interpretation.* The model can alternatively be thought of as a model of segmentation as in Bergemann et al. (2015).

Assumptions. We assume that the seller after a message offers only one product rather than a menu. This assumption is not without loss of generality but captures more realistically the behavior of online platforms; for a similar model see Hidir and Vellodi (2021). In addition, we assume that the density is monotonic. We are working on relaxing this assumption.

3 Analysis

The analysis centers on the consumer's fundamental trade-off when disclosing information about her tastes. By revealing more, she enables the seller to tailor the product

more closely to her preferences. The consumer values a better fit more highly and, all else equal, becomes more inclined to purchase. This is the *product steering* effect. However, more precise information about her preferred design allows the seller to appropriate a larger share of the consumer's surplus by offering the product at a personalized price—the *personalized pricing* effect.[3]

If the consumer were to fully reveal her preferred design, the seller would offer the design that the consumer prefers the most and charge price 1, thereby extracting all the surplus from the consumer. An equilibrium where the consumer fully reveals her type exists, and maximizes the total surplus, but leaves the consumer with no share of it.[4]

To obtain some surplus, the consumer must therefore preserve some privacy. This is achieved by revealing information coarsely through partitional equilibria (segmentation).

Without loss of generality, we can restrict attention to equilibrium outcomes that correspond to a countable partition. Such a partition is described by a sequence $(x_n) \subset [0,1]$ that partitions $[0,1]$, where $x_{i+1} \leq x_i$ for each i, and the interpretation that all the consumer types in $(x_{i+1}, x_i]$ send the same message.[5] Moreover, any consumer-surplus-maximizing partition must be *clearing*: if the seller knows that the consumer belongs to the partition cell $(x_{i+1}, x_i]$ for some i, it is optimal for the seller to cover the whole interval by offering the design $\ell = (x_{i+1} + x_i)/2$ and reach $\Delta = (x_i - x_{i+1})/2$.[6]

Notations: Seller's payoff from offering design $\ell = y - \Delta$ and reach Δ is denoted by $\pi(y; \Delta)$. Surplus of a consumer x that is offered a product with design $y - \Delta$ and reach Δ is denoted by $CS(x, y, \Delta)$ and total ex-ante consumer surplus attained by consumer types in $(a, b]$ when a seller offers a product with design $(a+b)/2$ and reach $(b-a)/2$ is denoted by $CS((a, b])$.

[3] Kim and Kos (2023) characterize the distribution of tastes that maximizes ex-ante consumer surplus in the absence of communication between consumer and seller.

[4] A type that lies would be offered a product different from their preferred product at price 1.

[5] If some other message is sent, the seller believes that the type is $x = 1$ with certainty and designs the product $\ell = 1$ with price $p = 1$. Payoff of the consumer type is zero in this case.

[6] If the seller does not to sell to all consumer types in the partition cell, consumer surplus can be increased by creating an additional partition cell for the non-purchasing consumer types.

3.1 Preliminary results

We start by characterizing some properties of the seller's best response.

Lemma 1. *For an increasing density f, suppose the consumer reveals that her preferred design is below some $x > 0$. Then the seller offers a product with design $x - \Delta$ for some $\Delta \in (0, 1/2)$.*

Proof. For any arbitrary Δ, the seller's payoff from offering design $\ell = y - \Delta$ and reach Δ is denoted by $\pi(y; \Delta)$ and is as follows.

$$\begin{aligned}\pi(y;\Delta) &= (1-\Delta)[F(\min\{y,x\}) - F(\min\{y-2\Delta,x\})] \\ &= (1-\Delta)\int_{\min\{y-2\Delta,x\}}^{\min\{y,x\}} f(y)dy\end{aligned}$$

Since f is increasing, the above expression is maximized when $y = x$. The optimal Δ is then determined by maximizing:

$$\pi(x;\Delta) = (1-\Delta)\int_{\max\{x-2\Delta,0\}}^{x} f(y)dy.$$

□

With increasing density f, if the seller learns the consumer's preferred design lies below a threshold x, then for any Δ he offers the highest possible design, $\ell = x - \Delta$, since consumer mass increases as the design approaches x. The seller's payoff from positioning at $x - \Delta$ and offering reach Δ is

$$\pi(\Delta; x) = (1-\Delta)[F(x) - F(x-2\Delta)]$$

The first order condition with respect to Δ can be written as

$$2(1-\Delta)f(x-2\Delta) = F(x) - F(x-2\Delta). \tag{1}$$

Totally differentiating the above equality with respect to x yields

$$2\frac{\partial \Delta}{\partial x} = \frac{2(1-\Delta)f'(x-2\Delta) - (f(x) - f(x-2\Delta))}{2(f(x-2\Delta) + (1-\Delta)f'(x-2\Delta))}.$$

The following result gives bounds for the reach offered by the seller as the disclosure threshold x varies.

Lemma 2. *Suppose that the density f is increasing, logarithmically concave and the consumer reveals that her preferred design is below x. The seller offers a product with design $x - \Delta$ and reach Δ, where,*

$$\frac{\partial \Delta}{\partial x} \geq 0$$

holds for all $x \in (0, 1)$.

Proof. To establish the statement, observe that the denominator of $\frac{\partial \Delta}{\partial x}$ is positive due to the assumption $f' > 0$. The numerator can be rewritten as

$$\begin{aligned}
2(1-\Delta)f'(x-2\Delta) - (f(x) - f(x-2\Delta)) &= \frac{F(x) - F(x-2\Delta)}{f(x-2\Delta)} f'(x-2\Delta) - (f(x) - f(x-2\Delta)) \\
&= (F(x) - F(x-2\Delta)) \left(\frac{f'(x-2\Delta)}{f(x-2\Delta)} - \frac{f(x) - f(x-2\Delta)}{F(x) - F(x-2\Delta)} \right) \\
&= (F(x) - F(x-2\Delta)) \left(\frac{f'(x-2\Delta)}{f(x-2\Delta)} - \frac{\int_{x-2\Delta}^{x} f'(t)dt}{\int_{x-2\Delta}^{x} f(t)dt} \right) \\
&\geq 0.
\end{aligned}$$

where the first equality follows by substituting the above optimality condition 1 and the inequality from $\frac{d^2 log f(x)}{dx^2} \leq 0$. Namely, the last inequality is equivalent to f'/f being decreasing, which in turn implies

$$\frac{f'(x-2\Delta)}{f(x-2\Delta)} \geq \frac{\int_{x-2\Delta}^{x} f'(t)dt}{\int_{x-2\Delta}^{x} f(t)dt}.$$

□

Thus, for any increasing, logarithmically concave density function, the seller offers a larger reach (smaller price) to a consumer that reveals a higher threshold x. We now

characterize some properties of ex-ante consumer surplus for such density functions.

Lemma 3. *For an arbitrary fixed y, suppose the seller offers a product with design $y-\Delta$ and reach Δ. For any increasing f, ex-ante consumer surplus increases as y increases.*

Proof. When the seller offers a product with design $y-\Delta$ and price $1-\Delta$, surplus obtained by a consumer located at $x \in (y-2\Delta, y]$ is $CS(x, y, \Delta)$, where

$$\begin{aligned} CS(x,y,\Delta) &= \Delta - |x-(y-\Delta)| \\ &= \begin{cases} x-y+2\Delta & \text{if } y-2\Delta < x \leq y-\Delta \\ y-x & \text{if } y-\Delta < x \leq y \end{cases} \end{aligned}$$

Ex-ante consumer surplus for consumer types in $(y-2\Delta, y]$ is $CS((y-2\Delta, y])$, where

$$CS((y-2\Delta, y]) = \int_{y-2\Delta}^{y-\Delta} (x-y+2\Delta) f(x)dx + \int_{y-\Delta}^{y} (y-x) f(x)dx.$$

Taking the first order derivative with respect to y and using the Leibniz integral rule gives

$$\frac{\partial CS((y-2\Delta, y])}{\partial y} = \int_{y-\Delta}^{y} f(x)dx - \int_{y-2\Delta}^{y-\Delta} f(x)dx \geq 0. \tag{2}$$

for any increasing density function f. □

For any fixed reach Δ, increasing y shifts the design $y-\Delta$ upward and this has two effects. The product becomes less appealing for the consumer types in $(y-2\Delta, y-\Delta]$ but more appealing for consumer types in $(y-\Delta, y]$. The overall effect on consumer surplus is weakly beneficial as there are weakly fewer consumer types in $(y-2\Delta, y-\Delta]$ relative to $(y-\Delta, y]$, for any increasing f.

The following lemma establishes how $CS((x-2\Delta, x])$ varies with Δ.

Lemma 4. *For an arbitrary fixed y, suppose the seller offers a product with design $y-\Delta$ and reach Δ. For any increasing f, ex-ante consumer surplus increases as Δ increases.*

Proof. The surplus attained by a consumer located at x, when offered a product with design $y - \Delta$ and reach Δ is $CS(x, y, \Delta)$, where

$$\begin{aligned} CS(x, y, \Delta) &= \Delta - |x - (y - \Delta)| \\ &= \begin{cases} x - y + 2\Delta & y - 2\Delta < x \leq y - \Delta \\ y - x & y - \Delta < x \leq y \end{cases} \end{aligned}$$

The ex-ante consumer surplus over the interval $(y - 2\Delta, y]$ is $CS((y - 2\Delta, y])$, where

$$CS((y - 2\Delta, y]) = \int_{y-2\Delta}^{y-\Delta} (x - y + 2\Delta) f(x) dx + \int_{y-\Delta}^{y} (y - x) f(x) dx$$

The partial derivative of the ex-ante consumer surplus with respect to Δ, using the Leibniz integral rule, is:

$$\frac{\partial CS((y - 2\Delta, y])}{\partial \Delta} = \int_{y-2\Delta}^{y-\Delta} 2 f(x) dx \geq 0$$

□

An increase in Δ has two effects. First, it shifts the design of the product, $y - \Delta$, downward. This makes the product less appealing for the types above $y - \Delta$ and brings the types below $y - \Delta$ closer to their preferred design. Second, increasing Δ corresponds to reducing the price for all types. For the types above $y - \Delta$, the two effects go in the opposite directions and cancel themselves out. For the consumers below $y - \Delta$ both effects are beneficial and thus reinforce each other.

4 Optimal consumer disclosure - greedy segmentation

4.1 Monotonic and log-concave density function

In this subsection we show that *the greedy partition* is the consumer optimal partition.

A countable partition is described by a sequence (x_n) in $[0,1]$ with $x_{k+1} \leq x_k$ for all k. Of particular interest is the greedy partition, denoted by (g_n). It is obtained recursively by setting $g_0 = 1$ and defining

$$g_n := 1 - 2\Delta_n,$$

where $\Delta_n = \arg\max_{\Delta \in [0,1]} (1-\Delta)(F(g_{n-1}) - F(g_{n-1} - 2\Delta))$ for $n \geq 1$. The choice of these partition points is made as follows. The sequence of design and reach corresponding to the greedy partition is (ℓ_n, Δ_n), with $\ell_n = g_n - \Delta_n$. For each Δ_n, the seller offers the highest possible design ℓ_n, as seen in Lemma 1. Each Δ_n is determined recursively; by starting from the highest types and grouping as many as possible in a segment, ensuring that all types buy the product. The following lemma establishes that the sequence (g_n) yields a partition of $(0,1]$.

Lemma 5. *Suppose f is an increasing and continuous density function with $f(x) > 0$ for all $x \in (0,1]$. Then, for every $x \in (0,1]$, there exists a g_n such that $g_n \leq x$.*

Proof. Suppose there was an $x > 0$ such that $g_n > x$, for all n. Then the sequence (g_n) is decreasing and bounded below by x. By the monotone convergence theorem there exists $\underline{x} := \lim_{n\to\infty} g_n \geq x$.

Consider the maximization problem $\pi(x,\Delta) = (1-\Delta)(F(x)-F(x-2\Delta))$ and notice that $\frac{d}{d\Delta}\pi(x,0) = 2f(x)$, and thus larger than 0 if and only if $f(x) > 0$. Moreover, the problem of maximizing $\pi(x,\Delta)$ over Δ is concave as f is increasing. Thus, as long as $f(x) > 0$, the optimal $\Delta(x)$ is unique and strictly positive. Moreover, by Lemma 2, $\Delta(x)$ is non-decreasing. By assumption $f(x) > 0$ for any $x \in (0,1]$ and thus, in particular, $f(\underline{x}) > 0$ and, finally, $\Delta(\underline{x}) > 0$. Since $\Delta(x)$ is non-decreasing, the greedy

partition covers the interval $[\underline{x}, 1]$ in at most $(1-\underline{x})/\lceil 2\Delta(\underline{x})\rceil$ steps. This contradicts the statement $\lim_{n\to\infty} g_n \geq x$, thereby proving the lemma.

□

Corollary 1. *The greedy partition covers the interval* $[0,1]$ *if and only if* $f(0) > 0$*, for any increasing and continuous density function* f*.*

Proof. The proof follows directly from the above lemma. □

In the greedy partition, the seller in each step naively optimizes only the profit from that step. In what follows, we show that the greedy partition achieves a higher consumer surplus than any other partition.

First, suppose that consumers disclose their valuation is in some interval $[y, x] \subseteq [0, 1]$. The seller then optimally chooses Δ to maximize $(1-\Delta)F([x-2\Delta, x] \cap [y, x])$, with Δ^* used to denote the maximizer[7]. The consumer surplus accruing from the interval $[y, x]$ is

$$\begin{aligned} CS([y,x]) &= \int_{x-2\Delta^*}^{x} (\Delta^* - |s - (x-\Delta^*)| f(s) ds \\ &= \int_0^{\Delta^*} s(f(x-2\Delta^*+s) + f(x-s)) ds \end{aligned}$$

where the first equality follows from the fact that consumer z's utility is $1-p-|z-\ell|$, where the price is $p = 1-\Delta$ and the design is $\ell = x-\Delta$, and the second by exchange of variables. The second equality demonstrates that each consumer's surplus is the difference between the largest disutility from the mismatch at which she would still buy and the actual disutility from the mismatch. In our case this is simply the distance to the closest boundary of the selling interval. Notice that there is one type of the consumer on each side of $x-\Delta^*$ with distance $s \in [0, \Delta^*)$ to the boundary.

Following Hidir and Vellodi (2022), we call a partition *clearing* if for every segment $[x_k, x_{k-1}]$, the seller optimally offers the product $\ell = (x_{k-1}+x_k)/2$ and reach $\Delta = (x_k - x_{k-1})/2$.[8] If a partition is not clearing, there is a partition cell $[x_j, x_{j-1}]$, such

[7] More generally, the seller chooses the product design and the reach to maximize $(1-\Delta)F([\ell - \Delta, \ell + \Delta] \cap [y, x])$.

[8] To avoid more cumbersome notation we will abuse terminology and write partition cells as closed intervals.

that upon learning the consumers are in this partition cell, the seller optimally sells only to consumers in the interval $[y, x_{j-1}]$ for some $y > x_j$. Since the consumers in $[x_j, y]$ do not buy and thus get consumer surplus 0, the consumers would be strictly better off disclosing whether they are in $[x_j, y]$ or $[y, x_{j-1}]$. A consumer surplus-maximizing partition is, therefore, clearing. Hereon we restrict attention to clearing partitions.

Lemma 6. *Suppose f is increasing and logarithmically concave. Let (x_n) be some countable clearing partition different from (g), let $k \in \mathbb{N}$ be the smallest j for which $x_j \neq g_j$, and let partition $(y_{k,n})$ be defined by:*

$$y_{k,i} = \begin{cases} g_i & \text{if } 1 \leq i \leq k, \\ x_i & \text{if } k < i. \end{cases}$$

Then $CS((y_{k,n})) \geq CS((x_n))$.

Proof. Since the partition (x_n) is clearing, the definition of the greedy partition implies that $g_k < x_k$, thus $g_{k-1} - 2\Delta((x_k, g_{k-1}]) = x_k$. Consider the partition $(y_{n,k})$, as defined in the statement of the result. The partition is well defined if the condition $x_{k+1} \leq g_k$ holds. Suppose $g_k < x_{k+1}$, then consumer surplus in (x_n) can be increased by considering another partition where the partition cells $(x_{k+1}, x_k]$ and $(x_k, x_{k-1}]$ are merged, keeping all other partition points the same. Thus, $x_{k+1} \leq g_k$ holds without loss of generality, making $(y_{k,n})$ a well defined clearing partition. We now prove the lemma by considering two cases.

Case 1: Suppose $x_k \in \left(x_{k+1}, \frac{x_{k+1}+g_{k-1}}{2}\right]$. Consider the consumer surplus from the partition cells $(x_{k+1}, x_k]$ and $(x_k, g_{k-1}]$):

$$CS((x_{k+1}, x_k]) + CS((x_k, g_{k-1}]) =$$
$$\int_{x_{k+1}}^{\frac{x_{k+1}+x_k}{2}} (x - x_{k+1}) f(x)dx + \int_{\frac{x_{k+1}+x_k}{2}}^{x_k} (x_k - x) f(x)dx + \int_{x_k}^{\frac{x_k+g_{k-1}}{2}} (x - x_k) f(x)dx + \int_{\frac{x_k+g_{k-1}}{2}}^{g_{k-1}} (g_{k-1} - x) f(x)dx.$$

The partial derivative with respect to x_k is

$$\frac{\partial}{\partial x_k}\left(CS((x_{k+1},x_k])+CS((x_k,g_{k-1}])\right)=\underbrace{\int_{\frac{x_{k+1}+x_k}{2}}^{x_k} f(x)\,dx}_{\text{length }\frac{x_k-x_{k+1}}{2}}-\underbrace{\int_{x_k}^{\frac{x_k+g_{k-1}}{2}} f(x)\,dx}_{\text{length }\frac{g_{k-1}-x_k}{2}}$$
$$\leq 0,$$

where the last inequality holds as for any $a\in\left(\frac{x_{k+1}+x_k}{2},x_k\right]$, for any $b\in\left(x_k,\frac{x_k+g_{k-1}}{2},x_k\right]$, $f(a)\leq f(b)$ holds since f increasing, and as $\frac{x_k-x_{k+1}}{2}\leq\frac{g_{k-1}-x_k}{2}$, by condition of Case 1. Moreover, since $g_k<x_k$ and $x_{k+1}\leq g_k\leq\frac{x_{k+1}+g_{k-1}}{2}$, the above inequality implies,

$$CS((x_{k+1},x_k])+CS((x_k,g_{k-1}])\leq CS((x_{k+1},g_k])+CS((g_k,g_{k-1}])$$

and consequently,

$$CS((y_{k,n}))-CS((x_n))=CS((x_{k+1},g_k])+CS((g_k,g_{k-1}])-(CS((x_{k+1},x_k])+CS((x_k,g_{k-1}]))$$
$$\geq 0.$$

holds, proving the lemma for Case 1.

Case 2: Suppose $x_k\in(\frac{x_{k+1}+g_{k-1}}{2},g_{k-1}]$. We argue that there exists a sequence (x'_n) that satisfies the condition of Case 1 and $CS((x'_n))>CS((x_n))$.

Consider the partition (x'_n), where,

$$x'_i=\begin{cases}x_i & \text{for all } i\neq k,\\ x_{k+1}+g_{k-1}-x_k & \text{for } i=k.\end{cases}$$

The term x'_k is interpreted as the mirror image of x_k in the interval $(x_{k+1},g_{k-1}]$ as $x'_k-x_{k+1}=g_{k-1}-x_k$. By Lemma 2, $x_k-x_{k+1}\leq g_{k-1}-g_k$ holds. This, in turn, implies $g_k\leq x'_k$, and thus the interval $(x'_k,g_{k-1}]$ is covered by the seller. Also, since (x_n) is a clearing partition, $(x_{k+1},x_k]$ is covered, which implies that the interval $(x_{k+1},x'_k]$ is also covered by the seller. Consider the difference in consumer surplus from the two

partitions.

$$CS((x'_n)) - CS((x_n)) = CS((x_{k+1}, x_k]) + CS((x_k, g_{k-1}]) - CS((x_{k+1}, x'_k] - CS((x'_k, g_{k-1}]). \tag{3}$$

The first two terms can be rewritten as

$$CS((x_{k+1}, x_k]) + CS((x_k, g_{k-1}]) = \tag{4}$$

$$= \int_0^{\frac{g_{k-1}-x_k}{2}} z\left(f(x_k+z) + f(g_{k-1}-z)\right) dz + \int_0^{\frac{x_k-x_{k+1}}{2}} z(f(x_{k+1}+z) + f(x_k-z))dz \tag{5}$$

$$= \int_0^{\frac{x_k-x_{k+1}}{2}} z h_1(z) dz, \tag{6}$$

where

$$h_1(z) = \begin{cases} f(x_k+z) + f(g_{k-1}-z) + f(x_{k+1}+z) + f(x_k-z), & \text{if } z \le \frac{g_{k-1}-x_k}{2}, \\ f(x_{k+1}+z) + f(x_k-z), & \text{if } z \in (\frac{g_{k-1}-x_k}{2}, \frac{x_k-x_{k+1}}{2}]. \end{cases}$$

Similarly,

$$CS((x_{k+1}, x'_k] + CS((x'_k, g_{k-1}]) = \tag{7}$$

$$= \left(\int_0^{\frac{x'_k-x_{k+1}}{2}} z(f(x_{k+1}+z) + f(x'_k-z))dz + \int_0^{\frac{g_{k-1}-x'_k}{2}} z(f(x'_k+z) + f(g_{k-1}-z))dz \right) \tag{8}$$

$$= \int_0^{\frac{x_k-x_{k+1}}{2}} z h_2(z) dz, \tag{9}$$

where

$$h_2(z) = \begin{cases} f(x'_k+z) + f(g_{k-1}-z) + f(x_{k+1}+z) + f(x'_k-z), & \text{if } z \le \frac{x'_k-x_{k+1}}{2}, \\ f(x'_k+z) + f(g_{k-1}-z), & \text{if } z \in (\frac{x'_k-x_{k+1}}{2}, \frac{g_{k-1}-x'_k}{2}]. \end{cases}$$

The definition of x'_k implies $\frac{g_{k-1}-x_k}{2} = \frac{x'_k-x_{k+1}}{2}$ and $\frac{x_k-x_{k+1}}{2} = \frac{g_{k-1}-x'_k}{2}$. Therefore for $z \leq \frac{x'_k-x_{k+1}}{2}$:

$$\begin{aligned} h_2(z) - h_1(z) &= f(x'_k + z) + f(x'_k - z) - (f(x_k + z) + f(x_k - z)) \\ &\leq 0, \end{aligned}$$

where the inequality follows from the fact that f is increasing and $x'_k < x_k$. Likewise, $h_2(z) - h_1(z) \geq 0$ if $z \in (\frac{x'_k-x_{k+1}}{2}, \frac{g_{k-1}-x'_k}{2}]$. The above two observations imply that $h_1(z)$ crosses $h_2(z)$ from above at $\frac{x'_k-x_{k+1}}{2}$. Now, one can define densities $\tilde{h}_i(z) = h_i(z)/(F(g_{k-1}) - F(x_{k+1}))$ for $i \in \{1, 2\}$ for all $z \in [x_{k+1}, g_{k-1}]$ I think for $z \in [0, \frac{x_k-x_{k+1}}{2}]$. Density $\tilde{h}_1$ crosses $\tilde{h}_2$ from above. Therefore, the cdf H_2 corresponding to $\tilde{h}_2$ first-order stochastically dominates H_1 corresponding to $\tilde{h}_1$. Consequently,

$$\int_0^{\frac{x_k-x_{k+1}}{2}} z h_1(z) dz = (F(g_{k-1}) - F(x_{k+1})) \int_0^{\frac{x_k-x_{k+1}}{2}} z \tilde{h}_1(z) dz \tag{10}$$

$$\leq (F(g_{k-1}) - F(x_{k+1})) \int_0^{\frac{x_k-x_{k+1}}{2}} z \tilde{h}_2(z) dz \tag{11}$$

$$= \int_0^{\frac{x_k-x_{k+1}}{2}} z h_2(z) dz. \tag{12}$$

Inequality eq. (10) together with eq. (4), eq. (7), and eq. (3) imply:

$$CS((x'_n)) - CS((x_n)) \geq 0.$$

Thus, consumer surplus under (x'_n) is higher than consumer surplus under (x_n). By Case 1, consumer surplus under the partition (y_k) is higher than the consumer surplus under the partition (x'_n). This completes the proof for Case 2 and thus for the lemma. □

The above lemma leads to the main result.

Theorem 1. *For any equilibrium outcome with disclosure strategy $\tilde{x}$, ex-ante consumer surplus is weakly higher under the greedy segmentation (g_n).*

Proof. Let (x_n) be some countable clearing partition. Construct a sequence of se-

quences $(y_{k,n})$ as follows:

$$y_{k,i} = \begin{cases} g_i, & \text{if } 1 \leq i \leq k, \\ x_i, & \text{if } k < i. \end{cases}$$

In the sequence $(y_{k,n})$ the first k terms are equal to the first k terms of the segmentation g while the rest of the sequence coincides with the sequence (x_n).

Since $CS((x_n)) = \sum_{i=1}^{\infty} CS((x_{i+1}, x_i])$ is a convergent series, for any $\epsilon/2 > 0$ there exists an N large enough so that $\sum_{i=j}^{\infty} CS((x_{i+1}, x_i]) < \epsilon/2$ for all $j > N$. A corresponding N' exists for the series $CS(g)$. By taking an $M = \max\{N, N'\}$, one has that for $k > M$:

$$\begin{aligned} |CS(y_k) - CS(g)| &= |\sum_{i=k}^{\infty} CS((y_{k,i+1}, y_{k,i}]) - \sum_{i=k}^{\infty} CS((g_{i+1}, g_i])| \\ &= |\sum_{i=k}^{\infty} CS((x_{i+1}, x_i]) - \sum_{i=k}^{\infty} CS((g_{i+1}, g_i])| \\ &\leq \sum_{i=k}^{\infty} CS((x_{i+1}, x_i]) + \sum_{i=k}^{\infty} CS((g_{i+1}, g_i]) \\ &\leq \epsilon \end{aligned}$$

where the first and the second equality follow from the definition of the sequence $(y_{k,i})_i$. Sequence $CS(y_k)$, therefore, converges to $CS(g)$ as k goes to infinity.

Lemma 5 implies

$$\begin{aligned} CS(y_k) - CS(y_{k-1}) &= CS((x_{k+1}, g_k]) + CS((g_k, g_{k-1}]) - CS((x_{k+1}, x_k]) - CS((x_k, g_{k-1}]) \\ &\geq 0. \end{aligned}$$

Thus $CS(y_k)$ is weakly increasing in k. For each k, $CS(y_k) \leq 1$ as ex-ante consumer surplus for any partition is bounded by 1. By the monotone convergence theorem, the sequence $\{CS(y_k)\}$ converges to $\sup_{k \in \mathbb{N}} CS(y_k)$.

The above two paragraphs imply that $CS(g) = \sup_{k\in\mathbb{N}} CS(y_k)$, and therefore

$$CS(x) = CS(y_1) \leq \sup_{k\in\mathbb{N}} CS(y_k) = CS(g).$$

This completes the proof of the theorem. □

5 Greedy is not optimal

In this section we show that the greedy segmentation is not always optimal for ex-ante consumer surplus.

5.1 Single peak density function

Consider a density function f_a given by

$$f_a(x) = \begin{cases} \dfrac{ax + \left(1 - \frac{a}{2}\right)}{1 - \frac{a}{4}}, & x \in \left[0, \frac{1}{2}\right], \\ \dfrac{a(1-x) + \left(1 - \frac{a}{2}\right)}{1 - \frac{a}{4}}, & x \in \left(\frac{1}{2}, 1\right], \end{cases}$$

where $a \in (0, 2]$. Then f_a is a symmetric, non-monotonic density function, with peak at $x = 1/2$. With slight abuse of notation, we drop subscript a and denote f_a by f.

A consumer of type x who receives offer with design ℓ and price p obtains utility,

$$u_C(x) \;=\; 1 - p \;-\; c\,|x - \ell|,$$

where $c > 1$ scales the disutility from mismatch $|x - \ell|$.

Define $\Delta := \frac{1-p}{c}$. The above utility then becomes,

$$u_C(x) \;=\; c\Delta \;-\; c\,|x - \ell|,$$

and the payoff of the seller is,

$$\pi(\Delta) \;=\; (1 - c\Delta),$$

if the consumer buys the product.

i) CS under the symmetric partition $\left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]$.

Consider the symmetric disclosure that splits the market exactly at $1/2$. For any $a \in (0, 2]$, choose $c(a)$ such that $\Delta = \frac{1}{4}$ is the seller optimal choice. To do so, maximize the seller's expected payoff,

$$\mathbb{E}(\pi(\Delta; x)) = (1 - c\Delta)\left[F(x) - F(x - 2\Delta)\right],$$

which gives the corresponding first order condition as,

$$2(1 - c\Delta)\, f(x - 2\Delta) = c\left[F(x) - F(x - 2\Delta)\right].$$

Substituting $x = \frac{1}{2}$ and $\Delta = \frac{1}{4}$,

$$2\Big(1 - \frac{c}{4}\Big)\, f(0) = c\left[F(\tfrac{1}{2}) - F(0)\right],$$

where,

$$f(0) = \frac{1 - \frac{a}{2}}{1 - \frac{a}{4}} = \frac{2(2 - a)}{4 - a},$$

and by symmetry,

$$F(\tfrac{1}{2}) - F(0) = \int_0^{1/2} f(t)\, dt = \frac{1}{2}.$$

The first order condition then becomes,

$$2\Big(1 - \frac{c}{4}\Big)\frac{2(2 - a)}{4 - a} = \frac{c}{2}.$$

Simplifying gives,

$$c(a) = \frac{8(2 - a)}{8 - 3a} = \frac{2 - a}{1 - \frac{3a}{8}}. \tag{13}$$

Thus, for $a = \frac{1}{4}$,

$$c(a) = \frac{8(2-a)}{8-3a} = \frac{56}{29}.$$

Further, for $a = 1/4$, the density simplifies to

$$f(x) = \frac{ax + (1 - \frac{a}{2})}{1 - \frac{a}{4}} = \frac{\frac{1}{4}x + \frac{7}{8}}{\frac{15}{16}} = \frac{4x + 14}{15}, \qquad x \in \left[0, \tfrac{1}{2}\right],$$

and by symmetry $f(x) = f(1-x)$ on $(\frac{1}{2}, 1]$. The ex-ante consumer surplus under this partition is as follows.

$$\begin{aligned}
\frac{CS\left([0, \frac{1}{2}], [\frac{1}{2}, 1]\right)}{2c} &= \int_0^{\frac{1}{4}} z[f(0+z) + f(\tfrac{1}{2} - z)]\, dz \\
&= \int_0^{\frac{1}{4}} z[2 - \tfrac{a}{2}]\, dz \\
&= \frac{1}{32}\left(2 - \frac{a}{2}\right) \\
CS &= \frac{c}{16}\left(2 - \frac{a}{2}\right)
\end{aligned}$$

where the second equality is obtained by substituting f. Thus, for $a = \frac{1}{4}$, $c = \frac{56}{29}$,

$$CS_{\text{2-part}} = \frac{\frac{56}{29}}{16}\left(2 - \frac{1}{8}\right) = \left(\frac{56}{464}\right)\left(\frac{15}{8}\right) = \frac{105}{464}.$$

ii) CS under greedy segmentation.

Under the greedy segmentation, the seller offers design $\ell = 1/2$ and the corresponding profit maximizing reach. For reach Δ, the partition cell is $\left[\frac{1}{2} - \Delta,\ \frac{1}{2} + \Delta\right]$ and the expected seller profit is,

$$\mathbb{E}(\pi(\Delta)) \ = \ \Big(1 - c\,\Delta\Big)\ \Big(F\Big(\tfrac{1}{2} + \Delta\Big) - F\Big(\tfrac{1}{2} - \Delta\Big)\ \Big),$$

For $\Delta \in [0, \frac{1}{2}]$,

$$F\Big(\tfrac{1}{2}+\Delta\Big) - F\Big(\tfrac{1}{2}-\Delta\Big) = \int_{\frac{1}{2}-\Delta}^{\frac{1}{2}+\Delta} f(t)\,dt,$$
$$= \frac{4\,\Delta\,\Big(2 - a\Delta\Big)}{4-a}.$$

Hence,

$$\mathbb{E}(\pi(\Delta)) = \Big(1 - c\Delta\Big)\,\frac{4\,\Delta\,\Big(2 - a\Delta\Big)}{4-a}.$$

The corresponding first-order condition can be written as the quadratic equation,

$$\frac{3}{2}\,ac\,\Delta^2 \;-\; (2c+a)\,\Delta \;+\; 1 \;=\; 0.$$

Solving yields,

$$\Delta^* = \frac{(2c+a) \;-\; \sqrt{(a+2c)^2 - 6ac}}{3ac}.$$

Plugging $a = \frac{1}{4}$, $c = \frac{56}{29}$ gives,

$$\frac{21}{29}\Delta^2 \;-\; \frac{477}{116}\Delta \;+\; 1 = 0.$$

The smaller root is the optimal reach Δ^*:

$$\Delta^* = \frac{\frac{477}{116} - \sqrt{\Big(\frac{477}{116}\Big)^2 - \frac{84}{29}}}{2 \cdot \frac{21}{29}}.$$

The consumer surplus for greedy segmentation can be written as

$$CS_{\text{greedy}} = 2c\,[I_1(\Delta^*) + I_2(\Delta^*)],$$

where,

$$I_1(\Delta) = \int_0^{\frac{1}{2}\left(\frac{1}{2}-\Delta\right)} z\Big[f(z) + f\left(\tfrac{1}{2} - \Delta - z\right)\Big]\, dz, \qquad I_2(\Delta) = \int_0^{\Delta} z\, f\left(\tfrac{1}{2} - \Delta + z\right)\, dz.$$

Computing the bracket $f(z) + f\left(\frac{1}{2} - \Delta - z\right)$:

$$f(z) = \frac{4z + 14}{15},$$
$$f\Big(\tfrac{1}{2} - \Delta - z\Big) = \frac{4(\frac{1}{2} - \Delta - z) + 14}{15} = \frac{2 - 4\Delta - 4z + 14}{15} = \frac{16 - 4\Delta - 4z}{15}.$$

Hence,

$$\begin{aligned} f(z) + f\Big(\tfrac{1}{2} - \Delta - z\Big) &= \frac{4z + 14}{15} + \frac{16 - 4\Delta - 4z}{15}, \\ &= \frac{30 - 4\Delta}{15} = 2 - \frac{4\Delta}{15}. \end{aligned}$$

Therefore,

$$\begin{aligned} I_1(\Delta) &= \int_0^{\frac{1}{2}(\frac{1}{2}-\Delta)} z\Big(2 - \tfrac{4\Delta}{15}\Big)\, dz, \\ &= \left(2 - \frac{4\Delta}{15}\right) \int_0^{\frac{1}{2}(\frac{1}{2}-\Delta)} z\, dz, \\ &= \frac{2 - \frac{4\Delta}{15}}{2} \left(\frac{1}{4} - \frac{\Delta}{2}\right)^2. \end{aligned}$$

Expanding this expression yields the cubic polynomial,

$$I_1(\Delta) = -\frac{1}{30}\Delta^3 + \frac{17}{60}\Delta^2 - \frac{31}{120}\Delta + \frac{1}{16}.$$

Simplifying the integrand in $I_2(\Delta)$:

$$\begin{aligned} f\left(\tfrac{1}{2}-\Delta+z\right) &= \frac{4(\frac{1}{2}-\Delta+z)+14}{15}, \\ &= \frac{2-4\Delta+4z+14}{15}, \\ &= \frac{16-4\Delta+4z}{15}. \end{aligned}$$

Thus,

$$I_2(\Delta) = \int_0^{\Delta} z\frac{16-4\Delta+4z}{15}\,dz = \frac{1}{15}\int_0^{\Delta}\left(16z-4\Delta z+4z^2\right)dz.$$

Compute term by term:

$$\begin{aligned} \int_0^{\Delta} 16z\,dz &= 8\Delta^2, \\ \int_0^{\Delta}(-4\Delta z)\,dz &= -2\Delta^3, \\ \int_0^{\Delta} 4z^2\,dz &= \frac{4}{3}\Delta^3. \end{aligned}$$

Hence,

$$\begin{aligned} I_2(\Delta) &= \frac{1}{15}\left(8\Delta^2-2\Delta^3+\tfrac{4}{3}\Delta^3\right), \\ &= \frac{1}{15}\left(8\Delta^2-\tfrac{2}{3}\Delta^3\right), \\ &= -\frac{2}{45}\Delta^3+\frac{8}{15}\Delta^2. \end{aligned}$$

Adding the two polynomials,

$$\begin{aligned} I_1(\Delta)+I_2(\Delta) &= \left(-\frac{1}{30}\Delta^3+\frac{17}{60}\Delta^2-\frac{31}{120}\Delta+\frac{1}{16}\right)+\left(-\frac{2}{45}\Delta^3+\frac{8}{15}\Delta^2\right), \\ &= -\frac{7}{90}\Delta^3+\frac{49}{60}\Delta^2-\frac{31}{120}\Delta+\frac{1}{16}. \end{aligned}$$

For $a=\frac{1}{4}$,

$$I_1(\Delta^*)+I_2(\Delta^*) = -\frac{7}{90}\Delta^{*3}+\frac{49}{60}\Delta^{*2}-\frac{31}{120}\Delta^*+\frac{1}{16},$$

Hence,

$$CS_{\text{greedy}} = \frac{112}{29}\left(-\frac{7}{90}\Delta^{*3} + \frac{49}{60}\Delta^{*2} - \frac{31}{120}\Delta^{*} + \frac{1}{16}\right). \\ = \frac{2914003}{146160} - \frac{59831}{1315440}\sqrt{188553}.$$

Thus,

$$CS_{\text{2-part}} = \frac{105}{464} \approx 0.2263, \qquad CS_{\text{greedy}}(\Delta^*) \approx 0.1869.$$

and

$$CS_{\text{greedy}} < CS_{\text{2-part}},$$

showing that for a non-monotonic and log concave density function f_a, $a = \frac{1}{4}$, ex-ante consumer surplus under the greedy segmentation is strictly lower than under the symmetric two-partition.

5.2 Monotonic not-log concave density function

Consider first the following observation.

Proposition 1. *Suppose the density, f, is strictly increasing. Let x_1 and x_2 be the first two thresholds created by the greedy partition. Then*

$$1 - F(x_1) \geq F(x_1) - F(x_2).$$

Proof. Since we have a greedy partition

$$(1 - F(x_1))(1 - c(\frac{1 - x_1}{2})) \geq (F(x_1) - F(x_2))(1 - c(\frac{x_1 - x_2}{2})).$$

In addition, the first order conditions for x_1 and x_2 are

$$-f(x_1)(1 - c(\frac{1 - x_1}{2})) + (1 - F(x_1))\frac{1}{2}c'(\frac{1 - x_1}{2})) = 0, \\ -f(x_2)(1 - c(\frac{x_1 - x_2}{2})) + (F(x_1) - F(x_2))\frac{1}{2}c'(\frac{x_1 - x_2}{2}) \leq 0.$$

Expressing the $1-c$ term from the FOCs and plugging it into the inequality yields

$$\frac{(1-F(x_1))^2}{f(x_1)}c'(\frac{1-x_1}{2}) \geq \frac{(F(x_1)-F(x_2))^2}{f(x_2)}c'(\frac{x_1-x_2}{2}).$$

We are assuming that c is weakly convex, thus c' nondecreasing. There are two cases to consider, if $x_1 - x_2 \geq 1 - x_1$, then $c'(\frac{x_1-x_2}{2}) \geq c'(\frac{1-x_1}{2})$, and since the density is increasing (and $x_1 \geq x_2$), for the above inequality to hold, it must be the case that $1 - F(x_1) \geq F(x_1) - F(x_2)$. On the other hand, if $1 - x_1 \geq x_1 - x_2$, then $1 - F(x_1) \geq F(x_1) - F(x_2)$ because the density is nondecreasing.

□

The above result establishes that as long as the density is nondecreasing, the greedy partition leads to a decreasing quantity: the first partition cell has the highest quantity, etc.

Notice, however, that this does not guarantee that the greedy partition leads to a welfare optimum. For that, one would need a stronger result, that $F((1+x_1)/2) - F(x_1) \geq F(x_1) - F((x_1+x_2)/2)$ always holds. The following example demonstrates that increasing density is not sufficient to guarantee that the greedy partition delivers a welfare optimal allocation.

Consider the density function

$$f(x) = be^{x^a}, \qquad x \in [0,1],$$

where $a = 5$ and $b \approx 0.8164$ is chosen so that $\int_0^1 f(x)\,dx = 1$. Then f is strictly increasing on $[0,1]$ and not log-concave, since $\log f(x) = \log b + x^5$ is strictly convex.

The seller's profit from covering the interval $[x - 2\Delta, x] \subset [0,1]$ is

$$\pi(\Delta; x) = (1 - c\Delta)\Big(F(x) - F(x - 2\Delta)\Big),$$

where F is the cdf associated with f, and $c > 0$ is a cost parameter. Fix $c = 1.8$. Starting from $x = 1$, the seller chooses Δ_1 to maximize $\pi(\Delta; 1)$. This yields an interior

optimum

$$\Delta_1 \approx 0.2302, \qquad x_1 = 1 - 2\Delta_1 \approx 0.5396,$$

so the first interval covered is $[x_1, 1]$.

Starting from x_1, the seller again maximizes $\pi(\Delta; x_1)$ over $\Delta \in (0, x_1/2]$. For the above parameters,

$$\pi'\Big(\frac{x_1}{2}; x_1\Big) > 0$$

so profit is increasing up to the boundary and the optimal choice is

$$\Delta_2 = \frac{x_1}{2}.$$

Therefore, $x_2 = 0$ and the greedy partition is $[0, x_1]$ and $[x_1, 1]$. The corresponding consumer surplus is

$$CS_{greedy} \approx CS(0, 0.54) + CS(0.54, 1) \approx 1.8[0.05977 + 0.06004] = 1.8(0.11982) = 0.215676.$$

Further, with these endogenous thresholds x_1 and x_2, we have

$$\begin{aligned}
\frac{\partial CS(x_2, 1)}{\partial x_1} &= \Big(F(x_1) - F\Big(\tfrac{x_1+x_2}{2}\Big)\Big) - \Big(F\Big(\tfrac{1+x_1}{2}\Big) - F(x_1)\Big) \\
&\approx F(0.5396) - F(0.5396/2) - (F(1.5396/2) - F(0.5396)) \\
&\approx F(0.5396) - F(0.2698) - \Big(F(0.7698) - F(0.5396)\Big) \\
&\approx 0.4439 - 0.2203 - (0.6590 - 0.4439) \\
&\approx 0.2236 - 0.2151 \\
&\approx 0.0085 \ > \ 0.
\end{aligned}$$

The above establishes that by slightly increasing x_1 from the greedy partition value, one obtains higher welfare; and since $\pi'\Big(\frac{x_1}{2}; x_1\Big) > 0$, shifting x_1 to the right keeps the

intervals covered by the seller. Specifically, take $x_1' = 0.55$. Then

$$\begin{aligned} CS_{higherx_1} &\approx CS(0, 0.55) + CS(0.55, 1) \\ &\approx 1.8[0.06203 + 0.05791] = (1.8)(0.11995) = 0.21591 > 0.215676 \approx CS_{greedy}. \end{aligned}$$

5.3 Characterization of equilibria

Proposition 2. *An equilibrium is necessarily a partition, with each partition cell being connected.*

Proof. Consider an equilibrium disclosure strategy that is not connected partitional. Then there exists an interval m and an interior point $x \in m$ such that x reports some $m' \neq m$. This implies

$$CS(m'; x) > CS(m; x).$$

Further, being an equilibrium, $CS(\inf m) = CS(\sup m) = 0$.

Recall that payoff of a type is its distance from the nearest end point. Since x is better off reporting m', it must be the case that either $\inf m' < \inf m$ or $\sup m < \sup m'$. Thus, either

$$CS(m'; \inf m) > 0$$

or

$$CS(m'; \sup m) > 0,$$

which implies that either $\inf m$ or $\sup m$ is better off deviating to m', thereby contradicting the initial assumption that an equilibrium disclosure strategy is being considered. Thus, any equilibrium strategy must be a connected partition.

□

6 Conclusion

Our analysis demonstrates how voluntary disclosure of consumer data shapes market outcomes when a monopolist personalizes both product design and pricing. We characterize the greedy partition, a sequential segmentation that begins with the highest-

demand region and expands by adding a segment whenever lowering the price to attract additional consumers yields the same profit as raising the price per consumer. This equilibrium maximizes ex-ante consumer surplus. Nevertheless, disclosure incentives differ systematically across consumer types: niche consumers disclose more information, receive products better aligned with their preferences, and face higher prices, leaving them worse off than mass consumers. These results reveal an inherent asymmetry in data-driven personalization: although all consumers ultimately purchase, the distribution of surplus is uneven. The findings underscore the importance of accounting for heterogeneity in consumer preferences when assessing the implications of data disclosure and the broader welfare effects of privacy regulation in digital markets.

References

S. N. Ali, G. Lewis, and S. Vasserman. Voluntary disclosure and personalized pricing. *The Review of Economic Studies*, 90(2):538–571, 2023.

D. Bergemann, B. Brooks, and S. Morris. The limits of price discrimination. *American Economic Review*, 105(3):921–957, 2015.

S. Hidir and N. Vellodi. Privacy, personalization, and price discrimination. *Journal of the European Economic Association*, 19(2):1342–1363, 2021.

K. Kim and N. Kos. Monopoly with product design. *Available at SSRN 4425071*, 2023.