Electronic Companion

Notation	
d	Dimensionality of the offering feature space
$\mathcal{D} \triangleq [0,1]^d$	Domain of the offering feature space
$\boldsymbol{\gamma}$	An offering in $\mathcal{D}$
$\mathbf{x}^*$	A user's optimal offering
$y(\mathbf{x}\|\mathbf{x}^*)$	Response function for a user with optimal offering $\mathbf{x}^*$
$f(\mathbf{x}^*)$	Probability density function of users' optimal offerings
Γ	A set of offerings with cardinality G
τ	The suitability threshold
$\boldsymbol{\gamma}(\mathbf{x}^*)$	A user's response-maximizing offering out of a set Γ
c	The fixed cost of an offering
$\mathcal{C}(\boldsymbol{\gamma})$	The domain covered by $\boldsymbol{\gamma}$
Section 4	
$\psi(x)$	Offering intensity function
ψ^τ	Suitability threshold for the smoothed model
$\bar{y}(\psi(x^*))$	Average response value of a user with optimal offering x^* and offering intensity $\psi(x^*)$ in the smoothed model
r	Critical density, defined as $\frac{c\psi^\tau}{P\bar{y}(\psi^\tau)}$
$\mathcal{R}_1$, $\mathcal{R}_2$	Regions with $\psi^*(x) > \psi^\tau$ and $\psi^*(x) = \psi^\tau$, respectively
l, h, s	Parameters of the tent distribution (Section 4.1 only)
$\tilde{F}(l)$	Highest-density size-l set in the domain $[0,1]$ (Section 4.2 only)
L^*	The natural level of inclusivity achieved by S-ODP (Section 4.2 only)
λ	Level of inclusivity required in S-ODP(λ) and S-ODP(λ,Ψ) (Section 4.2 only)
Ψ	Total intensity used under $\psi^*(x)$ (Section 4.2 only)

Table EC.1 Table of the main notation used in the paper.

EC.1. Proofs of Results

EC.1.1. Proof of Theorem 1

The hitting set problem is formulated as follows: given a universe of discrete elements $\mathcal{U}$ and a collection of sets $\mathcal{S}$, find a set $\mathcal{M}$ of minimum cardinality such that every element of $\mathcal{S}$ contains at least one element in $\mathcal{M}$. Mathematically this can be written as the following optimization problem, where $x_u \in \{0,1\}$ indicates whether element $u \in \mathcal{U}$ is chosen.

$$\begin{aligned} \min_{\{x_u\}_{u\in\mathcal{U}}} & \sum_{u\in\mathcal{U}} x_u \\ s.t. & \sum_{u\in s} x_u \geq 1 \,\forall\, s\in\mathcal{S} \\ & x_u \in \{0,1\} \end{aligned} \tag{EC.1}$$

This problem (mathematically equivalent to the set covering problem) is known to be NP-hard (Garey and Johnson 1979). We will show that I-ODP is equivalent to this problem. Consider an arbitrary instance of the set covering problem where $\mathcal{U}$ is an arbitrarily fine mesh of the interval $[0,1]$. Because we allow for an arbitrarily fine mesh, considering the discrete version of I-ODP is without loss of generality. Let the elements of $\mathcal{S}$ be indexed by i, so $s_i \in \mathcal{S}$ is the $i-$th set in $\mathcal{S}$. Let N be the cardinality of $\mathcal{S}$.

Now, define N response functions over the elements in $\mathcal{U}$, where $f_i(x) = \tau$ if $x \in s_i$ and $f_i(x) = 0$ otherwise. The discrete version of I-ODP can be written as

$$\begin{aligned}
&\min_{\mathbf{z}} c \sum_{u\in\mathcal{U}} \mathbb{1}\{\sum_{i=1}^{N} z_{u,i} > 0\} \\
&s.t. \sum_{u\in\mathcal{U}} z_{u,i} \geq 1 \ \forall\, i \in [\![N]\!] \\
&\qquad \sum_{u\in\mathcal{U}} f_i(u) z_{u,i} \geq \tau \ \forall\, i \in [\![N]\!] \\
&\qquad z_{u,i} \in \{0,1\} \ \forall\, g \in [\![G]\!],\, i \in [\![N]\!]
\end{aligned} \tag{EC.2}$$

where $z_{u,i} = 1$ if the ith user (with response function $f_i(x)$) is assigned to element u. Notice that we force each user to be assigned to *at least* one element. Because $f_i(x) \in \{0, \tau\}$, if $\sum_{u\in\mathcal{U}} f_i(u) z_{u,i} \geq \tau$ then there is at least one u for which $f_i(u) z_{u,i} \geq \tau$.

Let $\mathbf{z}^*$ be the optimal solution to Problem EC.2. Let $x_u = \mathbb{1}\{\sum_{i=1}^{N} z^*_{u,i} > 0\}$, which captures whether element u is used (i.e., assigned to any users). We will show that this solution is optimal for Problem EC.1. Therefore, if Problem I-ODP is not NP-hard, it must be the case that the hitting set problem is not NP-hard.

We will proceed by showing that the first two constraints of Problem EC.2 are equivalent to the first constraint of Problem EC.1. First, we prove the following two claims.

Claim 1. The second constraint of Problem EC.2 implies the first constraint. To see this, notice that if the first constraint does not hold, then $\sum_u z_{u,i} = 0$ and therefore $\sum_u f_i(u) z_{u,i} = 0$ implying that the second constraint cannot hold. Therefore, the first constraint is redundant and can be removed.

Claim 2. If the constraint $z_{u,i} \geq \mathbb{1}\{\sum_{j=1}^{N} z_{u,j} > 0\}$ for all u,i is added to Problem EC.2, the resulting problem is equivalent. This constraint says that if element u is "turned on" (i.e., assigned to at least one user), it should be activated for all users i. To see that adding this constraint results in an equivalent problem, notice that once $\sum_j z_{u,j}$ is non-zero, increasing the quantity has no impact on the objective function. Additionally, as $z_{u,i}$ always appears on the left-hand side of the inequalities in Problem EC.2, increasing it does not impose any new restrictions.

This new constraint ensures that $z_{u,i} = 1$ if $\mathbb{1}\{\sum_{j=1}^{N} z_{u,j} > 0\} \geq 1$. Additionally, it also guarantees that $z_{u,i} = 0$ if $\mathbb{1}\{\sum_{j=1}^{N} z_{u,j} > 0\} = 0$. This is because if $z_{u,i} = 1$, it must be the case that $\mathbb{1}\{\sum_{j=1}^{N} z_{u,j} > 0\} = \mathbb{1}\{\sum_{j\neq i} z_{u,j} + z_{u,i} > 0\} \geq 1$. Therefore, this new constraint is equivalent to $z_{u,i} = \mathbb{1}\{\sum_{j=1}^{N} z_{u,j} > 0\}$ for all u,i.

Now consider again the second constraint of Problem EC.2, given by $\sum_{u\in\mathcal{U}} f_i(u) z_{u,i} \geq \tau$ for all $i \in [\![N]\!]$. Recalling that $f_i(u) \in \{0,\tau\}$, this can be equivalently written as $\sum_{u\in\{u:\, f_i(u)\geq\tau\}} z_{u,i} \geq 1$. Now, employing the new constraint above allows us to write this as as $\sum_{u\in\{u:\, f_i(u)\geq\tau\}} \mathbb{1}\{\sum_{j=1}^{N} z_{u,j} > 0\} \geq 1$

Recalling that $s_i = \{u :\ f_i(u) \geq \tau\}$, and that $x_u = \mathbb{1}\{\sum_j z_{u,j} > 0\}$, this constraint is equivalent to $\sum_{u\in s_i} x_u \geq 1$—the first constraint of Problem EC.1. This also shows that the objective function of Problem EC.2 is equivalent to that of Problem EC.1.

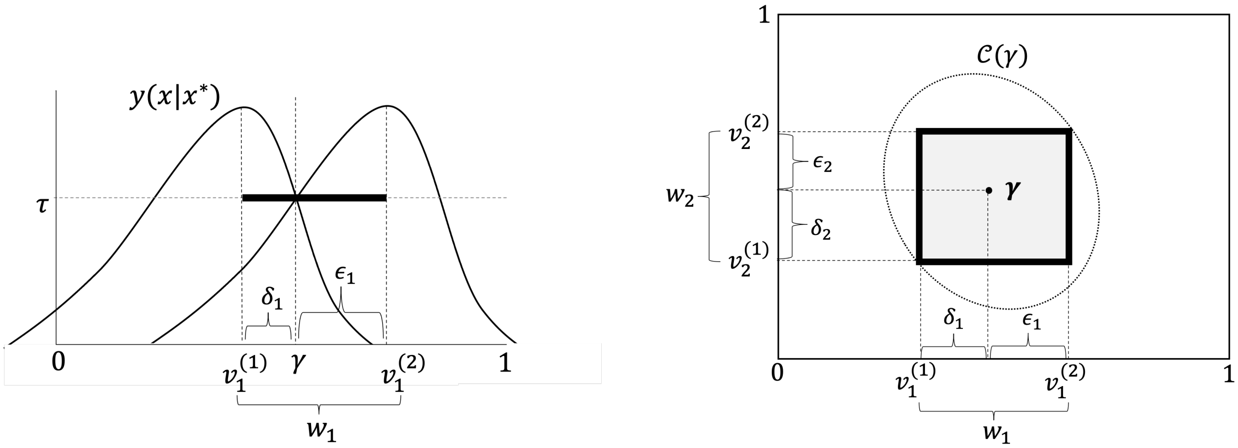
Therefore, given an arbitrary instance of the set hitting problem, it can be converted into an instance I-ODP. If I-ODP could be solved in polynomial time, it would imply that the set hitting problem could therefore be solved in polynomial time. □

EC.1.2. Proof of Proposition 1

Because $\mathcal{S}_k \subseteq \mathcal{C}(\boldsymbol{\gamma}_k)$, it follows that $\mathcal{D} = \cup_k \mathcal{C}(\boldsymbol{\gamma}_k)$. Therefore, for every $\mathbf{x}^* \in \mathcal{D}$, there exists a $\mathcal{C}(\boldsymbol{\gamma}_k)$ such that $\mathbf{x}^* \in \mathcal{C}(\boldsymbol{\gamma}_k)$. By definition of $\mathcal{C}(\boldsymbol{\gamma}_k)$, this implies that $\boldsymbol{\gamma}_k$ is a suitable offering for user with optimal offering $\mathbf{x}^*$. Since each $\boldsymbol{\gamma}_k \in \mathcal{D}$, $\{\boldsymbol{\gamma}_k\}$ is a feasible solution to I-ODP. □

EC.1.3. Proof of Corollary 1 and Corollary 2

Proof of Corollary 1. Let $\mathcal{S} \subseteq \mathcal{C}(\boldsymbol{\gamma})$ be the maximal volume hyperrectangle with side length w_j in the jth dimension, with minimal and maximal vertices $\mathbf{v}^{(1)}$ and $\mathbf{v}^{(2)}$ given by the solution to



(a) $d = 1$. (b) $d = 2$.

Figure EC.1 Illustration of notation used in Even Segmentation for $d = 1$ (left) and $d = 2$ (right). In both figures, the thick black line outlines $\mathcal{S}$.

Problem 5. Let $w_j \triangleq v_j^{(2)} - v_j^{(1)}$. Additionally, let $\epsilon_j \triangleq v_j^{(2)} - \gamma_j$, and $\delta_j \triangleq \gamma_j - v_j^{(1)}$. See Figure EC.1 for illustrations of this notation First, notice that $\epsilon_k + \delta_k = w_k$. Therefore, we the following also holds: $\left\lceil \frac{1-\epsilon_k-\delta_k}{w_k} \right\rceil + 1 = \left\lceil \frac{1}{w_k} \right\rceil$.

First we consider the case when $d = 1$. In this case, the covering region $\mathcal{C}(\gamma)$ is equivalent to the segment $[v_1^{(1)}, v_1^{(2)}]$, and both have length w_1. Suppose an optimal solution contains fewer than $\left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil + 1$ offerings. Let the offerings be $\Gamma = \{\gamma_1, ..., \gamma_G\}$ for $G \leq \left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil$, ordered so that $\gamma_1 \leq \gamma_2 \leq ... \leq \gamma_G$. First, we note that any solution to I-ODP must have $\gamma_1 \leq \delta_1$ and $\gamma_G \geq 1 - \epsilon_1$. To see this, suppose by contradiction that $\gamma_1 > \delta_1$. Then with optimal offering $x^* = 0$ is not covered by γ_1. To be covered by γ_1, we require $x^* \in [v_1^{(1)}, v_1^{(2)}]$. However, $v_1^{(1)} = \gamma_1 - \delta_1$ which is greater than zero. Therefore, $0 \notin [v_1^{(1)}, v_1^{(2)}]$ and this user is not covered. Thus, a person with optimal product equal to zero is not covered unless $\gamma_1 \leq \delta_1$. We can similarly show that γ_G must be greater than or equal to $1 - \epsilon_1$.

After setting $\gamma_1 \leq \delta_1$ and $\gamma_G \geq 1 - \epsilon_1$, by assumption there are less than or equal to $\left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil - 2$ offerings in Γ in the open interval (γ_1, γ_G). Our goal is to show that there must exist a pair of consecutive offerings with distance between them greater than w_1. If two offerings consecutive offerings have a distance greater than w_1 between them, there exists an x' such that a user with $x^* = x'$ is not covered by aby offering.

To minimize the maximum distance between two consecutive offerings on the domain (γ_1, γ_G) using less than or equal to $\left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil - 2$ offerings, the optimal solution is to set $\gamma_1 = \delta_1$, $\gamma_G = 1-\epsilon_1$, and equally space out exactly $\left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil - 2$ offerings on the domain (γ_1, γ_G). Following this solution, the distance between two consecutive offerings becomes

$$\frac{1-\epsilon_1-\delta_1}{\left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil - 1}.$$

Our goal is to show that this distance is greater than w_1. By definition of the ceiling function,

$$\frac{1-\epsilon_1-\delta_1}{w_1} > \left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil - 1.$$

After rearranging terms, this becomes $\frac{1-\epsilon_1-\delta_1}{\left\lceil \frac{1-\epsilon_1-\delta_1}{w_1} \right\rceil - 1} > w_1$, proving the result for the one-dimensional case.

We now consider the higher dimensional case. We must show that there exists a set of offerings $\{\gamma_k\}$ with cardinality $\prod_{k=1}^{d} 1 + \left\lceil \frac{1-\epsilon_k-\delta_k}{w_k} \right\rceil$ such that $\mathcal{D} = \cup_k \mathcal{C}(\gamma_k)$. As Even Segmentation exactly covers $\mathcal{D}$ by construction and uses exactly $\prod_{k=1}^{d} 1 + \left\lceil \frac{1-\epsilon_k-\delta_k}{w_k} \right\rceil$ offerings, the proof is complete. $\square$

Proof of Corollary 2 As Even Segmentation results in $\left\lceil \frac{1}{w_1} \right\rceil$ offerings when $d=1$, we know from Theorem 1 that there are no solutions with fewer offerings than Even Segmentation in the 1-D case. Additionally, by construction $\mathcal{D} = \cup_{\gamma \in \Gamma^{ES}} \mathcal{C}(\gamma)$. Therefore, Γ^{ES} is both feasible and optimal for I-ODP. $\square$

EC.1.4. Proof of Corollary 3

First, we will show that $\mathcal{D}$ can be covered in $\min\left\{2\left\lceil \frac{1}{2w+2l} \right\rceil, 2\left\lceil \frac{1-w}{2w+2l} \right\rceil + 1\right\} \cdot \left(\left\lceil \frac{1-h/2}{h} \right\rceil + 1\right)$ elongated hexagons with parameters h, l, and w as shown in Figure 2. Then, we will describe how to translate this covering into a valid set of offerings that lie within $\mathcal{D}$.

Consider placing the first hexagon so that it has top and bottom edges aligned with the axis, and so that the bottom left vertex is at the point $(0,0)$ (see the red hexagon in Figure EC.2).

First, we will determine how many columns of hexagons are needed. To do this, consider covering the line segment going from $(0,0)$ to $(1,0)$. The first hexagon (in red in Figure EC.2) covers a segment of length w. The next hexagon covers a segment of length $w+2l$, and so on. Thus, the question is: how many segments of alternating length w and $w+2l$ are needed to cover a length-1 line segment? With n segments, we cover length $\lfloor\frac{n+1}{2}\rfloor w+\lfloor\frac{n}{2}\rfloor(2w+l)$. We want to find the smallest n such that $\lfloor\frac{n+1}{2}\rfloor w+\lfloor\frac{n}{2}\rfloor(w+2l)\geq 1$. If n is constrained to be even, so that $n=2m$, then we have that the total length covered is $m(w+w+2l)$. Therefore, the smallest even n is $n=2\lceil\frac{1}{2w+2l}\rceil$. If m is constrained to be odd, so that $n=2m+1$, then the total length covered is $m(2w+2l)+w$. Therefore, the smallest odd n is $n=2\lceil\frac{1-w}{2w+2l}\rceil+1$. Putting this together, the optimal n is the smallest of the two options, and thus $n=\min\left\{2\lceil\frac{1}{2w+2l}\rceil,\, 2\lceil\frac{1-w}{2w+2l}\rceil+1\right\}$. This gives us the number of columns of hexagons needed.

It remains to find the number of rows of hexagons needed to cover $[0,1]^2$. This is accomplished by considering the number of stacked hexagons needed to reach a height of 1. The height of each hexagon is h. There are two "types" of columns that alternate: those whose bottoms are flush with the line segment connecting $(0,0)$ to $(1,0)$ (columns 1 and 3 in Figure EC.2) and those that are not flush, such that the line segment connecting $(0,0)$ to $(0,1)$ cuts through the middle of the bottom hexagon, such as column 2 of Figure EC.2. The second "type" of column will always require more hexagons to reach the line $y=1$. Therefore, we focus on the second column type. The base hexagon reaches a height of $h/2$. The number of hexagons needed in this type of column is therefore $\lceil\frac{1-h/2}{h}\rceil+1$. Multiplying the number of rows by the number of columns gives:

$$\min\left\{2\left\lceil\frac{1}{2w+2l}\right\rceil,\, 2\left\lceil\frac{1-w}{2w+2l}\right\rceil+1\right\}\cdot\left(\left\lceil\frac{1-h/2}{h}\right\rceil+1\right).$$

Now, it remains to show that there exists a set of $\min\left\{2\left\lceil\frac{1}{2w+2l}\right\rceil,\, 2\left\lceil\frac{1-w}{2w+2l}\right\rceil+1\right\}\cdot\left(\left\lceil\frac{1-h/2}{h}\right\rceil+1\right)$ offerings within $[0,1]^2$ such that the domain is covered. A set of offerings $\{\boldsymbol{\gamma}_i\}$ is feasible for I-ODP if and only if $[0,1]^2=\cup_k\,\mathcal{C}(\boldsymbol{\gamma}_i)$ and $\boldsymbol{\gamma}_k\in\mathcal{D}$ for all k.

Let $\{\mathbf{c}_i\}$ be the set of the centers of the hexagon produced using the tiling method described above, and let $\{H(\mathbf{c}_i)\}$ be the set of hexagons. Since $\{H(\mathbf{c}_i)\}$ covers $[0,1]^2$, selecting $\{\mathbf{c}_i\}$ as the offerings is a natural candidate to solve I-ODP. However, it may be the case some of the centers lie outside the domain $[0,1]^2$, making them infeasible offerings (as is the case for four of the centers of the gray hexagons in Figure EC.2 (left)—the top left hexagon, and the three along the right-hand

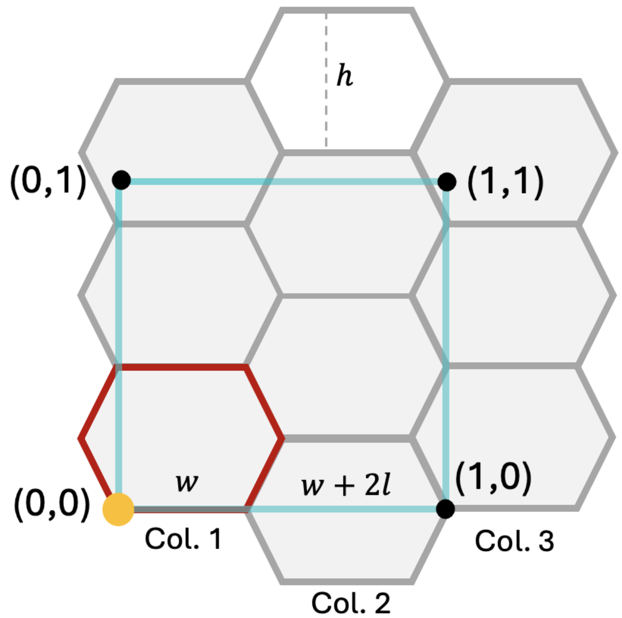


Figure EC.2 Hexagonal tiling example. All shaded hexagons intersect the domain $[0,1]^2$. The first hexagon (outlined in red) is placed so that its bottom left-hand vertex is at $(0,0)$, illustrated by the yellow dot, and so that its base is flush with the line $y=0$.

side). To solve this problem, we will map each center onto a point in $[0,1]^2$ such that the domain remains fully covered.

Consider the following set of offerings, defined for each $i \in 1,...,\min\left\{2\left\lceil\frac{1}{2w+2l}\right\rceil, 2\left\lceil\frac{1-w}{2w+2l}\right\rceil+1\right\}\cdot\left(\left\lceil\frac{1-h/2}{h}\right\rceil+1\right)$ as:

$$\gamma_i = \begin{cases} \mathbf{c}_i \text{ if } \mathbf{c}_i \in [0,1]^2 \\ [\min\{1, c_{i,1}\}, \min\{1, c_{i,2}\}] \text{ otherwise.} \end{cases} \tag{EC.3}$$

For $\mathbf{c}_1$ in the domain, γ_i is equivalent to $\mathbf{c}_i$. However, if $\mathbf{c}_i$ lies outside the domain, then γ_i is the closest point to $\mathbf{c}_i$ in the domain. Because the domain is the square $[0,1]^2$, the $x-$coordinate of this closest point is $\min\{1, c_{i,1}\}$ and the $y-$coordinate is $\min\{1, c_{i,2}\}$. We note that because of the specific tiling strategy used, it will never be the case that the x or y coordinate of a hexagon's center is below zero. Therefore, we only need to worry about the coordinates exceeding 1.

To show that $\cup_i \mathcal{C}(\gamma_i)$ covers $[0,1]^2$, we will show that $(H(\mathbf{c}_i)\cap[0,1]^2) \subseteq (H(\gamma_i)\cap[0,1]^2)$. In other words, each offering γ_i covers as much of the domain $[0,1]^2$ as the corresponding offering $\mathbf{c}_i$.

Without loss of generality we will consider a point $\mathbf{c}_i$ such that $c_{i,1} > 1$ but $c_{i,2} \in [0,1]$ and therefore $\boldsymbol{\gamma}_i = [1, c_{i,2}]$. $H(\boldsymbol{\gamma}_i)$ is thus a horizontally-shifted version of $H(\mathbf{c}_i)$ (see Figure EC.3). Because of the symmetry of the hexagon and the fact that it is shifted only until its vertical hemisphere falls on the line $x = 1$, it is clear that $(H(\mathbf{c}_i) \cap [0,1]^2) \subseteq (H(\boldsymbol{\gamma}_i) \cap [0,1]^2)$ (see Figure EC.3). The same logic holds when $\boldsymbol{\gamma}_i$ and $\mathbf{c}_i$ are vertically offset, and when they are diagonally offset (requiring both a vertical and horizontal shift). Thus, the shifting procedure always ensures that $(H(\mathbf{c}_i) \cap [0,1]^2) \subseteq (H(\boldsymbol{\gamma}_i) \cap [0,1]^2)$. Therefore, if $\mathcal{D} = \cup_i (H(\mathbf{c}_i) \cap [0,1]^2)$, then $\mathcal{D} = \cup_i (H(\boldsymbol{\gamma}_i) \cap [0,1]^2)$. □

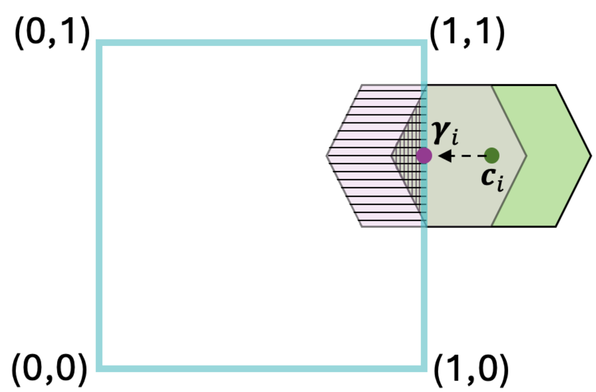


Figure EC.3 Example of vertically shifting a hexagon's center so that it lies within the boundary. The region shaded with vertical lines shows $(H(\mathbf{c}_i) \cap [0,1]^2)$ and the region shaded with horizontal lines shows $(H(\boldsymbol{\gamma}_i) \cap [0,1]^2)$. As is clear from this image, $(H(\mathbf{c}_i) \cap [0,1]^2) \subseteq (H(\boldsymbol{\gamma}_i) \cap [0,1]^2)$.

EC.1.5. Proof of Theorem 2 and Corollaries 4-5

First, we prove the following lemma.

Lemma EC.1. *$\bar{y}(\psi(x^*))$ is concave increasing in $\psi(x^*)$.*

Proof of Lemma EC.1. We can re-write $\bar{y}(\psi)$ as:

$$\bar{y}(\psi) = \psi \int_{\delta=-\Delta(\psi)}^{1/\psi - \Delta(\psi)} \tilde{y}(\delta) d\delta$$

where $\tilde{y}(\delta) \triangleq y(x^* - x | x^*)$ for all x such that $x^* - x = \delta$, and $\Delta(\psi)$ is such that $\tilde{y}(-\Delta(\psi)) = \tilde{y}(1/\psi - \Delta(\psi))$. By the quasi-convexity of $y(\cdot)$ (Assumption 2), $\tilde{y}(\delta)$ attains its maximum value when $\delta = 0$ and is strictly increasing for $\delta < 0$ and decreasing for $\delta > 0$.

Taking the second derivative with respect to $\psi(x^*)$, we have

$$\bar{y}''(\psi(x^*)) = \left(\psi\Delta'(\psi)^2 + \frac{2\Delta'(\psi)}{\psi} + \frac{1}{\psi^3}\right)\tilde{y}'\left(\frac{1}{\psi} - \Delta(\psi)\right) - \psi\Delta'(\psi)^2\tilde{y}'(-\Delta(\psi)) + \left(\tilde{y}(-\Delta(\psi)) - \tilde{y}\left(\frac{1}{\psi} - \Delta(\psi)\right)\right)(\psi\Delta''(\psi) + 2\Delta'(\psi)).$$

By definition of $\Delta(\psi)$, we know that $\left(\tilde{y}(-\Delta(\psi)) - \tilde{y}\left(\frac{1}{\psi} - \Delta(\psi)\right)\right) = 0$ and therefore the last term is equal to zero. The second term is negative because $\tilde{y}'(-\Delta) \geq 0$. To evaluate the sign of the first term, notice that $\psi\Delta'(\psi)^2 + \frac{2\Delta'(\psi)}{\psi} + \frac{1}{\psi^3} = \frac{\left(\psi^2\Delta'(\psi)+1\right)^2}{\psi^3}$ which is positive. Therefore, because $\tilde{y}'\left(\frac{1}{\psi} - \Delta(\psi)\right) \geq 0$, the first term is negative. Thus, we have that $\bar{y}''(\psi(x^*)) \leq 0$. A similar analysis of the first derivative shows that $\bar{y}'(\psi(x^*)) \geq 0$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. To find $\psi^*(x)$, consider the modified objective function

$$P\int_{x\in[0,1]} f(x)\cdot\bar{y}(\psi) - c\int_{x\in[0,1]}\psi(x). \tag{EC.4}$$

Because there are no constraints on $\psi(x)$, maximizing EC.4 is equivalent to maximizing $Pf(x)\cdot\bar{y}(\psi) - c\psi(x)$ for every x. First, we will show that this expression is concave in ψ. The second derivative is given by $P\bar{y}''(\psi)f(x)$ which is negative since $\bar{y}''(\psi)$ is negative (Lemma EC.1).

Thus, we can consider the first-order optimality condition, given by:

$$\bar{y}'(\psi) = \frac{c}{Pf(x^*)}.$$

Therefore, if the indicator function in the objective function of S-ODP were not present, the optimal function $\psi(x)$ would satisfy $\psi(x^*) = \bar{y}'^{-1}\left(\frac{c}{Pf(x^*)}\right) \triangleq \psi^\dagger(x^*)$.

The indicator function effectively places the following constraint on $\psi^*(x)$: either $\psi^*(x) = 0$ or $\psi^*(x) \geq \psi^\tau$. Thus, when $\psi^\dagger(x^*) \geq \psi^\tau$, $\psi^\dagger(x^*)$ is optimal. However, when $\psi^\dagger(x^*) < \psi^\tau$, $\psi^\dagger(x^*)$ cannot be optimal. For x^* such that $\psi^\dagger(x^*) < \psi^\tau$, $\psi(x^*) \in \{0, \psi^\tau\}$ because of the concavity of $Pf(x)\cdot\bar{y}(\psi) - c\psi(x)$ with respect to ψ. Therefore, for x^* such that $\psi^\dagger(x^*) < \psi^\tau$, we must determine whether setting $\psi(x^*) = \psi^\tau$ yields higher utility than $\psi(x^*) = 0$.

We will again use the separability of EC.4 with respect to x^*, and determine the values of x^* for which $\phi(x^*) = \psi^\tau$ yields higher utility than $\phi(x^*) = 0$. When $\psi(x^*) = 0$, the infinitesimal utility at x^* is zero. Therefore, it is worth covering all individuals that yield a positive utility when covered with an intensity ψ^τ. Thus, we are interested in determining which x^* satisfy the following condition: $P\bar{y}(\psi^\tau)f(x) - c\psi^\tau \geq 0$. This simplifies to $f(x^*) \geq \frac{c\psi^\tau}{P\bar{y}(\psi^\tau)}$. □

Proof of Corollary 4. This follows immediately from Theorem 2. □

Proof of Corollary 5 The fact that the coverage of $\psi^*(x)$ increases in P follows from Theorem 2.

To see that the coverage of Γ^* is also non-decreasing in P, notice that increasing P strictly increases the number of users that are covered by any offering and therefore strictly increases the marginal benefit of every offering without impact the cost. Let $z(G|P)$ be the profit obtained if the optimal solution to P-ODP is required to use G offerings and has population size P. We will show that $z(G|P_2) - z(G|P_1) \leq z(G+1|P_2) - z(G+1|P_1)$ for $P_2 > P_1$. In other words, the increase in profit due to a larger population size increases with the number of offerings. This implies that if $G^*(P)$ is the optimal number of offerings for a particular P, then $G^*(P') \geq G^*(P)$ for all $P' > P$, thus proving the result.

For a given G, let $\max_{\substack{\Gamma=\{\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_G\}\in\mathcal{D}^G \\ \{z_g(\cdot)\}_{g\in[\![G]\!]}}} \int_{\mathbf{x}^*\in\mathcal{D}} f(\mathbf{x}^*)\cdot o(\mathbf{x}^*)\cdot \mathbb{1}\{o(\mathbf{x}^*)\geq\tau\} \triangleq u^*(G)$ be the optimal profit using G offerings, and let $\int_{\mathbf{x}^*\in\mathcal{D}} f(\mathbf{x}^*)\cdot o(\mathbf{x}^*)\cdot \mathbb{1}\{o(\mathbf{x}^*)\geq\tau\} \triangleq u(\{\gamma_1,\ldots,\gamma_G\})$ be the profit using a particular offering set of size G. Thus, we can write $z(G|P) = Pu^*(G) - cG$. The difference $z(G+1|P_2) - z(G+1|P_1) - (z(G|P_2) - z(G|P_1))$ is therefore positive if $(P_2 - P_1)u^*(G+1) - (P_2 - P_1)u^*(G)$ which is positive as long as $u^*(G+1) \geq u^*(G)$. To see that $u^*(G+1) \geq u^*(G)$, notice that optimal set of offerings for G, $\gamma_1^*(G),\ldots,\gamma_G^*(G)$ can also be selected for $G+1$ offerings by duplicating one offering, e.g. by using the offering set $\{\gamma_1^*(G),\ldots.,\gamma^*(G),\gamma^*(G)\}$ which has cardinality $G+1$. Thus, we have that $u^*(G+1) \geq u(\{\gamma_1^*(G),\ldots.,\gamma^*(G),\gamma^*(G)\}) = u^*(G)$.

Lastly, we must show that when the number of offerings in the optimal solution to P-ODP increases, the coverage also eventually increases. Suppose for contradiction that the covered region *does not* increase, and consider increasing G. Let $Y(\gamma|\Gamma)$ be the profit contributed by offering γ when the set of offerings is Γ. This is given by

$$Y(\gamma|\Gamma) = \int_{x \in S(\gamma|\Gamma)} Pf(x)y(\gamma|x)dx,$$

where $S(\gamma|\Gamma) = \{x|y(\gamma|x) \geq \max_{\gamma \in \Gamma} y(\gamma|x)\}$ is the set of users assigned to γ. We will argue that there exists a γ for which $Y(\gamma|\Gamma) < t$ for Γ with large enough cardinality, where t is an arbitrary threshold.

First, notice that $Y(\gamma|\Gamma)$ is upper bounded by $y^{\max}P|S(\gamma|\Gamma)|$, where $y^{\max}$ is the maximum possible response value. Let L be the length of the initial covered region for some offering set Γ. For a larger offering set Γ' it holds by assumption that $L \leq \sum_\gamma |S(\gamma|\Gamma')|$. Consider $\min_{\gamma \in \Gamma'} Y(\gamma|\Gamma') = \min_{\gamma \in \Gamma'} y^{\max}P|S(\gamma|\Gamma')| \leq y^{\max}PL/|\Gamma'|$. This upper bound becomes arbitrarily small as Γ' increases in cardinality. Therefore, for Γ' large enough, there always exists a γ such that $Y(\gamma|\Gamma') < t$ for any t.

Now, let Γ^* be the optimal solution to P-ODP. It must hold that $Y(\gamma|\Gamma^*) \geq c$ for all $\gamma \in \Gamma^*$. In other words, each offering must be individually profitably. Otherwise, that offering could be removed and the objective function of P-ODP would increase.

By the above logic, it is impossible for the cardinality of Γ^* to increase and for the covering region to remain constant or decrease. If this occurs, then at least one offering would eventually become unprofitable (setting $t = c$), contradicting Γ^* being an optimal solution. Thus, if the cardinality of Γ^* increases, the covered region must also eventually increase. This completes the proof. □

EC.1.6. Proof of Lemmas 1-4

Proof of Lemma 1. This follows immediately from Theorem 2. □

Proof of Lemma 2 Computing the variance is straightforward. Now, notice that since $s \leq 1$, the term $s + 2l^2(1-s)$ is a weighted sum of one and $2l^2$. Since $l \leq 1/2$, this term is maximized when $s = 1$. When $s = 1$, it is necessarily the case that $l = 0$ and the density function becomes a uniform distribution on $[0,1]$. □

Proof of Lemma 3. Let $\kappa^{(1)} \triangleq \{x \in f^{-1}(r) : x \leq \max_x f(x)\}$ and $\kappa^{(2)} \triangleq \{x \in f^{-1}(r) : x \geq \max_x f(x)\}$. Thus, $\mathcal{R}_1 \cup \mathcal{R}_2 = [\kappa^{(2)}, \kappa^{(1)}]$. We can write $\kappa^{(1)}$ and $\kappa^{(2)}$ in closed form as $\kappa^{(1)} = \frac{1}{2} - \frac{l(h+s-r)}{h}$ and $\kappa^{(2)} = \frac{1}{2} + \frac{l(h+s-r)}{h}$ if $s \leq r \leq s+h$. Now consider the percentage of characteristics covered (Definition 4). This is simply the length $\kappa^{(2)} - \kappa^{(1)} = \frac{2l(h+s-r)}{h}$. After substituting $h = \frac{1-s}{l}$, this becomes $\frac{-2l(l(r-s)+s-1)}{1-s}$. Taking the derivatives of this quantity with respect to l and s proves the result. □

Proof of Lemma 4. Using the closed form expressions of $\kappa^{(1)}$ and $\kappa^{(2)}$ (see the proof of Lemma 3) we can characterize the density of the covered population by

$$\begin{cases} \frac{c^2 l^2 - ((l-1)s+1)^2}{s-1} & \text{if } s \leq c \leq s+h \\ 1 & \text{if } c < s \\ 0 & \text{if } c > s+h. \end{cases}$$

The function $\frac{c^2 l^2 - ((l-1)s+1)^2}{s-1}$ is concave in l, and its maximum value on the domain $[0, 1/2]$ is obtained at:

$$\begin{cases} l = s\frac{1-s}{c^2 - s^2} & \text{if } c^2 \geq 2s - s^2 \\ l = 1/2 & \text{otherwise.} \end{cases}$$

□

EC.1.7. Proof of Theorem 3.

Let $\mathcal{A}(l)$ be the highest-density set with cardinality l within $[0,1]$. As stated in the theorem, we use $\tilde{F}(l)$ to denote its density.

First, we will show that if $\lambda \geq L^*$, the optimal solution to S-ODP(λ) must satisfy $\mathbb{1}\{\psi(x) \geq \psi^\tau\} = 1$ if and only if $x \in \mathcal{A}(\lambda)$. Based on the objective function of S-ODP(λ), it is straightforward to see that if the optimal solution to S-ODP(λ) covers a region with cardinality l, then the covered region is $\mathcal{A}(l)$—the highest-density length-l region. Now it remains to show that the optimal solution to S-ODP(λ) covers a region with cardinality exactly λ. We will prove this by showing that for $l \geq L^*$,

$\int_{x\in\mathcal{A}(l)} Pf(x)\cdot\bar{y}(\psi^*(x)) - c\cdot\psi^*(x)$ is strictly decreasing in l. Therefore, when $\lambda \geq L^*$, it is optimal to cover a region with cardinality exactly λ.

Consider the following string of equalities:

$$\begin{aligned}
&\frac{\partial}{\partial l}\int_{x\in\mathcal{A}(l)} Pf(x)\cdot\bar{y}(\psi^*(x)) - c\cdot\psi^*(x)\\
=&\frac{\partial}{\partial l}\int_{x\in\mathcal{A}(L^*)} Pf(x)\cdot\bar{y}(\psi^*(x)) - c\cdot\psi^*(x) + \frac{\partial}{\partial l}\int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)} Pf(x)\cdot\bar{y}(\psi^*(x)) - c\cdot\psi^*(x)\\
=&\frac{\partial}{\partial l}\int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)} \big(Pf(x)\,\bar{y}(\psi^*(x)) - c\,\psi^*(x)\big)\,dx\\
=&\frac{\partial}{\partial l}\int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)} \big(Pf(x)\,\bar{y}(\psi^\tau) - c\,\psi^\tau\big)\,dx.
\end{aligned}$$

Let $t(s)$ be the density threshold that generates the "top–s" densest region, i.e.

$$t(s) := \max\big\{t : |\{x : f(x) > t\}| \geq s\big\}, \qquad s\in[L^*,1].$$

Consider the integrals $\int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)} dx$ and $\int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)} f(x)dx$. Instead of integrating over x, consider integrating over s—level set lengths. This gives us:

$$\int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)} f(x)dx = \int_{s=L^*}^{l} t(s)ds \qquad \text{and} \qquad \int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)} dx = \int_{s=L^*}^{l} ds.$$

We can therefore re-write $\frac{\partial}{\partial l}\int_{x\in\mathcal{A}(l)\setminus\mathcal{A}(L^*)}\big(Pf(x)\,\bar{y}(\psi^\tau) - c\,\psi^\tau\big)\,dx$ as

$$\frac{\partial}{\partial l}\int_{s=L^*}^{l}\big(Pt(s)\,\bar{y}(\psi^\tau) - c\,\psi^\tau\big)\,ds = Pt(l)y(\psi^\tau) - c\,\psi^\tau.$$

Now, notice that $Pf(x)y(\psi^\tau) - c\,\psi^\tau < 0$ for all $x \notin \mathcal{A}(L^*)$. Otherwise, if there was some x for which this expression were non-negative, that value of x would be covered under the unconstrained optimal solution. Finally, because $t(l) \leq f(x)$ for all $x\in\mathcal{A}(L^*)$, it follows that $Pt(l)y(\psi^\tau) - c\,\psi^\tau \leq 0$. Consequently, when $\lambda \geq L^*$ the optimal solution to S-ODP(λ) covers *exactly* λ characteristics, i.e. it selects $\mathcal{A}(\lambda)$.

Having established this, we can now return to our goal of upper bounding $\frac{z^*-z^*(\lambda)}{z^*}$. The difference $z^* - z^*(\lambda)$ is exactly the (negative of) utility minus the cost of covering an additional $\lambda - L^*$ characteristics. The utility gained is exactly $P(\tilde{F}(\lambda) - \tilde{F}(L^*))\bar{y}(\psi^\tau)$ and the cost is $c(\lambda - L^*)\psi^\tau$. Thus, the numerator $z^* - z^*(\lambda)$ is given by $-\Big(P(\tilde{F}(\lambda) - \tilde{F}(L^*))\bar{y}(\psi^\tau) - c(\lambda - L^*)\psi^\tau\Big)$.

Now consider the denominator, z^*, given by $P\int_{\mathcal{R}_1} f(x)\bar{y}(\psi^*(x)) + P\int_{\mathcal{R}_2} f(x)\bar{y}(\psi^\tau) - c\int_{\mathcal{R}_1}\psi^*(x) - c\int_{x\in\mathcal{R}_2}\psi^\tau$. A lower bound on z^* can be found by replacing $\psi^*(x)$ with ψ^τ, yielding $z^* \geq P\tilde{F}(L^*)\bar{y}(\psi^\tau) - cL^*\psi^\tau$. The ratio $\frac{z^*-z^*(\lambda)}{z^*}$ is therefore upper bounded by

$$-\frac{P(\tilde{F}(\lambda)-\tilde{F}(L^*))\bar{y}(\psi^\tau) - c(\lambda - L^*)\psi^\tau}{P\tilde{F}(L^*)\bar{y}(\psi^\tau) - cL^*\psi^\tau}$$

which is equivalent to the bound provided in the statement of the theorem. □

EC.1.8. Proof of Theorem 4.

To lower bound the loss in planner utility under S-ODP(λ,Ψ) compared to S-ODP, we will consider a specific feasible solution of S-ODP(λ,Ψ). This is accomplished by reducing the offering intensity over $\mathcal{R}_1$ in a specific fashion so that we can add offerings to an additional $\lambda - L^*$ characteristics while ensuring that the total volume of offering intensity remains unchanged.

Suppose that, to cover $\lambda - L^*$ additional characteristics, the planner reduces the offering intensity within $\mathcal{R}_1$ to $(\psi^*(x)-\psi^\tau)(1-\rho)+\psi^\tau$ where ρ is between zero and one. By reducing the intensity as such, the planner reduces the total intensity volume over $\mathcal{R}_1$ by $\int_{x\in\mathcal{R}_1}\rho(\psi^*(x)-\psi^\tau)dx$. To cover an additional $\lambda - L^*$ characteristics with intensity ψ^τ, a volume of $(\lambda - L^*)\psi^\tau$ is needed.

Therefore, the minimal value of ρ such that the intensity "saved" over $\mathcal{R}_1$ is enough to cover $\lambda - L^*$ additional characteristics is given by:

$$\rho = \frac{(\lambda - L^*)\psi^\tau}{\int_{x\in\mathcal{R}_1}(\psi^*(x)-\psi^\tau)dx}.$$

If this quantity is larger than 1 it is not feasible to cover λ characteristics without increasing the total intensity volume. This is the condition given in the theorem.

For $x \in \mathcal{R}_1$, the total change in the service provider's profit is given by $\int_{x\in\mathcal{R}_1} Pf(x)\left(\bar{y}\left(\psi^\tau + (1-\rho)(\psi^*(x)-\psi^\tau)\right) - \bar{y}(\psi^*(x))\right)$. We will focus on lower bounding this quantity.

Let $\psi' = \rho\psi^\tau + (1-\rho)\psi^*(x)$. Because $\bar{y}(\cdot)$ is concave increasing (Lemma EC.1), by Jensen's inequality we have that $\bar{y}(\psi') \geq \rho\bar{y}(\psi^\tau) + (1-\rho)\bar{y}(\psi^*(x))$. Rearranging terms, this implies

$$\bar{y}(\psi^*(x)) - \bar{y}(\psi') \leq \rho\left(\bar{y}(\psi^*(x)) - \bar{y}(\psi^\tau)\right).$$

We will upper bound $\bar{y}(\psi^*(x))$ by $\bar{y}(\psi^{\max})$, where $\psi^{\max} = \max_x \psi^*(x)$. We then have the inequality

$$\bar{y}(\psi^*(x)) - \bar{y}(\psi') \leq \rho\left(\bar{y}(\psi^{\max}) - \bar{y}(\psi^\tau)\right).$$

Recall from the proof of Theorem 3 that the utility gained by covering an additional $\lambda - L^*$ users with an intensity of ψ^τ is given by $P\bar{y}(\psi^\tau)(\tilde{F}(\lambda) - \tilde{F}(L^*))$. Thus, overall we can write

$$\begin{aligned}
&z^* - z^*(\lambda, \Psi) \\
&\quad = \int_{x\in\mathcal{R}_1} Pf(x)\left(\bar{y}(\psi^*(x)) - \bar{y}(\psi')\right) - P\bar{y}(\psi^\tau)(\tilde{F}(\lambda) - \tilde{F}(L^*)) \\
&\quad \leq \rho\left(\bar{y}(\psi^{\max}) - \bar{y}(\psi^\tau)\right)\int_{x\in\mathcal{R}_1} Pf(x)dx - P\bar{y}(\psi^\tau)(\tilde{F}(\lambda) - \tilde{F}(L^*)) \\
&\quad \leq \rho\left(\bar{y}(\psi^{\max}) - \bar{y}(\psi^\tau)\right)P\tilde{F}(L^*) - P\bar{y}(\psi^\tau)(\tilde{F}(\lambda) - \tilde{F}(L^*)) \\
&\quad = \frac{(\lambda - L^*)\psi^\tau}{\int_{x\in\mathcal{R}_1}(\psi^*(x) - \psi^\tau)dx}\left(\bar{y}(\psi^{\max}) - \bar{y}(\psi^\tau)\right)P\tilde{F}(L^*) - P\bar{y}(\psi^\tau)(\tilde{F}(\lambda) - \tilde{F}(L^*)) \\
&\quad = \frac{(\lambda - L^*)\psi^\tau}{\Psi - L^*\psi^\tau}\left(\bar{y}(\psi^{\max}) - \bar{y}(\psi^\tau)\right)P\tilde{F}(L^*) - P\bar{y}(\psi^\tau)(\tilde{F}(\lambda) - \tilde{F}(L^*))
\end{aligned}$$

where the final equality is due to the fact that $\Psi = \int_{\mathcal{R}_1}\psi^*(x) + |\mathcal{R}_2|\psi^\tau$.

Using the same lower bound on z^* as in the proof of Theorem 3, we obtain the following bound, equivalent to that stated in the theorem:

$$\frac{z^* - z^*(\lambda, \Psi)}{z^*} \leq \frac{\frac{(\lambda - L^*)\psi^\tau}{\Psi - L^*\psi^\tau}(\bar{y}(\psi^{\max}) - \bar{y}(\psi^\tau))\tilde{F}(L^*) - \bar{y}(\psi^\tau)(\tilde{F}(\lambda) - \tilde{F}(L^*))}{\bar{y}(\psi^\tau)\tilde{F}(L^*) - cL^*\psi^\tau/P}.$$

□

EC.2. Details on ILP Method

We present an asymptotically exact formulation of I-ODP and P-ODP and heuristics to solve these formulations. To begin, consider discretizing the domain $\mathcal{D}$ so that the new domain becomes $\mathscr{D} = \{0, s, 2s, ..., 1\}^d$ where s controls the granularity of the discretization. We now assume that users have optimal offering $\mathbf{x}^* \in \mathscr{D}$ and also that offerings must be chosen from $\mathscr{D}$.

EC.2.1. ILP for I-ODP

With a discrete domain, I-ODP can be formulated as a classic set covering problem. As $s \to 0$, this formulation is exact for I-ODP. The set covering formulation is given by

$$\min_{\mathbf{z} \in \{0,1\}^{|\mathscr{D}|}} \sum_{j \in 1,\ldots,|\mathscr{D}|} z_j \; s.t. \sum_j \mathscr{M}_{ij} z_j \geq 1 \; \forall i \in 1, \ldots, |\mathscr{D}| \tag{EC.5}$$

where $\mathscr{M} \in \{0,1\}^{|\mathscr{D}| \times |\mathscr{D}|}$ is an incidence matrix such that $\mathscr{M}_{ij} = 1$ if and only if offering $\mathbf{x}_j$ is suitable for a user with optimal offering $\mathbf{x}_i$. This occurs if and only if $\mathbf{x}_i \in \{\mathbf{x} \in \mathscr{D} : y(\mathbf{x}_j | \mathbf{x}^* = \mathbf{x}) \geq \tau\}$.

Solving Problem EC.5 presents two main challenges: (i) it is not guaranteed that a feasible solution to Problem EC.5 is also feasible for I-ODP, and (ii) the runtime required to solve Problem EC.5 may be prohibitively slow when $|\mathscr{D}|$—which grows with both the dimensionality d and the granularity of the discretization—is very large. To see why (i) is a challenge, notice that the optimal solution to Problem EC.5 covers every point in $\mathscr{D}$, but may not fully cover the continuous domain $\mathcal{D}$. We overcome this complication through a pre-processing step (Step 1, described below). Challenge (ii) is not unique to the specific problem studied here, and much of the recent literature on the location covering and set covering problem focuses on overcoming the scalability issue. While any off-the-shelf set covering problem could be used after employing Step 1, we propose two specific heuristics. First, we take advantage of the symmetry of our specific problem to reduce the domain (Step 2). Second, we employ a modified version of the heuristic proposed in Ceria et al. (1998) (Step 3). In what follows, we summarize each step. We note that while Step 1 is necessarily to guarantee feasibility of the solution for I-ODP, Steps 2-3 are only necessary when the problem is too large to solve exactly (and Steps 2-3 can be replaced with any set covering heuristic).

Step 1: Trim $\mathscr{C}_j$. Let $Y(\cdot)$ be the natural extension of $y(\cdot)$ to the domain $\mathbb{R}^d$, and let $\mathscr{D}^E$ be the extended discrete domain. This allows us to ignore issues caused by the boundary of the domain $\mathcal{D}$. Let $\mathscr{C}_j \triangleq \{\mathbf{x} \in \mathscr{D}^E : Y(\mathbf{x}_i | \mathbf{x}^* = \mathbf{x}) \geq \tau\}$. A point, $\mathbf{x}$, is *internal* to $\mathscr{C}_j$ if $\mathbf{x} + s\mathbf{h}$ is also in $\mathscr{C}_j$ for $\mathbf{h} \in \{1, -1\}^d$. The points $\mathbf{x} + s\mathbf{h}_k$ therefore represent all 2^d vertices of the hypercube with side length $2s$ and center $\mathbf{x}$. In two dimensions, we can slightly relax this definition and define an

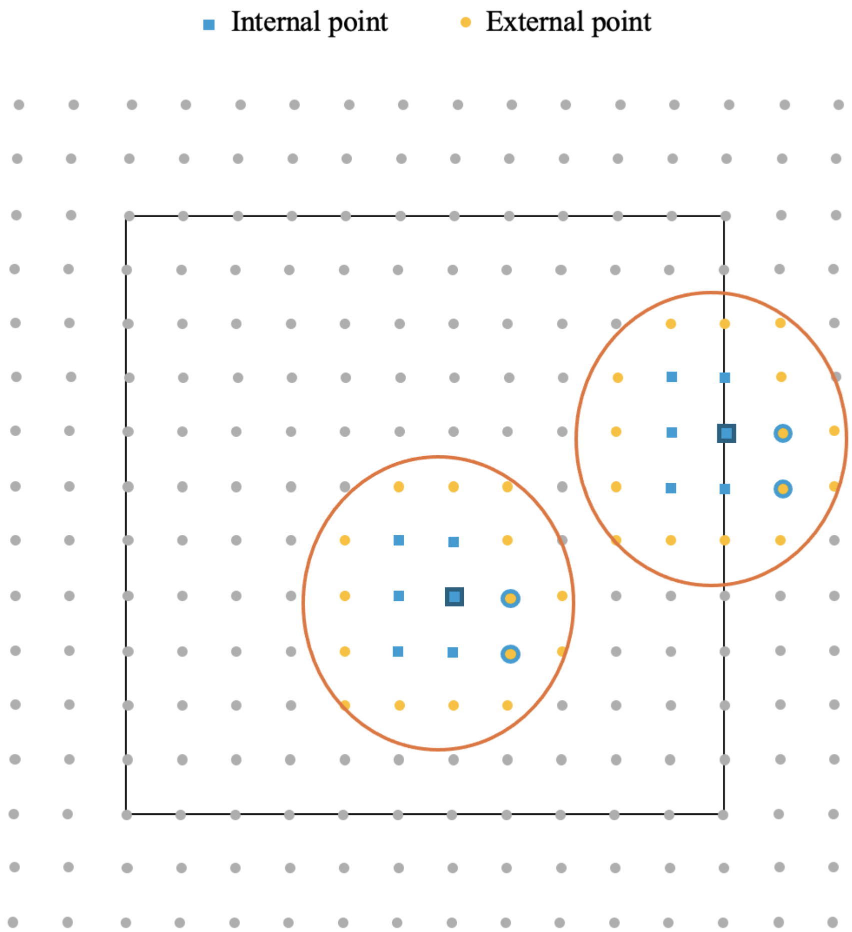


Figure EC.4 **Example of internal points in two dimensions. The orange ellipses represent the covering sets $\mathscr{C}_j$ for two offerings, where the offerings are depicted by the larger outlined blue squares. Within each covering set, the blue squares are internal to $\mathscr{C}_j$. The yellow points are within $\mathscr{C}_j$ but are not internal (and are external). This also demonstrates why it is important to consider the extended domain $\mathscr{D}^E$ when defining internal points. The yellow bounds outlined in blue are internal under the relaxed definition in 2-D, but not under the general definition.**

internal point as one in which $\mathbf{x}+s\mathbf{e}_k$ is also in $\mathscr{C}_j$, where $\mathbf{e}_k$ is the kth standard basis vector. See Figure EC.4 for an example in two dimensions.

Let $\mathscr{C}_j^I$ be the set of points that are internal to $\mathscr{C}_j$ and let $\mathscr{M}_{ij}^I = 1$ if and only if $\mathbf{x}_i \in \mathscr{C}_j^I$ and zero otherwise. Proposition EC.1 shows that by solving Problem EC.5 using $\mathscr{M}^I$ as the incidence matrix, we recover a feasible solution to I-ODP.

PROPOSITION EC.1. *The offering set* $\{\boldsymbol{\gamma}_j : z_j^* = 1\}$ *is feasible for Problem I-ODP, where* $\mathbf{z}^*$ *is the solution EC.6, given below.*

$$\min_{\mathbf{z}\in\{0,1\}^{|\mathscr{D}|}} \sum_{j\in 1,\ldots,|\mathscr{D}|} z_j \; s.t. \; \sum_j \mathscr{M}_{ij}^I z_j \geq 1 \; \forall i \in 1,\ldots,|\mathscr{D}|. \tag{EC.6}$$

Proof of Proposition EC.1 Recall the definition of $\mathcal{C}(\boldsymbol{\gamma})$ from Section 3. Let $\mathcal{C}_j \triangleq \mathcal{C}(\boldsymbol{\gamma}_j)$. Our goal is to show that all points in $\mathcal{D}$ are covered by the optimal solution to Problem EC.6. Letting $\mathbf{z}^*$ denote the optimal solution to Problem EC.6, we must show that $\mathcal{D} \subseteq \cup_{j:\, z_j^*=1} \mathcal{C}_j$. As $\mathbf{z}^*$ is optimal for Problem EC.6, we are only guaranteed *a priori* that $\mathscr{D} \subseteq \cup_{j:\, z_j^*=1} \mathscr{C}_j^I$.

First notice that $\mathcal{C}_j$ is convex because $y(\cdot)$ is quasi-convex. Let $\mathcal{S} \subseteq \mathscr{C}_j$ be a subset of points in $\mathscr{C}_j$. For any such $\mathcal{S}$, is necessarily true that $\text{Conv}(\mathcal{S}) \subseteq \mathcal{C}_j$ where $\text{Conv}(\cdot)$ denotes a convex hull (Fact 1). We will use this fact at the end of the proof.

Suppose, by contradiction, that some point $\mathbf{r} \in \mathcal{D}$ is not covered by $\cup_{i:\, z_j^*=1} \mathcal{C}_j$. Let $\mathcal{V}(H) \subseteq \mathscr{D}$ be the vertices of the hypercube H with side length s containing $\mathbf{r}$. Choose any vertex $\mathbf{v} \in \mathcal{V}$. Pick an index c such that $z_c^* = 1$ and $\mathbf{v} \in \mathscr{C}_c^I$. Such an index exists since $\mathbf{z}^*$ covers all points in $\mathscr{D}$ with sets $\mathscr{C}_j^I$.

By definition of an internal point, we have that $\mathbf{v} + s\mathbf{h} \in \mathscr{C}_c$ for all $\mathbf{h} \in \{-1,1\}^d$. Therefore, by Fact 1, $\text{Conv}\left(\cup_{k=1,\ldots,2^d} \{\mathbf{v} + s\mathbf{h}_k\}\right) \in \mathcal{C}_c$. Finally, since $r \in \text{Conv}\left(\cup_{k=1,\ldots,2^d} \{\mathbf{v} + s\mathbf{h}_k\}\right)$, $r \in \mathcal{C}_c$.

In 2-D, using the relaxed definition of an internal point we can define $\mathbf{v}$ as the vertex of the hypercube H such that the simplex oriented at vertex $\mathbf{v}$ contains $\mathbf{r}$. Let P denote that simplex. As above, let c denote an index such that $z_c^* = 1$ and $\mathbf{v} \in \mathscr{C}_c^I$. By the relaxed definition of an internal point and Fact 1, we have that $P \subseteq \mathcal{C}_c$. Since $r \in P$, it follows that $r \in \mathcal{C}_c$.

□

After applying Step 1, any heuristics for the large-scale set covering problem can be utilized to solve Problem EC.6 and will result in a feasible solution to I-ODP. We find that the following two steps work well to reduce runtime while taking advantage of the specific structure of I-ODP. Either or both can be applied.

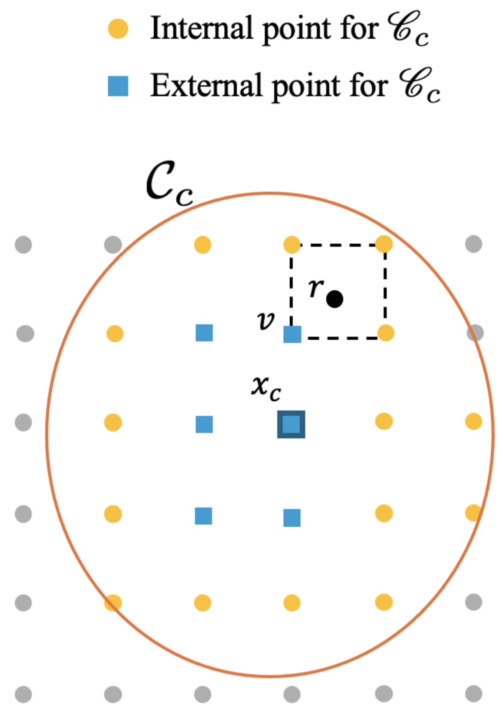


Figure EC.5 **Illustration of the notation and proof of Proposition EC.1**

Step 2. Scale down the domain. Consider a subset of $\mathcal{D}$, $\prod_{k=[\![d]\!]}[0, 1/r_k]$ where r_k is an integer. Because of the homogeneity of the utility functions and thus the covering sets $\mathcal{C}(\gamma)$, if a covering of $\prod_{k=[\![d]\!]}[0, 1/r_k]$ that uses m covering sets is found, the domain $[0, 1]^d$ can be covered with $m \prod_{k=[\![d]\!]} r_k$ sets by repeating the covering found over $\prod_{k=[\![d]\!]}[0, r_k]$. This is similar to the shifting idea of (Hochbaum and Maass 1985). We take advantage of this fact to reduce the size of the domain when needed. Choosing the right scaling factor is critical: if the domain is decreased too much, it limits the solver's ability to find good covering patterns. If the domain is too large, the runtime will be prohibitively slow. As a heuristic, we find that allowing for about 6 non-overlapping copies of covering set to fit along each dimension results in a domain that is sufficiently large. For example, if a 2-D covering set can be inscribed into a rectangle with width and height m_1 and m_2, respectively, we would choose r_k to be around $\min\{6m_k, 1\}$.

Step 3. Apply a dual-based heuristic. Ceria et al. (1998) propose a greedy heuristic that ranks possible sets (i.e., offerings) according to two values: the reduced costs that come from the linear programming relaxation of Problem EC.5 (Problem EC.7, below), and the shadow prices of the dual of the LP relaxation (Problem EC.8, below). In what follows we describe an adaption of this heuristic.

$$\min_{\mathbf{z}\geq 0} \sum_{j\in m} z_j \;\; s.t. \; \sum_{j} \mathscr{M}^{I'}_{ij} z_j \geq 1 \, \forall i \in n \tag{EC.7}$$

Let ρ_j be the reduced cost of an offering at $\mathbf{x}_j$ from Problem EC.7. Let λ_j be the dual variable associated with the jth constraint of Problem EC.8. Ceria et al. (1998) propose a heuristic that first limits the sets that are under consideration by determining a "core" set of sets. The core is found by ranking all sets based on ρ_j from lowest to highest and sequentially adding sets one by one until each row is covered by at least ten sets. Once each row is covered by at least ten sets, the resulting set of sets is the "core" and only sets in the core can be chosen in the final covering. This step typically reduces the problem size significantly.

$$\max_{\mathbf{y}\geq 0} \sum_{i\in n} y_i \;\; s.t. \; \sum_{i} \mathscr{M}^{I'}_{ij} y_i \leq 1 \, \forall j \in m \tag{EC.8}$$

Once the core is found, the heuristic proceeds by looping through the following steps until the matrix $\mathscr{M}^{I'}$ has no rows remaining. First, initialize Γ to be empty and initialize $\mathscr{M}^{I'}$ to be equal to $\mathscr{M}^{I}$ where only columns corresponding to indices of sets in the core are retained.

1. Determine ρ_j and λ_j by using the incidence matrix $\mathscr{M}^{I'}$ to solve Problems EC.7 and EC.8.
2. Rank sets according to $\lambda_j - \alpha\rho_j$ where α is a chosen weight.
3. Add the highest-ranked set to Γ and remove rows from $\mathscr{M}^{I'}$ that are covered by this set.

This procedure requires solving both Problems EC.7 and EC.8 in each iteration; however, these programs are linear and the problem size shrinks with each iteration. We find anecdotally, in agreement with Ceria et al. (1998), that the chosen value of α does not impact the results substantially.

EC.2.2. ILP for P-ODP

The discrete version of P-ODP is given by:

$$\begin{aligned}
\max_{\mathbf{z}\in\{0,1\}^{|\mathscr{D}|},\,\mathbf{y}\in\{0,1\}^{|\mathscr{D}|^2}} \quad & \sum_{i=1}^{|\mathscr{D}|} f(\mathbf{x}_i)\mathscr{O}_i - c\sum_{j=1}^{|\mathscr{D}|} z_j \\
\text{s.t.} \quad & \mathscr{O}_i = \sum_{j=1}^{|\mathscr{D}|} O_{ij} y_{ij} && \forall i = 1,\ldots,|\mathscr{D}| \\
& y_{ij} \le z_j && \forall i,j = 1,\ldots,|\mathscr{D}| \\
& \sum_{j=1}^{|\mathscr{D}|} y_{ij} = 1 && \forall i = 1,\ldots,|\mathscr{D}|
\end{aligned} \tag{EC.9}$$

where $O_{ij} = y(\mathbf{x}_j|\mathbf{x}^* = \mathbf{x}_i)\mathbb{1}\{y(\mathbf{x}_j|\mathbf{x}^* = \mathbf{x}_i) \ge \tau\}$ is the outcome of an individual with optimal offering $\mathbf{x}_i$ if assigned to offering $\mathbf{x}_j$. As in Problem EC.5, z_j determines whether offering j is selected. We now have additional variables y_{ij} that indicate whether individual i is assigned to offering j. Such an assignment is only possible when $z_j = 1$, and each individual can only be assigned to one offering.

An inclusivity constraint on the percentage of population covered can be added to Problem EC.9 by including additional binary variables $\mathbf{u} \in \{0,1\}^{|\mathscr{D}|}$ along with the following constraints:

$$u_i \ge \sum_j \mathscr{M}^I_{ij} z_j / M \; \forall\, i \in 1,...,|\mathscr{D}| \tag{EC.10}$$

$$u_i \le \sum_j \mathscr{M}^I_{ij} z_j \; \forall\, i \in 1,...,|\mathscr{D}| \tag{EC.11}$$

$$\frac{1}{|\mathscr{D}|}\sum_i u_i \ge \lambda \tag{EC.12}$$

where $\mathscr{M}^I$ is trimmed incidence matrix described in Section EC.2.1 and M is a large constant. Here, the variable u_i indicates whether individual i is covered, and the last constraint above ensures that at least λ-portion of the population is covered.

EC.3. Heterogeneous Response Functions

In this section we discuss solutions to P-ODP and I-ODP in the case of heterogeneous response functions. Practically, this means that user's sensitivity to the product features can vary as a function of their optimal product. For example, perhaps individuals with more common values of $\mathbf{x}^*$ are less sensitive to offerings' features, or vice versa. For tractability, we focus on the one dimensional setting.

EC.3.1. I-ODP

Although the most general version of I-ODP is NP-hard (Theorem 1), we have shown in Section 3 that I-ODP can be solved to optimality in polynomial time in one dimension when response functions are homogeneous and unimodal. In this section, we show that the heterogeneous-width and unimodal one-dimensional setting can also be solved to optimality in polynomial time. Thus, as long as Assumption 2 holds, I-ODP has a polynomial-time optimal solution in one dimension.

For heterogeneous response functions, the Adaptive Segmentation algorithm, described below, results in an optimal solution to I-ODP (Proposition EC.2). Intuitively, Adaptive Segmentation starts from $\gamma_1 = 0$ and moves across the domain $[0,1]$, adding new offerings to Γ as needed. Because we focus on the one dimensional setting, we will drop the subscript "1" for the remainder of this section.

Let $[a(x^*), b(x^*)]$ be the line segment such that for all $\gamma \in [a(x^*), b(x^*)]$, $y(\gamma|x^*) \geq \tau$. Let $f_{b(x^*)|a(x^*)}(b)$ denote the conditional PDF of $b(x^*)$ given $a(x^*)$. Given $f_{b(x^*)|a(x^*)}(b)$ as an input, Adaptive Segmentation is formally described in the following algorithm.

Algorithm 1: Adaptive Segmentation

```
Input: f_{b(x*)|a(x*)}(·)
Output: Γ
γ ← 0 ;
while γ < 1 do
    γ ← inf_b{b : f_{b(x*)|a(x*)}(b|a ≥ γ) > 0} ;
    Γ ← Γ ∪ γ;
end
Γ ← Γ ∪ {1} ;
return Γ
```

To build intuition, an example solution produced by Adaptive Segmentation is illustrated in Figure EC.6. Proposition EC.2 shows that Adaptive Segmentation results in an optimal solution to I-ODP.

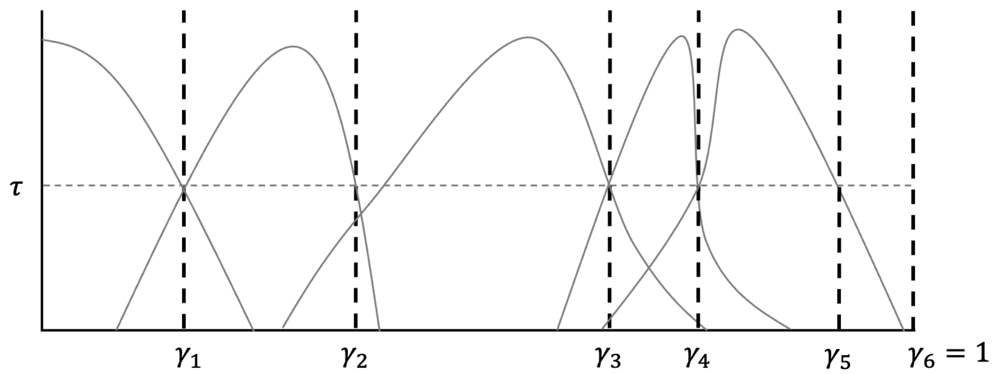


Figure EC.6 **Illustration of Adaptive Segmentation. For each product γ_g, $g = 1, ..., 5$, an example response function corresponding to $\inf_b\{b : f_{b(x^*)|a(x^*)}(b|a \geq \gamma_g) > 0\}$ is illustrated.**

PROPOSITION EC.2. ADAPTIVE SEGMENTATION *results in an optimal solution to I-ODP in one dimension.*

Proof of Proposition EC.2 Suppose that ADAPTIVE SEGMENTATION results in G offerings given by $\Gamma = \{\gamma_1, ..., \gamma_G\}$. Let $\mathcal{L}(x^*) = [a(x^*), b(x^*)]$. By construction, there is at least one x^* such that $\mathcal{L}(x^*) \subseteq [\gamma_{g-1}, \gamma_g]$ for each $g \in \{1, ..., G-1\}$ where $\gamma_0 \triangleq 0$. In particular, for such an x^*, we will necessarily have $b(x^*) = \gamma_g$.

Now suppose that there is a set of G' offerings, with $G' < G$ that results in a feasible solution for I-ODP. Let the set of offerings for this new policy be $\tilde{\gamma}_1$ through $\tilde{\gamma}_{G'}$. By the pigeonhole principle, there is at least one segment $[\gamma_{g-1}, \gamma_g]$ for $g \in \{1, ..., G-1\}$ that does not contain a single product from the set $\{\tilde{\gamma}_1, ..., \tilde{\gamma}_{G'}\}$. Let this segment be $[\gamma_{\ell-1}, \gamma_\ell]$. By the argument above, there is at least one x^* such that $\mathcal{L}(x^*) \subseteq [\gamma_{\ell-1}, \gamma_\ell]$. Let this x^* be x_ℓ. Therefore, G' does not cover x_ℓ and is not a feasible solution to I-ODP. □

Let $w^{min} = \inf_{x^* \in [0,1]}\{b(x^*) - a(x^*) : b(x^*) > 0, a(x^*) < 1\}$. The following corollary follows immediately.

COROLLARY EC.1. EVEN SEGMENTATION *using width w^{min} results in a feasible solution to I-ODP. Therefore, the number of offerings in the optimal solution to I-ODP with heterogeneous response functions is upper bounded by $\lceil \frac{1}{w^{min}} \rceil$.*

Thus, in both the homogeneous and heterogeneous response functions settings, the optimal solution to I-ODP can be readily obtained via either EVEN SEGMENTATION or ADAPTIVE SEGMENTATION, respectively.

EC.3.2. Solving S-ODP

To account for heterogeneity in S-ODP, we can write $\bar{y}(\psi)$ as $\bar{y}(\psi|x^*)$ to represent the outcome for a user with optimal offering x^* that receives an offering intensity of ψ. When $y(\cdot)$ was homogeneous, $\bar{y}(\psi|x^*)$ is invariant to x^* and thus x^* could be dropped from the notation. Additionally, for x^* to be covered we now require $\psi(x^*) \geq \psi^\tau(x^*)$ where $\psi^\tau(x^*)$ is the lowest offering intensity with which x^* is covered. Thus, the problem becomes

$$\max_{\psi(x)\geq 0} P \int_{x\in[0,1]} f(x)\cdot \bar{y}(\psi(x)|x)\cdot \mathbb{1}\{\psi(x)\geq \psi^\tau(x)\} - c\int_{x\in[0,1]} \psi(x). \tag{EC.13a}$$

Following the analysis from the homogeneous case, it is optimal to cover a certain user with optimal offering x^* if and only if $f(x^*)\cdot \bar{y}(\psi^\tau(x^*)|x^*) - c\psi^\tau(x^*) \geq 0$. Unlike before, the set of x^* for which $f(x^*) \geq \frac{c\psi^\tau(x^*)}{\bar{y}(\psi^\tau(x^*)|x^*)}$ need not be continuous, because $\psi^\tau(x^*)$ depends on x^*. As in the homogeneous case, we can determine that the optimal solution to EC.13 will still satisfy the following structure:

$$\psi^*(x)\begin{cases} \text{solves } \bar{y}'(\psi^*(x)|x) = \frac{c}{Pf(x)} \text{ if } x\in\mathcal{R}_1 \\ = \psi^\tau(x) \text{ if } x\in\mathcal{R}_2 \\ = 0 \text{ if } x\in[0,1]\setminus(\mathcal{R}_1\cup\mathcal{R}_2) \end{cases} \tag{EC.14}$$

where $\mathcal{R}_1 = \{x:\ f(x) \geq \frac{c}{\bar{y}'(\psi^\tau(x)|x)P}\}$, $\mathcal{R}_2 = \{x:\ f(x) \geq \frac{c\psi^\tau(x)}{\bar{y}(\psi^\tau(x)|x)},\ x\notin\mathcal{R}_1\}$.

EC.4. S-ODP in higher dimensions

To extend the analysis of S-ODP to higher dimensions, it it first necessary to extend our definition of $\bar{y}(\psi(x^*))$. Consider the following definition for $d\geq 2$:

$$\bar{y}(\psi) = \psi^d \int_{\mathbf{x}\in V_\psi(\boldsymbol{\gamma})} y(\gamma|\mathbf{x})d\mathbf{x}$$

for any $\boldsymbol{\gamma}\in\mathcal{D}$ such that $V_\psi(\boldsymbol{\gamma})\in\mathcal{D}$. Here,

$$V_\psi(\boldsymbol{\gamma}) \triangleq \{\mathbf{x} : y(\boldsymbol{\gamma}|\mathbf{x}) \geq y(\mathbf{g}|\mathbf{x})\ \forall\, \mathbf{g}\in\{\boldsymbol{\gamma}+\frac{1}{\psi}\mathbf{h}\}\ \text{ for }\ \mathbf{h}\in\{-1,1\}^d\}.$$

The set $V_\psi(\boldsymbol{\gamma})$ is the set of points for which $\boldsymbol{\gamma}$ is the best offering, given an offering intensity of ψ. To determine whether $\boldsymbol{\gamma}$ is the best offering for a user with optimal offering $\mathbf{x}^*$, we simply must compare $\boldsymbol{\gamma}$ to all adjacent offerings, captured by the set $\{\boldsymbol{\gamma} + \frac{1}{\psi}\mathbf{h}\}$. In one dimension, this definition is equivalent to that of Equation 6. As an offering intensity of ψ can be interpreted as having offerings evenly spaced out by $1/\psi$ in every dimension, forming a lattice.

The first-order conditions for optimizing $\psi(\mathbf{x})$ in higher dimensions follow the same procedure as in the one dimensional case. However, in order for the first-order conditions to be necessary and sufficient for optimality, we need $\bar{y}(\psi(\mathbf{x}^*))$ to be concave in $\psi(\mathbf{x}^*)$. Showing this is less straightforward in higher dimensions, and depends on the specific form of $y(\cdot)$.

EC.5. Additional Figures for Section 4.2

Figures EC.7 and EC.8 compare the upper bounds from Theorems 3 and 4 with the exact price of fairness as λ varies. Each panel shows a different population size. As seen in Figure EC.7, the bound of Theorem 3 is very close to the exact price of inclusivity. The bound of Theorem 4 is looser (Figure EC.8) as there are additional approximates made in constructing the bound (see the proof of Theorem 4). We also note that there is no analytical solution to S-ODP(λ, Ψ), so the exact bound in Figure EC.8 is based on a discrete approximation to this problem.

Figure EC.9 is analogous to Figure 8 but for different response functions. In Figure EC.9 (left), the steepness of $y(x|x^*)$ is increased. In EC.9 (right), $y(x|x^*)$ is quadratic in $|x^* - x|$ instead of linear. In both cases, the bounds are not substantively different than in the main text.

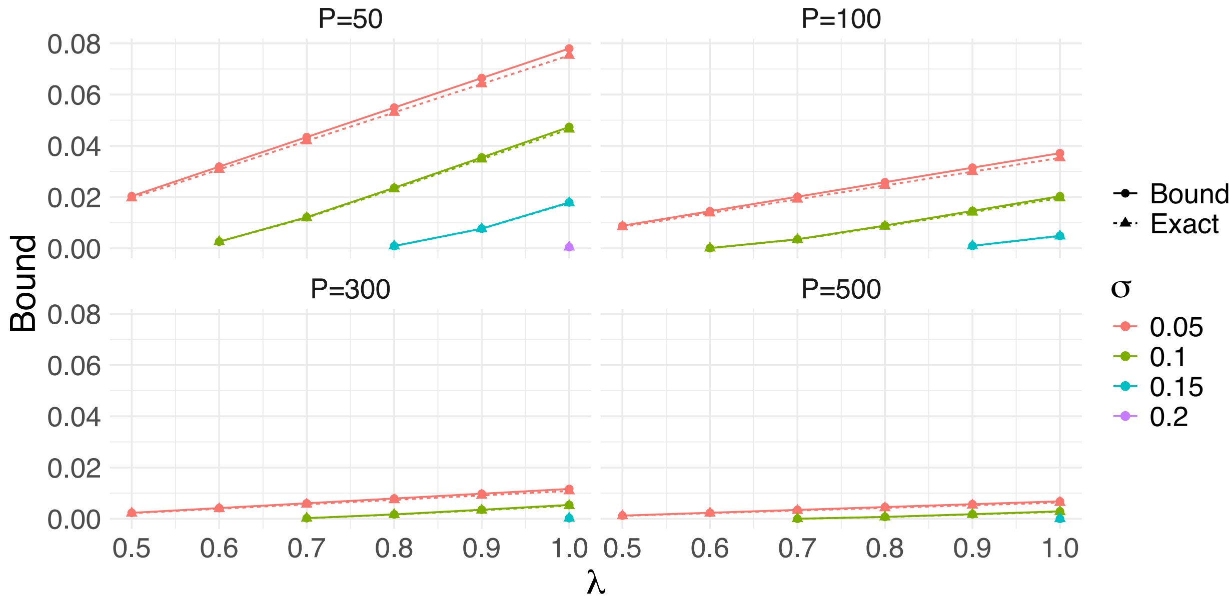


Figure EC.7 **Example of bound from Theorem 3 (solid line) and the exact price of inclusivity (dashed line) as λ varies. In this example, $f(x) \sim$ Normal$(.5, \sigma^2)$, $y(x|x^*) = 1 - 2|x^* - x|$, $c = 1$, and $\tau = 0.8$. Results are shown for instances where $\lambda \geq L^*$ and $L^* \leq 1$.**

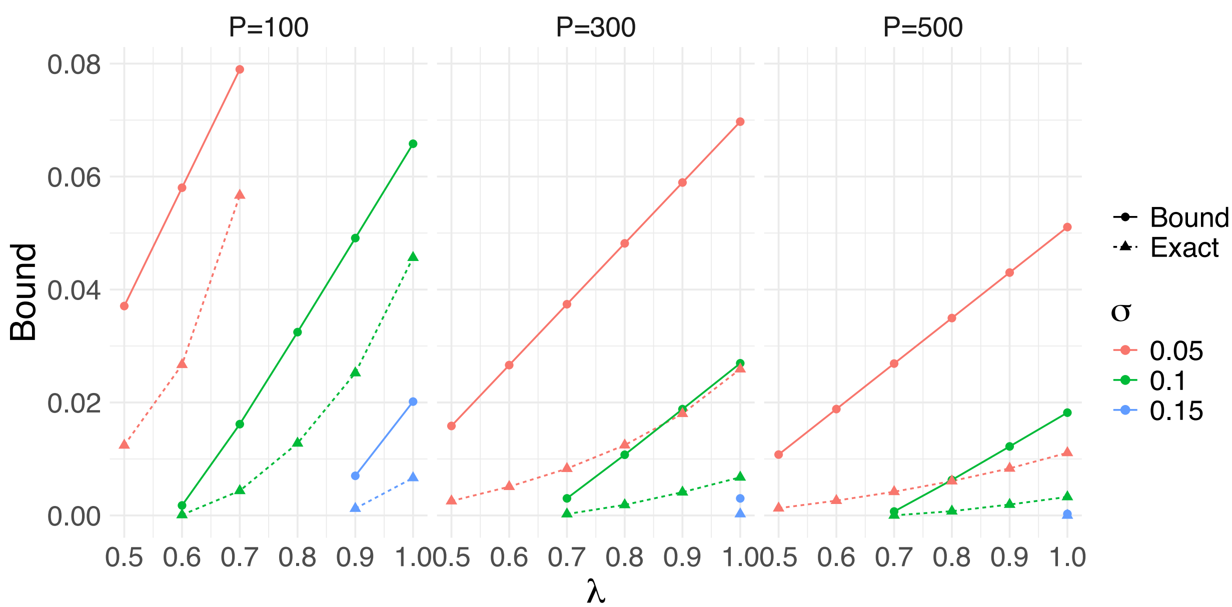


Figure EC.8 **Example of bound from Theorem 4 (solid line) and the exact price of inclusivity (dashed line) as λ varies. In this example, $f(x) \sim$ Normal$(.5, \sigma^2)$, $y(x|x^*) = 1 - 2|x^* - x|$, $c = 1$, and $\tau = 0.8$. Results are shown for instances where $\lambda \geq L^*$, $L^* \leq 1$, and λ is feasible for S-ODP(λ, Ψ).**

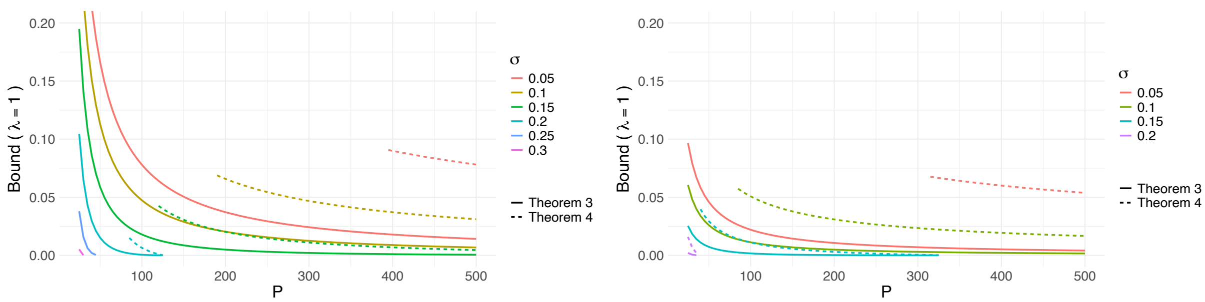


Figure EC.9 Example of bounds from Theorems 3 (solid lines) and 4 (dashed lines). In this example, $f(x) \sim$ TruncNormal$(.5, \sigma^2; [0,1])$, $\tau = 0.8$, $c = 1$, and $\lambda = 1$. On the left $y(x|x^*) = 1 - 4|x^* - x|$, and on the right $y(x|x^*) = 1 - 8|x^* - x|^2$. The bounds were calculated for $P \in [25, 500]$ and $\sigma \in [.05, .5]$. For Theorem 3, results are shown for instances where $L^* \leq 1$. For Theorem 4, results are shown for instances where $\lambda = 1$ is feasible and $L^* \leq 1$.

EC.6. Case Study Supplemental Material

This section provides more details on the foundation makeup case study. Table EC.2 provides details on the brands and foundation lines used in the case study, along with each brand's mission statement and number of shades offered.

Brand	About Us/Mission[10]	N shades
Fenty Beauty (Pro Filt'r Soft Matte Longwear Liquid Foundation)	"Rihanna was inspired to create the world of Fenty Beauty brands after years of partnering with the best of the best in the beauty industry—and still seeing a void for products that performed across all skintones + types and hair textures." https://fentybeauty.com/pages/about-the-brands	51
Haus Labs (Triclone Skin Tech Medium Coverage Foundation with Fermented Arnica)	"We are a collective of the Haus of Gaga, a collective of creatives, scientists, and innovators, who give you beauty tools for clean artistry. Kindness and inclusivity, forever." https://www.hauslabs.com/pages/about	51
ILIA Beauty (Super Serum Skin Tint SPF 40 Skincare Foundation)	"Protect + revive your skin." https://iliabeauty.com/pages/about	30
NARS (Light Reflecting Advanced Skincare Foundation)	"Image-maker. Instigator. Iconoclast." https://www.narscosmetics.com/USA/FN.html	36
Rare Beauty (Liquid Touch Weightless Foundation)	"Rare Beauty is breaking down unrealistic standards of perfection. This is makeup made to feel good in, without hiding what makes you unique—because Rare Beauty is not about being someone else, but being who you are." https://www.rarebeauty.com/pages/about	48
Armani Beauty (Luminous Silk Perfect Glow Flawless Oil-Free Foundation)	"Giorgio Armani Beauty's durability ambition covers a wide range of initiatives including refillable product formats, scalable boutiques with reusable modules and packaging made of PCR-recycled plastic and glass. This ambitious and lasting overall design strategy is thus infused from the very early stage of any project until it comes to life." https://www.giorgioarmanibeauty-usa.com/a-vision-for-the-future/a-vision-for-the-future.html	40
Dior Beauty (Backstage Face & Body Foundation)	"Partnering to regenerate biodiversity." https://www.dior.com/en_us/beauty/store-page-folder/our-commitments.html#christian-dior-parfums-x-wwf	43
Sephora (Best Skin Ever Liquid Foundation)	"We believe that beauty thrives in diversity and discovery. Our purpose is to expand the way the world sees beauty by empowering the extraordinary in each of us." https://www.inside-sephora.com/en/about-sephora	50
Estée Lauder (Double Wear Stay-in-Place Foundation)	"Creating the Future of Beauty Together." https://www.esteelauder.com/discover/our-mission-statement	56
Too Faced (Born This Way Natural Finish Longwear Liquid Foundation)	"At Too Faced, we believe makeup is power, giving women the freedom to express themselves and the confidence to take on the world." https://www.elcompanies.com/en/our-brands/too-faced	35
Urban Decay (Face Bond Self Setting Waterproof Foundation)	"Unapologetically colorful, expressive, 100% cruelty-free... it's been the Urban Decay way since 1996, and things have only gotten bolder and brighter from here. Our makeup products are the holy grail for lovers of high-pigmented, long-lasting color. We aim for reinvention over perfection. Conventional beauty standards were never our thing. It's time to BE FEARLESS." https://www.urbandecay.com/about-us.html	40

Table EC.2 Brands, foundation lines chosen, and mission statements. The rightmost column shows the number of shades offered by each line.

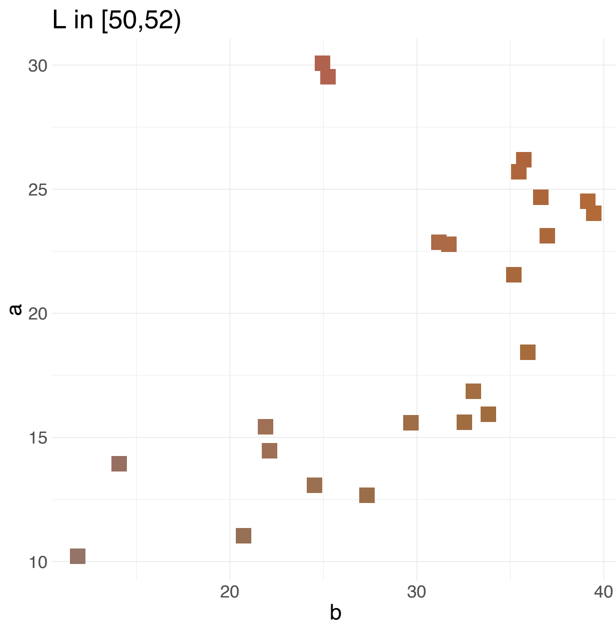


Figure EC.10 Impact of a and b for similar values of L.

To build intuition for the CIELAB (or $L*a*b$) color space, Figure EC.10 shows the impact of a and b dimensions. In this figure, all points have an L value between 50 and 52 but vary in a and b. The shades in the upper-right have a redder undertone compared to those in the bottom-left with a bluer undertone. In general, the value of a corresponds to the red-green spectrum and b corresponds to the blue-yellow spectrum.

EC.6.1. Brand color processing

Figure EC.11 shows a screenshot of Sephora's website after selecting a particular foundation. Each colored circle corresponds to a shade offered by that brand. Screenshots like this for each of the eleven brands considered are processed using Python's `skimage` package, and the median RGB color vector within each circle is extracted (after removing colors close to pure black and pure white). Once RGB color vectors are computed for each brand's foundation line, these vectors are converted to CIELAB space.

We note that it is possible that the colors depicted on Sephora's website are not accurate with respect to the actual make-up shade. However, because e-commerce comprises a significant portion of Sephora's business, it is likely that significant efforts are made to ensure that the shades depicted are accurate.

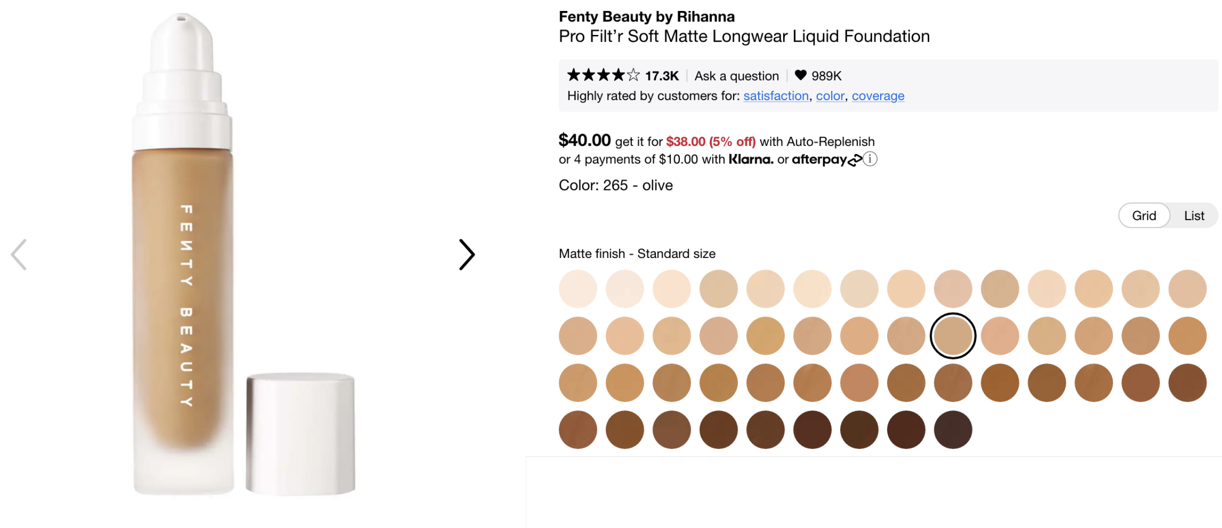


Figure EC.11 Screenshot of the Sephora.com page for Fenty Beauty's Pro Filt'r Soft Matte Longwear Liquid Foundation. The 51 shades offered are shown in thumbnails on the right. Date accessed: May 23, 2024.

EC.6.2. Population domain

When measuring each brand's inclusivity based on the percentage of characteristics covered, it is necessary to define a population domain. To do so, we consider the entire set of shades offered by each brand, along with two large skintone color palettes each containing 180 shades. Combining all of these skintone shades, we have 836 distinct shades. Let this be the set D. Then, across all 836 shades, we compute the maximum and minimum values of the L, a, and b dimensions. This gives us the following ranges: $L \in [9.02, 93.90]$, $a \in [0.49, 32.40]$, and $b \in [1.81, 43.26]$. To get a superset of all possible skintones, we consider the three-dimensional domain defined by the hyperrectangle $D^{super} \triangleq [9.02, 93.90] \times [0.49, 32.40] \times [1.81, 43.26]$.

The problem with using D^{super} as our population is that it may contain shades that are not realistic. Therefore, we keep all shades in D^{super} that are a distance of at most 5 from the nearest shade in D, where distances are computed as Euclidean distance in CIELAB space. This results in the population domain depicted in Figure 10.

EC.6.3. Impact of the chosen threshold

Figure 11 (right) and 12 use a threshold of $\tau = 10$ to define suitability. Figure 11 (left) shows the impact of different thresholds on the percentage of characteristics covered by each brand. As all

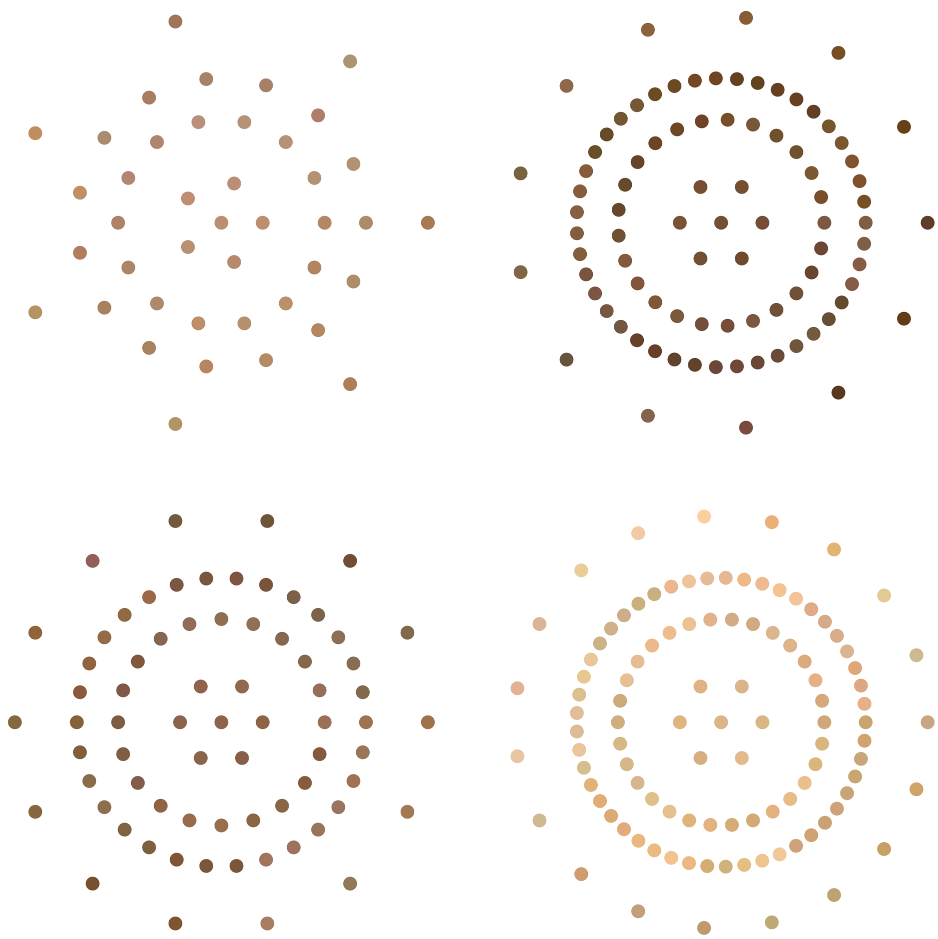


Figure EC.12 Illustration of distances between colors in the CIELAB space. The four rings around each center point show shades that are a Euclidean distance of 2, 5, 7, and 10, respectively from the center.

eleven brands claim that their foundation line is suitable for everyone, it is likely that distances of around 10-12 are considered reasonable for foundation makeup. If a smaller is threshold is chosen, for example a distance of 8, then many brands cover less than half of the population domain.

Figure EC.12 shows the impact on the chosen threshold on four different skin tones chosen at random. In each figure, the color in the center is the focal shade. The first ring around the focal shade is at a CIELAB distance of 2. The second, third, and fourth rings are a distance of 5, 7, and 10, respectively.

Figure EC.13 shows the number of offerings required to achieve inclusivity as the threshold varies. Of course, as τ decreases, the number of offerings required increases. Each point in Figure EC.13

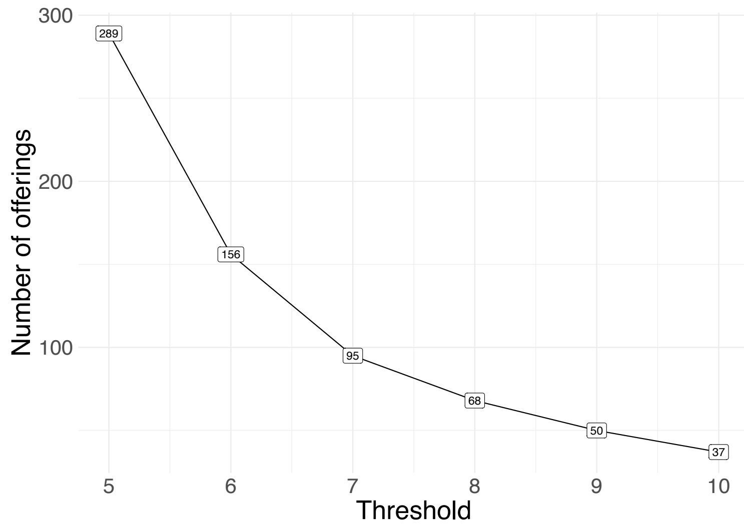


Figure EC.13 Number of offerings required to achieve inclusivity, where inclusivity is defined using the thresholds given on the x-axis. We note that these solutions are upper bounds on $G^{\text{I-ODP}}$, as they were obtained using the heuristics given in Steps (1)-(3) of Appendix EC.2.

was obtained by using the ILP heuristic described in Appendix EC.2 and are thus not necessarily the optimal number of offerings, but an upper bound on the number required for inclusivity.

EC.6.4. Details on the ILP solution

With a threshold of 10, ILP finds an inclusive set of 37 shades. Each point in Figure EC.14 shows one of those 37 shades. In Figure EC.14, the x-axis shows the distance between that shade and the closest shade in Fenty Beauty's foundation line, and the y-axis shows the distance to the closest shade in Haus Labs' foundation line. Shades that lie above $x = 10$ (resp. $y = 10$) are "missing" from Fenty Beauty's palette (resp. Haus Labs' palette). Interestingly, those missing from Haus Labs tend to have more yellow undertones, whereas those missing from Fenty Beauty tend to have bluer undertones.

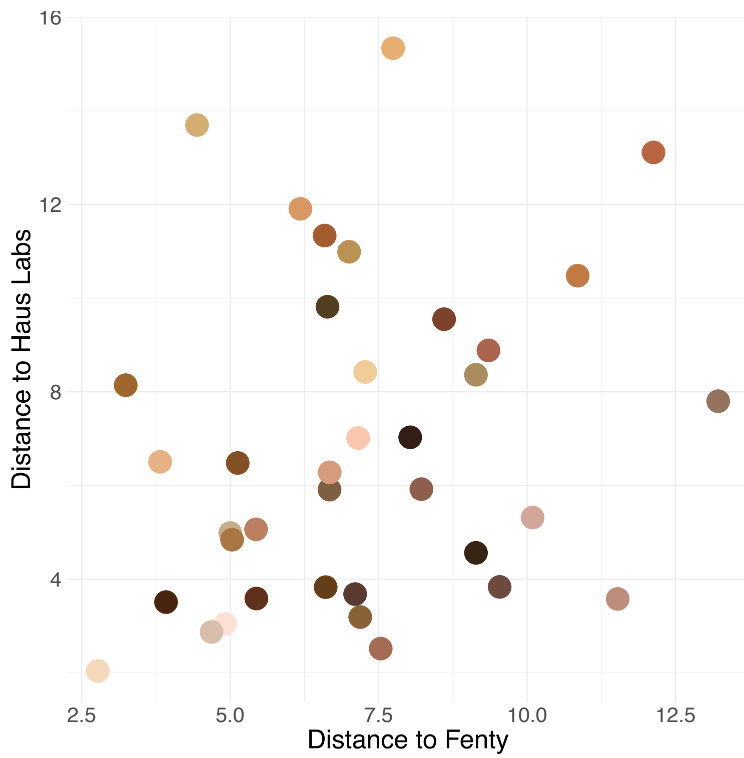


Figure EC.14 Each point corresponds to one of the shades present in the inclusive set of 37 shades determined by ILP using a threshold of 10. The x-axis shows the distance between that shade and the closest shade in Fenty Beauty's line, and the y-axis shows the distance to the closest shade in Haus Labs' line.