# Online Appendix for "Nested Bundling"

Frank Yang*

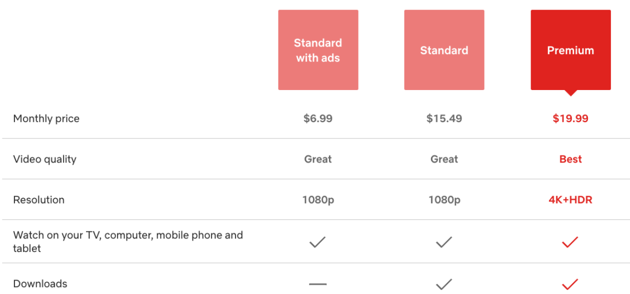
*Becker Friedman Institute, University of Chicago. Email: franky@uchicago.edu.

# B Online Appendix

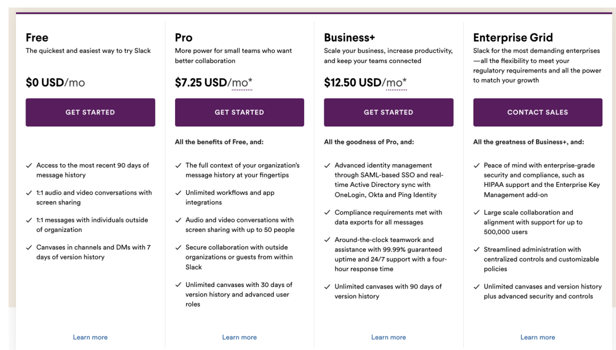
This appendix contains additional examples, results, and proofs that are supplementary to the main text. The sections are ordered by the order they are referenced in the main text.

## B.1 Empirical Examples of Nested Bundling

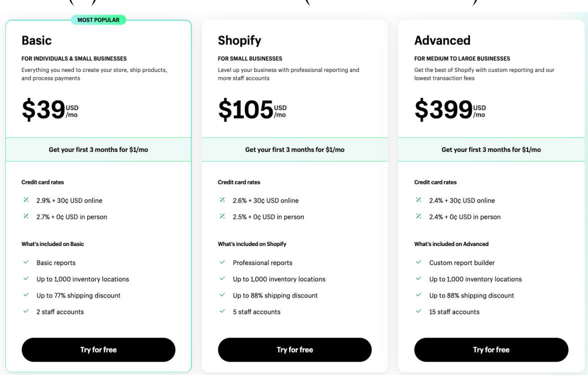
Figure 10 documents some empirical examples of nested bundling.



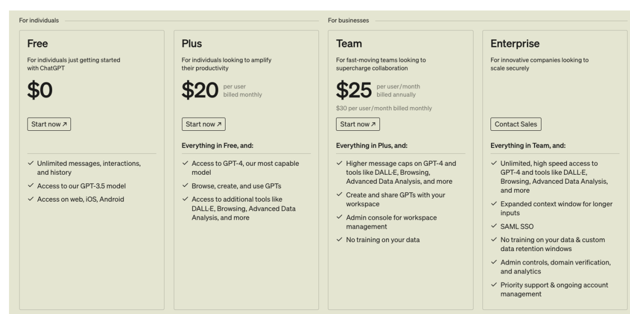

(a) Netflix's menu (Netflix 2023).




(b) Slack's menu (Slack 2023).




(c) Shopify's menu (Shopify 2023).




(d) OpenAI's menu (OpenAI 2024).

Figure 10: Empirical Examples of Nested Bundling

## B.2 General Procedure to Find the Optimal Menu

In practice, it might not be feasible to estimate the sold-alone quantities for all bundles when the seller must offer some base bundle (e.g., a "freemium" tier) to all consumers. In this section, we generalize our nesting condition to allow more ways to exclude bundles from consideration when finding the optimal menu.

For three bundles $b_0 \subset b_1 \subset b_2$, we say that $b_1$ ***is dominated by*** $b_2$ ***conditional on*** $b_0$ if

$$Q(b_1 \mid b_0) \leqslant Q(b_2 \mid b_0),$$

where for any $b \subset b'$, recall that $Q(b' \mid b)$ denotes the incremental quantity of $b'$ given $b$, i.e., the quantity at which the incremental profit function $\pi(b', q) - \pi(b, q)$ reaches its maximum in the interval $[0, \max\{Q(b'), Q(b)\}]$.

A bundle $b$ is ***strongly undominated*** if for all $b', b''$ such that $b' \subset b \subset b''$, we have

$$Q(b \mid b') > Q(b'' \mid b').$$

Clearly, a strongly undominated bundle must be an undominated bundle.

**Theorem 4** (General Nesting)**.** *Suppose that assumptions (A1) and (A2) hold. If the menu of strongly undominated bundles is nested, then it is a minimal optimal menu.*

Theorem 4 provides the following ***conditional sieve algorithm***:

**Step 1.** Pick any three bundles $b_0 \subset b_1 \subset b_2$.

**Step 2.** Remove $b_1$ from consideration if $Q(b_1 \mid b_0) \leqslant Q(b_2 \mid b_0)$.

**Step 3.** Repeat **Steps 1**-**2** until the remaining bundles are nested.

Theorem 4 implies that when **Step 3** stops, the remaining bundles always form an optimal menu regardless of how the bundles are chosen in **Step 1**. One can further apply Proposition 1 to the remaining bundles to find the minimal optimal menu. Theorem 4 generalizes Theorem 1 by allowing the removal of more bundles when checking the nesting condition. The proof again relies on Theorem 2 (monotone construction theorem) and Lemma 2 (switching lemma). In fact, the strongly undominated bundles are exactly the chain-essential elements in Theorem 2, when the partially ordered choice set is $(\mathcal{B}, \subseteq)$, and the objective function is the virtual surplus function.

*Proof of Theorem 4.* We follow the same proof strategy as in Theorem 1 and Proposition 1. In particular, we apply the monotone construction theorem (with the local single-crossing property), Theorem 3, to the partially ordered set $(\mathcal{B}, \subseteq)$ and virtual surplus function $\phi(b, t)$.

The result follows if we show that the strongly undominated bundles are exactly the chain-essential elements in Theorem 3 (see the proof of Theorem 1 and Proposition 1). We follow the notation as in the proof of Theorem 1. Because by assumption $Q(b) \in (0, 1)$, we have that $t(b) \in (\underline{t}, \overline{t})$ for all $b \neq \varnothing$. Therefore, for any $\varnothing \subset b_1 \subset b_2$, by Lemma 3, we have

$$t(b_2 \mid b_1) \geqslant \min\{t(b_1), t(b_2)\} > 0.$$

Thus, for any $b_1 \subset b_2$, we have $t(b_2 \mid b_1) > 0$.

Fix any strongly undominated bundle $b$. We show that $b$ is chain essential. If $b = \varnothing$ or $b = \overline{b}$, then we have that $b$ is a chain-essential element because *(i)* $\underline{t} < t(b')$ for all $b' \supset \varnothing$ and *(ii)* $t(\overline{b} \mid b') < \overline{t}$ for all $b' \subset \overline{b}$ (see the proof of Proposition 3). Thus, suppose that $\varnothing \subset b \subset \overline{b}$. Suppose for contradiction that $b$ is not chain essential. Then there exists $b_1 \subset b \subset b_2$ such that

$$t(b \mid b_1) \geqslant t(b_2 \mid b).$$

Since $t(b_2 \mid b_1) > 0$, by Lemma 6, this implies

$$t(b \mid b_1) \geqslant t(b_2 \mid b_1).$$

Now, by Lemma 3 and the proof of Proposition 1, we have

$$Q(b \mid b_1) \leqslant Q(b_2 \mid b_1),$$

contradicting to that $b$ is strongly undominated.

Now, fix any chain-essential element $b$. We show that $b$ is strongly undominated. Note that, for all $b_1 \subset b \subset b_2$, by the definition of chain-essential elements, we have

$$t(b \mid b_1) < t(b_2 \mid b).$$

Since $t(b_2 \mid b_1) > 0$, by Lemma 6, this implies that

$$t(b \mid b_1) < t(b_2 \mid b_1).$$

Then, by Lemma 3 and the proof of Proposition 1, we have

$$Q(b \mid b_1) > Q(b_2 \mid b_1).$$

Since this holds for all $b_1 \subset b \subset b_2$, we have that $b$ is strongly undominated. □

|  | $t_1$ | $t_2$ | $t_3$ |
|---|---|---|---|
| $\{1,2\}$ | 4 | 12 | 25 |
| $\{1\}$ | 2 | 6 | 8 |
| $\{2\}$ | 3 | 3 | 6 |

(a) Nesting condition fails

|  | $t_1$ | $t_2$ | $t_3$ |
|---|---|---|---|
| $\{1,2\}$ | 4 | 12 | 20 |
| $\{1\}$ | 2 | 6 | 8 |
| $\{2\}$ | 3 | 3 | 6 |

(b) Nesting condition holds

Table 2: Bundle values by types for Example 3. Circled are *sold-alone* monopoly prices. In case *(a)*, the nesting condition fails and the optimal mechanism is stochastic. In case *(b)*, the nesting condition holds, and the optimal mechanism is deterministic and given by the menu of undominated bundles.

## B.3 Additional Example

We provide an additional example that shows totally ordered types are not sufficient for the optimality of nested bundling. This example further illustrates our nesting condition. For simplicity, the example is discrete, but it can be made continuous by approximation.

**Example 3.** Suppose that there are two items $\{1,2\}$ and three types of consumers $\{t_1, t_2, t_3\}$ with mass $1/3$ each. Suppose that the costs are zero. We consider two cases.

**Case *(a)*.** The values are given by Table 2a. One can verify that the sold-alone quantities are given by $Q(\{1,2\}) = 1/3$, $Q(\{1\}) = 2/3$, and $Q(\{2\}) = 1$ (the sold-alone prices are circled in Table 2a). Thus, none of the bundles are dominated. So the nesting condition fails. Note that the nested menu $\{\{1\},\{1,2\}\}$ yields a profit $29/3$ (by pricing $\{1,2\}$ at 23, and $\{1\}$ at 6), and the nested menu $\{\{2\},\{1,2\}\}$ yields a profit $28/3$ (by pricing $\{1,2\}$ at 22, and $\{2\}$ at 3). The optimal deterministic menu in this case is not nested: it prices the bundle $\{1,2\}$ at 22, $\{1\}$ at 6, and $\{2\}$ at 3, which results in a profit

$$\frac{1}{3} \times \big(22 + 6 + 3\big) = \frac{31}{3} > \frac{29}{3}.$$

Moreover, the fully optimal mechanism is stochastic:

- price 22 for bundle $\{1,2\}$
- price $72/11$ for a lottery that puts probability $10/11$ on bundle $\{1\}$ and probability $1/11$ on bundle $\{1,2\}$
- price 3 for bundle $\{2\}$

which yields a profit

$$\frac{1}{3} \times \left(22 + \frac{72}{11} + 3\right) = \frac{347}{33} > \frac{31}{3}.$$

The suboptimality of nested bundling can be understood using Corollary 4 since the nested menu $\{\{2\},\{1,2\}\}$ that includes the best-selling bundle $\{2\}$ yields a strictly lower profit than the other nested menu $\{\{1\},\{1,2\}\}$.

**Case *(b)*.** The values are given by Table 2b, which is exactly the same as Table 2a except that type $t_3$'s value for bundle $\{1,2\}$ is lowered from 25 to 20. Given this change, when bundle $\{1,2\}$ is sold alone, the monopoly price would be 12 and the quantity $Q(\{1,2\})$ would be 2/3 rather than 1/3. Then, bundle $\{1\}$ is dominated, and the nesting condition holds. The optimal mechanism in this case is deterministic and given by the nested menu $\{\{2\},\{1,2\}\}$, which coincides with the undominated bundles. The optimality of nested bundling and the construction of optimal menu follow directly from Theorem 1 and Proposition 1.

## B.4 Implementability Lemma

For completeness, we also prove the following lemma about implementability which is used in Section III.

**Lemma 7.** *Suppose that assumption (A1) holds. Then, for any deterministic, monotone allocation rule $b(t)$, there exists a payment rule $p(t)$ such that $(b,p)$ satisfies all IC and IR constraints and that the lowest type $\underline{t}$ receives zero payoff under $(b,p)$.*

*Proof of Lemma 7.* Let $B = \{b(t)\}_{t\in\mathcal{T}}$ which is a nested menu. Without loss of generality, let $B = \{\varnothing, b_1,\ldots,b_m\}$ where $b_1 \subset \cdots \subset b_m$. For all $i = 1,\ldots,m$, let

$$s(b_i) := \inf\left\{t \in \mathcal{T} : b(t) \supseteq b_i\right\}.$$

We construct the bundle prices $\{p^\dagger(b)\}_{b\in B}$ by the following difference equation: for all $i = 1,\ldots,m$,

$$p^\dagger(b_i) - p^\dagger(b_{i-1}) := v(b_i, s(b_i)) - v(b_{i-1}, s(b_i)),$$

where we put $b_0 = \varnothing$ and $p^\dagger(b_0) = v(b_0,t) = 0$. To prove the result, it suffices to show that for all $t \in \mathcal{T}$, we have

$$b(t) \in \operatorname*{argmax}_{b'\in B}\left\{v(b',t) - p^\dagger(b')\right\}.$$

By assumption (A1), note that

$$U(b,t) := v(b,t) - p^\dagger(b)$$

has increasing differences, and hence single-crossing property, in $(b,t)$. Moreover, by construction, $s(b_i)$ is a crossing point of $U(b_i,t)$ and $U(b_{i-1},t)$. Since $b(\cdot)$ is monotone, we also have

$$s(b_1) \leqslant s(b_2) \leqslant \cdots \leqslant s(b_m).$$

Now fix any $i = 0,\ldots,m$ and any $t \in (s(b_i), s(b_{i+1}))$. For the edge cases, put $s(b_0) = \underline{t}$ and $s(b_{m+1}) = \bar{t}$. Now, observe that we have

$$U(b(t),t) = U(b_i,t) \geqslant U(b_{i+1},t) \geqslant U(b_{i+2},t) \geqslant \cdots \geqslant U(b_m,t),$$

and

$$U(b(t),t) = U(b_i,t) \geqslant U(b_{i-1},t) \geqslant U(b_{i-2},t) \geqslant \cdots \geqslant U(b_0,t).$$

Hence, for any such $t$, we have

$$U(b(t),t) = \max_{b' \in \mathcal{B}}\{U(b',t)\}.$$

Now, if $t = s(b_i)$ for some $i = 1,\ldots,m+1$, then by definition, we have that *(i)* $b(t) = b_i$ or $b(t) = b_{i-1}$, and *(ii)* $U(b_i,t) = U(b_{i-1},t)$. Hence, the above argument also implies that

$$U(b(t),t) = \max_{b' \in \mathcal{B}}\{U(b',t)\}.$$

Finally, suppose $t = s(b_0)$. If $s(b_0) < s(b_1)$, then the above argument holds for $t$. If $s(b_0) = s(b_1)$, then the above argument also holds for $t$ when applying to $i = 1$.

Thus, for all $t \in \mathcal{T}$, we have $U(b(t),t) = \max_{b' \in \mathcal{B}}\{U(b',t)\}$. Then, the payment rule defined by $p(t) := p^{\dagger}(b(t))$ implements the allocation rule $b(t)$, proving the result. $\square$

## B.5 Proof of Proposition 4

Suppose for contradiction that there are two elements $x_1$ and $x_2$ that are both chain essential but cannot be ordered. Then, we have

$$x_1 \wedge x_2 < x_1, x_2 < x_1 \vee x_2.$$

Since $x_1, x_2$ are chain essential, we have

$$t(x_1 \mid x_1 \wedge x_2) < t(x_1 \vee x_2 \mid x_1),$$

$$t(x_2 \mid x_1 \wedge x_2) < t(x_1 \vee x_2 \mid x_2).$$

Suppose without loss of generality that

$$t(x_1 \mid x_1 \wedge x_2) \leqslant t(x_2 \mid x_1 \wedge x_2).$$

Fix any $s$ such that

$$t(x_2 \mid x_1 \wedge x_2) < s < t(x_1 \vee x_2 \mid x_2).$$

Then we have

$$t(x_1 \mid x_1 \wedge x_2) \leqslant t(x_2 \mid x_1 \wedge x_2) < s < t(x_1 \vee x_2 \mid x_2),$$

which implies that

$$g(x_1, s) \geqslant g(x_1 \wedge x_2, s)$$

and

$$g(x_1 \vee x_2, s) < g(x_2, s),$$

contradicting that $g(\cdot, s)$ is quasisupermodular in $x$.

## B.6 Necessity of the Chain Condition

The chain condition is not only sufficient for monotone comparative statics but also *necessary* if one requires that the maximizer at each parameter can be found using only comparisons of the objective with ordered pairs (i.e., the pairs that satisfy the single-crossing property). That is, the iterative improvement arguments provided in Section II.D succeed *if and only if* the chain-essential elements are totally ordered.

**Proposition 11.** *Let $(\mathcal{X}, \leqslant)$ be a finite partially ordered set and $g : \mathcal{X} \times [0,1] \to \mathbb{R}$ be a function that is continuous in $t$ and satisfies the strict single-crossing property in $(x,t)$. The chain-essential elements for $g$ are totally ordered if and only if:*

- *(i) There exists a monotone selection $x(\cdot)$ such that $x(t) \in \operatorname{argmax}_{x \in \mathcal{X}} g(x,t)$ for all $t$.*
- *(ii) For all $t$ and $x_0$, there exists a sequence $(x_0, \ldots, x_n)$ such that $g(x_i, t) \leqslant g(x_{i+1}, t)$ for all $i$, $x_n = x(t)$, and each pair $(x_i, x_{i+1})$ satisfies either $x_i > x_{i+1}$ or $x_i < x_{i+1}$.*

*Proof of Proposition 11.* **($\implies$)** Suppose that the chain-essential elements are totally ordered. By Theorem 2, the existence of a monotone selection $x(\cdot)$ is immediate. We now show that for every $t \in [0,1]$ and every $x_0 \in \mathcal{X}$, there exists an ***improvement sequence*** $(x_0, \ldots, x_n)$ such that *(i)*

$$g(x_0, t) \leqslant g(x_1, t) \leqslant \cdots \leqslant g(x_n, t)$$

where $x_n = x(t)$ and that *(ii)* every pair $(x_i, x_{i+1})$ satisfies either $x_i < x_{i+1}$ or $x_i > x_{i+1}$.

We first prove this claim for all $s \notin \{t(x'' \mid x')\}_{x'<x''}$. Fix any such $s$. Let $\mathcal{Y}$ denote the set of chain-essential elements. Recall that $\mathcal{Y}$ is non-empty by definition. By Step 2 in the proof of Theorem 2, we know that if $x_0 \in \mathcal{Y}$, then there exists such an improvement sequence (moreover, the objective value is strictly increasing along the sequence). Now suppose $x_0 \notin \mathcal{Y}$. By Step 1 in the proof of Theorem 2, there exists $x_1 \in \mathcal{X}$ such that *(i)*

$$g(x_0, t) < g(x_1, t).$$

and that *(ii)* either $x_1 > x_0$ or $x_1 < x_0$.

If $x_1$ is in $\mathcal{Y}$, then we have found an improvement sequence by concatenating $(x_0, x_1)$ with an improvement sequence that starts with $x_1$ (which always exists since $x_1 \in \mathcal{Y}$).

If $x_1$ is not in $\mathcal{Y}$, then by Step 1 in the proof of Theorem 2 again, there exists $x_2 \in \mathcal{X}$ such that *(i)*

$$g(x_1, t) < g(x_2, t).$$

and that *(ii)* either $x_2 > x_1$ or $x_2 < x_1$.

Because $\mathcal{X}$ is a finite set, this process can be repeated at most $|\mathcal{X}|$ number of times until we find a sequence $(x_0, x_1, \ldots, x_n)$ such that *(i)*

$$g(x_0, t) < g(x_1, t) < \cdots < g(x_n, t).$$

and that *(ii)* every pair $(x_i, x_{i+1})$ satisfies either $x_i < x_{i+1}$ or $x_i > x_{i+1}$, and that *(iii)* $x_n \in \mathcal{Y}$. But then we may concatenate this improvement sequence with an improvement sequence that starts with $x_n$ (which always exists since $x_n \in \mathcal{Y}$). Moreover, the improvement sequence that starts with $x_n$ always ends with $x(t)$ by Step 2 in the proof of Theorem 2. Hence, our claim holds for all $s \notin \{t(x'' \mid x')\}_{x'<x''}$.

To show that this claim holds for all $s \in [0,1]$, we use a convergence argument as follows. Let $N = |\mathcal{X}|$ which is a finite number. Each improvement sequence can be viewed as a point in the finite-dimensional space $\{1, \ldots, N\}^N$ which is a compact subset of $\mathbb{R}^N$.

Fix any $s \in [0,1]$ and any $x_0 \in \mathcal{X}$. First, there exists a sequence $s_k$ converging to $s$ where $s_k \notin \{t(x'' \mid x')\}_{x'<x''}$ (since $\{t(x'' \mid x')\}_{x'<x''}$ has measure 0 in $[0,1]$). For each $s_k$, there exists an improvement sequence $Z_k \in \{1, \ldots, N\}^N$ by our previous step. Now, by the Bolzano-Weierstrass theorem, we know that there exists a converging subsequence $Z_{k_j}$ such that $Z_{k_j} \to Z \in \{1, \ldots, N\}^N$ as $j \to \infty$, where the convergence is with respect to the usual distance metric of $\mathbb{R}^N$. This implies that there exists some $J$ such that for all $j \geqslant J$, $Z_{k_j} = Z$. Therefore, for all $j \geqslant J$, each $s_{k_j}$ has the same improvement sequence $Z$. Denote

the improvement sequence by $(x_0, x_1, \ldots, x_n)$. Then, for all $j \geqslant J$, we have

$$g(x_0, s_{k_j}) \leqslant g(x_1, s_{k_j}) \leqslant \cdots \leqslant g(x_n, s_{k_j})$$

and $x_n = x(s_{k_j})$. By continuity of $g$ in $t$, we have

$$g(x_0, s) \leqslant g(x_1, s) \leqslant \cdots \leqslant g(x_n, s).$$

To ensure that $x_n = x(s)$, note that $x(\cdot)$ by construction is right-continuous at all $t \in [0,1)$ and left-continuous at $t = 1$. Hence, we may choose the approximating sequence $s_k$ to approximate $s$ from the right if $s < 1$ and to approximate $s$ from the left if $s = 1$. Then, we have $x_n = x(s)$, and hence $(x_0, x_1, \ldots, x_n)$ is a desired improvement sequence.

**(** $\Longleftarrow$ **)** Suppose for contradiction that the chain-essential elements cannot be totally ordered. Let

$$\mathcal{W} := \{x(t)\}_{t \in [0,1]}$$

which is totally ordered since $x(\cdot)$ is monotone. This implies that there must exist some chain-essential element $x^\dagger$ that is not in $\mathcal{W}$. By the definition of chain-essential elements, we have

$$\max_{x' : x' < x^\dagger} t(x^\dagger \mid x') < \min_{x'' : x'' > x^\dagger} t(x'' \mid x^\dagger),$$

where we put the left-hand side to be 0 if no $x' < x^\dagger$ exists and the right-hand side to be 1 if no $x'' > x^\dagger$ exists. Therefore, there exists some $s \in [0,1]$ such that for all $x' < x^\dagger$, we have

$$g(x^\dagger, s) > g(x', s)$$

and for all $x'' > x^\dagger$, we have

$$g(x^\dagger, s) > g(x'', s).$$

However, for such $s$ and $x_0 = x^\dagger$, we also know that there exists an improvement sequence. In particular, there exists some $x_1$ such that *(i)* either $x_1 > x^\dagger$ or $x_1 < x^\dagger$ and that *(ii)*

$$g(x^\dagger, s) \leqslant g(x_1, s).$$

But that is a contradiction. □

## B.7 Proof of Proposition 5

Since costs are zero, the price elasticity for any bundle $b$ at quantity $Q(b)$ must satisfy

$$\eta(b, Q(b)) = -1 .$$

Recall that

$$\mathrm{MR}(b,q) = P(b,q)\big[1 + \frac{1}{\eta(b,q)}\big].$$

This implies that the elasticity curve $\eta(b,\cdot)$ single-crosses $-1$ from below because the MR curve $\mathrm{MR}(b,\cdot)$ single-crosses 0 from above.

We claim that, under zero costs, the union elasticity condition implies the union quantity condition. Indeed, under zero costs and the union elasticity condition, for any $b_1, b_2$, because

$$\eta\big(b_1 \cup b_2, Q(b_1 \cup b_2)\big) = -1 ,$$

we have

$$\eta\big(b_1, Q(b_1 \cup b_2)\big) \geqslant -1 \quad \text{or} \quad \eta\big(b_2, Q(b_1 \cup b_2)\big) \geqslant -1 ,$$

and hence

$$Q(b_1 \cup b_2) \geqslant \min\big\{Q(b_1), Q(b_2)\big\} .$$

Thus, the union quantity condition holds. Thus, the nesting condition holds by Proposition 2.

## B.8 Proof of Proposition 6

Let $B$ be the proposed menu. By Proposition 5, the nesting condition holds. Hence, by Theorem 1, it suffices to show that any (non-empty) bundle $b \notin B$ is dominated. We start by showing that for all $i$, we have

$$Q(b_1^\star \cup \cdots \cup b_i^\star) \geqslant Q(b_i^\star) .$$

We prove this by induction on $i$. The base case $i = 1$ is trivial. For the inductive step, suppose that the claim holds for $i-1$. Now, observe that

$$Q(b_1^\star \cup \cdots \cup b_i^\star) \geqslant \min\Big\{Q(b_1^\star \cup \cdots \cup b_{i-1}^\star),\, Q(b_i^\star)\Big\} \geqslant \min\Big\{Q(b_{i-1}^\star),\, Q(b_i^\star)\Big\} = Q(b_i^\star),$$

where *(i)* the first inequality follows from that the union elasticity condition implies the union quantity condition (as shown in the proof of Proposition 5), *(ii)* the second inequality follows from the inductive hypothesis, and *(iii)* the last equality follows from the definition of $b_i^\star$ and $b_{i-1}^\star$. This proves the inductive step.

Now, fix any $b \notin B$. There exists some index $j$ such that $b = b_j^\star$. Since $b \notin B$, we have

$$b = b_j^\star \subset b_1^\star \cup \cdots \cup b_j^\star .$$

But, by the previous step, we also have

$$Q(b) = Q(b_j^\star) \leqslant Q(b_1^\star \cup \cdots \cup b_j^\star).$$

Thus, bundle $b$ is dominated, completing the proof.

## B.9 Proof of Proposition 7

Without loss of generality, suppose that there is a sequence of demand rotations for item 2. By Proposition 6, nested bundling is always optimal at any parameter $s$. By Proposition 1, the minimal optimal menu $B^{OPT}(s)$ equals the set of undominated bundles. To prove claim *(i)*, observe that it suffices to show that if $B^{OPT}(s) = \big\{\{1\},\{1,2\}\big\}$, then $B^{OPT}(s')$ must be $\big\{\{1\},\{1,2\}\big\}$ for any $s < s'$. Suppose not. Then, for some $s < s'$, we have

$$Q(\{1,2\};s') \geqslant Q(\{1\};s') = Q(\{1\};s) > Q(\{1,2\};s)\,,$$

which is impossible by our notion of demand rotations.

To prove claim *(ii)*, observe that it suffices to show that if $B^{OPT}(s') = \big\{\{2\},\{1,2\}\big\}$, then $B^{OPT}(s)$ must be $\big\{\{2\},\{1,2\}\big\}$ for any $s < s'$. Suppose not. Then, for some $s < s'$, we have

$$Q(\{2\};s) \leqslant Q(\{1,2\};s), \qquad Q(\{2\};s') > Q(\{1,2\};s'),$$

which is impossible by our notion of demand rotations.

To prove claim *(iii)*, observe that it suffices to show that it cannot be $|B^{OPT}(s)| = 1, |B^{OPT}(s')| = 2, |B^{OPT}(s'')| = 1$ for any $s < s' < s''$. To see why this is impossible, note that: if $B^{OPT}(s') = \big\{\{1\},\{1,2\}\big\}$, then $r_2(B^{OPT}(\cdot))$ cannot be nondecreasing, contradicting claim *(i)*; if $B^{OPT}(s') = \big\{\{2\},\{1,2\}\big\}$, then $r_1(B^{OPT}(\cdot))$ cannot be nonincreasing, contradicting claim *(ii)*.

### B.10 Proof of Proposition 8

By Theorem 1, it suffices to show that if $x^\dagger \notin X^\star$, then $x^\dagger$ is dominated by another quality level $x \neq x^\dagger$. Fix any $x^\dagger \notin X^\star$. Then $\widehat{Q}(x^\dagger) > Q(x^\dagger)$. Suppose, for contradiction, that there does not exist any $x > x^\dagger$ such that $Q(x) \geqslant Q(x^\dagger)$. Then, we have

$$\max_{x>x^\dagger}\left\{\widehat{Q}(x)\right\} = \max_{x>x^\dagger}\left\{Q(x)\right\} < Q(x^\dagger).$$

Let

$$\widetilde{Q}(x) := \begin{cases} \widehat{Q}(x) & \text{if } x \neq x^\dagger\,; \\ Q(x) & \text{otherwise}\,. \end{cases}$$

Then note that $\widetilde{Q}(x)$ is also nonincreasing and everywhere above $Q(x)$. Moreover, $\widetilde{Q}(x)$ is everywhere below $\widehat{Q}(x)$ with $\widetilde{Q}(x^\dagger) < \widehat{Q}(x^\dagger)$. Contradiction.

### B.11 Proof of Proposition 9

Note that since $Q(x) \in (0,1)$, we must have

$$\mathrm{MR}(Q(x)) \cdot x - C(x) = 0\,,$$

where the quality-adjusted marginal revenue curve $\mathrm{MR}(q) := \frac{\mathrm{d}}{\mathrm{d}q}(F^{-1}(1-q)\cdot q)$ is a strictly decreasing, continuous function. Therefore, we have

$$Q(x) = \mathrm{MR}^{-1}\Big(C_{avg}(x)\Big)\,.$$

Now, observe that for any function $h : \mathcal{X} \to \mathbb{R}$ and any strictly decreasing function $\Phi : \mathbb{R} \to \mathbb{R}$, we have

$$\mathbf{U}^{-}\big[\Phi \circ h\big] = \Phi \circ \mathbf{L}^{+}\big[h\big]\,,$$

where $\mathbf{U}^{-}[\,\cdot\,]$ denotes the upper decreasing envelope operator and $\mathbf{L}^{+}[\,\cdot\,]$ denotes the lower increasing envelope operator. Thus, we have

$$\widehat{Q}(x) = \mathrm{MR}^{-1}\Big(\breve{C}_{avg}(x)\Big)$$

because $\mathrm{MR}^{-1}(\,\cdot\,)$ is strictly decreasing. The claim follows from Proposition 8.

## B.12 Proof of Proposition 10

We map this problem into a bundling problem as follows. Consider a bundling problem with $n+m$ many items, where the first $n$ items represent ***quality upgrades*** exactly as in Section IV.B, and the remaining $m$ items represent the ***passes to avoid*** each of the $m$ costly activities. Specifically, for any $(x_i, y_j)$, we define

$$v\Big(\{1,\ldots,i\} \cup \{n+1,\ldots,n+m\} \backslash \{n+j\}, t\Big) := u(x_i,t) - c(y_j,t),$$

with $v(\{1,\ldots,i\} \cup \{n+1,\ldots,n+m\}, t) := u(x_i,t)$ being the value of quality $x_i$ without any costly action. We can map the production costs accordingly and let $v(b,t) = C(b) = 0$ for bundles $b$ that are not of the above form. With a slight abuse of notation, we also write $(x,y)$ as the bundle of quality $x$ and costly action $y$, and write $Q(x,y)$ as the corresponding sold-alone quantity for this *damaged* bundle, i.e., the unique quantity maximizing the profit function $\pi(x,q) - \pi(y,q)$.

($\Longleftarrow$) Suppose that $\min_{y>0} Q(y) < \max_{x>0} Q(x)$. Suppose for contradiction that there exists an optimal mechanism (among deterministic mechanisms) that does not use any costly instruments. Then, by Proposition 8, it is without loss of optimality to consider a menu $B$ such that

$$x^\star := \max\left\{\operatorname*{argmax}_{x>0} Q(x)\right\}$$

is the base-tier quality in menu $B$. Let $y^\star := \min\{\operatorname{argmin}_{y>0} Q(y)\}$. By assumption, $Q(y^\star) < Q(x^\star)$. Because $\pi(x^\star,q)$, $\pi(y^\star,q)$, and $\pi(x^\star,q) - \pi(y^\star,q)$ are strictly quasiconcave, we have

$$Q(y^\star) < Q(x^\star) \implies Q(y^\star) < Q(x^\star) < Q(x^\star,y^\star).$$

But this implies that

$$t(x^\star,y^\star) < t(x^\star),$$

where, as in the proof of Theorem 1, $t(\cdot)$ denotes the type at which the associated virtual surplus function crosses zero.[1] By Proposition 8 and the construction of Theorem 1, all types below $t(x^\star)$ consumes $\varnothing$ under the optimal mechanism. However, consider the perturbation of assigning the types $s \in [t(x^\star,y^\star), t(x^\star))$ the damaged bundle $(x^\star,y^\star)$. Because

$$u(x^\star,t) - (u(x^\star,t) - c(y^\star,t)) = c(y^\star,t)$$

[1] For any bundle $(x,y)$, the associated virtual function is given by $u(x,t) - C(x) - c(y,t) - \frac{1-F(t)}{f(t)}\big(u_t(x,t) - c_t(y,t)\big) = \phi(x,t) - \phi(y,t)$ where $\phi(x,t) := \frac{\mathrm{d}}{\mathrm{d}q}\pi(x,q)\,|_{q=1-F(t)}$ and $\phi(y,t) := \frac{\mathrm{d}}{\mathrm{d}q}\pi(y,q)\,|_{q=1-F(t)}$.

is strictly increasing in $t$, there exist prices to implement this change of the allocation. This change must increase the total profit by Lemma 1 since the virtual surplus function associated with bundle $(x^\star, y^\star)$ is strictly positive for all types $s > t(x^\star, y^\star)$. Contradiction.

($\implies$) Suppose that $\min_{y>0} Q(y) \geqslant \max_{x>0} Q(x)$. We show that there exists an optimal mechanism that does not use any costly instruments. Note that for all $x' > 0$ and $y' > 0$,

$$Q(y') \geqslant \min_{y>0} Q(y) \geqslant \max_{x>0} Q(x) \geqslant Q(x').$$

Because $\pi(x', q)$, $\pi(y', q)$, and $\pi(x', q) - \pi(y', q)$ are strictly quasiconcave, this implies that

$$Q(y') \geqslant Q(x') \geqslant Q(x', y').$$

Therefore, each damaged bundle $(x', y')$ is dominated by the undamaged version $x'$. Now, note that, by the quasiconcavity assumptions, *(i)* the virtual surplus function $\phi(x', t)$ single-crosses $\phi(x', t) - \phi(y', t)$ from below, *(ii)* both virtual surplus functions $\phi(x', t)$ and $\phi(x', t) - \phi(y', t)$ single-crosses 0 from below. Therefore, by the proof of Theorem 1, we have that $\phi(x', t) - \phi(y', t) \leqslant \max\{\phi(x', t), 0\}$ for all $t \in \mathcal{T}$.

Then, by Lemma 1, the optimal value of this screening problem is bounded from above by

$$\mathbb{E}\left[\max_{x \in \mathcal{X}} \phi(x, t)\right].$$

Since assumptions (A1) and (A2) hold for $\{u(x, t), C(x), F(t)\}$, by the proof of Theorem 1, the seller can attain the above profit by selling a deterministic menu of different qualities. Thus, costly screening is suboptimal.

## References


**Netflix.** 2023. "Pricing." *https://web.archive.org/web/20231009190817/https://www.netflix.com/signup/planform (Accessed November 2023).*

**OpenAI.** 2024. "Pricing." *https://web.archive.org/web/20240110213750/https://openai.com/chatgpt/pricing/ (Accessed July 2024).*

**Shopify.** 2023. "Pricing." *https://web.archive.org/web/20231025235039/https://www.shopify.com/pricing (Accessed November 2023).*

**Slack.** 2023. "Pricing." *https://web.archive.org/web/20231003031250/https://slack.com/pricing (Accessed November 2023).*