Online Appendix: A Characterization for Optimal Bundling of Products with Non-Additive Values

Soheil Ghili
Yale University

July 28, 2022

This online appendix provides an alternative proof to the main theorem in the manuscript. In this alternative environment, the setup and notations are also slightly different. As a result, I set up the problem again before giving the proof. I recommend to the reader to read this alternative setup and statement of the result before examining the proof.

Note that the original working paper includes further (non-peer-reviewed) analysis that do not appear in the paper or in this online appendix. Most importantly, the working paper provides an equivalent of the main result in a context of nonlinear pricing akin to Mussa and Rosen (1978). Additionally, it presents a simulation analysis which is used to discuss the implications of this paper for empirical research on bundling and second-degree price discrimination. To see these analyses, please visit my website at: https://sites.google.com/view/soheil-ghili/research.

The rest of this appendix is organized as follows. Section 1 sets up the model and formally presents the assumptions and the main result. Section 2 provides the proof. Note that there is a large overlap between the assumptions and lemmas in this proof and those of the proof provided in the main manuscript. In order for the reader to be able to follow the proof without referring to the main article, I do not drop the overlapping parts from this proof. For any assumption/result that is a repetition of (or is similar to) an assumption/result from the main text, the correspondence has been mentioned to aid the reader in comprehending the similarities between the two proofs.

1 Main Result

1.1 Setup and Notations

A monopolist has n products to sell, indexed 1 through n. Possible bundles of these products are denoted $b \subseteq \{1, ..., n\}$. Set $\mathcal{B} = \{b | b \subseteq \{1, ..., n\}\}$ represents the set of all possible bundles. By $\bar{b}$ denote the grand bundle $\{1, ..., n\}$. For any bundle b, denote $b^C = \bar{b} \setminus b$. There is a unit mass of customers whose types are represented by $t \in T \subset \mathbb{R}^m$ where T is compact. As will be shown later, one of the model assumptions will imply that types are one-dimensional (i.e., $m = 1$.)[1] Probability distribution over types $f(\cdot) > 0$ has no atoms.[2] The valuation by type t for bundle b is denoted $v(b, t)$. Assume $v(\emptyset, t) = 0$. Also, for all b, suppose that $v(b, t)$ is continuous in t. The per-unit cost of production for each product i is $c_i \geq 0$.

I consider deterministic selling procedures. The monopolist's problem has two components. First, she makes a bundling decision, choosing the optimal set B^* of bundles b among subsets B of $\mathcal{B}$ that satisfy $\emptyset \notin B$. Note that there are 2^{2^n-1} possible bundling strategies. Thus, characterizing the conditions under which the monopolist can simply choose $B^* = \{\bar{b}\}$ should indeed be of value.

Next, the monopolist chooses prices $p(\cdot) : B \to \mathbb{R}$ for the bundles offered.[3] Denote by $\mathcal{P}_B$ the set of all such pricing functions.

Once the firm decides on B and $p(\cdot)$, customers decide which bundles to purchase. Each customer t's decision $\beta(t|B, p) \subseteq B$ is determined by:

$$\beta(t|B, p) = \arg\max_{\hat{\beta} \subseteq B} v(\cup_{b \in \hat{\beta}} b, t) - \Sigma_{b \in \hat{\beta}} p(b) \tag{1}$$

Throughout, I assume customers break ties in favor of the seller. Also, note that equation 1 implies that customers want at most one unit of each product i and find additional units redundant.

Demand for bundle b is the measure of customers t who would purchase b:

[1] The reason why I start with general m and later show $m = 1$ is that this exposition clarifies that $m = 1$ is a an implication of the rest of the model assumptions, rather than a separate assumption itself.

[2] The main result should hold without these assumptions on f. But I expect the proof to be less clean.

[3] Note that, in principle, one could model the bundling decision through pricing: not offering a product would be equivalent to pricing it so high that no customer would purchase. As such, separating the bundling and pricing decisions in the model is redundant. Nevertheless, I decided to carry out this separation because it makes the notation easier.

$$D(b|B,p) = \int_t \mathbb{1}_{b\in\beta(t|B,p)} f(t)dt \tag{2}$$

Firm profit under strategy (B, p) is:

$$\pi(B,p) = \int_t \Sigma_{b\in\beta(t|B,p)} \Big(p(b) - \Sigma_{i\in b} c_i \Big) f(t)dt \tag{3}$$

The monopolist chooses (B^*, p^*) to maximize profit:

$$(B^*, p^*) = \arg\max_{B\in\mathcal{B}, p\in\mathcal{P}_B} \pi(B,p) \tag{4}$$

Unlike the proof in the main text, this setup only examines deterministic bundles.

Some additional definitions and notations: For disjoint bundles b and b', denote by $v(b, t|b')$ the valuation by type t for b conditional on possessing b'. Formally:

$$v(b, t|b') \equiv v(b \cup b', t) - v(b', t)$$

In a similar manner, denote $\beta(t|B, p, b') = \arg\max_{\hat\beta\subseteq(B\setminus\{b'\})} v(\cup_{b\in\hat\beta} b, t|b') - \Sigma_{b\in\hat\beta} p(b)$. Also denote $D(b|B, p, b') = \int_t 1_{b\in\beta(t|B,p,b')} f(t)dt$. Moreover, denote

$$\pi(B,p|b') = \int_t \Sigma_{b\in\beta(t|B,p,b')} \Big(p(b) - \Sigma_{i\in b} c_i \Big) f(t)dt.$$

Finally, for any set β of bundles, denote $\beta^\cup = \cup_{b\in\beta} b$. Similarly, I write $\beta^\cup(t|B, p, b')$ instead of $\cup_{b\in\beta(t|B,p,b')} b$. I next turn to the assumptions and the main result.

1.2 Assumptions and Characterization

I start by repeating the main assumptions from the main text and providing some definitions.

Assumption 1. ***Monotonicity:*** *For all b, $v(b, t)$ is increasing in $v(\bar b, t)$, and strictly so whenever $v(b, t) > 0$. The same applies to $v(b, t|b^C)$ for all b.*

Assumption 2. ***Quasi-concavity:*** *For any $b \in \mathcal{B}$, profit functions $\pi(b, p)$ and $\pi(b, p|b^C)$ are strictly quasiconcave in quantity as the monopolist varies p.*

Note that Assumptions 1 and 2 are the same as those from the main text. Overall, the assumptions here are similar to those in the main text except that $v(b, t)$ is only continuous and need not be differentiable in t here.

Definition 1. *By* $D^*(b)$ *denote the "optimal quantity sold" of bundle* b *if no other bundle were offered by the firm. Formally,* $D^*(b)$ *is defined as* $D(b|\{b\}, p_b^*)$ *where* p_b^* *is the optimal price for bundle* b *when* $B = \{b\}$*.*

Definition 2. *A given firm strategy* (B, p) *involves pure bundling if:*

$$\forall t : \beta^{\cup}(t|B, p) \in \{\emptyset, \bar{b}\}$$

Definitions of "optimal quantity sold" and "pure bundling" are conceptually similar to their counterparts in the main text but are defined in the posted-pricing setting rather than the mechanism-design setting.

We are now ready to state the main result.

Theorem 1. *Under assumptions 1 and 2, the optimal strategy* (B^*, p^*) *involves pure bundling if:*

$$D^*(\bar{b}) > \max_{b \in \mathcal{B} \setminus \{\bar{b}\}} D^*(b)$$

Conversely, the optimal strategy does not involve pure bundling if:

$$D^*(\bar{b}) < \max_{b \in \mathcal{B} \setminus \{\bar{b}\}} D^*(b)$$

In words, this result says that the firm should pure bundle if and only if it helps "sell more."

Note that the version of this result which was provided in the main text (and which, as the main text mentioned, was proposed by a referee at *American Economic Review: Insights*) is stronger than this version in two ways. First, the result in the main text applied to stochastic bundles as well. Second, the present result (unlike that in the main text) falls slightly short of a full if-and-only-if theorem since it is silent on whether pure bundling is optimal when $D^*(\bar{b}) = \max_{b \in \mathcal{B} \setminus \{\bar{b}\}} D^*(b)$.[4]

2 Proof of Theorem 1

I start by some remarks, definitions, and lemmas.

[4]One can show that under this last possibility, pure bundling is optimal if instead of assuming profits are strictly quasi-concave in each price, we assume they are strictly concave and differentiable at peak. Even though this would yield a full characterization, I decided that the ability to speak to the "measure-zero" case of $D^*(\bar{b}) = \max_{b \in \mathcal{B} \setminus \{\bar{b}\}} D^*(b)$ is too small a return to justify such a restrictive assumption as strict concavity. As such, I maintain the quasi-concavity assumption.

Lemma 1. *There is a mapping τ from the set T of types t on to the interval $[0,1]$ such that:*

1. $\forall t, t' \in T : v(\bar{b}, t) > v(\bar{b}, t') \Leftrightarrow \tau(t) > \tau(t')$.
2. *τ is a sufficient statistic: Once $\tau(t)$ is known, one can fully pin down all $v(b,t)$ without having to know t.*

Proof of Lemma 1. Set $\tau(t) \triangleq \frac{v(\bar{b},t)-\min_{t'\in T} v(\bar{b},t')}{\max_{t'\in T} v(\bar{b},t')-\min_{t'\in T} v(\bar{b},t')}$. By construction, it satisfies (1). To see why it satisfies (2), first note that by monotonicity, for any t, t' such that $v(\bar{b}, t) = v(\bar{b}, t')$, we have $v(b,t) = v(b,t')$ for any other b. As a result, $v(\bar{b}, t)$ is sufficient information for determining $v(b,t)$ for all b. Next, observe that one can recover $v(\bar{b}, t)$ from $\tau(t)$: $v(\bar{b}, t) = \min_{t'\in T} v(\bar{b}, t') + \tau(t) \times (\max_{t'\in T} v(\bar{b}, t') - \min_{t'\in T} v(\bar{b}, t'))$. As a result, once one knows $\tau(t)$, one would also know $v(b,t)$ for all b. **Q.E.D.**

Based on this lemma, it is without loss to think of t as $\tau(t)$ and, hence, the set of all possible t as $[0,1]$. Therefore, we can use expressions such as $t \geq t'$. Going forward, I assume $t \in [0,1]$. Note that this lemma is equivalent to Lemma 1 from the main text. I use it slightly differently here: I normalize the set of types to the $[0,1]$ interval rather than normalizing to $t = v(\bar{b}, t)$.

The next remark is also a repetition of one from the main text.

Remark 1. *Suppose functions $f_1(x)$, $f_2(x)$ and $f_1(x)+f_2(x)$ are all strictly quasi-concave over the interval $[a,b]$. Then either (i) $\arg\max f_1 \leq \arg\max f_2$ or (ii) $\arg\max f_1 \leq \arg\max(f_1 + f_2)$ will imply:*

$$\arg\max f_1 \leq \arg\max(f_1 + f_2) \leq \arg\max f_2.$$

The proof is left to the reader.

Definition 3. *For disjoint bundles b and b', denote by $D^*(b|b')$ the "optimal quantity sold" of bundle b if all customers are already endowed with b' and only b is offered by the firm at optimal price. Formally, $D^*(b|b') \equiv D(b|\{b\}, p^*_{b|b'}, b')$ where $p^*_{b|b'} : \{b\} \to \mathbb{R}$ is effectively one real number, and it is chosen among other possible p so that $\pi(\{b\}, p|b')$ is maximized.*

Next, I show that the problem of finding the optimal price for a bundle is equivalent to the problem of finding the right type t^* and sell to types $t \geq t^*$.

Definition 4. *Define by $t^*(b|b')$ the largest t such that $1 - F(t) \geq D^*(b|b')$. Also, for simplicity, denote $t^*(b|\emptyset)$ by $t^*(b)$.*

Lemma 2. *Consider bundles b and $b^C = \bar{b} \setminus b$. Suppose that all types are endowed with bundle b^C, and that the firm is selling only bundle b, optimally choosing $p^*_{b|b^C}$. The set of types who will buy the product at this price is the interval $[t^*(b|b^C), 1]$.*

Proof of Lemma 2. Follows directly from monotonicity. Monotonicity implies that the optimal sales volume $D^*(b|b^C)$ would be purchased by the highest types t with t weakly above some cutoff $\hat{t}$. Definition 4 says that for the demand volume to equal $D^*(b|b^C)$, the cutoff $\hat{t}$ has to equal $t^*(b|b^C)$. **Q.E.D.**

Lemma 2 is important in that it shows the problem of choosing $p^*_{b|b^C}$ can equivalently be thought of as the problem of choosing $t^*_{b|b^C}$. This allows us to set up the firm's problem based on t. Next definition introduces a necessary notation for this purpose.

Definition 5. *Consider disjoint bundles b and b'. Suppose that all types have already been endowed with b', and that the firm is to sell only bundle b. By $\pi_b(t|b')$ denote the profit to the firm if it chose a price for bundle b such that all types $t' \geq t$ would purchase bundle b:*

$$\pi_b(t|b') = (1 - F(t)) \times \big((v(t, b|b') - \Sigma_{i \in b} c_i\big)$$
$$= \pi(\{b\}, v(b, t|b')|b')$$

Note that $\pi_b(t|b^C)$ is strictly quasi-concave in t.

With the above definitions and lemmas in hand, we are ready to prove the main theorem. I start by the necessity condition (i.e., the condition that $D^*(\bar{b}) \geq D^*(b)$ for all b is necessary for pure bundling to optimal).

Proof of necessity. We want to show that if there is some b such that $D^*(b) > D^*(\bar{b})$, then pure bundling is sub-optimal. Specifically, I show that offering bundles b and $\bar{b}$ would be strictly more profitable to the firm compared to offering $\bar{b}$ alone. The argument follows.

Lemma 3. *$D^*(b) > D^*(\bar{b})$ implies $D^*(b) > D^*(b^C|b)$.*

Proof of Lemma 3. Suppose, on the contrary, that $D^*(b) \leq D^*(b^C|b)$. This means $t^*(b) \geq t^*(b^C|b)$. We know:

$$t^*(b) = \arg\max_t \pi_b(t)$$

and

$$t^*(b^C|b) = \arg\max_t \pi_{b^C}(t|b).$$

Also, given definition 5, it is straightforward to verify that:

$$\pi_{\bar{b}}(t) \equiv \pi_{b^C}(t|b) + \pi_b(t)$$

By strict quasi-concavity of all profits in t and by remark 1, it has to be that the argmax of $\pi_{\bar{b}}(t)$ falls in between the argmax values $t^*(b^C|b)$ and $t^*(b)$. Therefore, we get: $t^*(\bar{b}) \leq t^*(b)$, which implies $D^*(b) \leq D^*(\bar{b})$, contradicting a premise of the lemma. **Q.E.D.**

Lemma 4. *Selling $D^*(b^C|b)$ units of the grand bundle $\bar{b}$ along with $D^*(b) - D^*(b^C|b)$ units of bundle b would be strictly more profitable to the monopolist compared to selling $D^*(\bar{b})$ units of the grand bundle alone.*

Proof of Lemma 4. Note that given monotonicity and given Lemma 3, selling $D^*(b^C|b)$ units of the grand bundle $\bar{b}$ along with $D^*(b) - D^*(b^C|b)$ units of bundle b would simply mean selling b to types $[t^*(b), t^*(b^C|b))$ and selling $\bar{b}$ to types $[t^*(b^C|b), 1]$. This can be implemented by offering bundles b and $\bar{b}$ and pricing them at p_b^* and $p_b^* + p_{b^C|b}^*$ respectively.[5] This would deliver the following profit:

$$\pi_1 = \pi_{b^C}(t^*(b^C|b)|b) + \pi_b(t^*(b))$$

Again, by monotonicity, selling $D^*(\bar{b})$ units of the grand bundle can be thought of as selling $\bar{b}$ to types $t^*(\bar{b})$ and above. This would deliver a profit of $\pi_2 = \pi_{\bar{b}}(t^*(\bar{b}))$, which can be expanded and written as:

$$\pi_2 = \pi_{b^C}(t^*(\bar{b})|b) + \pi_b(t^*(\bar{b}))$$

Note that each term in π_2 is weakly less than its corresponding term in π_1 (due to the optimality of the terms in π_1). Also by the fact that $t^*(b^C|b) > t^*(b)$, then either $t^*(b^C|b) \neq t^*(\bar{b})$ or $t^*(b) \neq t^*(\bar{b})$. Thus, by strict quasi-concavity, at least one of the two inequalities between corresponding terms in π_1 and π_2 has to be strict, yielding $\pi_1 > \pi_2$. **Q.E.D.**

Given this lemma, the proof of the if side of the theorem is now complete. **Q.E.D.**

Note that the proof of the first side did *not* use the constant marginal costs assumption. Next, I turn to the proof of the sufficiency conditions (i.e., that $D^*(\bar{b}) > \max_{b \in \mathcal{B} \setminus \{\bar{b}\}} D^*(b)$ implies that pure bundling is optimal).

Proof of sufficiency. I start with some lemmas.

[5] Recall that I use the notation p_b^* for the optimal price of b when only b is offered. Similarly, $p_{b^C|b}^*$ is the optimal price of b^C when everyone is endowed by b and the monopolist is only selling b^C.

Lemma 5. *Under assumptions 1 and 2, there is a firm optimal strategy* (B^*, p^*) *such that non-measure-zero set of customers* t *we have* $\beta^{\cup}(t|B^*, p^*) = \bar{b}$.

Proof of Lemma 5. I start by assuming that no optimal strategy (B^*, p^*) involves selling $\bar{b}$ to a non-measure-zero set of consumers. Then I reach a contradiction by constructing a weakly profitable deviation from an assumed optimal (B^*, p^*) such that the deviation sells $\bar{b}$ to a non-measure-zero set of consumers.

By $D^*(\bar{b}) > D^*(b)$ for all $b \neq \bar{b}$, we get: $D^*(\bar{b}) > 0$, which in turn yields $t^*(\bar{b}) < 1$. Next, note that the number of possibilities for $\beta(t|B^*, p^*)$ is finite. By this finiteness and by continuity of value functions, there is some $\tilde{t} \geq t^*(\bar{b})$ such that all consumers with types higher than $\tilde{t}$ have the same purchase behavior.[6] Formally:

$$\forall t, t' \in (\tilde{t}, 1) : \beta^{\cup}(t|B^*, p^*) = \beta^{\cup}(t'|B^*, p^*)$$

Denote this commonly purchased bundle $\tilde{b}$. Given that our contrapositive assumption is that no non-measure-zero set of types purchases the grand bundle $\bar{b}$, it has to be that $\tilde{b} \neq \bar{b}$. Next, I construct a profitable deviation for the monopolist from (B^*, p^*). First, note that given that currently no consumer purchases $\bar{b}$ or constructs it from other bundles, it has to be that either $\bar{b}$ is not part of B^* or it is expensive enough for no customer to prefer to obtain it. Next, assume that the monopolist deviates from (B^*, p^*) by adding $\bar{b}$ to the set of bundles and pricing it at $p^*(\tilde{b}) + \Sigma_{i \in \tilde{b}^C} c_i + \epsilon$ where ϵ is chosen so that (i) type $\tilde{t}$ would weakly prefer $\tilde{b}$ over $\bar{b}$ but (ii) type 1 would weakly prefer $\bar{b}$ over $\tilde{b}$. Next, I proceed to show two things. First: finding such an ϵ is feasible. Second: with that ϵ, the monopolist will see a weak profit increase relative to (B^*, p^*) and sell $\bar{b}$ to a positive-measure set of types.

For type $\tilde{t}$ to weakly prefer $\tilde{b}$ over $\bar{b}$, it has to be that

$$v(\bar{b}, \tilde{t}) - p^*(\tilde{b}) - \Sigma_{i \in \tilde{b}^C} c_i - \epsilon \leq v(\tilde{b}, \tilde{t}) - p^*(\tilde{b})$$

$$\Leftrightarrow \epsilon \geq v(\tilde{b}^C, \tilde{t}|\tilde{b}) - \Sigma_{i \in \tilde{b}^C} c_i$$

Similarly, for type to 1 weakly to prefer to purchase $\bar{b}$, one can show that ϵ must satisfy:

$$\epsilon \leq v(\tilde{b}^C, 1|\tilde{b}) - \Sigma_{i \in \tilde{b}^C} c_i$$

But by monotonicity, we have $v(\tilde{b}^C, \tilde{t}|\tilde{b}) \leq v(\tilde{b}^C, 1|\tilde{b})$. Therefore, ϵ may be chosen within

[6] Perhaps with the exception of type $t = 1$; but that does not matter given its zero measure.

the following interval:

$$[v(\tilde{b}^C, \tilde{t}|\tilde{b}) - \Sigma_{i \in \tilde{b}^C} c_i, v(\tilde{b}^C, 1|\tilde{b}) - \Sigma_{i \in \tilde{b}^C} c_i]$$

If the interval is not a singleton, choose ϵ in the interior.

Next, I show that once such ϵ is chosen, the new bundling and pricing strategy by the firm will weakly increase the profit to the insurer while selling $\bar{b}$ to a non-measure zero set of consumers. To see the latter, denote by $\tilde{t}'$ the set of types who weakly prefer $\bar{b}$ over $\tilde{b}$ under this new strategy. By the choice of ϵ and by monotonicity, we have $\tilde{t} \leq \tilde{t}' < 1$. Therefore the new strategy will change the purchase behavior by types $t \geq \tilde{t}'$ (and only those types.) Next note that this behavior-change is weakly profitable to the monopolist. Prior to this change, the profit to the monopolist from these types was:

$$\pi_1 = (1 - \tilde{t}') \times (p^*(\tilde{b}) - \Sigma_{i \in \tilde{b}} c_i)$$

Under the new strategy (i.e., with the introduction of $\bar{b}$ at the price of $p^*(\tilde{b}) + \Sigma_{i \in \tilde{b}^C} c_i + \epsilon$,) the new profit level from these types is:

$$\pi_2 = (1 - \tilde{t}') \times \left((p^*(\tilde{b}) + \Sigma_{i \in \tilde{b}^C} c_i + \epsilon) - \Sigma_{i \in \bar{b}} c_i \right)$$

$$= \pi_1 + (1 - \tilde{t}') \times \epsilon$$

Thus, it remains to show $\epsilon \geq 0$. To this end, note that by $D^*(\bar{b}) > D^*(\tilde{b})$ we have $t^*(\bar{b}) \leq t^*(\tilde{b})$. This, by monotonicity, quasi-concavity, and remark 1, implies $t^*(\tilde{b}^C|\tilde{b}) \leq t^*(\bar{b})$ which in turn yields $t^*(\tilde{b}^C|\tilde{b}) \leq \tilde{t}$. That is, all types weakly above $\tilde{t}$ would purchase $\tilde{b}^C$ if (i) they were offered it at the optimal price for the monopolist and (ii) they were pre-endowed with $\tilde{b}$. This implies that $\forall t \geq \tilde{t} : v(\tilde{b}^C, t|\tilde{b}) \geq p^*_{\tilde{b}^C|\tilde{b}} \geq \Sigma_{i \in \tilde{b}^C} c_i$. But this means $\epsilon = v(\tilde{b}^C, \tilde{t}|\tilde{b}) - \Sigma_{i \in \tilde{b}^C} c_i \geq 0$. As a result, we get $\pi_2 \geq \pi_1$, which completes the proof of the lemma.**Q.E.D.**

Next, it is useful to observe that the monotonicity assumption imposes a vertical relationship not only on consumers' preferences, but also on their purchase behaviors.

Lemma 6. *Consider bundling strategy B and pricing strategy p. Consider types t and t' such that $\beta^{\cup}(t|B,p) \neq \beta^{\cup}(t'|B,p)$. Then the following statements hold:*

1. *If $\beta^{\cup}(t'|B,p) = \emptyset$, we have $t' < t$.*

2. If $\beta^{\cup}(t'|B,p) = \bar{b}$, we have $t' > t$.

This lemma says that if there are types who buy $\bar{b}$, they are the highest types. Also if there are types who buy nothing, they are the lowest types.

Proof of Lemma 6. I prove the second statement in the lemma. The first statement would be proved in a similar way. Suppose that $\beta^{\cup}(t|B,p) \neq \beta^{\cup}(t'|B,p) = \bar{b}$. For simplicity, denote $\beta^{\cup}(t|B,p) = \tilde{b}$. Now suppose, contrary to the statement of the lemma, that $t' \leq t$. Given $\beta(t|B,p)^{\cup} \neq \beta^{\cup}(t'|B,p)$, we know $t \neq t'$ which implies $t' < t$. Next, observe the following two inequalities:

First, note that under (B,p), type t prefers purchasing $\beta(t|B,p)$ and forming $\tilde{b}$ over purchasing $\beta(t'|B,p)$ and forming $\bar{b}$. Formally:

$$v(\tilde{b},t) - \Sigma_{b\in\beta(t|B,p)}p(b) \geq v(\bar{b},t) - \Sigma_{b\in\beta(t'|B,p)}p(b) \tag{5}$$

Similarly, type t' prefers to purchase $\beta(t'|B,p)$ and forming $\bar{b}$ over purchasing $\beta(t|B,p)$ and forming $\tilde{b}$. Formally:

$$v(\bar{b},t') - \Sigma_{b\in\beta(t'|B,p)}p(b) \geq v(\tilde{b},t') - \Sigma_{b\in\beta(t|B,p)}p(b) \tag{6}$$

At least one of the two inequalities above has to be strict (because if both types were indifferent between $\bar{b}$ and $\tilde{b}$, they would break this tie the same way.) Adding these two inequalities together, we get:

$$v(\bar{b},t') + v(\tilde{b},t) > v(\bar{b},t) + v(\tilde{b},t')$$

$$\Leftrightarrow v(\bar{b},t') - v(\tilde{b},t') > v(\bar{b},t) - v(\tilde{b},t)$$

$$v(\tilde{b}^C,t'|\tilde{b}) > v(\tilde{b}^C,t|\tilde{b})$$

This latter statement, combined with $t' < t$, contradicts monotonicity. **Q.E.D.**

In light of lemma 5, the following two corollaries of lemma 6 are useful.

Corollary 1. *Under (B^*,p^*), the set of types to for which $\beta^{\cup}(t|B^*,p^*) = \bar{b}$ takes the form of $[t_1,1]$ for some $t_1 < 1$.*

Corollary 2. *Under (B^*,p^*), the set of types to for which $\beta^{\cup}(t|B^*,p^*) = \emptyset$ takes the form of $[0,t_2)$ for some $t_2 < 1$.*

With these lemmas in hand, I next turn to the proof of the sufficiency conditions. The strategy is, again, contrapositive.

Assume on the contrary that we have, at the same time: (i) $\forall b \neq \bar{b} : D^*(b) < D^*(\bar{b})$ and (ii) the firm's optimal strategy does not involve pure bundling. This latter statement implies that the set of all distinct bundles chosen by customers under (B^*, t^*) includes members other than $\emptyset$ or $\bar{b}$. Formally, if we denote

$$\beta^* = \{b | \exists t : \beta^{\cup}(t|B^*, p^*) = b\}$$

then $\beta^* \setminus \{\emptyset, \bar{b}\} \neq \emptyset$. In other words, our contrapositive assumption implies that t_1 in corollary 1 is strictly larger than t_2 in corollary 2.

We now turn to the behavior of $\beta^{\cup}(t|B^*, p^*)$ in close vicinity of t_1 and t_2, starting with t_1. Take set B_1 to be the set of all $b \in \mathcal{B}$ such that no matter how close $t'_1 < t_1$ is to t_1, there is some $t \in [t'_1, t_1)$ with $\beta^{\cup}(t|B^*, p^*) = b$. That is, B_1 is the set of all bundles who are chosen by infinitely many types in $[t'_1, t_1)$ no matter how close t'_1 is to t_1 (note that B_1 may be a singleton). Pick some $\hat{t}_1 < t_1$ close enough to t_1 so that types between $\hat{t}_1$ and t_1 only choose bundles in B_1:

$$\forall t \in [\hat{t}_1, t_1) : \beta^{\cup}(t|B^*, p^*) \in B_1 \tag{7}$$

Among members of B_1, choose b_1 to be a bundle with the *"lowest complement net value"* at t_1. That is: $\forall b \in B_1 : v(b^C, t_1|b) - c_{b^C} \geq v(b_1^C, t_1|b_1) - c_{b_1^C}$.

Similarly, define B_2 to be the set of all bundles which, under (B^*, p^*), are chosen by infinitely many types in $[t_2, t'_2]$ no matter how close t'_2 is to t_2 (this set too may be a singleton). Pick some $\hat{t}_2 > t_2$ close enough to t_2 so that types between $\hat{t}_2$ and t_2 only choose bundles in B_2:

$$\forall t \in [t_2, \hat{t}_2] : \beta^{\cup}(t|B^*, p^*) \in B_2 \tag{8}$$

Among members of B_2, choose b_2 to be a bundle with the *"highest net value"* at t_2. That is: $\forall b \in B_2 : v(b, t_2) - c_b \leq v(b_2, t_2) - c_{b_2}$.

The rest of the proof of the sufficiency conditions of the theorem is organized as follows. I first make a series of claims (without proving them). Next I use the claims to prove the sufficiency conditions of the theorem. Finally, I will return to the proofs of the claims.

Claim 1. $t^*(b_1^C|b_1) \geq t_1$.

In words, claim 1 says that the set of customers who purchase the grand bundle $\beta(t|B^*, p^*) = \bar{b}$ under the firm optimal strategy (B^*, p^*) is a subset of those who purchase b_1^C and construct the grand bundle if (i) everyone is endowed with b_1 and (ii) the firm offers only b_1^C, pricing it optimally.

Claim 2. $t^*(b_2) \leq t_2$.

Claim 2 says that the set of customers who purchase $\emptyset$ under the firm optimal strategy (B^*, p^*) is a subset of those who purchase $\emptyset$ if the firm offers only b_2 and prices it optimally.

Next, note that the assumption $D^*(\bar{b}) > D^*(b_2)$, combined with monotonicity and claim 2, implies $t^*(\bar{b}) \leq t_2$. By $t_1 > t_2$, we get $t^*(\bar{b}) < t_1 \leq t^*(b_1^C|b_1)$. Also note that:

$$\forall t : \pi_{\bar{b}}(t) = \pi_{b_1^C}(t|b_1) + \pi_{b_1}(t)$$

As such, by strict quasi-concavity of profits, by $t^*(\bar{b}) < t^*(b_1^C|b_1)$, and by remark 1, the peak of $\pi_{\bar{b}}(t)$ should happen in between those of $\pi_{b_1^C}(t|b_1)$ and $\pi_{b_1}(t)$. Therefore, we should have: $t^*(b_1) \leq t^*(\bar{b}) \leq t^*(b_1^C|b_1)$. But $t^*(b_1) \leq t^*(\bar{b})$ implies:

$$D^*(b_1) \geq D^*(\bar{b})$$

which is a contradiction. Therefore, the sufficiency part of the theorem is true provided that claims 1 and 2 are true. I now turn to the proofs of these claims.

Proof of Claim 1. Suppose on the contrary that $t^*(b_1^C|b_1) < t_1$. In that case, it can be shown that the firm can strictly improve its profit upon the optimal strategy (B^*, p^*). The proof of this claim constructs such improvement. To this end, the following remark is useful to state.

Remark 2. *Construct the bundling strategy $(\hat{B}, \hat{p})$ from (B^*, p^*) in the following way:*

- $\hat{B} = \{\beta^{\cup}(t|B^*, p^*)\forall t\}$
- *For each t, or in other words for each $\hat{b} = \beta^{\cup}(t|B^*, p^*) \in \hat{B}$, set $\hat{p}(\hat{b}) = \Sigma_{b \in \hat{b}} p^*(b)$.*

For such $(\hat{B}, \hat{p})$, we have:

1. $\forall t : \beta(t|\hat{B}, \hat{p}) = \beta(t|B^*, p^*)$
2. $\pi(\hat{B}, \hat{p}) = \pi(B^*, p^*)$

This remark simply states that there is an optimal strategy by the seller under which each buyer type only purchases a single bundle instead of combining different bundles to construct her/his desired one. I skip the proof of this remark. Also, in order to save on notation, I assume from now on that it is (B^*, p^*) itself that has the feature of every $\beta(t|B^*, p^*)$ being a singleton. I now return to the proof of claim 1 and construct a strict improvement upon the profit of (B^*, p^*).

I do so by slightly adjusting the price of $\bar{b}$. That is, I show that there is a pricing strategy p with $p(b) = p^*(b)$ for all $b \neq \bar{b}$, but with $p(\bar{b}) \neq p^*(\bar{b})$, such that $\pi(B^*, p) > \pi(B^*, p^*)$.

To see why this is the case, first construct bundling strategy B' in the following way:

$$B' = \{b_1, \bar{b}\} \tag{9}$$

Also construct pricing strategy p' by fixing $p'(b_1) = \min_t v(b_1)$ and setting $p'(\bar{b}) = p'(b_1) + p^*(\bar{b}) - p^*(b_1)$. Note that under (B', p'), all types $t \geq t_1$ purchase $\bar{b}$ and all other types purchase b_1. This is because $p^*(\bar{b}) - p^*(b_1) = v(b_1^C, t_1|b_1)$ by construction.[7]

Assume, contrary to the claim, that $t^*(b_1^C|b_1) < t_1$. By quasi-concavity, this means that if the monopolist starts from (B', p') and then charges slightly less for $\bar{b}$, it will strictly increase its profit. That is, if the monopolist changes $p'(\bar{b})$ from $p'(b_1)+v(b_1^C, t_1|b_1)$ to $p'(b_1)+v(b_1^C, t|b_1)$ for any $t \in [t^*(b_1^C|b_1), t_1)$, the net profit change is strictly positive. Formally:

$$(F(t_1) - F(t)) \times \left[v(b_1^C, t|b_1) - c_{b_1^C}\right] - (1 - F(t_1)) \times \left[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\right] > 0 \tag{10}$$

The first term in this inequality indicates the fact that types $[t, t_1)$ now purchase $\bar{b}$ instead of b_1 and pay $v(b_1^C, t|b_1)$ more each, but at an increased cost of $c_{b_1^C}$ for each. The second term corresponds to the margin decrease by the amount of $v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)$ for all types in $[t_1, 1]$.

Using inequality 10 from the (B', p') setting, I now construct a strict improvement upon (B^*, p^*) by replacing it with (B^*, p), where $p(b) = p^*(b)$ for all $b \neq \bar{b}$ but $p(\bar{b}) \neq p^*(\bar{b})$. More specifically, choose $t \in [t^*(b_1^C|b_1), t_1)$ such that t is larger than $\hat{t}_1$ from equation 7 and that $\beta^{\cup}(t|B^*, p^*) = b_1$. Set the value of $p(\bar{b})$ as follows:

$$p(\bar{b}) = p^*(\bar{b}) - \left[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\right] \tag{11}$$

[7]This latter equality is given by continuity, the fact that b_1 is preferred over $\bar{b}$ under (B^*, p^*) by types arbitrarily close to t_1, and the fact that $\bar{b}$ is preferred over b_1 under (B^*, p^*) by type t_1.

This price change, too, has two effects on the monopolist's profit. First, it lowers the margin by $v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)$ for all types in $[t_1, 1]$. The resulting profit change is exactly equal to $-(1 - F(t_1)) \times \left[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\right]$, the latter term in equation 10.

The price change from p^* to p also leads all types $[t, t_1)$ to switch to $\bar{b}$.[8] Denoting this component of the profit change by Δ, we can write:

$$\Delta = \int_{\tau=t}^{t_1} \left(p^*(\bar{b}) - \left[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\right] - p^*(\beta^{\cup}(\tau|B^*, p^*)) - c_{[\beta^{\cup}(\tau|B^*,p^*)]^C} \right) f(\tau) d\tau$$

The is because each type τ will now (i) pay $p(\bar{b})$–which is replaced from equation 11– instead of $p^*(\beta^{\cup}(\tau|B^*, p^*))$ and (ii) also inflict an extra cost of $c_{[\beta^{\cup}(\tau|B^*,p^*)]^C}$ on the monopolist.

Next, observe that for all $\tau \in [t, t_1) : \beta^{\cup}(\tau|B^*, p^*) \in B_1$ and that by continuity, for any $b \in B_1$ we have: $p^*(\bar{b}) - p^*(b) = v(b^C, t_1|b)$. Plugging this into the above expression for Δ yields:

$$\Delta = \int_{\tau=t}^{t_1} \left(v\big([\beta^{\cup}(\tau|B^*, p^*)]^C, t_1|\beta^{\cup}(\tau|B^*, p^*)\big) - \left[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\right] - c_{[\beta^{\cup}(\tau|B^*,p^*)]^C} \right) f(\tau) d\tau$$

In words, for each type $\tau \in [t, t_1)$ with $\beta^{\cup}(\tau|B^*, t^*) = b \in B_1$, the margin change for the monopolist from this price decrease has two components: (i) the $-\left[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\right]$ component which was also present for types above t_1, and (ii) a $v(b^C, t_1|b) - c_{b^C}$ component caused by the switching from b to $\bar{b}$. But we know that this second component is exactly the "net complement value at t_1" which, by our choice of b_1, is weakly smaller for b_1 than for any other $b \in B_1$. That is: $\forall b \in B_1 : v(b^C, t_1|b) - c_{b^C} \geq v(b_1^C, t_1|b_1) - c_{b_1^C}$. Replacing this into the expression for Δ, we get:

$$\Delta \geq \int_{\tau=t}^{t_1} \left(v(b_1^C, t_1|b_1) - c_{b_1^C} - \left[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\right] \right) f(\tau) d\tau$$

$$= \int_{\tau=t}^{t_1} \left(v(b_1^C, t|b_1) - c_{b_1^C} \right) f(\tau) d\tau$$

$$= (F(t_1) - F(t)) \times \left(v(b_1^C, t|b_1) - c_{b_1^C} \right)$$

[8] To see why, note that type t will be indifferent between b_1 and $\bar{b}$ and chooses $\bar{b}$. By monotonicity, all lower types will prefer b_1 over $\bar{b}$ and hence, will not adopt $\bar{b}$ no matter whether b_1 is their optimal choice. All types larger than t will also buy $\bar{b}$ by Lemma 6.

But by construction, the total profit change due to a price change from p^* to p is given by:

$$\Delta - (1 - F(t_1)) \times \big[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\big]$$

Given $\Delta \geq (F(t_1) - F(t)) \times \Bigg(v(b_1^C, t|b_1) - c_{b_1^C}\Bigg)$, a lower bound for this total profit change is:

$$(F(t_1) - F(t)) \times \big[v(b_1^C, t|b_1) - c_{b_1^C}\big] - (1 - F(t_1)) \times \big[v(b_1^C, t_1|b_1) - v(b_1^C, t|b_1)\big]$$

But the latter expression is strictly positive by 10. Thus, changing the prices from p^* to p is strictly profitable. This is a contradiction that completes the proof of the claim. **Q.E.D.**

Proof of Claim 2. The proof of this claim is similar to that of the previous claim with some minor differences. Similar to the proof of the previous claim, assume that (B^*, p^*) takes the form described in Remark 2. I show that if $t_2 < t^*(b_2)$, then one can find pricing strategy $p(\cdot)$ such that $\pi(B^*, p) > \pi(B^*, p^*)$. To this end, consider seller strategies (B^*, p) and (B', p') where $B' = \{b_2\}$ and $p' = v(b_2, t_2)$. Note that given B' only includes one bundle, p' basically reduces to a real number rather than a function.

Observe that under (B', p'), types t_2 and above purchase b_2 while others purchase nothing. If $t_2 < t^*(b_2)$, then by quasi-concavity, the the seller will strictly profit by increasing p' from $v(b_2, t_2)$ to $v(b_2, t)$ for any $t \in (t_2, t^*(b_2)]$. Formally, the following profit change is strictly positive.

$$(1 - F(t)) \times \big[v(b_2, t) - v(b_2, t_2)\big] - (F(t) - F(t_2))\big[v(b_2, t_2) - c_{b_2}\big] > 0 \tag{12}$$

The first term in the above accounts for the now higher margin from types $[t, 1]$ while the second represents the set of types $[t_2, t)$ for whom the margin $v(b_2, t_2) - c_{b_2}$ is now lost due to the price increase. The above inequality holds for any $t \in (t_2, t^*(b_2)]$ but let us fix a t that is also smaller than $\hat{t}_2$ described in equation 8 and that satisfies $\beta(t|B^*, p^*) = b_2$.

Now let us construct p from p^* in the following way:

$$\forall b \in B^* : p(b) = p^*(b) + \big(v(b_2, t) - v(b_2, t_2)\big)$$

Before describing the profit change from this change in pricing, note that $p^*(b) = v(b, t_2)$ for any $b \in B_2$. This is by continuity, by the fact that all types to the left of t_2 prefer $\emptyset$ over any $b \in B_2$, and the fact that types arbitrarily close to t_2 from the right prefer b over $\emptyset$.

Now let us examine the profit change due to changing (B^*, p^*) to (B^*, p). First note that types $[t, 1]$ will not change their choices due to this price change. This is simply because: (i) preferences between any pair of non-empty bundles do not change given then uniform price increase; so those who buy anything after the price change will buy the same bundle as they did before the price change. And (ii) type t (as well as types above it) prefers b_2 over $\emptyset$ because $p(b_2) = p^*(b_2) + \big(v(b_2, t) - v(b_2, t_2)\big) = v(b_2, t)$.

As a result of choices not changing for types $[t, 1]$, the profit change from those type is given by: $(1 - F(t)) \times \big[v(b_2, t) - v(b_2, t_2)\big]$, exactly equal to its corresponding term from equation 12.

The second component of the profit change is due to the fact that types $[t_2, t)$ may no longer purchase any bundle due to the price increase. As the worst case scenario, I will work with the possibility that all types $[t_2, t)$ switch to $\emptyset$.[9] In this case, this profit change to the firm from this types, which we again denote Δ, will be given by:

$$\Delta = -\int_{\tau=t_2}^{t} \Big[p^*\big(\beta^{\cup}(\tau|B^*, p^*)\big) - c_{\beta^{\cup}(\tau|B^*, p^*)}\Big] f(\tau) d\tau$$

In words, for every type $\tau \in [t_2, t)$ who switches to $\emptyset$ from bundle $\beta^{\cup}(\tau|B^*, p^*) \in B_2$, the monopolist foregoes a margin of $p^*(b) - c_b$. Given that for all $b \in B_2$ we have $p^*(b) = v(b, t_2)$, this lost margin is in fact equal to the "net value" of b at t_2: $\forall b \in B_2 : v(b, t_2) - c_b \leq v(b_2, t_2) - c_{b_2}$. But we chose b_2 to be the bundle in B_2 with the highest such net value, and hence, the highest lost margin: $\forall b \in B_2 : v(b, t_2) - c_b \leq v(b_2, t_2) - c_{b_2}$. Thus, the profit loss is weakly smaller (i.e., net profit change is weakly larger) when types t switch to $\emptyset$ from different members of B_2, compared to the case where they all switch from b_2. This puts a lower bound on the net profit change Δ:

$$\Delta \geq -\int_{\tau=t_2}^{t} \Big[v(b_2, t_2) - c_{b_2}\Big] f(\tau) d\tau$$

$$= -(F(t) - F(t_2))\big[v(b_2, t_2) - c_{b_2}\big]$$

Thus, the overall profit change has the following lower bound:

$$(1 - F(t)) \times \big[v(b_2, t) - v(b_2, t_2)\big] - (F(t) - F(t_2))\big[v(b_2, t_2) - c_{b_2}\big]$$

[9]This is the "worst case scenario" because if there are types who still purchase some $b \in B^*$, the firm will profit from them.

But this is what we know to be strictly positive from equation 12. Thus, the change of strategy from (B^*, p^*) to (B^*, p) is guaranteed to deliver a strict profit improvement. This contradiction completes the proof of this claim and hence the theorem. **Q.E.D.**

References

Michael Mussa and Sherwin Rosen. 1978. Monopoly and product quality. *Journal of Economic theory* 18, 2 (1978), 301–317.