Multiproduct Price Discrimination with Two-Part Tariffs*

Ming Gao†

March 2010
(Preliminary version for conference review only.
Please email author for most up-to-date version.)

Abstract

This paper gives a new "multiproduct" explanation of the wide application of two-part tariffs, complementary to the classical "single-product" efficiency-related explanation. We consider a monopolist provider of n (> 1) products who uses two-part tariffs consisting of a membership fee common to all consumers and separate prices for different product bundles. We show that the change in demand for any bundle of $m \in [1, n]$ products caused by imposing an extra membership fee on top of any separate pricing strategy is proportional to the membership fee to the power of m. Therefore a small extra membership fee has no first-order impact on the demand for any multi-product bundles $(m > 1)$, which enables the firm to extract more consumer surplus. When this positive effect dominates the loss of single-product demand, two-part tariff dominates separate pricing. We present conditions that guarantee such an outcome, which generalize McAfee, McMillan and Whinston (1989)'s result from two products to multiple products. Our results suggest that two-part tariffs can achieve multidimensional price discrimination and should be subject to the same antitrust scrutiny as bundling strategies.

Key Words: two-part tariff, multiproduct pricing, price discrimination, bundling
JEL Codes: D42, L11, L12.

1 Introduction

In this paper, we uncover the mechanism through which a multiproduct monopolist can use two-part tariffs to achieve multidimensional price discrimination, and thereby provide a new explanation for the prevalence of two-part tariffs in real life. One classical explanation of *single-product* two-part tariffs is that it may minimize deadweight loss and hence achieve

*I thank Jean-Pierre Benoît, Emeric Henry, Preston McAfee and Joao Montez for valuable discussions and comments.

†Economics Department, London Business School, Regent's Park, London NW1 4SA, UK. Email: mgao.phd2005@london.edu.

efficiency (e.g. a firm may want to set the *unit price* of its product equal to marginal cost and capture all social surplus by an *entry fee*).

Now consider a monopolist provider of three products: 1, 2 and 3. Suppose the firm originally only sells them separately at respective prices p_1, p_2 and p_3. Then a buyer of products 2 and 3, say, needs to pay p_2+p_3. Suppose, in addition to p_1, p_2 and p_3, the firm now charges an extra "membership fee" ε to everyone who wants to buy any product at all (e.g. ε may be the entry fee to the shopping mall owned by the firm). Now a consumer has to pay total $q_2 = p_2 + \varepsilon$ for product 2 alone, and $q_3 = p_3 + \varepsilon$ for product 3 alone. But products 2 and 3 together now cost $p_2 + p_3 + \varepsilon = q_2 + q_3 - \varepsilon$, which is as if giving a "discount" of ε for purchasing the "bundle" of products 2 and 3. Moreover, buying all three products together now costs $p_1 + p_2 + p_3 + \varepsilon$, which gives an even higher "discount" of 2ε compared to separate purchases $(2\varepsilon = q_1 + q_2 + q_3 - (p_1 + p_2 + p_3 + \varepsilon))$. This is a form of multiproduct price discrimination, achieved by the two-part tariff consisting of the *membership fee* ε and the *separate product prices* p_1, p_2 and p_3.

In this paper we consider a monopolist provider of n (> 1) products and study when he would find it profitable to uses the kind of two-part tariff described in the previous example. In particular, we study the impact of imposing an extra membership fee on top of separate-product pricing strategies, which we call the *two-part-tariff effect*. We show that the change in demand for any bundle of $m \in [1, n]$ products due to the extra membership fee is proportional to the membership fee to the power of m. Therefore, a small extra membership fee has no first-order impact on the demand for any multi-product $(m > 1)$ bundles, and hence the firm extracts strictly more surplus from consumers of such bundles by imposing the extra membership fee. When this positive effect dominates the loss of single-product demand, two-part tariff dominates separate pricing. We present conditions that guarantee such an outcome.

Our results generalize McAfee, McMillan and Whinston (1989)'s result to the multi-product case. Their paper addresses the case of two products and provide conditions for *mixed bundling* to strictly dominate separate pricing. The two-part tariff we study can be seen as a particular way of mixed bundling, where membership fee and its multiples serve as the "bundle discounts" (as shown in the example of three products previously).

Although both two-part tariffs and mixed bundling can achieve multiproduct price discrimination, they work through different mechanisms. McAfee, McMillan and Whinston (1989) show that, offering a discount to the bundle of two products (down from the sum of their separate prices) will achieve the effect of increasing the demand for *both* products by just lowering *one* bundle price, thus increasing total profits. We show that imposing a small membership fee has *zero* first-order impact on the demand from *all* multi-product consumers, thereby enabling more surplus extraction from them.

Our results suggest that (multiproduct) two-part tariffs should be subject to the same antitrust scrutiny as other discriminatory pricing strategies, such as bundling.

This paper fits in the theoretical literature of multiproduct pricing. A few papers in this literature have identified important properties of *optimal* multiproduct pricing strategies. Armstrong (1996) finds that the optimal *non-linear* multiproduct pricing strategy will always exclude the lowest-typed consumers from the market. McAfee and McMillan (1988) and Manelli and Vincent (2006) identify the conditions that optimal *mixed bundling* strategies have to satisfy. A "drawback" of the optimal mechanisms studied in this literature is that they seldom appear very analogous to the strategies we observe in real life. For instance, Manelli and Vincent (2006) show that every multiproduct mixed bundling strategy may be dominated by a mechanism involving random assignments, which we rarely observe in real-life pricing.

Optimality is not the focus of this paper. Instead, we concentrate on two-part tariff as a particular form of pricing strategy. In this sense, this paper is closely related to Armstrong (1999), which shows *asymptotic* results that cost-based two-part tariffs can be "almost" optimal when the number of products is large. We show complementary results on the underlying mechanism of two-part tariffs, which hold for any number of products.

The remainder of this paper is organized as follows: Section 1 describes the model; Section 2 shows the effect of two-part tariffs on demand, i.e. the two-part-tariff effect; Section 3 compares profits under two-part tariffs and separate pricing, and provides conditions for the former to dominate the latter; Section 4 concludes.

2 Model[1]

There is only one firm, which produces $n(>1)$ different kinds of products. We use $j \in \{1, 2, ..., n\}$ to denote a **product**. There is no cost of production. The firm maximizes total profit.

There is a continuum of **consumers**, each of whom has a **valuation** for each of these products (i.e. the utility she derives from the product) and demands 0 or 1 unit of each product (i.e. 2 or more units of any one product will provide the exact same utility as 1 unit of that product). The total utility a consumer derives from consuming different products is simply the sum of her valuations for these products. No consumption results in zero utility. We denote a consumer's **type** by an n-dimensional real-valued parameter $\mathbf{x} \equiv (x_1, ..., x_n)$, where x_j is this consumer's valuation for product j.

The firm does not know each consumer's type. Rather, it has prior (p.d.f.) $f(\mathbf{x})$ of the distribution of $\mathbf{x}$ among consumers. The support of f is denoted by $\mathbf{S} \equiv \times_{j=1}^{n} S_j \subset \mathbb{R}^n$, where S_j is the support in dimension j (i.e. for product j).

Assumption 1 $\mathbf{S}$ *is weakly convex and has full dimension in* $\mathbb{R}^n$.

[1] Our notation mostly follows that of Manelli and Vincent (2006).

For expository simplicity, in the following discussion we focus on the case when $\mathbf{S} = [0,1]^n \equiv I^n$. All the results can easily be shown to hold for general $\mathbf{S}$ that satisfies Assumption 1.

Assumption 2 *f is continuous and $f(\mathbf{x}) > 0$ if and only if $\mathbf{x} \in I^n$.*

A **bundle** is a set of different products. We denote the full bundle of all n products as $N = \{1, 2, ..., n\}$. Any bundle, denoted by J, is therefore a subset of N, i.e. $J \subset N$. The empty bundle is $\varnothing \subset N$.

When it does not cause confusion, we also use j to represent the bundle $\{j\}$ (i.e. containing only product j). And j^c simply means $\{j\}^c$.

For bundle J, we denote the number of products in it by $|J|$.

A general rule we use in the notation below is: superscript represents dimensionality; subscript represents bundle or product.

Definition 1 (Price Schedule) *A **price schedule** $\mathbf{P}$ specifies the price for each possible bundle, $\mathbf{P} \equiv \{p_J\}_{J \subset N}$, where $p_J \in \mathbb{R}^+$ for any $J \subset N$.*

Note that $\mathbf{P}$ consists of 2^n prices since there are 2^n possible bundles (including the full bundle and the empty bundle).

Definition 2 (IC) *Consider any K different subsets of N, denoted $\{J_k\}_{k=1,2,...,K}$ where $J_k \subset N$ $\forall k = 1, 2, ..., K$. A price schedule $\mathbf{P} = \{p_J\}_{J \subset N}$ is **incentive compatible** (IC) if the following condition holds for all $K = 1, 2, 3, ..., 2^n$*

$$p_{\bigcup_{k=1}^{K} J_k} \leq \sum\nolimits_{k=1}^{K} p_{J_k}$$

Intuition: In an IC price schedule, the price of a bundle would not exceed the sum of the prices of any "profile" of its sub-bundles that forms a full "cover" of this bundle, otherwise no consumer would ever demand this bundle. IC is a necessary condition for each bundle to attract some demand.

Since a *partition* of a bundle is such a profile of sub-bundles, IC therefore implies that *the price of a bundle in an IC price schedule would not exceed the sum of the sub-bundle prices in any partition of this bundle.*

Our discussion from now on focuses only on IC price schedules.

Definition 3 (Additivity/Separate Pricing) *A price schedule $\mathbf{P} = \{p_J\}_{J \subset N}$ is **additive** (or **separate pricing**) if $p_\varnothing = 0$ and $p_J = \sum\limits_{j \in J} p_j$ for any non-empty $J \subset N$.*

$p_\varnothing = 0$ is actually a constraint on all price schedules that satisfy consumers' individual rationality. Note that this definition implies the following result.

Lemma 4 *If* $\mathbf{P} = \{p_J\}_{J \subset N}$ *is additive, for any two bundles* J *and* K*, it must be* $p_J + p_K = p_{J \cup K} + p_{J \cap K}$.

Lemma 5 *Additivity implies IC.*

Definition 6 (Demand Segment) *Given IC price schedule* $\mathbf{P} = \{p_J\}_{J \subset N}$*, the* ***demand segment*** *for any bundle* $J \subset N$*, denoted* A_J*, is the set of all the consumers that buy bundle* J*:*

$$A_J \equiv \{\mathbf{x} \in I^n | \sum_{j \in J} x_j - p_J \geq \sum_{k \in K} x_k - p_K, \forall K \subset N\}$$

Definition 7 (Allocation) *A (consumer)* ***allocation*** *given price schedule* $\mathbf{P}$ *is the profile of demand segments of all bundles induced by* $\mathbf{P}$*, denoted* $\{A_J\}_{J \subset N}$*.*[2]

Lemma 8 (Additive Allocation) *If* $\mathbf{P} = \{p_J\}_{J \subset N}$ *is additive, the allocation it induces* $\{A_J\}_{J \subset N}$ *must satisfy for any* $J \subset N$

$$A_J = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k < p_k, \forall k \in J^c\} \tag{1}$$

Intuition: An additive price schedule allocates all consumers into "cubes" delineated by orthogonal hyperplanes.

Definition 9 (Truncated Type) *Given any bundle* $J \subset N$*, a* $J-$***truncated type parameter*** *is denoted* $\mathbf{x}^J = \{x_j\}_{j \in J} \in I^J$*, where* $I^J \equiv \times_{j \in J} I_j$*.*

We sometimes use x_j^J (where $j \in J$) to denote the element of $\mathbf{x}^J$ pertaining to product j.

Notice $\mathbf{x}^J$ is a $|J|$-dimensional vector (or a point) in I^J. We use J instead of $|J|$ as the superscript of $\mathbf{x}^J$ to emphasize that $\mathbf{x}^J$ keeps the dimensions in I^n *according to* bundle J, rather than any $|J|$ dimensions of I^n. This distinction is important for the following definitions.

Definition 10 (Projection) *For any (consumer set)* $A \subset I^n$ *and any bundle* $K \neq \varnothing$*, define*

$$A^K \equiv \{\mathbf{x}^K \in I^K | (\mathbf{x}^K, \mathbf{y}) \in A, \text{ for some } \mathbf{y} \in I^{K^c}\}$$

which is the ***projection of set*** A *on the* $|K|$*-dimensional hyperplane defined by the following* $|K^c|$ *equations:*

$$\{x_j^{K^c} = 0\}_{j \in K^c} \tag{2}$$

Note that when $K = N$, because $|N^c| = 0$, the projection operation above is the identity mapping, i.e. $A^N = A$.

[2] To lighten notation, we do not carry $\mathbf{P}$ in A_J or $\{A_J\}_{J \subset N}$, but it is always implied that a demand segment or allocation is induced by some price schedule.

Definition 11 (Projection of A_J) *For any two bundles $K, J \neq \varnothing$, the set*

$$A_J^K \equiv \{\mathbf{x}^K \in I^K | (\mathbf{x}^K, \mathbf{y}) \in A_J, \textit{ for some } \mathbf{y} \in I^{K^c}\}$$

is the ***projection*** *of set A_J on the $|K|$-dimensional hyperplane defined by the $|K^c|$ equations of (2).*

Definition 12 (Probability Measure) *For any $A \subset I^n$, we define the probability measure of A as $M(A)$ which satisfies*

$$M(A) = \int_A f(\mathbf{x}) d\mathbf{x}$$

We use $M^J(A) = \int_A f(\mathbf{x}^J) d\mathbf{x}^J$ to denote the **marginal measure** in I^J of set A, for any $J \subset N$, which is particularly useful when A does not have full dimension in I^n but has full dimension in I^J.

3 Two-Part Tariffs and the Two-Part-Tariff Effect

Definition 13 (Two-Part Tariff) *Given any price schedule $\mathbf{P} = \{p_J\}_{J \subset N}$, a two-part tariff is the price schedule $\mathbf{Q} \equiv (\varepsilon, \mathbf{P}) = \{q_J\}_{J \subset N}$ where*

$$q_J = \begin{cases} p_J + \varepsilon & , \textit{ if } J \neq \varnothing \\ 0 & , \textit{ if } J = \varnothing \end{cases} ; \textit{ and } \varepsilon > 0.$$

Comment: Compared to $\mathbf{P}$, $\mathbf{Q}$ increases the prices of all *non-empty* bundles by the *same amount* ε, whilst giving the empty bundle for free. $\mathbf{Q}$ therefore imposes an additional membership fee ε on top of the individual prices of products or bundles specified by $\mathbf{P}$ (except for $\varnothing$).

We have chosen to define two-part tariffs this way to emphasize the additional membership fee. It is easy to see that, since $\mathbf{P}$ can be any price schedule in Definition 13, actually *any price schedule $\mathbf{Q} = \{q_J\}_{J \subset N}$ with $q_J > 0$ $\forall J \neq \varnothing$ and $q_\varnothing = 0$ is a two-part tariff.* This is a very broad range of price schedules. What we really need for the purpose of this paper is only a very small subset of such schedules.

Consider the following two price schedules:

$$\begin{array}{rl} \mathbf{P} = \{p_J\}_{J \subset N}: & \textbf{additive}, \\ & \text{and inducing demand } \{A_J\}_{J \subset N}; \\ \mathbf{Q} = \{q_J\}_{J \subset N}: & \text{where } q_J = \left\{ \substack{p_J + \varepsilon, J \neq \varnothing \\ 0, J = \varnothing} \right., \varepsilon > 0, \\ & \text{and inducing demand } \{C_J\}_{J \subset N}. \end{array} \tag{3}$$

The $\mathbf{Q}$ in (3) is a two-part tariff defined using additive $\mathbf{P}$ that induces strictly positive single-product demand segments. Since $\mathbf{P}$ is additive, $\mathbf{Q}$ will not be additive as long as $\varepsilon > 0$.

More importantly, ε also acts as a kind of "bundle discount" in this case. To see this, suppose under $\mathbf{Q}$ consumer 1 demands bundle $\{1\}$ (by paying $q_1 = p_1 + \varepsilon$) consumer 2 demands bundle $\{2\}$ (by paying $q_2 = p_2 + \varepsilon$) and consumer 3 demands bundle $\{1, 2\}$ (by paying $q_{\{1,2\}} = p_{\{1,2\}} + \varepsilon = p_1 + p_2 + \varepsilon = q_1 + q_2 - \varepsilon$). Compared to 1 and 2, it is as if 3 gets a "discount" of ε, since 3 only needs to pay membership fee ε *once* although she buys *two* products. Actually, it is easy to see that a consumer of any bundle J under $\mathbf{Q}$ will get a "discount" of $(|J| - 1) \cdot \varepsilon$ compared to the consumers of the $|J|$ individual products. This is a special feature of the two-part tariff we defined in (3).

The in the remaining parts of the paper we focus on the two-part tariffs defined in (3).

Theorem 1 *Consider the price schedules and allocations defined in (3). If* $\mathbf{x}$ *is* ***uniformly distributed****, then for* ***any*** $J \neq \varnothing$*, we have*

$$M(A_J \backslash C_J) = c(\mathbf{P}) \cdot \varepsilon^{|J|}$$

where $c(\mathbf{P})$ *is a function of* $\mathbf{P}$ *(but not of* ε*).*

Intuition: Starting from an additive price schedule, imposing an extra membership fee will lead to a decrease in the demand for any non-empty bundle that is of the same order as the number of products in the bundle.

Proof.

Definition 14

$$A_J(\varepsilon) \equiv \{\mathbf{x} \in A_J | 0 \leq \sum_{j \in J} x_j - p_J < \varepsilon\} \tag{4}$$

Since $\mathbf{P}$ is additive, by (1) we know:

$$A_J = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k < p_k, \forall k \in J^c\}$$

Therefore

$$\begin{aligned} A_J(\varepsilon) &= \{\mathbf{x} \in A_J | 0 \leq \sum_{j \in J} x_j - p_J < \varepsilon\} \qquad (5) \\ &= \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; \sum_{j \in J} x_j < \sum_{j \in J} p_j + \varepsilon; x_k < p_k, \forall k \in J^c\} \end{aligned}$$

which implies:

Lemma 15 $\forall J \subset N$*, we have* $A_J = C_J \cup A_J(\varepsilon)$*.*

Lemma 16 $\forall J \subset N$*, we have* $M(A_J) = M(C_J) + M(A_J(\varepsilon))$*.*

Therefore $M(A_J \backslash C_J) = M(A_J) - M(C_J) = M(A_J(\varepsilon))$

Denote $A_J^{J^c}(\varepsilon) \equiv (A_J(\varepsilon))^{J^c}, A_J^J(\varepsilon) \equiv (A_J(\varepsilon))^J$.

Then by Definition 11, (1) and (5) we know:

$$\begin{aligned} A_J^J(\varepsilon) &= \{\mathbf{x}^J \in I^J | (\mathbf{x}^J, \mathbf{y}) \in A_J(\varepsilon), \text{for some } \mathbf{y} \in I^{J^c}\} \\ &= \{\mathbf{x}^J \in I^J | x_j \geq p_j, \forall j \in J; \sum_{j \in J} x_j < \sum_{j \in J} p_j + \varepsilon\} \end{aligned} \tag{6}$$

From this expression we know $A_J^J(\varepsilon)$ has full dimension in I^J, and each of its "$|J|$ sides" has "length" exactly equal to ε.

By (5) we also know

$$\begin{aligned} A_J^{J^c}(\varepsilon) &= \{\mathbf{x}^{J^c} \in I^{J^c} | (\mathbf{x}^{J^c}, \mathbf{y}) \in A_J(\varepsilon), \text{for some } \mathbf{y} \in I^J\} \\ &= \{\mathbf{x}^{J^c} \in I^{J^c} | x_k < p_k, \forall k \in J^c\} \\ &= \{\mathbf{x}^{J^c} \in I^{J^c} | 0 \leq x_k < p_k, \forall k \in J^c\} \end{aligned} \tag{7}$$

which implies:

Lemma 17

$$A_J^{J^c}(\varepsilon) = A_J^{J^c} \tag{8}$$

Lemma 18

$$A_J(\varepsilon) = A_J^J(\varepsilon) \times A_J^{J^c} \tag{9}$$

Therefore by mutual independence among all x_j's ($j \in N$) (implied by uniform distribution), we have

$$M(A_J \backslash C_J) = M(A_J(\varepsilon)) = M^J(A_J^J(\varepsilon)) \cdot M^{J^c}(A_J^{J^c}) \tag{10}$$

Note: $M^J(\cdot)$ and $M^{J^c}(\cdot)$ are the marginal measures of M in dimensions J and dimensions J^c, respectively.

Now we need to find $M^J(A_J^J(\varepsilon))$ and $M^{J^c}(A_J^{J^c})$.

By (6) and $f(\mathbf{x}) = 1, \forall \mathbf{x} \in I^n$, we have

$$M^J(A_J^J(\varepsilon)) = \int_{A_J^J(\varepsilon)} d\mathbf{x}^J = \frac{\varepsilon^{|J|}}{|J|!} \tag{11}$$

By (7) we have

$$M^{J^c}(A_J^{J^c}(\varepsilon)) = \int_{A_J^{J^c}(\varepsilon)} d\mathbf{x}^{J^c} = \prod_{k \in J^c} p_k \tag{12}$$

Finally, putting (11) and (12) together, we have

$$M(A_J(\varepsilon)) = \frac{\prod_{k\in J^c} p_k}{|J|!} \cdot \varepsilon^{|J|} \tag{13}$$

where the first part $\frac{\prod_{k\in J^c} p_k}{|J|!} \equiv c(\mathbf{P})$ is a function of $\mathbf{P}$ only (more precisely it is a function of $\mathbf{P}^{J^c}$ only), and does not depend on ε. ■

Comment: By (7) we know $M^{J^c}(A_J^{J^c}(\varepsilon))$ only depends on $\mathbf{P}$ and distribution f, but does not depend on ε. By (6) we know each of $A_J^J(\varepsilon)$'s "$|J|$ sides" has "length" ε, which means $M^J(A_J^J(\varepsilon))$ will be proportional to $\varepsilon^{|J|}$. Therefore $M(A_J(\varepsilon))$ is also proportional to $\varepsilon^{|J|}$.

Theorem 2 *Consider the price schedules and allocations defined in (3). For* ***any general*** *f satisfying Assumption 2, we have*

(i) $\frac{\partial M(A_J\backslash C_J)}{\partial\varepsilon}|_{\varepsilon=0} \geq 0, \forall J$ *such that* $|J| = 1$;
(ii) $\frac{\partial M(A_J\backslash C_J)}{\partial\varepsilon}|_{\varepsilon=0} = 0, \forall J$ *such that* $|J| > 1$ (**Two-Part-Tariff Effect**).

Intuition: Similar to the results shown in Theorem 1, even when distribution f does not satisfy independence, when ε is very small, we can still think of $M(A_J(\varepsilon))$ as "proportional" to $\varepsilon^{|J|}$, and thus its first order derivative with respect to ε would be "proportional" to $\varepsilon^{|J|-1}$, which goes to 0 as $\varepsilon \to 0$ unless $|J| = 1$.

Proof. The expressions (5), (6) and (7) derived above hold for any general distribution f that satisfies Assumption 2. Although property (10) requires mutual independence among x_j's, we do not really need it here as all we care about is the first order derivative of $M(A_J(\varepsilon))$, not $M(A_J(\varepsilon))$ itself.

Part (i): when $|J| = 1$, i.e. $J = \{j\}$, $\forall j \in N$.

By (5) we know

$$\begin{aligned} A_j(\varepsilon) &= \{\mathbf{x} \in I^n | p_j \leq x_j < p_j + \varepsilon; x_k < p_k, \forall k \in j^c\} \\ &= \times_{k\in j^c}\{x_k \in I^k | 0 \leq x_k < p_k\} \times \{x_j \in I^j | p_j \leq x_j < p_j + \varepsilon\} \end{aligned}$$

where by (8) and (7) we know that $A_j^{j^c}(\varepsilon) = A_j^{j^c} = \times_{k\in j^c}\{x_k \in I^k | 0 \leq x_k < p_k\}$

Then by (9) we have

$$M(A_j(\varepsilon)) = \int_{p_j}^{p_j+\varepsilon} [\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}] dx_j$$

The integral in the brackets is a function of x_j only, which we define as

$$W_j(x_j) \equiv \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}$$

And rewrite

$$M(A_j(\varepsilon)) = \int_{p_j}^{p_j+\varepsilon} W_j(x_j)dx_j$$

Therefore

$$\frac{\partial M(A_j(\varepsilon))}{\partial \varepsilon}|_{\varepsilon=0} = W_j(p_j+\varepsilon)|_{\varepsilon=0} = W_j(p_j) = \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j)d\mathbf{x}^{j^c} \geq 0$$

Notice that the last inequality above will be strict if $\mathbf{P}$ in (3) satisfies $p_j > 0\ \forall j \in N$.

Part (ii): when $|J| > 1$. Since all n dimensions are "symmetric", without loss of generality, consider $J = \{1, 2, ..., |J|\}$.

By (8) and (9) we have

$$M(A_J(\varepsilon)) = \int_{A_J(\varepsilon)} f(\mathbf{x})d\mathbf{x} = \int_{A_J^{J^c}} [\int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J]d\mathbf{x}^{J^c}$$

Since $A_J^{J^c}$ does not depend on ε (by (8) and (7)), we have

$$\frac{\partial M(A_J(\varepsilon))}{\partial \varepsilon} = \int_{A_J^{J^c}} [\frac{\partial}{\partial \varepsilon} \int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J]d\mathbf{x}^{J^c} \tag{14}$$

Now we focus on $\frac{\partial}{\partial \varepsilon} \int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J$.

First notice that by (6) above we know

$$A_J^J(\varepsilon) = \{\mathbf{x}^J \in I^J | x_j \geq p_j, \forall j \in J; \sum_{j\in J} x_j < \sum_{j \in J} p_j + \varepsilon\}$$

In the following expression we write out $\int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J$ in all $|J|$ dimensions, in ascending order of product indices from inside outwards.

$$\begin{aligned} &\int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J \\ = &\int_{p_{|J|}}^{\varepsilon+p_{|J|}} \int_{p_{|J|-1}}^{\varepsilon+p_{|J|-1}+p_{|J|}-x_{|J|}} ... \int_{p_k}^{\varepsilon+\sum\limits_{j\geq k} p_j - \sum\limits_{j>k} x_j} ... \int_{p_1}^{\varepsilon+\sum\limits_{j\geq 1} p_j - \sum\limits_{j>1} x_j} f(\mathbf{x})dx_1...dx_k...dx_{|J|-1}dx_{|J|} \end{aligned}$$

Notice because $|J| > 1$, this expression will have at least two "layers". We focus on the first (outmost) layer. Denote all the parts inside the first layer of integration as

$$V(|J|-1, \mathbf{P}, \varepsilon, x_{|J|}) \equiv \int_{p_{|J|-1}}^{\varepsilon+p_{|J|-1}+p_{|J|}-x_{|J|}} ... \int_{p_k}^{\varepsilon+\sum\limits_{j\geq k} p_j - \sum\limits_{j>k} x_j} ... \int_{p_1}^{\varepsilon+\sum\limits_{j\geq 1} p_j - \sum\limits_{j>1} x_j} f(\mathbf{x})dx_1...dx_k...dx_{|J|-1} \tag{15}$$

Therefore we can rewrite $\int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J) d\mathbf{x}^J$ as

$$\int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J) d\mathbf{x}^J = \int_{p_{|J|}}^{\varepsilon + p_{|J|}} V(|J| - 1, \mathbf{P}, \varepsilon, x_{|J|}) dx_{|J|}$$

Therefore

$$\begin{aligned} & \tfrac{\partial}{\partial \varepsilon} \textstyle\int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J) d\mathbf{x}^J \\ = \quad & V(|J| - 1, \mathbf{P}, \varepsilon, x_{|J|} = \varepsilon + p_{|J|}) + \textstyle\int_{p_{|J|}}^{\varepsilon + p_{|J|}} \tfrac{\partial}{\partial \varepsilon} V(|J| - 1, \mathbf{P}, \varepsilon, x_{|J|}) dx_{|J|} \end{aligned} \tag{16}$$

Now examine the first part of (16), $V(|J| - 1, \mathbf{P}, \varepsilon, x_{|J|} = \varepsilon + p_{|J|})$, which is found by letting $x_{|J|} = \varepsilon + p_{|J|}$ in (15). We only need to focus on the upper bound of integration in (15), which is $\varepsilon + p_{|J|-1} + p_{|J|} - x_{|J|}$. We immediately see that it is equal to $p_{|J|-1}$ when $x_{|J|} = \varepsilon + p_{|J|}$. But this means the upper and lower bounds of integration of (15) are the same. Therefore

$$V(|J| - 1, \mathbf{P}, \varepsilon, x_{|J|} = \varepsilon + p_{|J|}) = 0$$

Now consider the second part of (16), which clearly equals 0 when $\varepsilon = 0$.

Therefore we have

$$\frac{\partial}{\partial \varepsilon} \int_{A_J^J(\varepsilon)} f(\mathbf{x}^{J^c}, \mathbf{x}^J) d\mathbf{x}^J |_{\varepsilon = 0} = 0$$

Substitute back to (14) and we are done. ■

Comment: Part (i) of Theorem 2 says that, imposing a small extra membership fee on top of an additive price schedule will cause a first-order decrease in the demand for single-product bundles. Part (ii) says that such a price manipulation has no first-order impact on the demand for all multi-product bundles (consisting of two or more products). We call part (ii) the two-part-tariff effect as it is crucial for the profitability of two-part tariffs.

When the firm charges everyone an extra membership fee, part (ii) implies that this will lead to a pure gain in profit from all multi-product consumers, as their demand does not decrease as a result. This gives rise to the possibility of higher overall profit for the firm. In the next section we discuss when this gain will dominate the loss from the decreased demand for single products.

4 Two-Part Tariffs vs. Separate Pricing

4.1 Separate Pricing

In Definition 2, we have defined additivity to be a synonym of separate pricing to capture the idea that a separate pricing strategy does not involve "funny" manipulation of the prices of different combinations of products.

Note that, due to additivity, a separate pricing strategy $\mathbf{P} = \{p_J\}_{J \subset N}$ can also be written as $\mathbf{P} = \{p_j\}_{j \in N}$ which only lists the prices of single products, as they uniquely and completely determine all the other prices in schedule $\mathbf{P}$.

Among all the separate pricing strategies, the following two kinds present interesting benchmarks.

Definition 19 (Monopoly Separate Pricing/MSP) *A price schedule $\mathbf{P} = \{p_J\}_{J \subset N}$ is a monopoly separate pricing strategy if it is additive and $\forall j \in N$,*

$$p_j = \arg\max_{p'_j \geq 0} p'_j \cdot \Pr[x_j \geq p'_j] \tag{17}$$

Intuition: A monopoly separate pricing strategy $\mathbf{P}$ is an additive price schedule comprised of all the "individually optimal" prices that a single-product monopolist would charge (i.e. the monopolist sets the optimal price for each product **irrespective** of all the other products and their prices).

Lemma 20 *There exists monopoly separate pricing strategy $\mathbf{P} = \{p_J\}_{J \subset N}$ such that $p_j > 0$ $\forall j \in N$ which yields strictly positive profit.*

Proof. This result is implied by Assumption 2. ■

Definition 21 (Contingent Separate Pricing/CSP) *A price schedule $\mathbf{P} = \{p_J\}_{J \subset N}$ is a contingent separate pricing strategy if it is additive and $\forall j \in N$,*

$$p_j = \arg\max_{p'_j \geq 0} p'_j \cdot M(A_j(p'_j, \mathbf{P}^{j^c})) \tag{18}$$

Intuition: A contingent separate pricing strategy $\mathbf{P}$ is also additive, while each of its component prices p_j is optimal for single product j **given** the prices for all the other products ($\mathbf{P}^{j^c}$). Therefore, each p_j is actually a function of $\mathbf{P}^{j^c}$, say $p_j = p_j(\mathbf{P}^{j^c})$, and the whole profile $\mathbf{P}$ is the "solution" of a system of n such equations. Of course, such equation systems may not always have a solution, depending on distribution f.

Lemma 22 *When all x_j's ($j \in N$) are mutually* ***independent****, any MSP strategy is also a CSP strategy, and any strictly none-zero CSP strategy (i.e. all its component prices are positive except for $p_\varnothing$) is also an MSP strategy.*

Proof. First consider MSP $\mathbf{T} = (t_J)_{J \subset N}$. $\forall j \in N$, denote F_j and f_j the marginal distribution and density of x_j, respectively. $\mathbf{T}$ must satisfy (17), that is $\forall j \in N$

$$t_j = \arg\max_{t'_j \geq 0} t'_j \cdot (1 - F_j(t'_j))$$

And the first order condition is

$$1 - F_j(t_j) = t_j f_j(t_j) \tag{19}$$

Second consider CSP $\mathbf{P} = (p_J)_{J \subset N}$. Note by additivity $\mathbf{P}$ must satisfy (1) of Lemma 8, that is $\forall j \in N$

$$A_j = \{\mathbf{x} \in I^n | x_j \geq p_j; x_k < p_k, \forall k \in j^c\}$$

And therefore

$$M(A_j) = \int_{p_j}^1 [\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}] dx_j$$

Since $\mathbf{P}$ is CSP, it must satisfy (18), whose first order condition for is

$$\int_{p_j}^1 [\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}] dx_j = p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c} \tag{20}$$

When all x_j's ($j \in N$) are mutually independent, we know $f(\mathbf{x}) = \prod_{j \in N} f_j(x_j)$ and therefore the condition above reduce to

$$\prod_{k \neq j} F_k(p_k) \cdot [1 - F_j(p_j) - p_j \cdot f_j(p_j)] = 0$$

which in turn reduce to $1 - F_j(p_j) = p_j \cdot f_j(p_j)$ as $\prod_{k \neq j} F_k(p_k) > 0$ (since $p_k > 0 \ \forall k \neq \varnothing$), which is exactly the same condition as (19). ■

Lemma 23 *When all x_j's ($j \in N$) are mutually* ***independent****, the optimal separate pricing strategy is both an MSP and a CSP.*

Proof. Suppose $\mathbf{P} = \{p_J\}_{J \subset N}$ is the optimal separate pricing strategy, then it must be additive. Thus by (1) of Lemma 8, we have

$$A_J = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k < p_k, \forall k \in J^c\}$$

Since all x_j's ($j \in N$) are mutually independent, we have

$$M(A_J) = \prod_{j \in J} (1 - F_j(p_j)) \cdot \prod_{k \in J^c} F_k(p_k) \tag{21}$$

Now define profit function under $\mathbf{P}$:

$$\pi(\mathbf{P}) \equiv \sum_{J \subset N} p_J \cdot M(A_J) \tag{22}$$

By additivity of $\mathbf{P}$, we have

$$\pi(\mathbf{P}) = \sum_{J \subset N} (\sum_{j \in J} p_j) \cdot M(A_J)$$

Substituting (21), we can reduce the profit function to

$$\pi(\mathbf{P}) = \sum_{j \in N} p_j \cdot (1 - F_j(p_j))$$

Since $\mathbf{P}$ is optimal, and $\mathbf{P} = \{p_j\}_{j \in N}$, we must have

$$\begin{aligned} \mathbf{P} &= \arg \max_{\{p'_j \geq 0\}_{j \in N}} \sum_{j \in N} p'_j \cdot (1 - F_j(p'_j)) \\ &= \{\arg \max_{p'_j \geq 0} p'_j \cdot (1 - F_j(p'_j))\}_{j \in N} \end{aligned}$$

which is exactly (17). Therefore $\mathbf{P}$ is MSP.

And by Lemma 22, we know $\mathbf{P}$ is also CSP. ∎

Theorem 3 *No contingent separate pricing strategy is optimal. Any contingent separate pricing strategy* $\mathbf{P}$ *is always strictly dominated by the two-part tariff* $\mathbf{Q}$ *defined in (3) using* $\mathbf{P}$*.*

Proof. Consider the price schedules and allocations defined in (3). Our strategy is to find first the difference in profits from $\mathbf{P}$ and $\mathbf{Q}$, and then show that when $\varepsilon \to 0$, $\mathbf{Q}$ yields strictly higher profit than $\mathbf{P}$ if $\mathbf{P}$ is a contingent separate pricing strategy.

First define profit functions:

$$\begin{aligned} \pi(\mathbf{P}) &\equiv \sum_{J \subset N} p_J \cdot M(A_J) \\ \pi(\mathbf{Q}) &\equiv \sum_{J \subset N} q_J \cdot M(C_J) \end{aligned}$$

Notice by (3) and (4), we have

$$\begin{aligned} \Delta\pi &\equiv \pi(\mathbf{Q}) - \pi(\mathbf{P}) = \sum_{J \subset N} q_J \cdot M(C_J) - \sum_{J \subset N} p_J \cdot M(A_J) \\ &= \sum_{J \neq \varnothing} \{(p_J + \varepsilon) \cdot [M(A_J) - M(A_J(\varepsilon))] - p_J \cdot M(A_J)\} \\ &= \varepsilon \cdot \sum_{J \neq \varnothing} M(A_J) - \sum_{J \neq \varnothing} (p_J + \varepsilon) \cdot M(A_J(\varepsilon)) \end{aligned}$$

Notice by definition $M(A_J)$ does not depend on ε. Therefore

$$\begin{aligned}\frac{\partial\Delta\pi}{\partial\varepsilon} &= \sum_{J\neq\varnothing} M(A_J) - \sum_{J\neq\varnothing}[M(A_J(\varepsilon)) + (p_J{+}\varepsilon)\cdot\frac{\partial M(A_J(\varepsilon))}{\partial\varepsilon}] \\ &= \sum_{J\neq\varnothing} M(A_J) - \sum_{J\neq\varnothing} M(A_J(\varepsilon)) - \sum_{J\neq\varnothing}(p_J{+}\varepsilon)\cdot\frac{\partial M(A_J(\varepsilon))}{\partial\varepsilon}\end{aligned}$$

Thus

$$\frac{\partial\Delta\pi}{\partial\varepsilon}|_{\varepsilon=0} = \sum_{J\neq\varnothing} M(A_J) - \sum_{J\neq\varnothing}(p_J)\cdot\frac{\partial M(A_J(\varepsilon))}{\partial\varepsilon}|_{\varepsilon=0} \tag{23}$$

Now we study $\frac{\partial M(A_J(\varepsilon))}{\partial\varepsilon}|_{\varepsilon=0}$.
When $|J|>1$, by Theorem 2 we know $\frac{\partial M(A_J(\varepsilon))}{\partial\varepsilon}|_{\varepsilon=0} = -\frac{\partial M(A_J\backslash C_J)}{\partial\varepsilon}|_{\varepsilon=0} = 0$.
When $|J|=1$, i.e. $J=\{j\}, j\in N$, we first focus on $M(A_j(\varepsilon))$. By (9) we know

$$\begin{aligned}A_j(\varepsilon) &= \{\mathbf{x}\in I^n|p_j\leq x_j{<}p_j+\varepsilon; x_k<p_k, \forall k\in j^c\} \\ &= \times_{k\in j^c}\{x_k\in I^k|0\leq x_k<p_k\}\times\{x_j\in I^j|p_j\leq x_j{<}p_j\}\end{aligned}$$

where by (8) we know that $A_j^{j^c}(\varepsilon) = A_j^{j^c} = \times_{k\in j^c}\{x_k\in I^k|0\leq x_k<p_k\}$
Therefore

$$M(A_j(\varepsilon)) = \int_{p_j}^{p_j+\varepsilon}[\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j)d\mathbf{x}^{j^c}]dx_j$$

The integral in the brackets only depends on x_j, which we define as

$$W_j(x_j)\equiv\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j)d\mathbf{x}^{j^c}$$

And rewrite

$$M(A_j(\varepsilon)) = \int_{p_j}^{p_j+\varepsilon} W_j(x_j)dx_j$$

Therefore

$$\frac{\partial M(A_j(\varepsilon))}{\partial\varepsilon}|_{\varepsilon=0} = W_j(p_j+\varepsilon)|_{\varepsilon=0} = W_j(p_j)$$

Substituting in (23) we get

$$\begin{aligned}\tfrac{\partial\Delta\pi}{\partial\varepsilon}|_{\varepsilon=0} = &\ \textstyle\sum_{J\neq\varnothing} M(A_J) - \sum_{j=1}^{n} p_j\cdot W_j(p_j) \\ = &\ \textstyle\sum_{j=1}^{n}[M(A_j) - p_j\cdot W_j(p_j)] + \sum_{J\subset N,|J|>1} M(A_J)\end{aligned} \tag{24}$$

Notice that our proof up until this point applies to all general price schedules that satisfies (3).

Now we use the fact that $\mathbf{P}$ is CSP to show that $M(A_j) = p_j\cdot W_j(p_j)$. To see this, notice that the first order condition for (18) is

$$M(A_j) + p_j \cdot \frac{\partial M(A_j)}{\partial p_j} = 0$$

where $M(A_j) = \int_{p_j}^1 [\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}] dx_j = \int_{p_j}^1 W_j(x_j) dx_j$
thus $\frac{\partial M(A_j)}{\partial p_j} = -W_j(p_j)$ and therefore $M(A_j) - p_j \cdot W_j(p_j) = 0$, $\forall j \in N$.
Therefore we have

$$\frac{\partial \Delta\pi}{\partial \varepsilon}|_{\varepsilon=0} = \sum_{J,|J|>1} M(A_J) > 0$$

where the last strict inequality is because **P** is CSP. This implies **Q** yields strictly higher profit than **P**. ■

Theorem 4 *When all x_j's ($j \in N$) are mutually **independent**, **any** separate pricing strategy* **P** *(including MSP and CSP) is always strictly dominated by the two-part tariff* **Q** *defined in (3) using* **P**.

Proof. We only need to show that **Q** strictly dominates the optimal separate pricing strategy.

Suppose **P** is the optimal separate pricing strategy, then by Lemma 23 we know **P** is CSP, and therefore from Theorem 3 we immediately know **P** is strictly dominated by **Q**. ■

Interpretation: Theorem 4 says that*, under independence, imposing an extra membership fee on top of any separate pricing strategy P will strictly increase profit* (that is, raising all the prices in P except for the price of the empty bundle by a same small amount $\varepsilon > 0$).

Theorem 5 *With **any general** f satisfying Assumption 2, **any** separate pricing strategy* **P** *(including MSP and CSP) is strictly dominated by the two-part tariff* **Q** *defined in (3) using* **P** *if the following condition holds at* **P**

$$\sum_{j=1}^{n} [M(A_j) - p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}] + \sum_{J \subset N, |J|>1} M(A_J) > 0 \tag{25}$$

where A_J is defined in (1).

Proof. Consider the price schedules defined in (3). We need to show that when condition (25) holds, **Q** yields strictly higher profit than **P**.

In exactly the same way as in the proof of Theorem 3 we can get result (24). Substitute $W_j(p_j) = \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}$ in (24) and we see that condition (25) is exactly $\frac{\partial \Delta\pi}{\partial \varepsilon}|_{\varepsilon=0} > 0$. Therefore **P** is strictly dominated by **Q** when condition (25) holds. ■

Intuition:

As we have discussed in the comment of Theorem 2, imposing a small extra membership fee on top of a separate pricing strategy will only decrease the demand for single-product bundles, but has no first-order impact on the demand for all multi-product bundles. Since the firm charges everyone an extra membership fee, it gains from each and every one of multi-product consumers. This gain is represented by the term $\sum_{J\subset N,|J|>1} M(A_J)$ in condition (25) (which is exactly the "number" of all multi-product consumers). From single-product consumers, the firm also charges a higher price, but it also loses some demand. The net effect from single-products is represented by the term $\sum_{j=1}^{n}[M(A_j) - p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}]$, which may be positive or negative, depending on the distribution of valuations. The overall impact on profit from the whole market is therefore captured by the left-hand side of condition (25).

Comment:

Theorem 5 generalizes Proposition 1 of McAfee, McMillan and Whinston (1989) to the multiproduct case. The latter addresses the case of two products and provides a condition for *mixed bundling* to strictly dominate separate pricing. It can be shown that when $n = 2$, our condition (25) reduces to their condition (1).

As we have discussed in the comment of the two-part tariff defined in (3), the two-part tariff $\mathbf{Q}$ can also be seen as a particular way of mixed bundling, where it is as if a consumer of any bundle $J \subset N$ gets a "bundle discount" of $(|J| - 1) \cdot \varepsilon$.

5 Conclusion

Two-part tariffs are prevalent in life. In many cases, they involve more than one product provided by the same firm, such as the landline telephone and broadband services one gets from a telecommunication company, or one's bank account through which other services such as credit card, mortgage and travel insurance are also provided.

In this paper we have shown that these two-part tariffs can be understood as a means of price discrimination by a multiproduct firm. This new explanation has nothing to do with production cost or efficiency, but only requires two or more dimensions of consumer types (i.e. two or more products).

From this new perspective, we argue that two-part tariffs should be subject to the same regulatory scrutiny as other discriminatory pricing strategies.

References

[1] Armstrong, M. (1996): “Multiproduct Nonlinear Pricing”, *Econometrica*, 64(1), pp. 51-75.

[2] Armstrong, M. (1999): "Price Discrimination by a Many-Product Firm", *Review of Economic Studies*, 66(1), pp. 151-168.

[3] Gao, M. (2009): "When to Allow Buyers to Sell? Bundling in Mixed Two-Sided Markets", working paper.

[4] Long, J.B. Jr. (1984): "Comments on Gaussian Demand and Commodity Bundling", *Journal of Business*, 57(1), Part 2, pp. S235-S246.

[5] Manelli, A. M. and Vincent, D. (2006): "Bundling as an optimal selling mechanism for a multiple-good monopolist", Journal of Economic Theory, 127, pp. 1-35.

[6] McAfee, R. P. and McMillan, J. (1988): "Multidimensional Incentive Compatibility and Mechanism Design", *Journal of Economic Theory*, 46, pp. 335-354.

[7] McAfee, R. P., McMillan, J. and Whinston, M. D. (1989): "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values", *Quarterly Journal of Economics*, 104, pp. 371-383.