

Menu Implementability and Practical Pricing Schemes

Adam Chi Leung Wong*

February 20, 2012

First draft: August 24, 2011

Abstract

In the context of monopolistic nonlinear pricing, we compare the maximum profits of bundling, incremental discounts, and all-units discounts. When the number of pricing blocks is unrestricted, incremental discounts perform weakly the worst. However, if the performance of incremental discounts is not strictly worse when the number of blocks is unrestricted, then it performs the best when the number of blocks is restricted. It is because incremental discounts have the smallest "implementation power" and the largest "approximation power". These results are applications of our general theory of menu implementability, which characterizes the set of outcomes implementable by a menu of tariff options, where the set of admissible tariff options and the number of tariff options in the menu are pre-determined.

Keywords: Nonlinear pricing, Bundling, Incremental discounts, All-units discounts, Incentive-compatible mechanisms

JEL Classification Numbers: D42, D82, D86, L12.

1 Introduction

We revisit Maskin and Riley (1984) monopolistic nonlinear pricing problem, in which a monopolist faces heterogeneous consumers with one-dimensional continuous types,

*School of International Business Administration, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, China 200433, wongchileung@gmail.com.

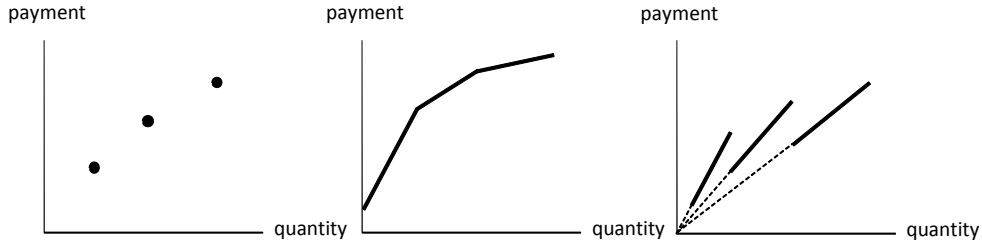


Figure 1: Bundle (left), Incremental discounts (middle) and All-units discounts (right)

and the consumers' types are private information. If the monopolist is free to adopt any pricing scheme to maximize profit, the optimal (or second best) solution is now well known. However, in this continuous type model, the optimal nonlinear pricing scheme is complicated, at least far more complicated than what we observe in reality. What if the monopolist has to use a pricing scheme that is in some "practical" form (e.g. offering a menu of two-part tariffs)? Given a practical form, what are the restrictions it put on the set of feasible outcomes? How should the monopolist choose among different practical forms of pricing schemes?

To tackle these issues, we need to be more precise about what forms of pricing schemes we consider practical. Our treatment to this has two levels. At the first level, the application part, we consider three forms of pricing schemes to be practical: *bundling (BD)* schemes, *incremental discounts (ID)* schemes, and *all-units discounts (AD)* schemes. These three forms are illustrated in Figure 1. Under a BD scheme, illustrated in the left panel, only several quantities, each associated with a gross price, are offered for consumers to choose. Under an ID scheme, illustrated in the middle panel, progressive discounts apply to incremental units when the order size exceeds certain thresholds. We also allow a fixed fee under an ID scheme.¹ Under an AD scheme, illustrated in the right panel, progressive discounts apply to all units when the order size exceeds certain thresholds. We also allow a minimum purchase under an AD scheme.² In our terminology, an AD scheme does not have a fixed fee; if it does (i.e. the vertical intercept of the right panel becomes non-zero), then we call it an *all-units discounts with fixed fee (ADF)* scheme. Notice that the three pricing

¹If one insists that ID has no fixed fee, we can mimic the fixed fee by letting the marginal price huge for the very first units.

²If one insists that AD has no minimum purchase, we can mimic the minimum purchase by letting the per-unit price huge for the order size below the first threshold.

schemes in Figure 1 have arguably the same level of complexity: each of them has three blocks, and can be characterized by six parameters. Indeed, under any of the three forms, a pricing scheme with n blocks requires $2n$ parameters to characterize. (Of course, ADF schemes require one more parameter for the fixed fee.) It then makes much sense to ask, with the same number of blocks, which of the three forms yields the highest profit for the monopolist.

The second level of our treatment is the general theory part. In this part we more generally deal with principal-agent model, where the meaning of practicality is flexible, and we develop a general theory that suffices to analyze any particular form of pricing schemes, provided that it is in the "menu class" that we explain below. In order to motivate our general theory, notice that each of the aforementioned three forms of pricing schemes can be regarded as offering a menu of simple tariff options for consumers to select. Offering a BD scheme with n blocks (or n points here) is equivalent to offering n quantity-payment bundles. Offering an ID scheme with n blocks is equivalent to offering n two-part tariffs.³ Offering an AD scheme with n blocks is equivalent to offering n "minimum purchase tariffs" (see Section 3). That makes it possible to build an elegant unifying framework to analyze all the three forms.

In the application part (Sections 2 – 4), we first characterize outcomes implementable by BD, ID and ADF schemes, given any number of blocks n . (Theorems 1, 2 and 3) Those characterizations are applications of the general theory part. Then we derive many results with our characterizations. First, when the number of blocks is unrestricted, i.e. $n = \infty$, every incentive-compatible and individually rational outcome is implementable by BD schemes, and by ADF schemes. On the other hand, AD schemes (without fixed fee) can implement at least as many outcomes as ID schemes do. In this sense we say that BD and ADF have the largest and ID has the smallest "implementation power" among the three forms. (Theorem 4) It follows that when the number of blocks is unrestricted, ADF can attain the Maskin-Riley second best monopoly profit, and ID can never perform better than AD. (Corollary 1)

Second, we derive the condition under which ID or AD can attain the second best profit, when the number of blocks is unrestricted. (Corollary 2) It amounts to derive the condition under which the Maskin-Riley second best outcome is implementable by ID or AD schemes.

³An ID pricing scheme with n blocks can be regarded as the lower envelope of n two-part tariffs.

Third, if ID can attain the second best profit when the number of blocks is unrestricted, then it yields strictly higher profit than BD and ADF when the number of blocks is restricted, i.e. $n < \infty$. Proving this result is harder because, given a finite number n of blocks, the sets of outcomes implementable by BD, ID and ADF are not comparable (i.e. do not contain one another). However, it can be done by showing that, under the premise and given any outcome implementable by BD or ADF scheme with n blocks, there exists an outcome implementable by ID scheme with n blocks that better approximate the Maskin-Riley second best solution. In this sense, we say that ID has larger "approximation power" than BD and ADF. (Theorem 5) The variations of implementation power and approximation power are closely related to trade-offs between control and flexibility, which we discuss in the concluding remarks.

The general theory part (Sections 5 and 6) generally analyzes forms of pricing schemes that can be described as a menu of tariff options, where the set of admissible tariff options \mathcal{C} and the number of tariff options n in the menu are pre-determined. We introduce the concept of *tariff single crossing property* and *tariff increasing differences* to characterize the set of outcomes implementable by such a menu given the number n and the set \mathcal{C} . (Theorems 6 – 8). Such outcomes are said to be *menu implementable* with respect to (n, \mathcal{C}) . In general, restricting to offering a menu makes the issue of incentive compatibility more severe. Hence menu implementability is strictly stronger than the ordinary incentive compatibility.

Our application part is related to the literature of monopolistic nonlinear pricing, which originated with Mussa and Rosen (1978) and Maskin and Riley (1984). Few papers are concerned with comparison among suboptimal forms of pricing schemes, although Kolay and Shaffer (2003) and Kolay, Shaffer, and Ordover (2004) are exceptions. Both of these two papers assume that the monopolistic seller has constant marginal cost and the privately informed buyers have only two types. Kolay and Shaffer (2003) show that the best BD scheme yields strictly higher profit than the best ID scheme. Kolay, Shaffer, and Ordover (2004) show that the best AD scheme yields higher profit than the best ID scheme. While their results are in spirit consistent with our Corollary 1, our analyses differ from theirs most significantly in that we assume a continuum of buyers' types but restrict the complexity level (i.e. number of blocks) of pricing schemes.

Although the use of AD in intermediate-goods markets is common,⁴ there is little

⁴According to Kolay, Shaffer, and Ordover (2004), AD is used by Coca-Cola, Irish Sugar British

theoretical analysis on it in the literature, except Kolay, Shaffer, and Ordover (2004). Why does a seller use AD rather than ID? It has been informally argued that the use of AD is anticompetitive or exclusionary (to exclude entry of competitors, or induce downstream retailers to promote the products at the expense of other substitute products).⁵ So one might wonder whether a monopolist without fear of competition would find AD superior. The implication of our results on this has two sides. First, in principle both AD and ID could perform better for a monopolist without fear of competition. But second, under certain conditions (which could be plausible in certain contexts but not in others) ID must outperform AD for such a monopolist.⁶

Our general theory part is related to the literature of principal-agent models and mechanism design. Offering simple menus as a practical scheme is relevant in other principal-agent contexts (e.g. the ones in Laffont and Martimort (2002)), where our general theory of menu implementability applies equally well. We list some works on this line. In the context of procurement contracting, Rogerson (2003) considers "Fixed Price Cost Reimbursement (FPCR) menus", that is, two-item menus where one item is a cost-reimbursement contract and the other item is a fixed-price contract, of which the principal allows the agent to pick one. He shows that, if the agent's utility is quadratic and the agent's type is distributed uniformly, then "the optimal FPCR menu always captures at least three-quarters of the gain that the optimal complex menu achieves". Chu and Sappington (2007) relax the assumption of uniform distribution, and show that a menu of two options, namely, a cost-reimbursement contract and a linear cost sharing contract, can always secure at least 73 percent of the gain. In the context of nonlinear pricing, Wilson (1993) claims that the loss due to limiting the number n of two-part tariffs is of order $1/n^2$. Bergemann, Shen, Xu, and Yeh (2011) consider Mussa and Rosen (1978) quality differentiation setting and show under "linear-quadratic specification" that the loss resulting from the usage of a finite n -class menu is of order $1/n^2$. Wong (2009) also considers Mussa and Rosen (1978) setting and shows that the marginal gain of increasing the number n is diminishing, and of order $1/n^3$. Miravete (2007) uses a large sample of independent

Airways, and Michelin.

⁵See, for example, Tom, Balto, and Averitt (1999).

⁶For example, ID must outperform AD if marginal cost is constant, the hazard rate of types' distribution is nondecreasing, and consumers' utility takes the form $\theta s(q) - t$ where θ is type, t is payment, and s is a concave function of quantity q . But AD could outperform ID under the above conditions except that marginal cost is increasing. Consider the specification in Example 1 with $1/2 < \alpha \leq \sqrt{2}/2$ and large enough n .

cellular telephone markets to structurally estimate a monopolistic nonlinear pricing model. His estimates suggests that "firms should only offer few tariff options if the product development costs of designing them are non-negligible."

The rest of the paper is organized as follows. Sections 2 – 4 are the application part. Section 2 formulates the monopolistic nonlinear pricing problem and presents some basic facts about it. Section 3 characterizes the set of outcomes implementable by BD, ID and ADF schemes. Section 4 compares the performances of these forms of pricing schemes. Sections 5 – 6 are the general theory part. Section 5 generalizes the model in Section 2 into one that can be adapted to many other principal-agent settings. Section 6 presents our characterization of menu implementability. Section 7 concludes. The proofs not in the main text are in Appendix.

2 Monopolistic nonlinear pricing problem

Consider a monopolistic nonlinear pricing problem, in which each consumer's utility function is

$$S(q, \theta) - t$$

where $S : \mathbb{R}_+ \times \Theta \rightarrow \mathbb{R}$ is the consumer's gross utility function and $t \in \mathbb{R}$ denotes the payment from the consumer to the monopolist. The argument q denotes the quantity consumed; the argument θ denotes the consumer's type (or preference parameter), whose domain Θ is an interval $[\underline{\theta}, \bar{\theta}]$. A consumer's type is her private information. The monopolist only knows the cumulative distribution function F of consumers' types, which has a positive density f on the support Θ . Each consumer has an outside option $(q, t) = (0, 0)$, i.e. buying nothing and paying nothing.

Given q and t , the monopolist's ex post (per-customer) profit is given by $t - c(q)$, where $c : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the monopolist's cost function. If each consumer of type θ buys quantity $Q(\theta)$ and pays $T(\theta)$, then the monopolist's ex ante (per-customer) profit is

$$\int_{\underline{\theta}}^{\bar{\theta}} [T(\theta) - c(Q(\theta))] dF(\theta). \quad (1)$$

Assumption 1 *S and c are twice differentiable. $S(0, \theta) = 0$ for all $\theta \in \Theta$. If $q > 0$, then $S_\theta(q, \theta) > 0$. Strict Spence-Mirrlees single crossing condition: $S_{q\theta}(q, \theta) > 0$.*⁷

⁷In particular, the gross utility S is not required to be increasing or concave in quantity q . Hence,

From the revelation principle, it is without loss of generality to restrict attention to outcomes (which can be thought of as direct revelation mechanisms) that are incentive-compatible and individually rational. Formally, an *outcome*, written as (Q, T) , is a pair of functions $Q : \Theta \rightarrow \mathbb{R}_+$ and $T : \Theta \rightarrow \mathbb{R}$. Q is called the quantity function and T the payment function of (Q, T) . An outcome (Q, T) is said to be *incentive-compatible (IC)* if for any $\theta, \theta' \in \Theta$, we have

$$S(Q(\theta), \theta) - T(\theta) \geq S(Q(\theta'), \theta) - T(\theta').$$

An outcome (Q, T) is said to be *individually rational (IR)* if for any $\theta \in \Theta$, we have

$$U(\theta) \equiv S(Q(\theta), \theta) - T(\theta) \geq 0.$$

Fact 1 and Fact 2 below are well known and we state them without proof for future reference.

Fact 1 *An outcome (Q, T) is IC if and only if $Q(\cdot)$ is nondecreasing and $U(\theta_2) - U(\theta_1) = \int_{\theta_1}^{\theta_2} S_\theta(Q(x), x) dx$ for any $\theta_1, \theta_2 \in \Theta$.*

It is well known that the profit (1) can be rewritten using IC of (Q, T) as

$$\int_{\underline{\theta}}^{\bar{\theta}} H(Q(\theta), \theta) dF(\theta) - U(\underline{\theta}), \quad (2)$$

where H is the "virtual surplus function" defined as

$$H(q, \theta) \equiv S(q, \theta) - S_\theta(q, \theta) \frac{1 - F(\theta)}{f(\theta)} - c(q).$$

If an outcome (Q^*, T^*) maximizes the monopolist's profit (1) subject to IC and IR, we say (Q^*, T^*) is a *second best outcome*, Q^* is a *second best quantity function*, and the associated profit, denoted as Π^* , is the *second best profit*. We assume that a second best outcome exists.

the consumer can be for example interpreted as a down-stream retailer, and the gross utility function as the retailer's revenue function, as in e.g. Kolay, Shaffer, and Ordover (2004).

Fact 2 *The second best profit Π^* can be written as*

$$\Pi^* = \max_{Q(\cdot) \geq 0} \left\{ \int_{\underline{\theta}}^{\bar{\theta}} H(Q(\theta), \theta) dF(\theta) \right. \\ \left. \text{s.t. } Q(\cdot) \text{ is nondecreasing} \right\}. \quad (3)$$

An outcome (Q^, T^*) is second best if and only if $Q^*(\cdot)$ solves problem (3) and*

$$T^*(\theta) = S(Q^*(\theta), \theta) - \int_{\underline{\theta}}^{\theta} S_{\theta}(Q^*(x), x) dx. \quad (4)$$

3 Bundling, incremental discounts, and all-units discounts

If the monopolist is free to adopt any pricing scheme, then any IC and IR outcome is implementable and hence the monopolist maximum profit is the second best profit Π^* . This section studies the sets of implementable outcomes when the monopolist restricts itself to adopt bundling schemes, incremental discounts schemes, or all-units discounts schemes. The results in this and the next sections are proved (in Appendix) after our general theory of menu implementability (Section 6) is developed.

A *bundling (BD) scheme* is a menu of options offered by the monopolist, with each option composed of a purchase quantity $q \geq 0$ and a total payment $t \in \mathbb{R}$. Each consumer has to pick either one option in the menu, or the outside option. We say a BD scheme implements an outcome (Q, T) if there is a best response of consumers (which for each consumers' type assigns an option in the menu or the outside option to maximize consumers' utility) such that the outcome of this best response coincides with (Q, T) . If a BD scheme has at most n options (where $n \in \mathbb{Z}_+ \cup \{\infty\}$, i.e. n is a nonnegative integer or the infinity), then it is called an *n -BD scheme*.

Theorem 1 *An outcome (Q, T) can be implemented by some n -BD scheme if and only if it is IC and IR, and Q takes at most n values except 0.*

If an outcome can be implemented by some n -BD scheme, we may call it *n -BD-implementable*.

We turn to incremental discounts schemes. A *two-part tariff* characterized by $(p, \phi) \in \mathbb{R}_+ \times \mathbb{R}$ is a function that assigns each quantity $q \in \mathbb{R}_+$ the payment $pq + \phi$. Of course, p and ϕ are the marginal price and the fixed fee of this two-part tariff.

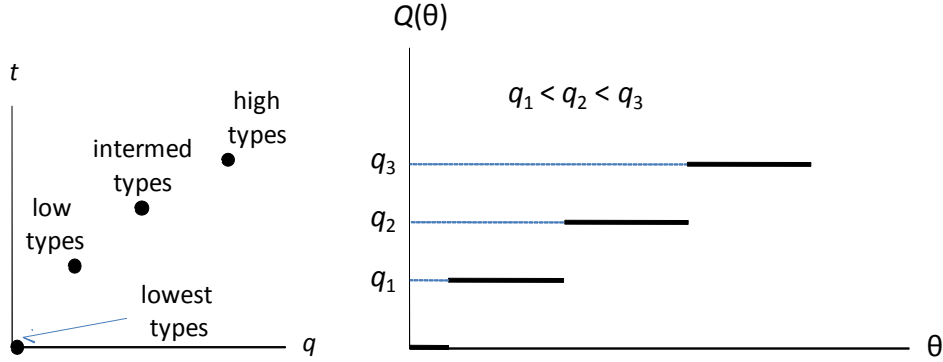


Figure 2: Quantity function implemented by a 3-BD scheme

An *incremental discounts (ID) scheme* is defined as a menu of two-part tariffs offered by the monopolist. Each consumer has to pick either a two-part tariff in the menu and a purchase quantity (and hence also a payment), or the outside option. We say an ID scheme implements an outcome (Q, T) if there is a best response of consumers (which for each consumers' type assigns a two-part tariff in the menu and a purchase quantity or the outside option to maximize consumers' utility) such that the outcome of this best response coincides with (Q, T) . If an ID scheme has at most n two-part tariffs (where $n \in \mathbb{Z}_+ \cup \{\infty\}$), then it is called an n -ID scheme.

Theorem 2 *An outcome (Q, T) can be implemented by some n -ID scheme if and only if it is IC and IR, and there exists a function $P : A \rightarrow \mathbb{R}_+$, where $A = \{\theta \in \Theta : (Q(\theta), T(\theta)) \neq (0, 0)\}$, such that*

1. $Q(\theta) \in \arg \max_{q \geq 0} \{S(q, \theta) - P(\theta)q\}$ for every $\theta \in A$,
2. P is nonincreasing, and
3. P takes at most n values.⁸

(Type θ consumers are meant to pick the outside option if $\theta \in \Theta \setminus A$, and pick the two-part tariff characterized by $(P(\theta), T(\theta) - P(\theta)Q(\theta))$ if $\theta \in A$.)

If an outcome can be implemented by some n -ID scheme, we may call it n -ID-implementable.

⁸Since $S(\cdot, \theta)$ is assumed to be differentiable, a necessary condition is that $S_q(Q(\theta), \theta)$ is nonincreasing and takes at most n values on $\{\theta \in \Theta : Q(\theta) > 0\}$.

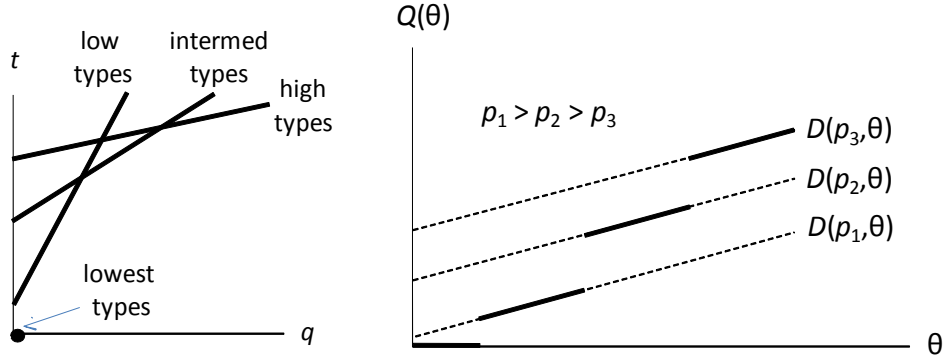


Figure 3: Quantity function implemented by a 3-ID scheme

We turn to all-units discounts schemes. A *minimum purchase tariff* characterized by $(m, p, \phi) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ is a function that assigns each quantity $q \geq m$ the payment $pq + \phi$. The parameters m , p and ϕ are called the minimum purchase, the marginal price and the fixed fee of this minimum purchase tariff. An *all-units discounts with fixed fee (ADF) scheme* is defined as a menu of minimum purchase tariffs offered by the monopolist, such that the fixed fees of those minimum purchase tariffs are the same. Each consumer has to pick either a minimum purchase tariff in the menu and a purchase quantity (and hence also a payment) that is no smaller than the associated minimum purchase, or the outside option. We say an ADF scheme implements an outcome (Q, T) if there is a best response of consumers such that the outcome of this best response coincides with (Q, T) . If an ADF scheme has zero fixed fee, we simply call it an *all-units discounts (AD) scheme*. If an ADF (or AD) scheme has at most n minimum purchase tariffs (where $n \in \mathbb{Z}_+ \cup \{\infty\}$), then it is called an *n-ADF (or n-AD) scheme*.

Theorem 3 *Given $\phi \in \mathbb{R}$, an outcome (Q, T) can be implemented by some n-ADF scheme with fixed fee ϕ if and only if it is IC and IR, and there exist functions $M : A \rightarrow \mathbb{R}_+$ and $P : A \rightarrow \mathbb{R}_+$, where $A = \{\theta \in \Theta : (Q(\theta), T(\theta)) \neq (0, 0)\}$, such that*

1. $T(\theta) = P(\theta)Q(\theta) + \phi$ for every $\theta \in A$,
2. $Q(\theta) \in \arg \max_{q \geq M(\theta)} \{S(q, \theta) - P(\theta)q\}$ for every $\theta \in A$,
3. M is nondecreasing and P is nonincreasing, and

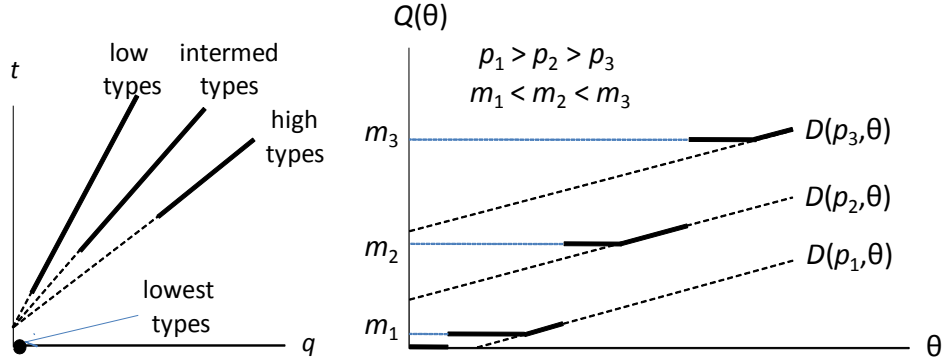


Figure 4: Quantity function implemented by a 3-ADF scheme

4. (M, P) takes at most n values.

(Type θ consumers are meant to pick the outside option if $\theta \in \Theta \setminus A$, and pick the minimum purchase tariff characterized by $(M(\theta), P(\theta), \phi)$ if $\theta \in A$.) (If $n = \infty$, the function M can be chosen as the restriction of Q on A .)

If an outcome can be implemented by some n -ADF (or AD) scheme, we may call it n -ADF-implementable (or n -AD-implementable).

Theorems 1 – 3 imply restrictions on outcomes that can be implemented by n -BD, n -ID and n -ADF (and hence n -AD) schemes. The left panels of Figures 2 – 4 illustrate how the options in the menu has to be distributed to different types; the right panels illustrate the patterns of the associated quantity functions, where $D(p, \theta)$ denotes the demand $\arg \max_{q \geq 0} \{S(q, \theta) - pq\}$. The following observations are important. Adopting an ID scheme, the induced quantity function cannot respond to type too little, because (i) within each block (corresponding to a two-part tariff), the quantity function has to follow a "type-demand curve", which is typically strictly increasing,⁹ and (ii) across blocks, the relevant "type-demand curve" can only shift up, since marginal price has to be nonincreasing in type. In contrast, adopting a BD scheme, the induced quantity function has to be flat within each block. (Of course the jumps have to be upward by incentive compatibility.) From this viewpoint, ADF (and also AD) stands between ID and BD. Adopting an ADF scheme, within a block, the induced quantity function can have both a flat portion when the minimum purchase

⁹Assumption 3 guarantees that any selection of demand correspondence with respect to type is nondecreasing.

is binding, and a "type-demand curve portion" when the minimum purchase is not binding.

Comparing Figures 2 – 4, one might suspect that, given the number of options n to be included in the menu, AD (even without fixed fee) can implement more general quantity functions than ID and BD can. But it is not true, because we are not totally free to choose the combination of the minimum purchases (m_i 's), the marginal prices (p_i 's) and the thresholds of AD schemes. The combination of those parameters has to make every marginal type between two blocks indifferent between picking the two minimum purchase tariffs corresponding to the two blocks. After all, under AD we have the same degrees of freedom, namely $2n$, as ID and BD. The comparison among the maximum profits of the three forms is studied in the next section.

4 Comparison among the three forms of pricing schemes

For each $n \in \mathbb{Z}_+ \cup \{\infty\}$, let Π_n^{BD} denote the maximum monopolist profit that can be made by offering some n -BD scheme. We define Π_n^{ID} , Π_n^{AD} and Π_n^{ADF} similarly. For these maximum profits to be well defined, we assume that corresponding optimal outcomes (i.e. outcomes that makes the highest profit among those implementable by the corresponding schemes) exist. Those optimal outcomes are called n -BD-optimal outcome, n -ID-optimal outcome, etc.; and those maximum profits are called n -BD-maximum profit, n -ID-maximum profit, etc. If $n = \infty$, we might simply call them BD-optimal outcome, ID-optimal outcome, BD-maximum profit, ID-maximum profit, etc.

From Theorem 1, any BD-optimal outcome is also a second best outcome, and Π_∞^{BD} is also the second best profit Π^* . From Theorems 1 – 3 one can immediately see that $\lim_{n \rightarrow \infty} \Pi_n^{BD} = \Pi_\infty^{BD} = \Pi^*$, $\lim_{n \rightarrow \infty} \Pi_n^{ID} = \Pi_\infty^{ID}$, $\lim_{n \rightarrow \infty} \Pi_n^{ADF} = \Pi_\infty^{ADF}$ and $\lim_{n \rightarrow \infty} \Pi_n^{AD} = \Pi_\infty^{AD}$.¹⁰

When the number of options in the menu is unrestricted (i.e. $n = \infty$), we can rank the maximum profits of BD, ID, ADF and AD because we can rank the corresponding sets of implementable outcomes. First of all, the set of outcomes implementable by BD schemes must be the largest: when the number of quantity-payment bundles is

¹⁰It is because a monotonic function (e.g. Q in Theorem 1, P in Theorem 2, M and P in Theorem 3) can be arbitrarily well approximated by a step function.

unrestricted, BD puts no restriction on implementable outcome except IC and IR. It follows that $\Pi^* = \Pi_\infty^{BD}$. The following results, which are implications of Theorems 2 and 3, are less obvious.

Theorem 4 *Let (Q, T) be an IC and IR outcome that satisfies $S(Q(\underline{\theta}), \underline{\theta}) - T(\underline{\theta}) = 0$. (a) (Q, T) can be implemented by some ADF scheme. (b) If (Q, T) can be implemented by some ID scheme, then it can also be implemented by some AD scheme.*

Part (a) of Theorem 4 says that it is without loss to use ADF schemes if we only consider outcomes with nonnegative payment. Indeed, any such outcome (Q, T) can be implemented by an ADF scheme as follows: Set a large transfer payment $-\phi$ to buyers (i.e. negative fixed fee ϕ) to induce consumers to participate; for every type θ with $Q(\theta) > 0$, put in the menu a minimum-purchase tariff with $Q(\theta)$ being its minimum purchase, and with some $P(\theta)$ being its marginal price such that the resulting gross payment is $T(\theta)$ (i.e. $P(\theta)Q(\theta) + \phi = T(\theta)$). Note that since the upfront transfer $-\phi$ is large, the marginal price $P(\theta)$ must be large, so that the minimum-purchase tariff characterized by $(Q(\theta), P(\theta), \phi)$ is effectively the single quantity-payment bundle $(Q(\theta), T(\theta))$. Therefore the above ADF scheme is effectively the BD scheme that implements (Q, T) .

Part (b) of Theorem 4 can be informally understood as follows. Under any ID scheme, the effective tariff function (i.e. minimum payment as a function of quantity) must be concave. In contrast, for an effective tariff function to be generated by an AD scheme, it only has to have nonincreasing average per-unit price. If we only consider tariff functions with nonnegative payment for zero order size (which is natural in this monopolist context and is formally guaranteed by the condition $S(Q(\underline{\theta}), \underline{\theta}) - T(\underline{\theta}) = 0$), then it is geometrically easy to see that concavity is strictly stronger than nonincreasing average. That is, when the number of blocks is unrestricted, AD can generate strictly more tariff functions than ID can.

An alternative way to understand part (b) of Theorem 4 is recalling the insights from Figures 2 – 4. Under ID, the induced quantity function cannot be too flat. When the number of blocks n tends to infinity, the induced quantity function can become smooth, but still cannot be flatter than type-demand curve. In contrast, under BD any nondecreasing quantity function can be induced when n tends to infinity. Under AD, which stands between BD and ID, induced quantity functions with portions

flatter than type-demand curve are possible. In this sense, we say ID has the smallest "implementation power" among the three forms, while BD has the largest.

Corollary 1 $\Pi^* = \Pi_\infty^{BD} = \Pi_\infty^{ADF} \geq \Pi_\infty^{AD} \geq \Pi_\infty^{ID}$.

Proof. We have already claimed that $\Pi^* = \Pi_\infty^{BD}$. Since AD schemes are special cases of ADF schemes, we also have $\Pi_\infty^{ADF} \geq \Pi_\infty^{AD}$.

If an outcome (Q, T) is second best, then it is IC and IR, and satisfies $S(Q(\underline{\theta}), \underline{\theta}) - T(\underline{\theta}) = 0$, then from part (a) of Theorem 4, (Q, T) can be implemented by some ADF scheme. Therefore $\Pi^* = \Pi_\infty^{ADF}$.

Now suppose that (Q, T) is ID-optimal. Then it is implementable by some ID scheme and hence IC and IR. It also satisfies $S(Q(\underline{\theta}), \underline{\theta}) - T(\underline{\theta}) = 0$, for if $S(Q(\underline{\theta}), \underline{\theta}) - T(\underline{\theta}) > 0$, the monopolist could increase profit by raising the fixed fees of all two-part tariffs. From part (b) of Theorem 4, (Q, T) can be implemented by some AD scheme. Therefore $\Pi_\infty^{AD} \geq \Pi_\infty^{ID}$. ■

Can ID or AD attain the second best profit? It amounts to check whether some second best outcome can be implemented by ID or AD scheme. In particular, if the second best outcome involves bunching, then typically it cannot be implemented by ID scheme because type-demand curves are typically strictly increasing.

Corollary 2 (a) $\Pi^* = \Pi_\infty^{ID}$ if and only if there exist some second best quantity function Q^* and some nonincreasing function $P^* : \{\theta \in \Theta : Q^*(\theta) > 0\} \rightarrow \mathbb{R}_+$ such that $Q^*(\theta) \in \arg \max_{q \geq 0} \{S(q, \theta) - P^*(\theta)q\}$ whenever $Q^*(\theta) > 0$.¹¹ (b) $\Pi^* = \Pi_\infty^{AD}$ if and only if there exist some second best outcome (Q^*, T^*) and some nonincreasing function $P^* : \{\theta \in \Theta : Q^*(\theta) > 0\} \rightarrow \mathbb{R}_+$ such that $Q^*(\theta) \in \arg \max_{q \geq Q^*(\theta)} \{S(q, \theta) - P^*(\theta)q\}$ and $T^*(\theta) = P^*(\theta)Q^*(\theta)$ whenever $Q^*(\theta) > 0$.

Proof. To see the sufficiency part of (a), suppose that Q^* and P^* satisfy the condition. Denote the domain of P^* as $A^* = \{\theta \in \Theta : Q^*(\theta) > 0\}$. Define $T^* : \Theta \rightarrow \mathbb{R}$ as in (4). It is straightforward to verify that (Q^*, T^*, A^*, P^*) satisfies conditions 1–2 in Theorem 2 and that $A^* = \{\theta \in \Theta : (Q^*(\theta), T^*(\theta)) \neq (0, 0)\}$. Therefore, the second best (and hence IC and IR) outcome (Q^*, T^*) can be implemented by ID scheme, so that $\Pi^* = \Pi_\infty^{ID}$. To see the necessity part of (a), suppose that $\Pi^* = \Pi_\infty^{ID}$. Then some second best outcome (Q^*, T^*) can be implemented by ID scheme. Then Theorem 2

¹¹Clearly, this function P^* , if exists, must be given by $P^*(\theta) = S_q(Q^*(\theta), \theta)$.

implies the condition. (Note that $(Q^*(\theta), T^*(\theta)) \neq (0, 0)$ is equivalent to $Q^*(\theta) > 0$, from (4).)

Part (b) can be proved similarly by applying Theorem 3, with defining $M^*(\theta) \equiv Q^*(\theta)$ for any $\theta \in A^* \equiv \{\theta \in \Theta : Q^*(\theta) > 0\}$. ■

Example 1 Suppose that $S(q, \theta) = \theta s(q)$ and θ is uniformly distributed on $[0, 1]$. Let $s(\cdot)$ and $c(\cdot)$ take the following forms:

$$s(q) = \frac{q^\alpha}{\alpha}, \quad c(q) = \frac{q^{\alpha+1}}{\alpha + 1},$$

where $\alpha \in (0, 1)$ is a parameter. Then the second best outcome is

$$(Q^*(\theta), T^*(\theta)) = \begin{cases} \left(2\theta - 1, \frac{(2\theta-1)^\alpha(2\alpha\theta+1)}{2\alpha(1+\alpha)} \right) & \text{if } \theta \geq 1/2 \\ (0, 0) & \text{if } \theta < 1/2 \end{cases}.$$

Applying Corollary 2, one can verify that $\Pi^* = \Pi_\infty^{ID}$ if and only if $\alpha \leq 1/2$, while $\Pi^* = \Pi_\infty^{AD}$ if and only if $\alpha \leq \sqrt{2}/2$.

We now turn to the comparison among the three forms when the number of blocks is restricted to be no larger than some finite number. It is a harder job because when n is finite, the concepts of n -ID-implementability, n -ADF-implementability and n -BD-implementability do not imply one another. While ID has the smallest implementation power (i.e. has the smallest set of implementable outcomes) when the number of blocks is unrestricted, we will nonetheless see that ID has the following advantage when the number of blocks is restricted. If a second best outcome can be implemented by some ID scheme with infinite blocks, then ID schemes with finite blocks can be constructed to better approximate the second best outcome than AD schemes or BD schemes with the same number of blocks do. In this sense, we say ID has the largest "approximation power".

In the rest of this section we impose the following regularity assumptions on the virtual surplus function H , which ensure that second best outcome is essentially unique, and better approximating the second best outcome raises profit.

Assumption 2 H is continuous. $\arg \max_{q \geq 0} H(q, \theta)$ is single-valued and nondecreasing in θ . $H(\cdot, \theta)$ is single-peaked (i.e. $H(q, \theta)$ gets weakly higher when q gets closer to the unique maximizer).

Lemma 1 *Let Q^* be a second best quantity function. Let (Q^1, T^1) and (Q^2, T^2) be two IC outcomes such that, for almost every $\theta \in \Theta$, either $Q^2(\theta) \geq Q^1(\theta) \geq Q^*(\theta)$ or $Q^2(\theta) \leq Q^1(\theta) \leq Q^*(\theta)$, and $S(Q^1(\underline{\theta}), \underline{\theta}) - T^1(\underline{\theta}) \leq S(Q^2(\underline{\theta}), \underline{\theta}) - T^2(\underline{\theta})$. Then the profit (1) generated by (Q^1, T^1) is weakly higher than that generated by (Q^2, T^2) . The last inequality is strict unless $S(Q^1(\underline{\theta}), \underline{\theta}) - T^1(\underline{\theta}) = S(Q^2(\underline{\theta}), \underline{\theta}) - T^2(\underline{\theta})$ and $Q^1(\theta) = Q^2(\theta)$ for almost every $\theta \in \Theta$.*

Proof. Under Assumption 2, $\{Q^*(\theta)\} = \arg \max_{q \geq 0} H(q, \theta)$ for all $\theta \in \Theta$ except possibly $\underline{\theta}$ and $\bar{\theta}$. Apply formula (2). ■

Theorem 5 *Suppose that $\Pi^* = \Pi_\infty^{ID}$ and n is finite. (a) $\Pi_n^{ID} \geq \Pi_n^{ADF}$, and this inequality is strict unless some n -ID-optimal outcome coincides with some n -ADF-optimal outcome almost everywhere. (b) $\Pi_n^{ID} \geq \Pi_n^{BD}$, and this inequality is strict unless some n -ID-optimal outcome coincides with some n -BD-optimal outcome almost everywhere.*

The idea of the proof of Theorem 5 is the following. To show $\Pi_n^{ID} \geq \Pi_n^{ADF}$, it amounts to show that, given an n -ADF-optimal outcome, there exists an n -ID-implementable outcome whose quantity function is uniformly closer to the second best than the quantity function of the n -ADF-optimal outcome. In fact, we do not need to know the characteristics of the n -ADF-optimal outcome except knowing that it is n -ADF-implementable. Let us fix $n = 3$ for example and start with some 3-ADF-implementable outcome. The bold curve in Figure 5 illustrates the quantity function Q^{ADF} of such a 3-ADF-implementable outcome, and the thin curve illustrates the second best quantity function Q^* . Under the assumption $\Pi^* = \Pi_\infty^{ID}$, Q^* is drawn to cross every type-demand curve from below. (Recall that, for the second best outcome to be implementable by some ID scheme, Q^* cannot be flatter than type-demand curves.) Then we can draw a quantity function Q^{ID} of a 3-ID-implementable outcome as shown by the dashed curve in Figure 5. Notice that Q^{ID} is uniformly closer to Q^* than Q^{ADF} . Now according to Lemma 1, the 3-ID-implementable outcome (Q^{ID}, T^{ID}) that leaves the lowest type of consumer a zero rent would make a higher profit than the original 3-ADF-implementable outcome does. Therefore $\Pi_n^{ID} \geq \Pi_n^{ADF}$. We can similarly argue that $\Pi_n^{ID} \geq \Pi_n^{BD}$ under the assumption $\Pi^* = \Pi_\infty^{ID}$.

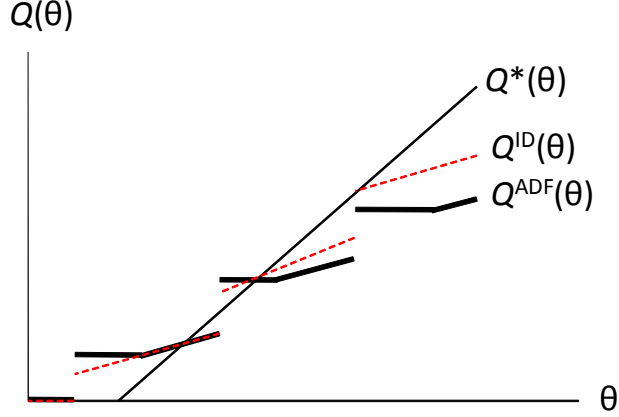


Figure 5: Improving upon ADF by ID

5 Generalized environment

In this and the next sections, we develop a unifying framework that can simultaneously analyze all the practical forms of pricing schemes studied in this paper. Moreover, this framework works in a much more general environment, which we describe in this section. The results in the next section, which are applied to prove the results in Section 3, relies only on the assumptions made in this section.

Consider a more general monopolistic nonlinear pricing problem. The domain for quantities is denoted as χ , which can be any closed subset of \mathbb{R} that includes 0. The monopolist's cost function $c : \chi \rightarrow \mathbb{R}$ and the cumulative distribution function F of consumers' types are irrelevant, except that the support of F still has to be an interval $\Theta = [\underline{\theta}, \bar{\theta}]$. Consumers' utility is still in the form $S(q, \theta) - t$, and consumers' outside option is still buying nothing and paying nothing. For the gross utility function $S : \chi \times \Theta \rightarrow \mathbb{R}$, all we need are the following assumptions.

Assumption 3 $S(0, \theta) = 0$ for all $\theta \in \Theta$. For any $q \in \chi$, $S(q, \theta)$ is absolutely continuous (and hence almost everywhere differentiable) in θ . For any $\theta \in \Theta$, $S(q, \theta)$ is continuous in q . Moreover, S satisfies strictly increasing differences in $(q; \theta)$, i.e. $q_1 \leq q_2$ and $\theta_1 \leq \theta_2$ imply

$$S(q_2, \theta_2) - S(q_1, \theta_2) \geq S(q_2, \theta_1) - S(q_1, \theta_1),$$

and the inequality becomes strict whenever $q_1 < q_2$ and $\theta_1 < \theta_2$.

Note that this general framework allows quantity to be discrete or continuous, and Assumption 3 is very weak, so that our results can be adapted to many other principal-agent settings. The only essential restriction in Assumption 3 is its last sentence, which is the strict Spence-Mirrlees single-crossing property. Also note that Assumption 3 is enough to imply Fact 1.

We now formalize the concepts of tariff option. A *tariff option* τ is a function that assigns each quantity $q \in \chi$ a total payment $\tau(q) \in \mathbb{R} \cup \{\infty\}$ for purchasing q units, such that $\tau(q) < \infty$ for some $q \in \chi$, and $\liminf_{x \rightarrow q} \tau(x) > -\infty$ for all $q \in \chi$.¹²

The interpretation of $\tau(q) = \infty$ is that purchasing q units is not allowed by the tariff option τ . The set

$$\chi(\tau) \equiv \{q \in \chi : \tau(q) < \infty\}$$

is called the *domain* of τ , whose interpretation is the set of order sizes allowed by τ .

Notice that our concept of tariff option is a general one: we formally allow the payment associated with an order size to be infinity to forbid that order size; we also allow the payment to be negative, provided that it never converges to negative infinity. In this terminology, the consumers' outside option is also a tariff option. We use τ_{out} to denote the tariff option that represents the outside option, i.e.

$$\tau_{out}(q) \equiv \begin{cases} 0 & \text{if } q = 0 \\ \infty & \text{if } q \in \chi \setminus \{0\} \end{cases} .$$

6 Menu implementability

Let \mathcal{C} be a class of tariff options, which is meant to be the class of "admissible" tariff options that the monopolist can put into a menu for the consumers to choose.

Definition 1 *We say an outcome (Q, T) is menu implementable with respect to (n, \mathcal{C}) (where $n \in \mathbb{Z}_+ \cup \{\infty\}$ and \mathcal{C} is a class of tariff options) if it is induced by a menu of at most n tariff options in \mathcal{C} , i.e. there exist some set (menu) $\mathcal{M} \subset \mathcal{C}$ with $|\mathcal{M}| \leq n$ and some mapping (best response of consumers) that assigns each consumers' type $\theta \in \Theta$ a tariff option $\tau_\theta \in \mathcal{M} \cup \{\tau_{out}\}$ such that*

¹²As usual, $\liminf_{x \rightarrow q} \tau(x)$ is defined by

$$\lim_{\varepsilon \downarrow 0} (\inf \{\tau(x) : |x - q| < \varepsilon \text{ and } x \in \chi\}) .$$

1. $T(\theta) = \tau_\theta(Q(\theta))$ for any $\theta \in \Theta$,
2. $S(Q(\theta), \theta) - T(\theta) \geq S(q, \theta) - \tau(q)$ for any $\theta \in \Theta$, $q \in \chi$, $\tau \in \mathcal{M} \cup \{\tau_{out}\}$.

Implementability by n -BD schemes, by n -ID schemes, and by n -ADF schemes are all special cases of the above menu implementability with respect to (n, \mathcal{C}) . For n -BD schemes, \mathcal{C} is the set of all tariff options with singleton domain. For n -ID schemes, \mathcal{C} is the set of all two-part tariffs. For n -ADF schemes with fixed fee ϕ , \mathcal{C} is the set of all minimum purchase tariffs with common fixed fee ϕ .

In order to characterize our concept of menu implementability, we introduce some notions. For any two tariff options τ_1 and τ_2 , we say (τ_1, τ_2) satisfies *tariff single crossing property* if for any $q_1, q_2 \in \chi(\tau_1) \cup \chi(\tau_2)$ with $q_1 < q_2$, we have

$$\tau_1(q_2) \leq \tau_2(q_2) \Rightarrow \tau_1(q_1) \leq \tau_2(q_1),$$

$$\tau_1(q_1) \geq \tau_2(q_1) \Rightarrow \tau_1(q_2) \geq \tau_2(q_2).$$

That is, $\tau_1(q) - \tau_2(q)$, regarded as a function of q and restricted on $\chi(\tau_1) \cup \chi(\tau_2)$, crosses or touches zero only once and only from below. The interpretation is that τ_2 is more favorable to purchasing large quantities than τ_1 .

For any two tariff options τ_1 and τ_2 , we say (τ_1, τ_2) satisfies *tariff increasing differences* if $\tau_1(q) - \tau_2(q)$ is nondecreasing in q on $\chi(\tau_1) \cup \chi(\tau_2)$. Another way to say that is: (τ_1, τ_2) satisfies tariff increasing differences iff $(\tau_1, \tau_2 + x)$ satisfies tariff single crossing property for any $x \in \mathbb{R}$. The interpretation is that τ_2 is more favorable to purchasing incremental units than τ_1 .

It is easy to see that if (τ_1, τ_2) satisfies tariff increasing differences, then it satisfies tariff single crossing property.

Theorem 6 *An outcome (Q, T) is menu implementable with respect to (n, \mathcal{C}) if it is IC and IR, and there exists a mapping that assigns each consumers' type $\theta \in \Theta$ a tariff option $\tau_\theta \in \mathcal{C} \cup \{\tau_{out}\}$ such that*

1. $T(\theta) = \tau_\theta(Q(\theta))$ for any $\theta \in \Theta$,
2. $S(Q(\theta), \theta) - T(\theta) \geq S(q, \theta) - \tau_\theta(q)$ for any $\theta \in \Theta$, $q \in \chi$,
3. *increasing differences monotonicity: if $\theta_1 < \theta_2$ then $(\tau_{\theta_1}, \tau_{\theta_2})$ satisfies tariff increasing differences,*

$$4. |\{\tau_\theta\}_{\theta \in \Theta} \setminus \{\tau_{out}\}| \leq n.$$

Theorem 7 *Theorem 6 still holds if condition 3 is replaced by*

3'a. *single crossing monotonicity: if $\theta_1 < \theta_2$ then $(\tau_{\theta_1}, \tau_{\theta_2})$ satisfies tariff single crossing property,*

3'b. *$T(\theta) \leq \tau_{\theta'}(Q(\theta))$ for any $\theta, \theta' \in \Theta$.*

Theorems 6 and 7 can be informally understood as follows. If the monopolist is free to offer any pricing scheme, it is without loss of generality to restrict attention to offering a direct revelation mechanism (Q, T) (which asks consumers to report types and specifies a quantity-payment bundle for each reported type) that satisfies incentive compatibility constraint and individual rationality constraint (so that it is equivalent to BD schemes). Now the monopolist is restricted to offering a menu of tariff options in \mathcal{C} , then it is only without loss of generality to restrict attention to offering a revelation mechanism (the mapping $\theta \mapsto \tau_\theta$ in Theorem 6) that specifies for each reported type a tariff option in $\mathcal{C} \cup \{\tau_{out}\}$, after which the consumer is free to choose any order size allowed by the chosen tariff option.

The issue of individual rationality simply does not change. The issue of incentive compatibility now has two parts. First, the consumer should be willing to choose the desired order size given the tariff option designated for her. This part is condition 2 in Theorem 6. Second, the consumer should be willing to choose the tariff option designated for her (i.e. report the true type). This second part is stronger than requiring (Q, T) to be incentive-compatible. Since a direct revelation mechanism specifies a particular quantity-payment bundle once a consumer's type has been reported, so a consumer has an incentive to report a false type only if she prefers the quantity-payment bundle designated for that false type. In contrast, under a revelation mechanism $\theta \mapsto \tau_\theta$ which specifies a tariff option after reporting, a consumer has some flexibility to choose the quantity-payment bundle after a type has been reported, as long as the tariff option specified for that reported type is not degenerate. Hence a consumer has an incentive to report a false type whenever she prefers any quantity-payment bundle allowed by the tariff option designated for that false type.

Then what condition can guarantee this second part? Under a direct revelation mechanism, it is well known that incentive compatibility requires monotonicity: higher type consumers should purchase more. Under the aforementioned revelation

mechanism $\theta \mapsto \tau_\theta$, it turns out that a natural extension of this monotonicity suffices: higher type consumers should pick tariff options that are favorable to purchasing more. The latter is formalized by increasing differences monotonicity (i.e. condition 3 in Theorem 6). If we replace increasing differences monotonicity by the related but weaker condition single crossing monotonicity (i.e. condition 3'a in Theorem 7), then the additional condition 3'b in Theorem 7 is needed.

Remark 1 *Notice that in Theorems 6 and 7, condition 3'b clearly has to be satisfied anyway (for τ_θ to be the consumers' best response). So actually condition 3, together with other conditions, implies condition 3' (including 3'a and 3'b). However, the sufficiency of condition 3 is worth knowing on top of knowing the sufficiency of condition 3', because for some applications condition 3 is more useful. Indeed, in our applications in Section 3, checking increasing differences monotonicity is not harder than checking single crossing monotonicity and hence we do not need to worry about condition 3'b.*

The necessity counterparts of Theorems 6 and 7 require some restrictions on $\mathcal{C} \cup \{\tau_{out}\}$. We say $\mathcal{C} \cup \{\tau_{out}\}$ is *closed* if it is closed under the product topology in the space of tariff options (i.e. whenever a net in $\mathcal{C} \cup \{\tau_{out}\}$ pointwise converges to a tariff option, this tariff option is also in $\mathcal{C} \cup \{\tau_{out}\}$.) For any two tariff options τ_1 and τ_2 , we say τ_1 *dominates* τ_2 if $\tau_1(q) \leq \tau_2(q)$ for any $q \in \chi$, and the inequality is strict for some $q \in \chi$. We say $\mathcal{C} \cup \{\tau_{out}\}$ is *increasing differences comparable* if for any two tariff options τ_1 and τ_2 in $\mathcal{C} \cup \{\tau_{out}\}$ which are not dominated by each other, either (τ_1, τ_2) or (τ_2, τ_1) satisfies tariff increasing differences. Similarly, we say $\mathcal{C} \cup \{\tau_{out}\}$ is *single crossing comparable* if for any two tariff options τ_1 and τ_2 in $\mathcal{C} \cup \{\tau_{out}\}$ which are not dominated by each other, either (τ_1, τ_2) or (τ_2, τ_1) satisfies tariff single crossing property.

Theorem 8 *For an outcome (Q, T) to be menu implementable with respect to (n, \mathcal{C}) , the sufficient conditions provided by Theorem 6 are also necessary if (i) n is finite or $\mathcal{C} \cup \{\tau_{out}\}$ is closed, and (ii) $\mathcal{C} \cup \{\tau_{out}\}$ is increasing differences comparable. Similarly, for an outcome (Q, T) to be menu implementable with respect to (n, \mathcal{C}) , the sufficient conditions provided by Theorem 7 are also necessary if (i) n is finite or $\mathcal{C} \cup \{\tau_{out}\}$ is closed, and (ii') $\mathcal{C} \cup \{\tau_{out}\}$ is single crossing comparable.*

Theorems 6 and 8 can be translated into necessary and sufficient conditions for implementability by n -BD schemes, by n -ID schemes, and by n -ADF schemes. In

particular, the increasing differences monotonicity condition in Theorem 6 translates into the monotonicity of marginal price (i.e. condition 2 in Theorem 2) in the context of n -ID schemes, and translates into the monotonicity of minimum purchase and marginal price (i.e. condition 3 in Theorem 3) in the context of n -ADF schemes.

7 Concluding remarks

We compare maximum profits of bundling, incremental discounts and all-units discounts in the context of monopolistic nonlinear pricing. The comparison hinges on whether the number of pricing blocks is restricted. We also develop a theory of menu implementability for general principal-agent settings.

Our comparison among bundling (BD), incremental discounts (ID) and all-units discounts with fixed fee (ADF) in the context of nonlinear pricing sheds lights on a general issue: how should an uninformed principal choose among different practical contract forms to offer to an informed agent? If the level of contract complexity and communication between the principal and the agent are unlimited, it is well known that restricting to direct revelation mechanisms is without loss. A heuristic reason is that direct revelation mechanisms have full control over the agent's action once the agent's private information is reported. Hence, under appropriately chosen direct revelation mechanism, the harm of private information is minimal. However, when contract complexity or communication is limited, it might pay to leave certain kind of flexibility or discretion power to the agent.

The above trade-off between control and flexibility explains the ranking among BD, ID and ADF. BD exhibits the largest control: each option specifies a single quantity. ID exhibits the smallest control: a two-part tariff does not directly control quantity, but only control indirectly through a marginal price. ADF is somehow in the middle: a minimum purchase tariff controls quantity through both a direct instrument, minimum purchase, and an indirect instrument, marginal price. This is why when the number of blocks (contract complexity) is unrestricted, BD has the largest and ID the smallest implementation power (Theorem 4 and Corollary 1). When the number of blocks is restricted, ID leaves certain kind of flexibility (one constrained by price) to the agent. Leaving flexibility to the agent, or giving up some control, in general may or may not be good since the agent has different interest from the principal. However, if the control loss implied by such kind of flexibility can be

fully overcome when the number of blocks is unrestricted (i.e. $\Pi^* = \Pi_\infty^{ID}$), then this kind of flexibility must help the approximation to second best. This is why ID has the largest approximation power (Theorem 5 and the paragraph thereafter).

The lesson is: when contract complexity is limited for practical concerns, the principal might gain from leaving to the agent some kind of flexibility that is nonbinding under unlimited contract complexity.

Appendix

The proof of Theorem 6 requires the following notation and lemma.

For any tariff option τ , we let $\tau^{\text{inf}} : \chi \rightarrow \mathbb{R} \cup \{\infty\}$ denote the highest lower semi-continuous function that is weakly lower than τ . That is, for each $q \in \chi$, $\tau^{\text{inf}}(q) = \liminf_{x \rightarrow q} \tau(x)$. Clearly, τ^{inf} is also a tariff option; and $\tau^{\text{inf}} = \tau$ if and only if τ is lower semi-continuous.

Lemma 2 *If (τ_1, τ_2) satisfies tariff increasing differences, then $(\tau_1^{\text{inf}}, \tau_2^{\text{inf}})$ satisfies tariff increasing differences.*

Proof. We must show that if $\tau_1(q) - \tau_2(q)$ is nondecreasing in q on $\chi(\tau_1) \cup \chi(\tau_2)$, then $\tau_1^{\text{inf}}(q) - \tau_2^{\text{inf}}(q)$ is nondecreasing in q on $\chi(\tau_1^{\text{inf}}) \cup \chi(\tau_2^{\text{inf}})$. Let $\chi_1 \equiv \chi(\tau_1)$ and $\chi_2 \equiv \chi(\tau_2)$. Notice that $\chi(\tau_1^{\text{inf}}) = \text{cl}(\chi_1)$ and $\chi(\tau_2^{\text{inf}}) = \text{cl}(\chi_2)$.¹³

Suppose that $\tau_1^{\text{inf}}(q) - \tau_2^{\text{inf}}(q)$ is not nondecreasing in q on $\text{cl}(\chi_1) \cup \text{cl}(\chi_2)$. Then there exist $q_1, q_2 \in \text{cl}(\chi_1) \cup \text{cl}(\chi_2)$ such that $q_1 < q_2$ and

$$\tau_1^{\text{inf}}(q_1) - \tau_2^{\text{inf}}(q_1) > \tau_1^{\text{inf}}(q_2) - \tau_2^{\text{inf}}(q_2). \quad (5)$$

(Since $q_1, q_2 \in \chi(\tau_1^{\text{inf}}) \cup \chi(\tau_2^{\text{inf}})$, both sides of (5) are not $\infty - \infty$ and hence are well defined.) (5) implies $\tau_2^{\text{inf}}(q_1) < \infty$ and $\tau_1^{\text{inf}}(q_2) < \infty$. Since $\tau_2(x) < \infty$ iff $x \in \chi_2$, we have

$$\tau_2^{\text{inf}}(q_1) = \liminf_{x \rightarrow q_1} \tau_2(x) = \liminf_{x \in \chi_2 \text{ \& } x \rightarrow q_1} \tau_2(x).$$

Since $\tau_1(x) < \infty$ iff $x \in \chi_1$, we have

$$\tau_1^{\text{inf}}(q_2) = \liminf_{x \rightarrow q_2} \tau_1(x) = \liminf_{x \in \chi_1 \text{ \& } x \rightarrow q_2} \tau_1(x).$$

¹³As usual, cl denotes closure.

Then,

$$\begin{aligned}
\tau_1^{\text{inf}}(q_1) - \tau_2^{\text{inf}}(q_1) &= \liminf_{x \rightarrow q_1} \tau_1(x) - \liminf_{x \in \chi_2 \ \& \ x \rightarrow q_1} \tau_2(x) \\
&\leq \liminf_{x \in \chi_2 \ \& \ x \rightarrow q_1} \tau_1(x) - \liminf_{x \in \chi_2 \ \& \ x \rightarrow q_1} \tau_2(x) \\
&\leq \limsup_{x \in \chi_2 \ \& \ x \rightarrow q_1} \tau_1(x) - \liminf_{x \in \chi_2 \ \& \ x \rightarrow q_1} \tau_2(x) \\
&= \limsup_{x \in \chi_2 \ \& \ x \rightarrow q_1} (\tau_1(x) - \tau_2(x)),
\end{aligned}$$

and

$$\begin{aligned}
\tau_1^{\text{inf}}(q_2) - \tau_2^{\text{inf}}(q_2) &= \liminf_{x \in \chi_1 \ \& \ x \rightarrow q_2} \tau_1(x) - \liminf_{x \rightarrow q_2} \tau_2(x) \\
&\geq \liminf_{x \in \chi_1 \ \& \ x \rightarrow q_2} \tau_1(x) - \liminf_{x \in \chi_1 \ \& \ x \rightarrow q_2} \tau_2(x) \\
&\geq \liminf_{x \in \chi_1 \ \& \ x \rightarrow q_2} \tau_1(x) - \limsup_{x \in \chi_1 \ \& \ x \rightarrow q_2} \tau_2(x) \\
&= \liminf_{x \in \chi_1 \ \& \ x \rightarrow q_2} (\tau_1(x) - \tau_2(x)).
\end{aligned}$$

Now, (5) implies that

$$\limsup_{x \in \chi_2 \ \& \ x \rightarrow q_1} (\tau_1(x) - \tau_2(x)) > \liminf_{x \in \chi_1 \ \& \ x \rightarrow q_2} (\tau_1(x) - \tau_2(x)).$$

Thus, there exist sequences $q_1^m \rightarrow q_1$ on χ_2 and $q_2^m \rightarrow q_2$ on χ_1 such that, for all m , $\tau_1(q_1^m) - \tau_2(q_1^m) > \tau_1(q_2^m) - \tau_2(q_2^m)$ and $q_1^m < q_2^m$, and hence $\tau_1(q) - \tau_2(q)$ is not nondecreasing in q on $\chi(\tau_1) \cup \chi(\tau_2)$. ■

Proof of Theorem 6. Fix any (n, \mathcal{C}) and any outcome (Q, T) . Suppose that the conditions in the theorem are satisfied. Consider the menu $\mathcal{M} = \{\tau_\theta\}_{\theta \in \Theta} \setminus \{\tau_{out}\} \subset \mathcal{C}$. Then $|\mathcal{M}| \leq n$ (from condition 4) and condition 1 in Definition 1 are satisfied (from condition 1). Suppose by way of contradiction that (Q, T) is not menu implementable with respect to (n, \mathcal{C}) . Then condition 2 in Definition 1 is violated, i.e.

$$S(Q(\theta_1), \theta_1) - T(\theta_1) < S(q', \theta_1) - \tau'(q') \tag{6}$$

for some $\theta_1 \in \Theta$, $q' \in \chi$, $\tau' \in \{\tau_\theta\}_{\theta \in \Theta} \cup \{\tau_{out}\}$.

In (6), if $\tau' = \tau_{out}$, then $q' = 0$ (otherwise the right-hand side is $-\infty$), then the

right-hand side is 0, contradicting to the IR of (Q, T) . Therefore $\tau' = \tau_{\theta_2}$ for some $\theta_2 \in \Theta$.

Let $\chi' \equiv \chi \cap ([Q(\underline{\theta}), Q(\bar{\theta})] \cup \{q'\})$. For any $\theta \in \Theta$ and any tariff option τ , we define

$$V(\tau, \theta) \equiv \sup_{q \in \chi} \{S(q, \theta) - \tau(q)\}, \quad V^+(\tau, \theta) \equiv \sup_{q \in \chi'} \{S(q, \theta) - \tau(q)\},$$

$$D(\tau, \theta) \equiv \arg \max_{q \in \chi} \{S(q, \theta) - \tau(q)\}, \quad D^+(\tau, \theta) \equiv \arg \max_{q \in \chi'} \{S(q, \theta) - \tau(q)\}.$$

For any $\theta \in \Theta$, we have $Q(\theta) \in D(\tau_\theta, \theta)$ and hence $U(\theta) \equiv S(Q(\theta), \theta) - T(\theta) = V(\tau_\theta, \theta)$, due to conditions 1 and 2. Then $S(q, \theta) - \tau_\theta(q)$ must be upper semi-continuous in q at $Q(\theta)$, and hence $\tau_\theta(Q(\theta)) = \tau_\theta^{\text{inf}}(Q(\theta))$. Then $Q(\theta) \in D(\tau_\theta^{\text{inf}}, \theta)$. Since $Q(\theta) \in \chi'$, we also have $Q(\theta) \in D^+(\tau_\theta^{\text{inf}}, \theta)$. It follows that

$$U(\theta) = V(\tau_\theta, \theta) = V(\tau_\theta^{\text{inf}}, \theta) = V^+(\tau_\theta^{\text{inf}}, \theta). \quad (7)$$

For any $\theta, \theta' \in \Theta$, $D^+(\tau_{\theta'}^{\text{inf}}, \theta)$ is nonempty since χ' is compact and the objective function is upper semi-continuous in q (from the continuity of $S(\cdot, \theta)$ and the lower semi-continuity of $\tau_{\theta'}^{\text{inf}}$). Also, clearly $V^+(\tau_{\theta'}^{\text{inf}}, \theta) = V^+(\tau_{\theta'}^{\text{inf}}, \theta)$.

Since $q' \in \chi'$, (6) implies $U(\theta_1) < S(q', \theta_1) - \tau_{\theta_2}^{\text{inf}}(q') \leq V^+(\tau_{\theta_2}^{\text{inf}}, \theta_1)$. Moreover, (7) implies $U(\theta_2) = V^+(\tau_{\theta_2}^{\text{inf}}, \theta_2)$. In other words, the value of $U(\cdot)$ is strictly below the value of $V^+(\tau_{\theta_2}^{\text{inf}}, \cdot)$ at θ_1 , but the values of these two functions are equal at θ_2 . By IC of (Q, T) and Fact 1, $U(\cdot)$ is absolutely continuous, and $U'(x) = S_\theta(Q(x), x)$ for almost every $x \in \Theta$. By Envelope Theorem (in the version of Milgrom and Segal (2002)), $V^+(\tau_{\theta_2}^{\text{inf}}, \cdot)$ is absolutely continuous, and $\partial V^+(\tau_{\theta_2}^{\text{inf}}, x) / \partial x = S_\theta(d^+(\tau_{\theta_2}^{\text{inf}}, x), x)$ for almost every $x \in \Theta$, where $d^+(\tau_{\theta_2}^{\text{inf}}, \theta)$ is any selection from $D^+(\tau_{\theta_2}^{\text{inf}}, \theta)$.

It follows that, if $\theta_1 < \theta_2$, then there is some $x_0 \in [\theta_1, \theta_2]$ such that $U(x_0) < V^+(\tau_{\theta_2}^{\text{inf}}, x_0)$ and $U'(x_0) > \partial V^+(\tau_{\theta_2}^{\text{inf}}, x_0) / \partial x_0$, and then $S_\theta(Q(x_0), x_0) > S_\theta(d^+(\tau_{\theta_2}^{\text{inf}}, x_0), x_0)$, and then $Q(x_0) > d^+(\tau_{\theta_2}^{\text{inf}}, x_0)$. (Notice that Assumption 3 implies that $S_\theta(q, \theta)$ is nondecreasing in q .) Similarly, if $\theta_1 > \theta_2$, then there is some $x_0 \in [\theta_2, \theta_1]$ such that $U(x_0) < V^+(\tau_{\theta_2}^{\text{inf}}, x_0)$ and $Q(x_0) < d^+(\tau_{\theta_2}^{\text{inf}}, x_0)$.

Increasing differences monotonicity (condition 3) and Lemma 2 imply that $S(q, \theta) - \tau_{\theta'}^{\text{inf}}(q)$ satisfies increasing differences in $(q; \theta')$. By monotone comparative statics (see Topkis (1978) or Milgrom and Shannon (1994)), we obtain that, given any $\theta \in \Theta$,

$D^+(\tau_{\theta'}^{\text{inf}}, \theta)$ is nondecreasing in θ' in the strong set order \leq_s . (For any $D_1, D_2 \subset \mathbb{R}$, $D_1 \leq_s D_2$ iff $d_1 \in D_1$ and $d_2 \in D_2$ imply $\min\{d_1, d_2\} \in D_1$ and $\max\{d_1, d_2\} \in D_2$.) If $\theta' < \theta$, then $D^+(\tau_{\theta'}^{\text{inf}}, \theta) \leq_s D^+(\tau_{\theta}^{\text{inf}}, \theta)$, then there is some $d \in D^+(\tau_{\theta'}^{\text{inf}}, \theta)$ such that $d \leq Q(\theta)$. (Recall that $Q(\theta) \in D^+(\tau_{\theta}^{\text{inf}}, \theta)$ and $D^+(\tau_{\theta}^{\text{inf}}, \theta)$ is nonempty.) Similarly, if $\theta' > \theta$, then there is some $d \in D^+(\tau_{\theta'}^{\text{inf}}, \theta)$ such that $d \geq Q(\theta)$. Now, we pick a selection $d^+(\tau_{\theta'}^{\text{inf}}, \theta)$ from $D^+(\tau_{\theta'}^{\text{inf}}, \theta)$ such that

$$\begin{aligned} & \leq & & < \\ d^+(\tau_{\theta'}^{\text{inf}}, \theta) & = & Q(\theta) & \text{ if } \theta' = \theta. \\ & \geq & & > \end{aligned}$$

It contradicts to our previous claim when we take θ' as θ_2 and θ as x_0 . ■

Proof of Theorem 7. The proof of Theorem 6 still goes through except the last paragraph, which is now replaced by the following.

Since $d^+(\tau_{\theta_2}^{\text{inf}}, x)$ is some selection of $D^+(\tau_{\theta_2}^{\text{inf}}, x_0)$, there exists a sequence $\{d^i\}$ on $\chi(\tau_{\theta_2})$ such that $d^i \rightarrow d^+(\tau_{\theta_2}^{\text{inf}}, x_0)$ and $S(d^i, x_0) - \tau_{\theta_2}(d^i) \rightarrow V^+(\tau_{\theta_2}^{\text{inf}}, x_0)$. Consider $\theta_1 < \theta_2$. From our previous claim (in the proof of Theorem 6), for all large enough i , we have

$$U(x_0) < S(d^i, x_0) - \tau_{\theta_2}(d^i) \text{ and } Q(x_0) > d^i.$$

Let i be large enough so that the above properties hold. Since $U(x_0) \geq S(d^i, x_0) - \tau_{x_0}(d^i)$, we have $S(d^i, x_0) - \tau_{x_0}(d^i) < S(d^i, x_0) - \tau_{\theta_2}(d^i)$ and then $\tau_{x_0}(d^i) > \tau_{\theta_2}(d^i)$. On the other hand, condition 3'b implies $\tau_{x_0}(Q(x_0)) = T(x_0) \leq \tau_{\theta_2}(Q(x_0))$. Then single crossing monotonicity (condition 3'a) implies $\tau_{x_0}(d^i) \leq \tau_{\theta_2}(d^i)$, because $Q(x_0) \in \chi(\tau_{x_0})$ (since $\tau_{x_0}(Q(x_0)) = T(x_0) < \infty$), $d^i \in \chi(\tau_{\theta_2})$, $x_0 \leq \theta_2$ and $Q(x_0) > d^i$. We have a contradiction. By an analogous argument, the case of $\theta_1 > \theta_2$ also yields a contradiction. ■

The proof of Theorem 8 requires the following two lemmas.

Lemma 3 *If an outcome (Q, T) is menu implementable with respect to (n, \mathcal{C}) , with either n finite or $\mathcal{C} \cup \{\tau_{\text{out}}\}$ closed, then the associated menu \mathcal{M} (that is contained in \mathcal{C} , has at most n elements, and induces (Q, T)) can be chosen such that*

1. for any $q \in \chi$, $\arg \min_{\tau \in \mathcal{M} \cup \{\tau_{\text{out}}\}} \tau(q)$ is nonempty, and

2. the tariff options in \mathcal{M} do not dominate one another.

Proof. Suppose that (Q, T) is menu implementable with respect to (n, \mathcal{C}) , i.e. (Q, T) is induced by some menu $\mathcal{M} \subset \mathcal{C}$ with $|\mathcal{M}| \leq n$. Let $\underline{\tau} : \chi \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$ denote the lower envelope of $\mathcal{M} \cup \{\tau_{out}\}$, i.e. $\underline{\tau}(q) \equiv \inf_{\tau \in \mathcal{M} \cup \{\tau_{out}\}} \tau(q)$.

If n is finite, then $\mathcal{M} \cup \{\tau_{out}\}$ is a nonempty finite set. We delete any dominated tariff options in \mathcal{M} so that condition 2 holds. After the deletion, $\mathcal{M} \cup \{\tau_{out}\}$ is still a nonempty finite set so that condition 1 holds, and the lower envelope of $\mathcal{M} \cup \{\tau_{out}\}$ is the same as before so that (Q, T) is still induced.

We hereafter suppose that $n = \infty$ and $\mathcal{C} \cup \{\tau_{out}\}$ is closed. Let $\Psi \equiv \{\tau \in \mathcal{C} \cup \{\tau_{out}\} : \tau \geq \underline{\tau}\}$. Also, for any $q \in \chi$, we define the following partial order \leq_q over tariff options: for any tariff options τ_1, τ_2 , we say $\tau_1 \leq_q \tau_2$ iff either $\tau_1(q) < \tau_2(q)$, or $\tau_1(q) = \tau_2(q)$ & $\tau_1 \leq \tau_2$ on $\chi \setminus \{q\}$. Then (Ψ, \leq_q) is a partially ordered set. Notice that Ψ is closed since $\mathcal{C} \cup \{\tau_{out}\}$ is closed. We claim that every chain in (Ψ, \leq_q) has a lower bound in Ψ . Indeed, any chain in (Ψ, \leq_q) , when regarded as a net directed downward by \leq_q , must have a pointwise limit in Ψ , and this pointwise limit is a lower bound of the chain. By Zorn's Lemma, (Ψ, \leq_q) has a minimal element ψ_q . By Axiom of Choice, there exists a family $\{\psi_q\}_{q \in \chi}$ of tariff options such that every ψ_q is a minimal element of (Ψ, \leq_q) .

Regard $\{\psi_q\}_{q \in \chi} \setminus \{\tau_{out}\}$ as a menu. By our construction, this menu is contained in \mathcal{C} . For any $q, q' \in \chi$, $\psi_q(q) = \underline{\tau}(q) \leq \psi_{q'}(q)$. Thus this menu satisfies condition 1. Since $\{\psi_q\}_{q \in \chi}$ and $\mathcal{M} \cup \{\tau_{out}\}$ share a common lower envelope $\underline{\tau}$, $\{\psi_q\}_{q \in \chi} \setminus \{\tau_{out}\}$ induces the same outcome as \mathcal{M} does. Finally, if ψ_q is dominated by $\psi_{q'}$, then ψ_q is not a minimal element of Ψ , a contradiction. Therefore, our menu satisfies condition 2. ■

Lemma 4 *If tariff options τ_1, \dots, τ_n do not dominate one another, and (τ_1, τ_2) , (τ_2, τ_3) , \dots , (τ_{n-1}, τ_n) and (τ_n, τ_1) satisfy tariff single crossing property or tariff increasing differences, then $\tau_1 = \dots = \tau_n$.*

Proof. We will use the induction argument. Let $n = 2$ first. Suppose that $\tau_1 \neq \tau_2$, and both (τ_1, τ_2) and (τ_2, τ_1) satisfy tariff single crossing property. Then pick some $x \in \chi$ such that $\tau_1(x) \neq \tau_2(x)$. Without loss of generality, assume $\tau_1(x) < \tau_2(x)$. Since (τ_1, τ_2) satisfies tariff single crossing property, $\tau_1(q) \leq \tau_2(q)$ for all $q \in \chi$ with $q < x$. Since (τ_2, τ_1) satisfies tariff single crossing property, $\tau_1(q) \leq \tau_2(q)$ for all

$q \in \chi$ with $q > x$. Therefore τ_2 is dominated by τ_1 . (If $\tau_1(x) > \tau_2(x)$, one can prove that τ_1 is dominated by τ_2 .) Hence, the lemma holds for $n = 2$.

Assume the induction hypothesis: the lemma holds for $n = 2, 3, \dots, k$. Suppose that $\tau_1, \tau_2, \dots, \tau_{k+1}$ do not dominate one another, and $(\tau_1, \tau_2), (\tau_2, \tau_3), \dots, (\tau_k, \tau_{k+1})$ and (τ_{k+1}, τ_1) satisfy tariff single crossing property. If $\tau_i = \tau_j$ for some $i, j \in \{1, \dots, k+1\}$ with $i \neq j$, then the induction hypothesis implies that $\tau_1 = \dots = \tau_{k+1}$, and we are done. So suppose that $\tau_1, \dots, \tau_{k+1}$ are all distinct. Consider any $i \in \{1, \dots, k\}$. Since τ_i and τ_{i+1} are distinct and do not dominate each other, there exist $x_i, y_i \in \chi$ such that $\tau_i(x_i) < \tau_{i+1}(x_i)$ and $\tau_i(y_i) > \tau_{i+1}(y_i)$. Since (τ_i, τ_{i+1}) satisfies tariff single crossing property, $x_i < y_i$. Let $I \in \arg \min_{i \in \{1, \dots, k\}} x_i$ and $J \in \arg \max_{i \in \{1, \dots, k\}} y_i$. Because $(\tau_1, \tau_2), \dots, (\tau_k, \tau_{k+1})$ satisfy tariff single crossing property, we have $\tau_1(x_I) \leq \dots \leq \tau_I(x_I) < \tau_{I+1}(x_I) \leq \dots \leq \tau_{k+1}(x_I)$ and $\tau_1(y_J) \geq \dots \geq \tau_J(y_J) > \tau_{J+1}(y_J) \geq \dots \geq \tau_{k+1}(y_J)$. But we also have $x_I < y_J$, hence (τ_{k+1}, τ_1) does not satisfy tariff single crossing property. ■

Remark 2 *In fact, the proof of Lemma 4 only requires a weaker version of tariff single crossing property, namely, (τ_1, τ_2) is said to satisfy this weaker version of tariff single crossing property if for any $q_1, q_2 \in \chi$ with $q_1 < q_2$, we have $\tau_1(q_1) \leq \tau_2(q_1)$ whenever $\tau_1(q_2) < \tau_2(q_2)$.*

Proof of Theorem 8. Suppose that conditions (i) and (ii) hold and an outcome (Q, T) is menu implementable with respect to (n, \mathcal{C}) . Then there exist some menu $\mathcal{M} \subset \mathcal{C}$ with $|\mathcal{M}| \leq n$ and some mapping $\theta \mapsto \tau_\theta$ in $(\mathcal{M} \cup \{\tau_{out}\})^\Theta$ such that conditions 1 and 2 in Definition 1 are satisfied. Then clearly (Q, T) is IC and IR.

Let $\tau_\Theta \equiv \{\tau_\theta\}_{\theta \in \Theta} \subset \mathcal{M} \cup \{\tau_{out}\}$. By condition (i) and Lemma 3, we can without loss of generality assume that the tariff options in τ_Θ do not dominate one another.

To introduce convenient notation, we write $\tau_1 \preceq \tau_2$ if (τ_1, τ_2) satisfies tariff increasing differences. Obviously, \preceq is a reflexive binary relation over tariff options. Increasing differences comparability of $\mathcal{C} \cup \{\tau_{out}\}$ (condition (ii)) carries over to τ_Θ (because $\tau_\Theta \subset \mathcal{C} \cup \{\tau_{out}\}$). Thus, \preceq is complete on τ_Θ . From Lemma 4, \preceq is anti-symmetric on τ_Θ . From reflexivity, completeness and Lemma 4, \preceq is transitive on τ_Θ . Therefore, \preceq is a linear order on τ_Θ . It makes τ_Θ a chain.

Define $\mathcal{T}(\theta)$ as the set $\{\tau \in \tau_\Theta : \tau(Q(\theta)) \leq \tau'(Q(\theta)) \text{ for any } \tau' \in \tau_\Theta\}$. $\mathcal{T}(\theta)$ is nonempty because $\tau_\theta \in \mathcal{T}(\theta)$ for every $\theta \in \Theta$ (due to conditions 1 and 2 in Definition 1). Let \preceq_s be the strong set order on τ_Θ induced by \preceq . That is, for any two subsets

$\mathcal{T}_1, \mathcal{T}_2$ of τ_Θ , we say $\mathcal{T}_1 \preceq_s \mathcal{T}_2$ if $\tau_1 \in \mathcal{T}_1$ and $\tau_2 \in \mathcal{T}_2$ and $\tau_2 \preceq \tau_1$ imply $\tau_1 \in \mathcal{T}_2$ and $\tau_2 \in \mathcal{T}_1$. Then \preceq_s is a partial order on the set $\mathcal{P}(\tau_\Theta)$ of nonempty subsets of τ_Θ .

We want to show that the mapping \mathcal{T} is nondecreasing on $((\Theta, \leq), (\mathcal{P}(\tau_\Theta), \preceq_s))$. Consider any $\theta_1, \theta_2 \in \Theta$ such that $\theta_1 < \theta_2$, and we need to show $\mathcal{T}(\theta_1) \preceq_s \mathcal{T}(\theta_2)$. Suppose $\tau_1 \in \mathcal{T}(\theta_1)$ and $\tau_2 \in \mathcal{T}(\theta_2)$ and $\tau_2 \preceq \tau_1$. Let $q_1 \equiv Q(\theta_1)$ and $q_2 \equiv Q(\theta_2)$. From IC of (Q, T) , we have $q_1 \leq q_2$. From IR of (Q, T) , we have $S(q_1, \theta_1) - \tau_1(q_1) \geq S(q_1, \theta_1) - \tau_{\theta_1}(q_1) = S(q_1, \theta_1) - T(\theta_1) \geq 0$ so that $\tau_1(q_1) < \infty$, and similarly $\tau_2(q_2) < \infty$. Thus $q_1 \in \chi(\tau_1)$ and $q_2 \in \chi(\tau_2)$. Since $\tau_1 \in \mathcal{T}(\theta_1)$ and $\tau_2 \in \mathcal{T}(\theta_2)$, we have $\tau_1(q_1) \leq \tau_2(q_1)$ and $\tau_2(q_2) \leq \tau_1(q_2)$. Now it follows from $\tau_2 \preceq \tau_1$ and $q_1 \leq q_2$ that $\tau_1(q_2) \leq \tau_2(q_2)$ and $\tau_2(q_1) \leq \tau_1(q_1)$, and hence $\tau_1 \in \mathcal{T}(\theta_2)$ and $\tau_2 \in \mathcal{T}(\theta_1)$. Therefore, $\mathcal{T}(\theta_1) \preceq_s \mathcal{T}(\theta_2)$. We conclude that the mapping \mathcal{T} is nondecreasing on $((\Theta, \leq), (\mathcal{P}(\tau_\Theta), \preceq_s))$.

Regard \mathcal{T} as a correspondence from Θ to τ_Θ . Our previous results show that \mathcal{T} is nonempty-valued and nondecreasing (with respect to the strong set order induced by \preceq), so that it has a nondecreasing selection $\theta \mapsto \tau_\theta^*$ (see Milgrom and Shannon (1994) Theorem A2).

Now this mapping $\theta \mapsto \tau_\theta^*$ satisfies conditions 3 and 4 in Theorem 6. It satisfies condition 1 in Theorem 6 because $T(\theta) = \tau_\theta(Q(\theta)) = \min_{\theta' \in \Theta} \tau_{\theta'}(Q(\theta)) = \tau_\theta^*(Q(\theta))$ for any $\theta \in \Theta$. It also satisfies condition 2 in Theorem 6 because

$$S(Q(\theta), \theta) - \tau_\theta^*(Q(\theta)) = S(Q(\theta), \theta) - \tau_\theta(Q(\theta)) \geq S(q, \theta) - \tau_\theta^*(q)$$

for any $\theta \in \Theta$ and any $q \in \chi$. It completes the proof of the first part.

For the second part, assume condition (ii') instead of condition (ii). The above proof still goes through except that the binary relation \preceq requires to be redefined: $\tau_1 \preceq \tau_2$ if and only if (τ_1, τ_2) satisfies tariff single crossing property. ■

Proof of Theorem 1. Simply notice that, from Fact 1, if (Q, T) is IC and $Q(\cdot)$ is constant over an interval, then $T(\cdot)$ is also constant over that interval. ■

The proofs of Theorems 2 and 3 require the following lemmas, which characterize the closures¹⁴ of the set of two-part tariffs and the set of common-fixed-fee minimum purchase tariffs.¹⁵

¹⁴As usual, the closure $\bar{\mathcal{C}}$ of a set \mathcal{C} of tariff options is defined as the smallest closed set of tariff options that contains \mathcal{C} , i.e. $\bar{\mathcal{C}}$ comprises all tariff options that is a pointwise limit of a net in \mathcal{C} .

¹⁵Lemma 5 is needed only for the necessity part of Theorem 2 when $n = \infty$. Similarly Lemma 6

Lemma 5 *A tariff option $\hat{\tau}$ is the pointwise limit of some net of two-part tariffs if and only if either $\chi(\hat{\tau}) = \{0\}$ (i.e. the domain of $\hat{\tau}$ is $\{0\}$), or $\hat{\tau}$ is a two-part tariff.*

Proof. The "if" part is quite obvious, so we only prove the "only if" part. Let $\{\tau_\beta\}$ be a net of two-part tariffs, with each τ_β characterized by (p_β, ϕ_β) . Suppose that it pointwise converges to a tariff option $\hat{\tau}$. Since each τ_β is nondecreasing, $\hat{\tau}$ is too. Then $\hat{\tau}(0) \neq -\infty$ (otherwise $\hat{\tau}$ is not a tariff option) and $\hat{\tau}(0) \neq \infty$ (otherwise $\hat{\tau}$ is ∞ everywhere and hence not a tariff option). If $\hat{\tau}(1) = \infty$, then $\hat{\tau}(q) = \infty$ for all $q > 0$ (since each τ_β is linear), so that $\chi(\hat{\tau}) = \{0\}$. Now suppose that $\hat{\tau}(1)$ is finite. Then both $\hat{\phi} \equiv \hat{\tau}(0)$ and $\hat{p} \equiv \hat{\tau}(1) - \hat{\tau}(0)$ are finite. Moreover, we have $\phi_\beta = \tau_\beta(0) \rightarrow \hat{\tau}(0) = \hat{\phi}$ and $p_\beta = \tau_\beta(1) - \tau_\beta(0) \rightarrow \hat{\tau}(1) - \hat{\tau}(0) = \hat{\phi}$. Then, for any $q \geq 0$,

$$\tau_\beta(q) - (\hat{p}q + \hat{\phi}) = (p_\beta - \hat{p})q + (\phi_\beta - \hat{\phi}) \rightarrow 0.$$

Therefore, τ_β pointwise converges to $q \mapsto \hat{p}q + \hat{\phi}$, which is a two-part tariff. ■

Lemma 6 *Given $\phi \in \mathbb{R}$, a tariff option $\hat{\tau}$ is the pointwise limit of some net of minimum purchase tariffs with common fixed fee ϕ if and only if either $\chi(\hat{\tau}) = \{0\}$ and $\hat{\tau}(0) = \phi$, or $\hat{\tau}$ is a minimum purchase tariff with fixed fee ϕ , or $\hat{\tau}$ is a "strict-minimum purchase tariff with fixed fee ϕ " of the form*

$$q \mapsto \begin{cases} pq + \phi & \text{if } q > m \\ \infty & \text{if } 0 \leq q \leq m \end{cases}.$$

(Strict-minimum purchase tariffs differ from minimum purchase tariffs in that the order size has to be strictly larger than the minimum purchase.)

Proof. The "if" part is quite obvious, so we only prove the "only if" part. Let $\{\tau_\beta\}$ be a net of minimum purchase tariffs, with each τ_β characterized by (m_β, p_β, ϕ) . Suppose that it pointwise converges to a tariff option $\hat{\tau}$. Since $\hat{\tau}$ is a tariff option, its domain $\chi(\hat{\tau})$ is nonempty. If $\chi(\hat{\tau}) = \{0\}$, clearly $\hat{\tau}(0) = \phi$. Suppose that $\chi(\hat{\tau}) \neq \{0\}$. Let $\hat{m} \equiv \inf \chi(\hat{\tau}) \geq 0$. Then $m_\beta \rightarrow \hat{m}$. For any $q > \hat{m}$, $\hat{\tau}(q)$ is the limit of $\{p_\beta q + \phi\}$, and then $\hat{\tau}(q)$ cannot be ∞ , for otherwise $p_\beta \rightarrow \infty$ and hence $\chi(\hat{\tau})$ is either $\{0\}$ or empty. Now $\hat{p} \equiv \hat{\tau}(\hat{m} + 2) - \hat{\tau}(\hat{m} + 1)$ is finite. Moreover, for large β ,

$$p_\beta = \tau_\beta(\hat{m} + 2) - \tau_\beta(\hat{m} + 1) \rightarrow \hat{\tau}(\hat{m} + 2) - \hat{\tau}(\hat{m} + 1) = \hat{p}.$$

is needed only for the necessity part of Theorem 3 when $n = \infty$.

Therefore, for any $q > \hat{m}$, we have $\hat{\tau}(q) = \lim \tau_\beta(q) = \lim (p_\beta q + \phi) = \hat{p}q + \phi$. For any $q < \hat{m}$, we have $\hat{\tau}(q) = \lim \tau_\beta(q) = \infty$. The limit $\hat{\tau}(q)$ of $\{\tau_\beta(\hat{m})\}$ must be either $\lim p_\beta q + \phi = \hat{p}q + \phi$ or ∞ . We conclude that $\hat{\tau}$ is the minimum purchase tariff or the strict-minimum purchase tariff characterized by (\hat{m}, \hat{p}, ϕ) . ■

Proof of Theorem 2. Let \mathcal{C} be the set of all two-part tariffs, and $\bar{\mathcal{C}}$ be the closure of \mathcal{C} , which is characterized by Lemma 5.

Sufficiency. Suppose that $P : A \rightarrow \mathbb{R}_+$ satisfies the conditions in this theorem. For each $\theta \in \Theta$, let τ_θ be τ_{out} if $\theta \in A$, and be the two-part tariff characterized by $(P(\theta), \Phi(\theta))$ if $\theta \in \Theta \setminus A$, where $\Phi(\theta) \equiv T(\theta) - P(\theta)Q(\theta)$. It suffices to verify that the mapping $\theta \mapsto \tau_\theta$ satisfies the sufficient conditions provided in Theorem 6. Notice that $\theta_1 < \theta_2$ and $Q(\theta_2) = T(\theta_2) = 0$ imply $Q(\theta_1) = T(\theta_1) = 0$, by IC of (Q, T) and Fact 1. Hence A is an increasing subset of Θ . (That is, $\theta_1 \in A$ and $\theta_1 < \theta_2 \in \Theta$ imply $\theta_2 \in A$.) It, together with condition 2, implies that $\theta \mapsto \tau_\theta$ satisfies condition 3 in Theorem 6. Condition 1 and the definition of A and Φ imply conditions 1 and 2 in Theorem 6. IC of (Q, T) and Fact 1 imply that, for any $\theta \in A$ and $\theta_* \in \Theta$,

$$\Phi(\theta) = S(Q(\theta), \theta) - P(\theta)Q(\theta) - \int_{\theta_*}^{\theta} S_\theta(Q(x), x) dx - U(\theta_*),$$

where $U(\theta_*) \equiv S(Q(\theta_*), \theta_*) - T(\theta_*)$. If $P(\cdot)$ is some constant \bar{p} over an interval in A , then, for any θ, θ_* in that interval,

$$\begin{aligned} \Phi(\theta) &= v(P(\theta), \theta) - \int_{\theta_*}^{\theta} v_\theta(P(x), x) dx - U(\theta_*) \\ &= v(\bar{p}, \theta) - \int_{\theta_*}^{\theta} v_\theta(\bar{p}, x) dx - U(\theta_*) = v(\bar{p}, \theta_*) - U(\theta_*), \end{aligned}$$

where $v(p, \theta) \equiv \sup_{q \geq 0} \{S(q, \theta) - pq\}$. Hence $\Phi(\cdot)$ is constant over that interval. Therefore, condition 3 implies that condition 4 in Theorem 6 holds.

Necessity. Suppose that (Q, T) can be implemented by some n -ID scheme. In other words, (Q, T) is menu implementable with respect to (n, \mathcal{C}) . Then (Q, T) is also menu implementable with respect to $(n, \bar{\mathcal{C}})$.¹⁶ Notice that $\bar{\mathcal{C}} \cup \{\tau_{out}\} = \bar{\mathcal{C}}$, which is closed. Take any $\tau_1, \tau_2 \in \bar{\mathcal{C}}$. If $\chi(\tau_1) = \{0\}$, then (τ_1, τ_2) satisfies tariff increasing differences. If $\chi(\tau_1) \neq \{0\}$ and $\chi(\tau_2) \neq \{0\}$ (so that both τ_1 and τ_2 are two-part

¹⁶If we only consider the case of finite n , we do not need to work with $\bar{\mathcal{C}}$, and this proof would be simpler. The same remark can be made for the necessity proof of Theorem 3.

tariffs), then (τ_1, τ_2) (respectively (τ_2, τ_1)) satisfies tariff increasing differences if and only if $p_1 \geq p_2$ (respectively $p_2 \geq p_1$), where p_i denotes the marginal price of τ_i . Hence, $\bar{\mathcal{C}} \cup \{\tau_{out}\}$ is increasing differences comparable. By Theorem 8, the conditions provided by Theorem 6 are satisfied by some mapping $\theta \mapsto \tau_\theta$ in $(\bar{\mathcal{C}} \cup \{\tau_{out}\})^\Theta$. For any $\theta \in \Theta$ such that $\tau_\theta \notin \mathcal{C} \cup \{\tau_{out}\}$, by Lemma 5, $\chi(\tau_\theta) = \{0\}$, and hence $Q(\theta) = 0$. Let $\Theta_0 \equiv \{\theta \in \Theta : Q(\theta) = 0\}$. By IC of (Q, T) and Fact 1, every $\theta \in \Theta_0$ has the same $T(\theta)$, which we denote as t_0 . If Θ_0 is nonempty and $t_0 = 0$, then we redefine every τ_θ with $\theta \in \Theta_0$ as τ_{out} . If Θ_0 is nonempty and $t_0 \neq 0$, then we redefine every τ_θ with $\theta \in \Theta_0$ as a two-part tariff characterized by (p_0, t_0) , such that consumers of type $\theta \in \Theta_0$ would pick quantity $q = 0$ (this two-part tariff exists because (Q, T) can be implemented by some n -ID scheme). Clearly this p_0 can be chosen to be higher than the marginal price of every other two-part tariff in $\{\tau_\theta\}_{\theta \in \Theta \setminus \Theta_0}$. Now the new mapping $\theta \mapsto \tau_\theta$ is in $(\mathcal{C} \cup \{\tau_{out}\})^\Theta$ and still satisfies the conditions provided by Theorem 6. Since $\{\tau_\theta\}_{\theta \in \Theta} \setminus \{\tau_{out}\} \subset \mathcal{C}$, we can for each $\theta \in A$ define $P(\theta)$ as the marginal price of τ_θ . Then conditions 1–3 follow from conditions 1–4 in Theorem 6. ■

Proof of Theorem 3. Let \mathcal{C} be the set of all minimum purchase tariffs with fixed fee ϕ , and $\bar{\mathcal{C}}$ be the closure of \mathcal{C} , which is characterized by Lemma 6.

Sufficiency. Suppose that $M : A \rightarrow \mathbb{R}_+$ and $P : A \rightarrow \mathbb{R}_+$ satisfies the conditions in this theorem. For each $\theta \in \Theta$, let τ_θ be τ_{out} if $\theta \in A$, and be the minimum purchase tariff characterized by $(M(\theta), P(\theta), \phi)$ if $\theta \in \Theta \setminus A$. It suffices to verify that the mapping $\theta \mapsto \tau_\theta$ satisfies the sufficient conditions provided in Theorem 6. Notice that $\theta_1 < \theta_2$ and $Q(\theta_2) = T(\theta_2) = 0$ imply $Q(\theta_1) = T(\theta_1) = 0$, by IC of (Q, T) and Fact 1. Hence A is an increasing subset of Θ . It, together with condition 3, implies that $\theta \mapsto \tau_\theta$ satisfies condition 3 in Theorem 6. Conditions 1 and 2 and the definition of A imply conditions 1 and 2 in Theorem 6. Condition 4 implies that condition 4 in Theorem 6 holds. If in addition $n = \infty$, clearly the function M can be chosen as the restriction of Q on A without affecting the validity of conditions 1 and 3.

Necessity. Suppose that (Q, T) can be implemented by some n -ADF scheme with fixed fee ϕ . In other words, (Q, T) is menu implementable with respect to (n, \mathcal{C}) . Then (Q, T) is also menu implementable with respect to $(n, \bar{\mathcal{C}})$. Notice that $\bar{\mathcal{C}} \cup \{\tau_{out}\}$ is closed since $\bar{\mathcal{C}}$ is. Take any $\tau_1, \tau_2 \in \bar{\mathcal{C}} \cup \{\tau_{out}\}$. If $\chi(\tau_1) = \{0\}$, then (τ_1, τ_2) satisfies tariff increasing differences. If $\chi(\tau_1) \neq \{0\}$ and $\chi(\tau_2) \neq \{0\}$ (so that both τ_1 and τ_2 are minimum purchase tariffs or strict-minimum purchase tariffs with

fixed fee ϕ), then they do not dominate each other if and only if either $\chi(\tau_2)$ is a proper subset of $\chi(\tau_1)$ and $p_2 < p_1$, or $\chi(\tau_1)$ is a proper subset of $\chi(\tau_2)$ and $p_1 < p_2$, where p_i denotes the marginal price of τ_i . In the first case (respectively second case), (τ_1, τ_2) (respectively (τ_2, τ_1)) satisfies tariff increasing differences. Hence, $\bar{\mathcal{C}} \cup \{\tau_{out}\}$ is increasing differences comparable. By Theorem 8, the conditions provided by Theorem 6 are satisfied by some mapping $\theta \mapsto \tau_\theta$ in $(\bar{\mathcal{C}} \cup \{\tau_{out}\})^\Theta$. For any $\theta \in \Theta$ such that $\tau_\theta \notin \mathcal{C} \cup \{\tau_{out}\}$, by Lemma 6, either τ_θ is a strict-minimum purchase tariff, or $\chi(\tau_\theta) = \{0\}$ and hence $Q(\theta) = 0$. For any $\theta \in \Theta$ such that τ_θ is a strict-minimum purchase tariff characterized by $(m_\theta, p_\theta, \phi)$, we redefine τ_θ as the minimum purchase tariff characterized by $(m_\theta, p_\theta, \phi)$, so that $\tau_\theta \in \mathcal{C}$ and picking the original quantity $q = Q(\theta)$ is still optimal under the new τ_θ for consumers of type θ (since $S(q, \theta)$ is continuous in q). Let $\Theta_0 \equiv \{\theta \in \Theta : Q(\theta) = 0\}$. By IC of (Q, T) and Fact 1, every $\theta \in \Theta_0$ has the same $T(\theta)$, which we denote as t_0 . Clearly $t_0 \in \{0, \phi\}$. If Θ_0 is nonempty and $t_0 = 0$, then we redefine every τ_θ with $\theta \in \Theta_0$ as τ_{out} . If Θ_0 is nonempty and $t_0 = \phi$, then we redefine every τ_θ with $\theta \in \Theta_0$ as a minimum purchase tariff characterized by $(0, p_0, \phi)$, such that consumers of type $\theta \in \Theta_0$ would pick quantity $q = 0$ (this minimum purchase tariff exists because (Q, T) can be implemented by some n -ADF scheme with fixed fee ϕ). Clearly this p_0 can be chosen to be higher than the marginal price of every other minimum purchase tariff in $\{\tau_\theta\}_{\theta \in \Theta \setminus \Theta_0}$. Now the new mapping $\theta \mapsto \tau_\theta$ is in $(\mathcal{C} \cup \{\tau_{out}\})^\Theta$ and still satisfies the conditions provided by Theorem 6. Since $\{\tau_\theta\}_{\theta \in \Theta} \setminus \{\tau_{out}\} \subset \mathcal{C}$, we can for each $\theta \in A$ define $M(\theta)$ and $P(\theta)$ as the minimum purchase and the marginal price of τ_θ . Then conditions 1–4 follow from conditions 1–4 in Theorem 6. ■

Proof of Theorem 4. It follows from IC of (Q, T) , Fact 1 and $U(\underline{\theta}) \equiv S(Q(\underline{\theta}), \underline{\theta}) - T(\underline{\theta}) = 0$ that $T(\theta) = 0$ whenever $Q(\theta) = 0$. Therefore, $A \equiv \{\theta \in \Theta : (Q(\theta), T(\theta)) \neq (0, 0)\} = \{\theta \in \Theta : Q(\theta) > 0\}$.

Part (a). It suffices to construct $\phi \in \mathbb{R}$, $M : A \rightarrow \mathbb{R}_+$ and $P : A \rightarrow \mathbb{R}_+$ such that conditions 1–3 in Theorem 3 hold. (The number n is taken as ∞ so that condition 4 in Theorem 3 trivially holds.) For each $\theta \in A$, we let $M(\theta) \equiv Q(\theta)$ and $P(\theta) \equiv (T(\theta) - \phi) / Q(\theta)$, where the parameter $\phi \in \mathbb{R}$ is chosen to be small (i.e. negative with large magnitude). Then condition 1 in Theorem 3 holds and M is nondecreasing. Clearly, when ϕ is small enough so that $P(\theta)$ is large enough, condition 2 in Theorem 3 holds. Using IC of (Q, T) and Fact 1, it is straightforward to verify that P is nonincreasing when ϕ is small enough. Hence condition 3 in Theorem

3 also holds.

Part (b). Suppose that (Q, T) can be implemented by some ID scheme. By Theorem 2, there exists a nonincreasing function $P^{ID} : A \rightarrow \mathbb{R}_+$ that takes at most n values such that

$$Q(\theta) \in \arg \max_{q \geq 0} \{S(q, \theta) - P^{ID}(\theta)q\} \quad \forall \theta \in A. \quad (8)$$

For each $\theta \in \Theta$, consumers of type θ are meant to pick the two-part tariff characterized $(P^{ID}(\theta), \Phi(\theta))$, where $\Phi(\theta) \equiv T(\theta) - P^{ID}(\theta)Q(\theta)$. For each $\theta \in A$, define $M(\theta) \equiv Q(\theta)$ and $P^{AD}(\theta) \equiv T(\theta)/Q(\theta)$.

It remains to verify that the tuple (M, P^{AD}) satisfies conditions 1–3 in Theorem 3 when $\phi = 0$. Clearly condition 1 in Theorem 3 holds.

Since consumers of type $\underline{\theta}$ has no incentive to deviate to pick the two-part tariff for any type $\theta \in A$, we have

$$0 = U(\underline{\theta}) \geq \max_{q \geq 0} \{S(q, \underline{\theta}) - P(\theta)q - \Phi(\theta)\} \geq -\Phi(\theta) \quad \forall \theta \in A.$$

That is, Φ is nonnegative.

Pick any $\theta \in A$. By nonnegativity of Φ ,

$$P^{AD}(\theta) = \frac{T(\theta)}{Q(\theta)} = \frac{P^{ID}(\theta)Q(\theta) + \Phi(\theta)}{Q(\theta)} \geq P^{ID}(\theta).$$

It, together with (8), implies

$$Q(\theta) \in \arg \max_{q \geq Q(\theta)} \{S(q, \theta) - P^{AD}(\theta)q\} \quad \forall \theta \in A.$$

That is, condition 2 in Theorem 3 holds.

Since (Q, T) is IC, by Fact 1, Q is nondecreasing. Hence M is nondecreasing. Pick any $\theta_1, \theta_2 \in A$ with $\theta_1 < \theta_2$. Since consumers of type θ_2 has no incentive to deviate to pick the two-part tariff for type θ_1 , we have $T(\theta_2) \leq P^{ID}(\theta_1)Q(\theta_2) + \Phi(\theta_1)$. It, together with the monotonicity of Q and the nonnegativity of Φ , implies that

$$P^{AD}(\theta_2) = \frac{T(\theta_2)}{Q(\theta_2)} \leq P^{ID}(\theta_1) + \frac{\Phi(\theta_1)}{Q(\theta_2)} \leq P^{ID}(\theta_1) + \frac{\Phi(\theta_1)}{Q(\theta_1)} = \frac{T(\theta_1)}{Q(\theta_1)} = P^{AD}(\theta_1).$$

That is, P^{AD} is nonincreasing. Therefore, condition 3 in Theorem 3 holds. ■

Proof of Theorem 5.

In the following we prove part (a). The proof of part (b) is similar to but easier than the proof of part (a), and is omitted.

Suppose that $\Pi^* = \Pi_\infty^{ID}$. By Corollary 2, there exist some second best quantity function Q^* and some nonincreasing function $P^* : \{\theta \in \Theta : Q^*(\theta) > 0\} \rightarrow \mathbb{R}$ such that

$$Q^*(\theta) \in \arg \max_{q \geq 0} \{S(q, \theta) - P^*(\theta)q\} \quad (9)$$

whenever $Q^*(\theta) > 0$. By Assumption 2, $Q^*(\theta)$ is the unique solution of $\max_{q \geq 0} H(q, \theta)$ for all $\theta \in \Theta$ except possibly $\underline{\theta}$ and $\bar{\theta}$. By Berge Maximum Theorem, Q^* is continuous on $\Theta \setminus \{\underline{\theta}, \bar{\theta}\}$.

Let (Q_n^{ADF}, T_n^{ADF}) be an n -ADF-optimal outcome with a finite n . Apply Theorem 3. (Q_n^{ADF}, T_n^{ADF}) is associated with some ϕ, A, M, P as described in Theorem 3. Since M is nondecreasing, P is nonincreasing, and (M, P) takes at most n values, we can let

$$(M(\theta), P(\theta)) = \begin{cases} (m_1, p_1) & \text{if } \theta \in \Theta_1 \\ \vdots & \vdots \\ (m_{n'}, p_{n'}) & \text{if } \theta \in \Theta_{n'} \end{cases}$$

where $\{\Theta_1, \dots, \Theta_{n'}\}$ is a partition of A , $n' \leq n$, each Θ_i ($i = 1, \dots, n'$) is nonempty, and

$$\theta_i < \theta_j \text{ whenever } 1 \leq i < j \leq n' \text{ and } \theta_i \in \Theta_i \text{ and } \theta_j \in \Theta_j,$$

$$m_1 \leq m_2 \leq \dots \leq m_{n'},$$

$$p_1 \geq p_2 \geq \dots \geq p_{n'}.$$

Moreover, for $\theta \in \Theta_i$,

$$Q_n^{ADF}(\theta) \in \arg \max_{q \geq m_i} \{S(q, \theta) - p_i q\}. \quad (10)$$

Claim 1: For every $i = 1, \dots, n'$, exactly one of the following occurs: (I) $Q_n^{ADF}(\theta) > Q^*(\theta)$ for all $\theta \in \Theta_i$; (II) $Q_n^{ADF}(\theta) < Q^*(\theta)$ for all $\theta \in \Theta_i$; (III) there exists some $\mu_i \in \Theta_i$ such that $Q_n^{ADF}(\mu_i) = Q^*(\mu_i)$.

To prove Claim 1, suppose first that both (I) and (II) are false. Then $Q_n^{ADF}(\theta_1) > Q^*(\theta_1)$ and $Q_n^{ADF}(\theta_2) < Q^*(\theta_2)$ for some $\theta_1, \theta_2 \in \Theta_i$. Notice that Q^* is continuous (by Assumption 2) and nondecreasing, and Q_n^{ADF} is nondecreasing. If $\theta_1 < \theta_2$, then

there exists some $\mu_i \in (\theta_1, \theta_2)$ such that $Q_n^{ADF}(\mu_i) = Q^*(\mu_i)$. If $\theta_2 < \theta_1$, then $p_i \leq P^*(\theta_1)$ (otherwise $Q_n^{ADF}(\theta_1) \leq Q^*(\theta_1)$, a contradiction), and $p_i \geq P^*(\theta_2)$ (otherwise $Q_n^{ADF}(\theta_2) \geq Q^*(\theta_2)$, a contradiction), and $P^*(\theta_2) \geq P^*(\theta_1)$ (because P^* is nonincreasing), and then $P^*(\cdot) = p_i$ on $[\theta_2, \theta_1]$, and then $Q_n^{ADF}(\theta) = Q^*(\theta)$ for almost all $\theta \in (\theta_2, \theta_1)$ (otherwise Q_n^{ADF} is not n -ADF-optimal, a contradiction). Therefore, Claim 1 is true.

For every $i = 1, \dots, n'$, we take $\mu_i \equiv \sup \Theta_i$ if case I in Claim 1 occurs, and take $\mu_i \equiv \inf \Theta_i$ if case II in Claim 1 occurs, and take μ_i such that $Q_n^{ADF}(\mu_i) = Q^*(\mu_i)$ if case III in Claim 1 occurs.

Now we define

$$P_n^{ID}(\theta) \equiv \begin{cases} P^*(\mu_1) & \text{if } \theta \in \Theta_1 \\ \vdots & \vdots \\ P^*(\mu_{n'}) & \text{if } \theta \in \Theta_{n'} \end{cases}.$$

Clearly, $P_n^{ID} : A \rightarrow \mathbb{R}$ is nonincreasing because P^* is.

We will construct an outcome (Q_n^{ID}, T_n^{ID}) such that

$$Q_n^{ID}(\theta) = 0 \text{ for } \theta \in \Theta \setminus A,$$

$$Q_n^{ID}(\theta) \in \arg \max_{q \geq 0} \{S(q, \theta) - P^*(\mu_i)q\} \text{ for } \theta \in \Theta_i, \quad (11)$$

$$T_n^{ID}(\theta) = S(Q_n^{ID}(\theta), \theta) - \int_{\theta}^{\theta} S_{\theta}(Q_n^{ID}(x), x) dx \text{ for } \theta \in \Theta.$$

By Theorem 2, any such (Q_n^{ID}, T_n^{ID}) is n -ID-implementable.

Notice that $S(Q_n^{ID}(\underline{\theta}), \underline{\theta}) - T_n^{ID}(\underline{\theta}) = 0 \leq S(Q_n^{ADF}(\underline{\theta}), \underline{\theta}) - T_n^{ADF}(\underline{\theta})$. In order to apply Lemma 1, we will prove, for every $i = 1, \dots, n'$, that $\inf \Theta_i < \theta < \mu_i$ implies $Q_n^{ADF}(\theta) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$, and that $\mu_i < \theta < \sup \Theta_i$ implies $Q_n^{ADF}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$.

Pick any $i = 1, \dots, n'$ and any $\theta \in \Theta_i$. Suppose that case I in Claim 1 occurs. Then $\theta \leq \sup \Theta_i = \mu_i$. If the constraint $q \geq m_i$ is not binding in problem (10) for type μ_i , then $p_i \leq P^*(\mu_i) \leq P^*(\theta)$. Comparing (9), (10) and (11), we can select $Q_n^{ID}(\theta)$ such that $Q_n^{ADF}(\theta) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$. If the constraint $q \geq m_i$ is binding in problem (10) for type μ_i , then $Q_n^{ADF}(\theta) = m_i = Q_n^{ADF}(\mu_i) \geq Q^*(\mu_i)$. Since $P^*(\mu_i) \leq P^*(\theta)$, comparing (9) and (10), we can select $Q_n^{ID}(\theta)$ such that $Q^*(\mu_i) = Q_n^{ID}(\mu_i) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$.

Suppose that case II in Claim 1 occurs. Then $\theta \geq \inf \Theta_i = \mu_i$. Then $p_i \geq P^*(\mu_i) \geq P^*(\theta)$. Comparing (9), (10) and (11), we can select $Q_n^{ID}(\theta)$ such that $Q_n^{ADDF}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$.

Suppose that case III in Claim 1 occurs and the constraint $q \geq m_i$ is not binding in problem (10) for type μ_i . Then $p_i = S_q(Q_n^{ADDF}(\mu_i), \mu_i) = S_q(Q^*(\mu_i), \mu_i) = P^*(\mu_i)$. For $\inf \Theta_i < \theta < \mu_i$, we have $p_i = P^*(\mu_i) \leq P^*(\theta)$, so that we can select $Q_n^{ID}(\theta)$ such that $Q_n^{ADDF}(\theta) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$ whenever $Q_n^{ADDF}(\theta) \geq Q^*(\theta)$. If $Q_n^{ADDF}(\theta) < Q^*(\theta)$, then, for small $\varepsilon > 0$ and $x \in (\theta - \varepsilon, \theta]$, we have $Q_n^{ADDF}(x) < Q^*(x)$ (since Q^* is continuous and Q_n^{ADDF} is nondecreasing) and $p_i = P^*(x)$, and then Q_n^{ADDF} is not n -ADF-optimal because $Q_n^{ADDF}(x)$ can be reselected as $Q^*(x)$ for all $x \in (\theta - \varepsilon, \theta]$ to raise profit, a contradiction. For $\mu_i < \theta < \sup \Theta_i$, we have $p_i = P^*(\mu_i) \geq P^*(\theta)$, so that we can select $Q_n^{ID}(\theta)$ such that $Q_n^{ADDF}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$ whenever $Q_n^{ADDF}(\theta) \leq Q^*(\theta)$. If $Q_n^{ADDF}(\theta) > Q^*(\theta)$, then, for small $\varepsilon > 0$ and $x \in [\theta, \theta + \varepsilon)$, we have $Q_n^{ADDF}(x) > Q^*(x)$ (again since Q^* is continuous and Q_n^{ADDF} is nondecreasing) and $p_i = P^*(x)$, and then Q_n^{ADDF} is not n -ADF-optimal because $Q_n^{ADDF}(x)$ can be reselected as $Q^*(x)$ for all $x \in [\theta, \theta + \varepsilon)$ to raise profit, a contradiction.

Suppose that case III in Claim 1 occurs and the constraint $q \geq m_i$ is binding in problem (10) for type μ_i . Then $p_i \geq P^*(\mu_i)$ and $m_i = Q_n^{ADDF}(\mu_i) = Q^*(\mu_i)$. For $\inf \Theta_i < \theta < \mu_i$, we have $Q_n^{ADDF}(\theta) = m_i = Q^*(\mu_i)$. Since $P^*(\mu_i) \leq P^*(\theta)$, we can select $Q_n^{ID}(\theta)$ such that $Q^*(\mu_i) = Q_n^{ID}(\mu_i) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$. For $\mu_i < \theta < \sup \Theta_i$, we have $p_i \geq P^*(\mu_i) \geq P^*(\theta)$, so that we can select $Q_n^{ID}(\theta)$ such that $Q_n^{ADDF}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$ whenever $Q_n^{ADDF}(\theta) \leq Q^*(\theta)$. Repeating our previous logic, one can show that $Q_n^{ADDF}(\theta) > Q^*(\theta)$ is impossible.

Apply Lemma 1, we see that $\Pi_n^{ID} \geq \Pi_n^{ADDF}$. If $\Pi_n^{ID} = \Pi_n^{ADDF}$, then it must be the case that (Q_n^{ADDF}, T_n^{ADDF}) and (Q_n^{ID}, T_n^{ID}) make the same profit and (Q_n^{ID}, T_n^{ID}) is n -ID-optimal. Then Lemma 1 implies that $(Q_n^{ADDF}(\theta), T_n^{ADDF}(\theta)) = (Q_n^{ID}(\theta), T_n^{ID}(\theta))$ for almost all $\theta \in \Theta$. It completes the proof of part (a). ■

References

BERGEMANN, D., J. SHEN, Y. XU, AND E. M. YEH (2011): ‘‘Mechanism Design with Limited Information: The Case of Nonlinear Pricing,’’ 2nd International ICST Conference on Game Theory for Networks, Shanghai, 2011.

- CHU, L. Y., AND D. E. M. SAPPINGTON (2007): “Simple cost-sharing contracts,” *American Economic Review*, pp. 419–428.
- KOLAY, S., AND G. SHAFFER (2003): “Bundling and Menus of Two-Part Tariffs,” *The Journal of Industrial Economics*, 51(3), 383–403.
- KOLAY, S., G. SHAFFER, AND J. A. ORDOVER (2004): “All-units discounts in retail contracts,” *Journal of Economics & Management Strategy*, 13(3), 429–459.
- LAFFONT, J.-J., AND D. MARTIMORT (2002): *The Theory of Incentives: the Principal-agent Model*. Princeton University Press.
- MASKIN, E., AND J. RILEY (1984): “Monopoly with Incomplete Information,” *RAND Journal of Economics*, 15(2), 171–196.
- MILGROM, P., AND I. SEGAL (2002): “Envelope Theorems for Arbitrary Choice Sets,” *Econometrica*, 70(2), 583–601.
- MILGROM, P., AND C. SHANNON (1994): “Monotone comparative statics,” *Econometrica*, 62(1), 157–180.
- MIRAVETE, E. J. (2007): “The Limited Gains From Complex Tariffs,” <http://www.eugeniomiravete.com/papers/EJM-Gains.pdf>.
- MUSSA, M., AND S. ROSEN (1978): “Monopoly and Product Quality,” *Journal of Economic Theory*, 18(2), 301–317.
- ROGERSON, W. P. (2003): “Simple menus of contracts in cost-based procurement and regulation,” *American Economic Review*, 93(3), 919–926.
- TOM, W. K., D. A. BALTO, AND N. W. AVERITT (1999): “Anticompetitive Aspects of Market-Share Discounts and Other Incentives to Exclusive Dealing,” *Antitrust Law Journal*, 67, 615.
- TOPKIS, D. M. (1978): “Minimizing a submodular function on a lattice,” *Operations Research*, 26(2), 305–321.
- WILSON, R. (1993): *Nonlinear Pricing*. Oxford University Press New York.
- WONG, A. C. L. (2009): “The Choice of the Number of Varieties: Justifying Simple Mechanisms,” <http://ssrn.com/abstract=1366228>.