Menu Implementability and Practical Pricing Schemes

Adam Chi Leung Wong*

August 24, 2011

Abstract

We develop a general theory to characterize the set of direct mechanisms of nonlinear pricing that are implementable by practical pricing schemes. Here a practical pricing scheme is required to take the form of a menu of tariff options, where the set of admissible tariff options and the number of tariff options in the menu are pre-specified. Our results are applied to compare the maximum profits of three forms of pricing schemes: incremental discounts, all-units discounts, and quantity forcing. We found that, when the number of 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".

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

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.

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 direct selling mechanisms? 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 general theory level, 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. At the second level, the application level, we consider three forms of pricing schemes to be practical: incremental discounts (ID), all-units discounts (AUD), and quantity forcing (QF). These three forms are illustrated in Figure 1. Under ID, illustrated in the left panel, marginal prices of successive units decline in steps. We also allow a fixed fee under ID.[1] Under AUD, illustrated in the middle panel, the per-unit price progressively drops when the order size exceeds certain thresholds. We also allow a minimum purchase under AUD.[2] Under QF, illustrated in the right panel, only several quantities, each associated with a gross price, are offered for consumers to choose.[3]

In order to motivate our general theory, let us notice two things from Figure 1. First, the three pricing 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. 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.

Second, each of the three forms can be regarded as offering a menu of simple tariff options for consumers to select. Offering an ID pricing scheme with n blocks is equiv-

[1] If 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.

[2] If one insists that AUD 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.

[3] I was originally interested in comparing ID and AUD, but I later found that it is theoretically meaningful to add QF into the picture. See e.g. the paragraph right after Theorem 8.

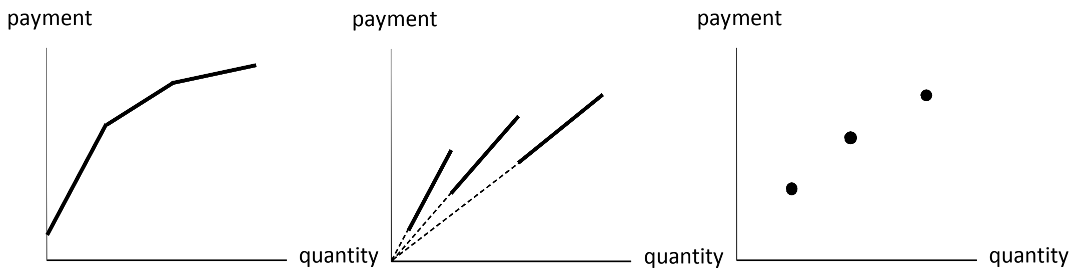

Figure 1: Incremental discounts (left), All-units discounts (middle) and Quantity forcing (right)

alent to offering n two-part tariffs.[4] Offering an AUD pricing scheme with n blocks is equivalent to offering n "minimum purchase tariffs" (see Section 6). Offering a QF pricing scheme with n blocks (or n points here) is equivalent to offering n quantity-payment pairs. That makes it possible to build an elegant unifying framework to analyze all the three forms.

Our theory 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-specified. From revelation principle, any pricing scheme can be represented as a direct (selling) mechanism, under which consumers are asked to report their types and each reported type is associated with a quantity-payment pair. The main theorems (Theorems 4 and 5) of our general theory characterize the set of direct mechanisms that is implementable by a menu of tariff options given the number n and the set $\mathcal{C}$. We say the direct mechanisms in this set is *menu implementable* with respect to $(n, \mathcal{C})$, or simply $(n, \mathcal{C})$-*implementable*. 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.

Applying our theorems of menu implementability to ID, AUD and QF, we are able to characterize the set of direct mechanisms implementable by each of the three forms, given any number of blocks n. (Theorems 6, 7 and 8) Those mechanisms are called n-ID implementable, n-AUD implementable, and n-QF implementable. When the number of blocks is unrestricted, i.e. $n = \infty$, we simply call them ID, AUD or QF implementable. Then we are able to derive many results with our characterizations.

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

First, ID implementability is the most restrictive, and in this sense we say that ID has the smallest "implementation power" among the three forms.[5] (Theorem 9) It follows that when the number of blocks is unrestricted, ID can never perform better than AUD.

Second, we derive the condition under which ID or AUD can attain the Maskin-Riley second best profit, when the number of blocks is unrestricted. It amounts to derive the condition under which the second best direct mechanism is ID (or AUD) implementable.

Third, if ID can attain the second best profit when the number of blocks is unrestricted, then it performs the best among the three forms when the number of blocks is restricted, i.e. $n < \infty$. Proving this result is harder because, for finite n, the three concepts n-ID, n-AUD and n-QF implementability do not imply one another. However, it can be done by showing that, under the premise and given any n-AUD or n-QF implementable direct mechanism, there exists an n-ID direct mechanism that better approximate the Maskin-Riley second best solution. In this sense, we say that ID has the largest "approximation power" among the three forms. (Theorem 10) 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.

Although the use of AUD in intermediate-goods markets is common,[6] there is little theoretical analysis on it in the literature.[7] Why does a seller use AUD rather than ID? It has been informally argued that the use of AUD is anticompetitive or exclusionary (to exclude entry of competitors, or induce downstream retailers to promote the products at the expense of other substitute products).[8] So one might wonder whether a monopolist without fear of competition would find AUD superior. The implication of our results on this has two sides. First, in principle both AUD 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 AUD for such a monopolist.[9]

[5] It is trivial that QF implementability is the least restrictive. Indeed, if the number of blocks is infinity, QF allows any nonlinear pricing scheme so that we are back to Maskin-Riley.

[6] According to Kolay, Shaffer, and Ordover (2004), AUD is used by Coca-Cola, Irish Sugar British Airways, and Michelin.

[7] Kolay, Shaffer, and Ordover (2004) is an exception. They concentrate on a situation where an upstream firm faces a downstream firm with only two types, and find that AUD is always better than ID under incomplete information.

[8] See, for example, Tom, Balto, and Averitt (1999).

[9] For example, ID must outperform AUD if marginal cost is constant, the hazard rate of types'

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 related 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 (2010) 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 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. Section 2 describes the environment. Sections 3 – 5 are the general theory part. Sections 3 and 4 build the basic concepts. Section 5 presents our characterization of menu implementability. Sections 6 – 7 are the application part. Section 6 characterizes the set of direct mechanisms that is implementable by each of the three forms of pricing schemes. Section 7 compares the performances of the three forms. Section 8 concludes. The proofs that are not in the main text are in Appendix.

---

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 AUD 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.

2 Environment

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

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

where $S:\chi\times\Theta\to\mathbb{R}$ 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, whose domain χ can be any closed subset of $\mathbb{R}$ that includes 0. The argument θ denotes the consumer's type (or preference parameter), whose domain Θ is an interval $\left[\underline{\theta},\bar{\theta}\right]$. 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 $c:\chi\to\mathbb{R}$ 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

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

The following assumptions are all we need for the next three sections (general theory part).

Assumption 1 *$S(0,\theta)=0$ for all $\theta\in\Theta$. For any $q\in\chi$, $S(q,\theta)$ is absolutely continuous, strictly increasing, and 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 our general theory allows quantity to be discrete or continuous, and Assumption 1 is very weak, so that our results can be adapted to many other principal-agent settings. The only essential restriction in Assumption 1 is its last sentence, which is the strict Spence-Mirrlees property.

3 Direct mechanisms, menus, and menu-direct mechanisms

The following definitions about direct mechanisms (or direct selling mechanisms, or direct revelation mechanisms) are standard. Theorem 1 is also well known and we state it without proof for future reference.

Definition 1 *A* direct mechanism*, written as (Q,T), is a pair of functions $Q:\Theta\to\chi$ and $T:\Theta\to\mathbb{R}$. Q is called the quantity function and T the payment function of (Q,T). A direct mechanism (Q,T) is said to be* incentive-compatible (IC) *if for any $\theta,\theta'\in\Theta$, we have*

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

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

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

If a direct mechanism (Q^{*},T^{*}) maximizes the monopolist's profit (1) subject to IC and IR, we say (Q^{*},T^{*}) is a *second best direct mechanism*, Q^{*} is a *second best quantity function*, and the associated profit, denoted as Π^{*}, is the *second best profit*.

Theorem 1 *A direct mechanism (Q,T) is IC if and only if $Q\left(\cdot\right)$ is nondecreasing and $U\left(\theta_2\right)-U\left(\theta_1\right)=\int_{\theta_1}^{\theta_2}S_\theta\left(Q\left(x\right),x\right)dx$ for any $\theta_1,\theta_2\in\Theta$.*

We now formalize the concepts of tariff option and menu of tariff options.

Definition 2 *A tariff option τ is a function that assigns each quantity $q\in\chi$ a total payment $\tau\left(q\right)\in\mathbb{R}\cup\{\infty\}$ for purchasing q units, such that $\tau\left(q\right)<\infty$ for some $q\in\chi$, and $\liminf_{x\to q}\tau\left(x\right)>-\infty$ for all $q\in\chi$.*[10] *(The interpretation of $\tau\left(q\right)=\infty$ is that purchasing q units is not allowed by τ.) The set $\chi\left(\tau\right)\equiv\{q\in\chi:\tau\left(q\right)<\infty\}$ is called the domain of τ. (The interpretation of $\chi\left(\tau\right)$ is the set of order sizes allowed by τ.)*

[10] As usual, $\liminf_{x\to q}\tau\left(x\right)$ is defined by

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

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, i.e. to buy nothing and pay nothing, is also a tariff option. We use τ_{out} to denote the tariff option that represents the outside option, i.e.

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

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

Definition 3 *A* menu *is a set of tariff options. A* menu-direct mechanism*, written as* $\{\tau_\theta\}_{\theta\in\Theta}$ *or simply* τ_Θ*, is an indexed set of tariff options, in which every tariff option is indexed by* $\theta \in \Theta$.[11] *We say a menu is on* $\mathcal{C}$ *if it is a subset of* $\mathcal{C}$*. We say a menu-direct mechanism is on* $\mathcal{C}$ *if it is a subset of* $\mathcal{C} \cup \{\tau_{out}\}$*.*

Intuitively, a menu is a list of admissible tariff options for each consumer to select one, or none. Once a tariff option is selected, the consumer is free to choose an order size, which determines the corresponding payment according to the tariff option. The concept of menu-direct mechanism is derived from the concept of menu in the spirit of revelation principle. Intuitively, under a menu-direct mechanism, a tariff option (or the outside option) is designed for each type of consumers, and each consumer is asked to report her type in order to determine her tariff option.

Having chosen a tariff option τ, a consumer's problem is to choose order size q to maximize $S(q,\theta) - \tau(q)$. For any tariff option τ and any $\theta \in \Theta$, we write

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

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

Definition 4 *We say a menu* $\tau_B \equiv \{\tau_\beta\}_{\beta\in B}$ induces *a direct mechanism* (Q,T) *if there exists a mapping* $\tau^* : \Theta \to \tau_B \cup \{\tau_{out}\}$*, written as* $\theta \mapsto \tau^*_\theta$*, such that* $T(\theta) =$

[11] The possibility that $\theta_1 \neq \theta_2$ and $\tau_{\theta_1} = \tau_{\theta_2}$ is allowed.

$\tau^*_\theta(Q(\theta))$ *and*

$$S(Q(\theta),\theta)-\tau^*_\theta(Q(\theta))\geq S(q,\theta)-\tau(q)$$

for any $\theta\in\Theta$*, any* $q\in\chi$*, and any* $\tau\in\tau_B\cup\{\tau_{out}\}$*. (We may call* τ^* *a consumers' best response to* τ_B*.)*

Definition 5 *We say a menu-direct mechanism* τ_Θ induces *a direct mechanism* (Q,T) *if* $T(\theta)=\tau_\theta(Q(\theta))$ *and* $Q(\theta)\in D(\tau_\theta,\theta)$ *for any* $\theta\in\Theta$*.*

Like direct mechanisms, menu-direct mechanisms can have the properties of incentive compatibility and individual rationality.

Definition 6 *We say a menu-direct mechanism* τ_Θ *is* IC *if for each* $\theta\in\Theta$*, there exists a* $Q(\theta)$ *such that*

$$S(Q(\theta),\theta)-\tau_\theta(Q(\theta))\geq S(q',\theta)-\tau_{\theta'}(q') \text{ for any } (q',\theta')\in\chi\times\Theta.$$

In other words, a menu-direct mechanism τ_Θ *is IC iff*

$$D(\tau_\theta,\theta)\neq\varnothing \text{ and } V(\tau_\theta,\theta)\geq V(\tau_{\theta'},\theta) \text{ for any } \theta,\theta'\in\Theta.$$

We say a menu-direct mechanism τ_Θ *is* IR *if for each* $\theta\in\Theta$*, there exists a* $Q(\theta)$ *such that* $S(Q(\theta),\theta)-\tau_\theta(Q(\theta))\geq 0$*. In other words, a menu-direct mechanism* τ_Θ *is IR iff for any* $\theta\in\Theta$*, either* $D(\tau_\theta,\theta)\neq\varnothing$ *and* $V(\tau_\theta,\theta)\geq 0$*, or* $V(\tau_\theta,\theta)>0$*.*

Remark 1 *In our analysis, a menu-direct mechanism* τ_Θ *always induces some direct mechanism, so that* $D(\tau_\theta,\theta)$ *is always nonempty, and then the IC of* τ_Θ *is simply* $V(\tau_\theta,\theta)\geq V(\tau_{\theta'},\theta)$ *for any* $\theta,\theta'\in\Theta$*, and the IR of* τ_Θ *is simply* $V(\tau_\theta,\theta)\geq 0$ *for any* $\theta\in\Theta$*.*

It is easy to see that a direct mechanism (Q,T) is induced by some menu on $\mathcal{C}$ if and only if (Q,T) is induced by some IC and IR menu-direct mechanism on $\mathcal{C}$.

4 Domination and single crossing properties for tariff options

Definition 7 *For any two tariff options* τ_1 *and* τ_2*, we say* τ_1 *is* dominated *by* τ_2 *(or* τ_2 dominates τ_1*) if* $\tau_1(q)\geq\tau_2(q)$ *for any* $q\in\chi$*, and the inequality is strict for* *some $q \in \chi$.*

Definition 8 *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*

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

and

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

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_1 is less favorable to high quantities than τ_2.

Definition 9 *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_1 is less favorable to incremental quantities than τ_2.*

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

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

Lemma 1 *If (τ_1, τ_2) satisfies tariff increasing differences, then $\left(\tau_1^{\inf}, \tau_2^{\inf}\right)$ satisfies tariff increasing differences.*

5 Menu implementability

The following is the central concept of this paper.

Definition 10 *We say a direct mechanism is $(n, \mathcal{C})$-implementable (where n is a nonnegative integer or ∞) if it is induced by some menu on $\mathcal{C}$ that has at most n tariff options. We say a direct mechanism is $\mathcal{C}$-implementable if it is $(\infty, \mathcal{C})$-implementable (i.e. it is induced by some menu on $\mathcal{C}$).*

From the logic of revelation principle, one can see the following proposition, which we state without proof.

Proposition 1 *An $(n, \mathcal{C})$-implementable direct mechanism must be IC and IR. A direct mechanism is $(n, \mathcal{C})$-implementable if and only if it is induced by some IC and IR menu-direct mechanism on $\mathcal{C}$ that has at most n distinct tariff options except τ_{out}.*

The above proposition reveals that any $\mathcal{C}$-implementable direct mechanism is IC and IR. But the converse is not true: an IC and IR direct mechanism might not be $\mathcal{C}$-implementable. We have the following result though.

Theorem 2 *Suppose that a menu-direct mechanism τ_Θ induces a direct mechanism (Q, T). Then τ_Θ is IR if and only if (Q, T) is IR. Moreover, if τ_Θ is IC, then (Q, T) is IC.*

Proof. Since τ_Θ induces (Q, T),

$$V(\tau_\theta, \theta) = \sup_{q \in \mathcal{X}} \{S(q, \theta) - \tau_\theta(q)\} = S(Q(\theta), \theta) - \tau_\theta(Q(\theta)) = S(Q(\theta), \theta) - T(\theta),$$

$$V(\tau_{\theta'}, \theta) = \sup_{q \in \mathcal{X}} \{S(q, \theta) - \tau_{\theta'}(q)\} \geq S(Q(\theta'), \theta) - \tau_{\theta'}(Q(\theta')) = S(Q(\theta'), \theta) - T(\theta').$$

Therefore, $S(Q(\theta), \theta) - T(\theta) \geq 0$ if and only if $V(\tau_\theta, \theta) \geq 0$. It follows that (Q, T) is IR if and only if τ_Θ is IR. Moreover, if τ_Θ is IC, then, for any $\theta, \theta' \in \Theta$, we have $V(\tau_\theta, \theta) \geq V(\tau_{\theta'}, \theta)$, and hence $S(Q(\theta), \theta) - T(\theta) \geq S(Q(\theta'), \theta) - T(\theta')$. Therefore (Q, T) is IC if τ_Θ is IC. ■

If the monopolist is free to offer any pricing scheme, it is well understood that the monopolist faces only incentive compatibility constraint and individual rationality constraint when designing a direct mechanism. Theorem 2 reveals that restricting to offering a menu, rather than any pricing scheme, does not have an impact on individual rationality constraint. However, restricting to offering a menu generally strengthens incentive compatibility constraint, in the sense that IC for a menu-direct mechanism is stronger than IC for the corresponding direct mechanism. The last result is intuitive. A direct mechanism specifies a particular quantity-payment pair 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 pair designated for that false type. In contrast, under a menu-direct mechanism a tariff option is specified after reporting. A consumer has some flexibility to choose the quantity-payment pair 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 pair allowed by the tariff option designated for that false type.

Then what is the additional restriction of the IC of a menu-direct mechanism on top of the ordinary IC (i.e. IC of the corresponding direct mechanism)? The following definition and theorem provide partial answers.

Definition 11 *We say a menu-direct mechanism $\tau_\Theta = \{\tau_\theta\}_{\theta\in\Theta}$ is* increasing differences monotonic *if for any $\theta_1, \theta_2 \in \Theta$ with $\theta_1 < \theta_2$, $(\tau_{\theta_1}, \tau_{\theta_2})$ satisfies tariff increasing differences. Similarly, we say τ_Θ is* single crossing monotonic *if for any $\theta_1, \theta_2 \in \Theta$ with $\theta_1 < \theta_2$, $(\tau_{\theta_1}, \tau_{\theta_2})$ satisfies tariff single crossing property.*

Theorem 3 *Suppose that a menu-direct mechanism τ_Θ induces an IC direct mechanism (Q, T). Then τ_Θ is IC if either one of the following two conditions holds.*

1. *τ_Θ is increasing differences monotonic.*
2. *(i) τ_Θ is single crossing monotonic; and (ii) for any $\theta, \theta' \in \Theta$, $\tau_\theta(Q(\theta)) \leq \tau_{\theta'}(Q(\theta))$.*

Proof. Suppose that τ_Θ induces (Q, T) and (Q, T) is IC, but τ_Θ is not IC. Then there exist some $\theta_1, \theta_2 \in \Theta$ and some $q' \in \chi$ such that $U(\theta_1) = V(\tau_{\theta_1}, \theta_1) < S(q', \theta_1) - \tau_{\theta_2}(q') \leq S(q', \theta_1) - \tau_{\theta_2}^{\inf}(q')$. (Recall that $U(\theta) \equiv S(Q(\theta), \theta) - T(\theta)$ and $\tau_\theta^{\inf}(q) \equiv \liminf_{x\to q} \tau_\theta(x)$ for any $(q, \theta) \in \chi \times \Theta$.)

Let $\chi' \equiv \chi \cap \left(\left[Q(\underline{\theta}), Q(\bar{\theta})\right] \cup \{q'\}\right)$. For any $\theta, \theta' \in \Theta$, the set

$$D\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right) \equiv \arg\max_{q\in\chi'} \left\{S(q, \theta) - \tau_{\theta'}^{\inf}(q)\right\}$$

is nonempty since χ' is compact and the objective function is upper semi-continuous in q (from the lower semi-continuity of $\tau_{\theta'}^{\inf}$ and the continuity of $S(\cdot, \theta)$).

For any $\theta \in \Theta$, we have $Q(\theta) \in D(\tau_\theta, \theta)$, since τ_Θ induces (Q, T). Then $S(q, \theta) - \tau_\theta(q)$ must be upper semi-continuous in q at $Q(\theta)$, and hence $\tau_\theta(Q(\theta)) = \tau_\theta^{\inf}(Q(\theta))$. Then $Q(\theta) \in D\left(\tau_\theta^{\inf}, \theta\right)$. Since $Q(\theta) \in \chi'$, we also have $Q(\theta) \in D\left(\tau_\theta^{\inf}, \theta; \chi'\right)$.

For any $\theta \in \Theta$, let

$$V\left(\tau_{\theta_2}^{\inf}, \theta; \chi'\right) \equiv \sup_{q \in \chi'} \left\{S(q,\theta) - \tau_{\theta_2}^{\inf}(q)\right\} = \sup_{q \in \chi'} \left\{S(q,\theta) - \tau_{\theta_2}(q)\right\}.$$

Since $q' \in \chi'$, we have $U(\theta_1) < S(q', \theta_1) - \tau_{\theta_2}^{\inf}(q') \leq V\left(\tau_{\theta_2}^{\inf}, \theta_1; \chi'\right)$. Since $Q(\theta_2) \in \chi'$, we have $U(\theta_2) = V(\tau_{\theta_2}, \theta_2) = V\left(\tau_{\theta_2}^{\inf}, \theta_2\right) = V\left(\tau_{\theta_2}^{\inf}, \theta_2; \chi'\right)$. In other words, the value of $U(\cdot)$ is strictly below the the value of $V\left(\tau_{\theta_2}^{\inf}, \cdot; \chi'\right)$ at θ_1, but the values of these two functions are equal at θ_2. By IC of (Q, T) and Theorem 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)), $V\left(\tau_{\theta_2}^{\inf}, \cdot; \chi'\right)$ is absolutely continuous, and $\partial V\left(\tau_{\theta_2}^{\inf}, x; \chi'\right)/\partial x = S_\theta\left(d\left(\tau_{\theta_2}^{\inf}, x; \chi'\right), x\right)$ for almost every $x \in \Theta$, where $d\left(\tau_{\theta_2}^{\inf}, \theta; \chi'\right)$ is any selection from $D\left(\tau_{\theta_2}^{\inf}, \theta; \chi'\right)$.

It follows that, if $\theta_1 < \theta_2$, then there is some $x_0 \in [\theta_1, \theta_2]$ such that $U(x_0) < V\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$ and $U'(x_0) > \partial V\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)/\partial x_0$, and then $S_\theta(Q(x_0), x_0) > S_\theta\left(d\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right), x_0\right)$, and then $Q(x_0) > d\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$. (Notice that Assumption 1 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 $U(x_0) < V\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$ and $Q(x_0) < d\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$.

So far neither condition 1 nor condition 2 is imposed yet. In the following we show that either one would derive contradiction.

Case of condition 1. Increasing differences monotonicity and Lemma 1 imply that $S(q, \theta) - \tau_{\theta'}^{\inf}(q)$ satisfies increasing differences in $(q; \theta')$. By monotone comparative statics (see Topkis (1978) or Milgrom and Shannon (1994)), we obtain that, given any $\theta \in \Theta$, $D\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right)$ 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 $D\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right) \leq_s D\left(\tau_{\theta}^{\inf}, \theta; \chi'\right)$, then there is some $d \in D\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right)$ such that $d \leq Q(\theta)$. (Recall that $Q(\theta) \in D\left(\tau_{\theta}^{\inf}, \theta; \chi'\right)$ and $D\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right)$ is nonempty.) Similarly, if $\theta' > \theta$, then there is some $d \in D\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right)$ such that $d \geq Q(\theta)$. Now, we pick a selection $d\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right)$ from $D\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right)$ such that

$$d\left(\tau_{\theta'}^{\inf}, \theta; \chi'\right) \begin{array}{c} \leq \\ = \\ \geq \end{array} Q(\theta) \text{ if } \theta' \begin{array}{c} < \\ = \\ > \end{array} \theta.$$

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

Case of condition 2. Since $d\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$ is some selection of $D\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$, there exists a sequence $\{d^i\}$ on $\chi(\tau_{\theta_2})$ such that $d^i \to d\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$ and $S(d^i, x_0) - \tau_{\theta_2}(d^i) \to V\left(\tau_{\theta_2}^{\inf}, x_0; \chi'\right)$. Consider $\theta_1 < \theta_2$. From our previous claim, for all large enough i, we have

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

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 2(ii) implies $\tau_{x_0}(Q(x_0)) \leq \tau_{\theta_2}(Q(x_0))$. Then single crossing monotonicity 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. ■

Theorem 3 can be informally understood as follows. Working with direct mechanism, it is well known that IC requires higher type consumers to purchase more. Working with menu-direct mechanism, we have a natural extension: IC requires higher type consumers to pick menus that are favorable to purchasing more. The latter is formalized by increasing differences monotonicity. If we replace increasing differences monotonicity by the related but weaker condition single crossing monotonicity, then an additional condition, condition 2(ii) in Theorem 3, is needed.

Remark 2 *Notice that in Theorem 3, condition 2(ii) clearly has to be satisfied anyway, as long as τ_Θ is IC and induces (Q, T). So actually condition 1, together with IC of (Q, T), implies condition 2. However, the sufficiency of condition 1 is worth knowing on top of knowing the sufficiency of condition 2, because for some applications condition 1 is more useful. Indeed, in our applications in Section 6, checking increasing differences monotonicity is not harder than checking single crossing monotonicity and hence we do not need to worry about condition 2(ii).*

Proposition 1, Theorem 2 and Theorem 3, put together, provide two sets of sufficient conditions for a direct mechanism to be $(n, \mathcal{C})$-implementable. The corresponding necessity requires some restrictions on $\mathcal{C}$.

Definition 12 *We say a set $\mathcal{C}$ of tariff options is* closed *if it is closed under the* *product topology in the space of tariff options, i.e. whenever a net in $\mathcal{C}$ pointwise converges to a tariff option, this tariff option is also in $\mathcal{C}$.*

Definition 13 *We say a set $\mathcal{C}$ of tariff options is* increasing differences comparable *if for any two tariff options τ_1 and τ_2 in $\mathcal{C}$ which are not dominated by each other, either (τ_1, τ_2) or (τ_2, τ_1) satisfies tariff increasing differences. Similarly, we say $\mathcal{C}$ is* single crossing comparable *if for any two tariff options τ_1 and τ_2 in $\mathcal{C}$ which are not dominated by each other, either (τ_1, τ_2) or (τ_2, τ_1) satisfies tariff single crossing property.*

The proof of our main results requires the following two lemmas, which are proved in Appendix.

Lemma 2 *If a direct mechanism (Q, T) is $(n, \mathcal{C})$-implementable with either n finite or $\mathcal{C} \cup \{\tau_{out}\}$ closed, then the associated menu $\tau_B \equiv \{\tau_\beta\}_{\beta \in B}$ (that is on $\mathcal{C}$, has at most n elements, and induces (Q, T)) can be chosen such that*

1. *for any $q \in \chi$, $\arg\min_{\tau \in \tau_B \cup \{\tau_{out}\}} \tau(q)$ is nonempty, and*
2. *the tariff options in τ_B do not dominate one another.*

Lemma 3 *If tariff options $\tau_1, \ldots, \tau_n$ do not dominate one another, and (τ_1, τ_2), (τ_2, τ_3), $\ldots$, (τ_{n-1}, τ_n) and (τ_n, τ_1) satisfy tariff single crossing property or tariff increasing differences, then $\tau_1 = \cdots = \tau_n$.*

Now we are ready to state and prove the main results of our general theory.

Theorem 4 *A direct mechanism (Q, T) is $(n, \mathcal{C})$-implementable if it is IC and IR, and is induced by some menu-direct mechanism τ_Θ on $\mathcal{C}$ such that:*

1. *τ_Θ is increasing differences monotonic, and*
2. *$|\tau_\Theta \backslash \{\tau_{out}\}| \leq n$.*

Moreover, if (i) n is finite or $\mathcal{C} \cup \{\tau_{out}\}$ is closed, and (ii) $\mathcal{C}$ is increasing differences comparable, then the above set of sufficient conditions for $(n, \mathcal{C})$-implementability is also necessary, and the menu-direct mechanism τ_Θ can be chosen such that the tariff options in τ_Θ do not dominate one another.

Proof. Sufficiency. It is a straightforward corollary of Proposition 1, Theorem 2 and Theorem 3.

Necessity. Suppose that the assumptions (i) and (ii) hold and a direct mechanism (Q,T) is $(n,\mathcal{C})$-implementable. By Proposition 1, (Q,T) is IC and IR. By definition of $(n,\mathcal{C})$-implementability, (Q,T) is induced by some menu $\tau_B \equiv \{\tau_\beta\}_{\beta\in B} \subset \mathcal{C}$ with $|\tau_B| \leq n$. That is, there exists a consumers' best response $\tau^* : \Theta \to \tau_B \cup \{\tau_{out}\}$, written as $\theta \mapsto \tau^*_\theta$, such that $T(\theta) = \tau^*_\theta(Q(\theta))$ and

$$S(Q(\theta),\theta) - \tau^*_\theta(Q(\theta)) \geq S(q,\theta) - \tau(q)$$

for any $\theta \in \Theta$, any $q \in \chi$, and any $\tau \in \tau_B \cup \{\tau_{out}\}$. Let τ^*_Θ be the range of τ^*. By assumption (i) and Lemma 2, 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 $\tau_1 \trianglelefteq \tau_2$ if (τ_1,τ_2) satisfies tariff increasing differences. Obviously, $\trianglelefteq$ is a reflexive binary relation over tariff options. Increasing differences comparability of $\mathcal{C}$ (assumption (ii)) carries over to $\mathcal{C}\cup\{\tau_{out}\}$ (because $\tau_{out} \trianglelefteq \tau$ for any tariff option τ) and hence also carries over to τ^*_Θ (because $\tau^*_\Theta \subset \mathcal{C}\cup\{\tau_{out}\}$). Thus, $\trianglelefteq$ is complete on τ^*_Θ. From Lemma 3, $\trianglelefteq$ is antisymmetric on τ^*_Θ. From reflexivity, completeness and Lemma 3, $\trianglelefteq$ is transitive on τ^*_Θ. Therefore, $\trianglelefteq$ 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 $\theta \in \Theta$ (otherwise τ^*_Θ is not IC). Let $\trianglelefteq_s$ be the strong set order on τ^*_Θ induced by $\trianglelefteq$. That is, for any two subsets $\mathcal{T}_1, \mathcal{T}_2$ of τ^*_Θ, we say $\mathcal{T}_1 \trianglelefteq_s \mathcal{T}_2$ if $\tau_1 \in \mathcal{T}_1$ and $\tau_2 \in \mathcal{T}_2$ and $\tau_2 \trianglelefteq \tau_1$ imply $\tau_1 \in \mathcal{T}_2$ and $\tau_2 \in \mathcal{T}_1$. Then $\trianglelefteq_s$ 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 $((\Theta,\leq),(\mathcal{P}(\tau^*_\Theta),\trianglelefteq_s))$. Consider any $\theta_1,\theta_2 \in \Theta$ such that $\theta_1 < \theta_2$, and we need to show $\mathcal{T}(\theta_1) \trianglelefteq_s \mathcal{T}(\theta_2)$. Suppose $\tau_1 \in \mathcal{T}(\theta_1)$ and $\tau_2 \in \mathcal{T}(\theta_2)$ and $\tau_2 \trianglelefteq \tau_1$. 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 $\tau_2 \trianglelefteq \tau_1$ 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, $\mathcal{T}(\theta_1) \trianglelefteq_s \mathcal{T}(\theta_2)$. We conclude that the mapping $\mathcal{T}$ is nondecreasing on

$((\Theta,\leq),(\mathcal{P}(\tau^*_\Theta),\trianglelefteq_s))$.

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 $\trianglelefteq$), so that it has a nondecreasing selection $\theta \mapsto \tau_\theta$ (see Milgrom and Shannon (1994) Theorem A2). Now $\tau_\Theta \equiv \{\tau_\theta\}_{\theta\in\Theta}$, regarded as a menu-direct mechanism on $\mathcal{C}$, satisfies conditions 1 and 2, and is such that the tariff options in it do not dominate one another. Finally, this τ_Θ induces (Q,T) because $T(\theta) = \tau^*_\theta(Q(\theta)) = \min_{\theta'\in\Theta}\tau^*_{\theta'}(Q(\theta)) = \tau_\theta(Q(\theta))$ for any $\theta\in\Theta$, and

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

for any $\theta\in\Theta$ and any $q\in\chi$. ■

Theorem 5 *A direct mechanism (Q,T) is $(n,\mathcal{C})$-implementable if it is IC and IR, and is induced by some menu-direct mechanism τ_Θ on $\mathcal{C}$ such that:*

1. *τ_Θ is single crossing monotonic,*
2. *for any $\theta,\theta'\in\Theta$, $\tau_\theta(Q(\theta)) \leq \tau_{\theta'}(Q(\theta))$, and*
3. *$|\tau_\Theta\backslash\{\tau_{out}\}| \leq n$.*

Moreover, if (i) n is finite or $\mathcal{C}\cup\{\tau_{out}\}$ is closed, and (ii) $\mathcal{C}$ is single crossing comparable, then the above set of sufficient conditions for $(n,\mathcal{C})$-implementability is also necessary, and the menu-direct mechanism τ_Θ can be chosen such that the tariff options in τ_Θ do not dominate one another.

Proof. **Sufficiency**. It is a straightforward corollary of Proposition 1, Theorem 2 and Theorem 3.

Necessity. Essentially repeat the proof of Theorem 4 except that the binary relation $\trianglelefteq$ is redefined: $\tau_1 \trianglelefteq \tau_2$ if and only if (τ_1,τ_2) satisfies tariff single crossing property. ■

The monopolist's problem is to maximize the profit (1) by choosing a direct mechanism (Q,T), subject to the constraint that (Q,T) is $(n,\mathcal{C})$-implementable, given some exogenous n and $\mathcal{C}$.

Definition 14 *We say a direct mechanism* (Q, T) *is* $(n, \mathcal{C})$*-optimal (where* n *is a nonnegative integer or* ∞*) if it maximizes the profit* (1) *subject to the constraint that* (Q, T) *is* $(n, \mathcal{C})$*-implementable. If* (Q, T) *is an* $(n, \mathcal{C})$*-optimal direct mechanism, we call* Q *an* $(n, \mathcal{C})$*-optimal quantity function, and call the associated profit the* $(n, \mathcal{C})$*-maximum profit. We say a direct mechanism* (Q, T) *is* $\mathcal{C}$*-optimal if it is* $(\infty, \mathcal{C})$*-optimal (i.e. it maximizes the profit subject to the constraint that* (Q, T) *is* $\mathcal{C}$*-implementable). If* (Q, T) *is a* $\mathcal{C}$*-optimal direct mechanism, we call* Q *a* $\mathcal{C}$*-optimal quantity function, and call the associated profit the* $\mathcal{C}$*-maximum profit.*

Results particularly for $(n, \mathcal{C})$-optimal direct mechanism or $(n, \mathcal{C})$-maximum profit are not available, until $\mathcal{C}$ is specified in context of applications, which we do next.

6 Incremental discounts, all-units discounts, and quantity forcing

In this and the next sections (application part), we add the following to supplement Assumption 1.

Assumption 2 *The domain* χ *of quantity is* $\mathbb{R}_+$*. For every* $\theta \in \Theta$*,* $S(q, \theta)$ *is differentiable in* q*.*

It in particular does not require the gross utility S to be increasing or concave in quantity q. Hence, 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).

A *two-part tariff* is a tariff option that is affine. That is, a two-part tariff is of the form $q \mapsto pq + \phi$, where $p \in \mathbb{R}$ and $\phi \in \mathbb{R}$ are the marginal price and the fixed fee associated with this two-part tariff. An *incremental discounts (ID) scheme* is a menu of two-part tariffs offered by the monopolist. If a type θ consumer chooses a two-part tariff characterized by (p, ϕ) and a purchase quantity $q \geq 0$, then her utility is $S(q, \theta) - pq - \phi$. If any consumer chooses none, then her outside option is buying nothing and paying nothing, and her utility is 0. If an ID scheme has at most n options (or blocks), where n could be finite or infinite, then it is called an *n-incremental discounts scheme (n-ID scheme).*

The form of ID schemes is simply a special case of our general concept of menu in Sections 3 – 5 where the set of admissible tariff options $\mathcal{C}$ comprises all two-part tariffs. With such a $\mathcal{C}$, a $\mathcal{C}$-implementable (respectively $(n, \mathcal{C})$-implementable, or $\mathcal{C}$-optimal, or $(n, \mathcal{C})$-optimal) direct mechanism is also said to be ID-implementable (respectively n-ID-implementable, or ID-optimal, or n-ID-optimal), and the $\mathcal{C}$-maximum (respectively $(n, \mathcal{C})$-maximum) profit is also called the ID-maximum (respectively n-ID-maximum) profit. Theorem 4 translates into the following theorem, which we prove in Appendix. In particular, the increasing differences monotonicity condition in Theorem 4 translates into the monotonicity of marginal price (i.e. condition 2 in Theorem 6).

Theorem 6 *A direct mechanism (Q, T) is n-ID-implementable if and only if it is IC and IR, and there exists a function $P : A \to \mathbb{R}$, 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.*[12]

If the above conditions hold, then (Q, T) is induced by the IC and IR menu-direct mechanism $\tau_\Theta \equiv \{\tau_\theta\}_{\theta \in \Theta}$ with τ_θ being τ_{out} for $\theta \in \Theta \backslash A$, and with τ_θ being $q \mapsto P(\theta) q + \Phi(\theta)$ for $\theta \in A$, where $\Phi(\theta) = T(\theta) - P(\theta) Q(\theta)$ for every $\theta \in A$.

We can similarly analyze all-units discounts. A *minimum purchase tariff* is a tariff option that is of the form

$$q \mapsto \begin{cases} pq & \text{if } q \geq m \\ \infty & \text{if } 0 \leq q < m \end{cases},$$

where $m \in \mathbb{R}_+$ and $p \in \mathbb{R}$ are the minimum purchase and the per-unit (or average) price associated with this minimum purchase tariff. An *all-units discounts (AUD) scheme* is a menu of minimum purchase tariffs offered by the monopolist. If a type θ consumer chooses a minimum purchase tariff characterized by (m, p), she has to

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

choose an order size $q \geq m$, and then her utility is $S(q, \theta) - pq$. If any consumer chooses none, then her outside option is buying nothing and paying nothing, and her utility is 0. If an AUD scheme has at most n options (or blocks), where n could be finite or infinite, then it is called an *n-all-units discounts scheme (n-AUD scheme).*

The form of AUD schemes is simply a special case of our general concept of menu in Sections 3 – 5 where the set of admissible tariff options $\mathcal{C}$ comprises all minimum purchase tariffs. The concepts of AUD-implementability, n-AUD-implementability, AUD-optimality, n-AUD-optimality, AUD-maximum profit and n-AUD-maximum profit are all defined analogously. Theorem 4 translates into the following theorem, which we prove in Appendix. In particular, the increasing differences monotonicity condition in Theorem 4 translates into the monotonicity of minimum purchase and per-unit price (i.e. condition 3 in Theorem 7).

Theorem 7 *A direct mechanism (Q, T) is n-AUD-implementable if and only if it is IC and IR, and there exist functions $M : A \to \mathbb{R}_+$ and $P : A \to \mathbb{R}$, where $A = \{\theta \in \Theta : (Q(\theta), T(\theta)) \neq (0, 0)\}$, such that*

1. *$Q(\theta) \in \arg\max_{q \geq M(\theta)} \{S(q, \theta) - P(\theta) q\}$ for every $\theta \in A$,*

2. *$T(\theta) = P(\theta) Q(\theta)$ for every $\theta \in A$,*

3. *M is nondecreasing and P is nonincreasing, and*

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

If the above conditions hold, then (Q, T) is induced by the IC and IR menu-direct mechanism $\tau_\Theta \equiv \{\tau_\theta\}_{\theta \in \Theta}$ with τ_θ being τ_{out} for $\theta \in \Theta \backslash A$, and with τ_θ being

$$q \mapsto \begin{cases} P(\theta) q & \text{if } q \geq M(\theta) \\ \infty & \text{if } 0 \leq q < M(\theta) \end{cases}$$

for $\theta \in A$. (If $n = \infty$, the function M can be chosen as the restriction of Q on A.)

A *quantity forcing (QF) scheme* (or a menu of quantity-payment pairs) is a menu of options offered by the monopolist, with each option composed of a purchase quantity $q \geq 0$ and a total payment $t \in \mathbb{R}$. If a type θ consumer chooses a quantity-payment pair (q, t), then her utility is $S(q, \theta) - t$. If any consumer chooses none, then her outside option is buying nothing and paying nothing, and her utility is 0. If a QF scheme has at most n options, then it is called an *n-quantity forcing scheme (n-QF scheme)*.

The form of QF schemes is simply a special case of our general concept of menu in Sections 3 – 5 where the set of admissible tariff options $\mathcal{C}$ comprises all tariff options with singleton domain. The concepts of QF-implementability, n-QF-implementability, QF-optimality, n-QF-optimality, QF-maximum profit and n-QF-maximum profit are all defined analogously. However, the structure of QF schemes is so simple that one does not need to invoke our general theory to analyze. The following theorem is obvious, and we state it without proof.[13]

Theorem 8 *A direct mechanism (Q, T) is n-QF-implementable if and only if it is IC and IR, and Q takes at most n values except* 0.

For each nonnegative integer n, let Π_n^{ID} (respectively Π_n^{AUD}, or Π_n^{QF}) denote the n-ID-maximum (respectively n-AUD-maximum, or n-QF-maximum) profit, and let Π_∞^{ID} (respectively Π_∞^{AUD}, or Π_∞^{QF}) denote the ID-maximum (respectively AUD-maximum, or QF-maximum) profit. From now on we assume that the corresponding optimal solutions exist, so that all these maximum profits are well defined.

From Theorem 8, any QF-optimal direct mechanism is also a second best direct mechanism, and Π_∞^{QF} is also the second best profit Π^*. From Theorems 6 – 8 one can immediately see that $\lim_{n\to\infty} \Pi_n^{QF} = \Pi_\infty^{QF} = \Pi^*$, $\lim_{n\to\infty} \Pi_n^{ID} = \Pi_\infty^{ID}$, and $\lim_{n\to\infty} \Pi_n^{AUD} = \Pi_\infty^{AUD}$.[14]

Theorems 6 – 8 imply restrictions on quantity functions that can be induced by n-ID, n-AUD and n-QF schemes. Figures 2 and 3 illustrate the patterns of those quantity functions. The following observations are important. Adopting an ID scheme, the induced quantity function cannot respond to type too little, because (i) within each block, the quantity function has to follow a "type-demand curve", which is typically strictly increasing,[15] and (ii) across blocks, the relevant "type-demand curve" can only shift up, since marginal price has to be nonincreasing in type. In contrast,

[13] Simply notice that if (Q, T) is IC and $Q(\cdot)$ is constant over an interval, then $T(\cdot)$ is also constant over that interval.

[14] It is because a monotonic function (e.g. Q in Theorem 8, P in Theorem 6, M and P in Theorem 7) can be arbitrarily well approximated by a step function.

[15] Assumption 1 guarantees that any selection of demand correspondence with respect to type is nondecreasing.

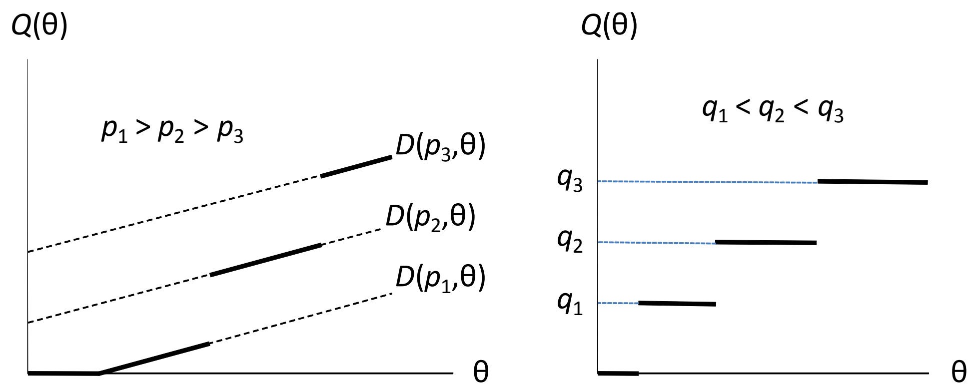

Figure 2: Quantity functions induced by a 3-ID scheme (left) and a 3-QF scheme (right)

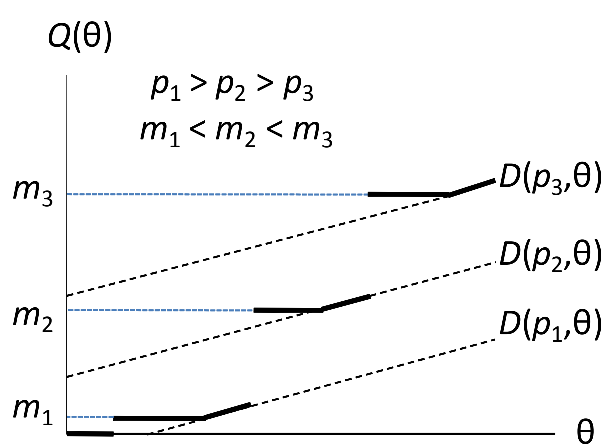

Figure 3: A quantity function induced by a 3-AUD scheme

adopting a QF scheme, the induced quantity function has to be flat within each block. (Of course the jumps have to be upward by ordinary IC.) From this viewpoint, AUD stands between ID and QF. (This is why I include QF into our analysis.) Adopting an AUD scheme, within a block, the induced quantity function can have both a flat portion when the minimum purchase is binding, and a "type-demand curve portion" when the minimum purchase is not binding.

Comparing Figures 2 and 3, one might suspect that, given the number of blocks, AUD can induce more general quantity functions than ID and QF can. But it is not true, because we are not totally free to choose the combination of the minimum purchases (m_i's), the per-unit prices (p_i's) and the thresholds. 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, we have $2n$ degrees of freedom if the number of blocks is n, under any of the three forms of pricing schemes. The comparison among the maximum profits of the three forms is studied in the next section.

7 Comparison among the three forms of pricing schemes

When the number of blocks is unrestricted, we can rank the maximum profits of ID, AUD and QF, because we can rank the concepts of ID-implementability, AUD-implementability and QF-implementability. The proof of the following theorem, which applies Theorems 6, 7 and 8, is provided in Appendix.

Theorem 9 *Any ID-implementable or AUD-implementable direct mechanism is also QF-implementable. If an ID-implementable direct mechanism* (Q, T) *satisfies* $S\left(Q\left(\underline{\theta}\right), \underline{\theta}\right) - T\left(\underline{\theta}\right) = 0$*, then it is also AUD-implementable.*

The first statement of Theorem 9 is trivial. When the number of quantity-payment pairs is unrestricted, QF puts no restriction on implementable direct mechanism except the ordinary IC and IR. The second statement can be informally understood as follows. Under any ID scheme, the tariff function (i.e. payment as a function of quantity) must be concave. In contrast, for a tariff function to be generated by an AUD 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 $S\left(Q\left(\underline{\theta}\right),\underline{\theta}\right)-T\left(\underline{\theta}\right)=0$), then it is geometrically easy to see that concavity is strictly stronger nonincreasing average. That is, when the number of blocks is unrestricted, AUD can generate strictly more tariff functions than ID can.

An alternative way to understand Theorem 9 is recalling the insights from Figures 2 and 3. 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 QF any nondecreasing quantity function can be induced when n tends to infinity. Under AUD, which stands in middle, 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 QF has the largest.

Corollary 1 *(a) Any ID-optimal direct mechanism is AUD-implementable. (b)* $\Pi^*=\Pi_\infty^{QF}\geq\Pi_\infty^{AUD}\geq\Pi_\infty^{ID}$.

Proof. Suppose that a direct mechanism (Q,T) is ID-optimal. Then it is ID-implementable and hence IC and IR. It also satisfies $S\left(Q\left(\underline{\theta}\right),\underline{\theta}\right)-T\left(\underline{\theta}\right)=0$, for if $S\left(Q\left(\underline{\theta}\right),\underline{\theta}\right)-T\left(\underline{\theta}\right)>0$, the monopolist could increase profit by raising the fixed fees of all two-part tariffs. Then part (a) follows from Theorem 9. Part (b) follows from part (a) and Theorem 8. ■

Can ID or AUD attain the second best profit? It amounts to check whether some second best direct mechanism is ID-implementable or AUD-implementable. In particular, if the second best direct mechanism involves bunching, then typically it is not ID-implementable 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* $P^*:\{\theta\in\Theta:Q^*\left(\theta\right)>0\}\rightarrow\mathbb{R}$ *such that* $Q^*\left(\theta\right)\in\arg\max_{q\geq0}\left\{S\left(q,\theta\right)-P^*\left(\theta\right)q\right\}$ *whenever* $Q^*\left(\theta\right)>0$. *(b)* $\Pi^*=\Pi_\infty^{AUD}$ *if and only if there exist some second best direct mechanism* (Q^*,T^*) *and some nonincreasing function* $P^*:\{\theta\in\Theta:Q^*\left(\theta\right)>0\}\rightarrow\mathbb{R}$ *such that* $Q^*\left(\theta\right)\in\arg\max_{q\geq Q^*(\theta)}\left\{S\left(q,\theta\right)-P^*\left(\theta\right)q\right\}$ *and* $T^*\left(\theta\right)=P^*\left(\theta\right)Q^*\left(\theta\right)$ *whenever* $Q^*\left(\theta\right)>0$.

Proof. To see the sufficiency part of (a), suppose that Q^* and P^* satisfy the conditions. Denote the domain of P^* as $A^*=\{\theta\in\Theta:Q^*\left(\theta\right)>0\}$. Define $T^*:\Theta\rightarrow\mathbb{R}$ as in (4) below. It is straightforward to verify that (Q^*, T^*, A^*, P^*) satisfies conditions 1-2 in Theorem 6 and that $A^* = \{\theta \in \Theta : (Q^*(\theta), T^*(\theta)) \neq (0,0)\}$. Therefore, the second best (and hence IC and IR) direct mechanism (Q^*, T^*) is ID-implementable, so that $\Pi^* = \Pi_\infty^{ID}$. To see the necessity part of (a), suppose that $\Pi^* = \Pi_\infty^{ID}$. Then some second best direct mechanism (Q^*, T^*) is ID-implementable. Then Theorem 6 implies the condition in the second statement. ($(Q^*(\theta), T^*(\theta)) \neq (0,0)$ is equivalent to $Q^*(\theta) > 0$, because of the IC of (Q^*, T^*).)

Part (b) can be proved similarly by applying Theorem 7, with defining $M^*(\theta) \equiv Q^*(\theta)$ for any $\theta \in A^* \equiv \{\theta \in \Theta : Q^*(\theta) > 0\}$. ■

Remark 3 *The function P^* in Corollary 2(a), if exists, must be given by $P^*(\theta) = S_q(Q^*(\theta), \theta)$. The function P^* in Corollary 2(b), if exists, must be given by $P^*(\theta) = T^*(\theta)/Q^*(\theta)$, where T^* is given by (4) below.*

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

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

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 $\mathcal{C}$ is unrestricted, then $U(\underline{\theta}) = 0$ at optimum, and the maximum monopolist profit can be written as

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

A direct mechanism (Q^*, T^*) is second best if and only if $Q^*(\cdot)$ solves problem (3) and

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

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 mechanism is

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

One can verify that (Q^*, T^*) is ID-implementable if and only if $\alpha \leq 1/2$, while it is AUD-implementable if and only if $\alpha \leq \sqrt{2}/2$. (Of course, it is always QF-implementable.)

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-AUD-implementability and n-QF-implementability do not imply one another. While ID has the smallest implementation power (i.e. has the smallest set of implementable direct mechanisms) when the number of blocks is unrestricted, we will nonetheless see that ID has the following advantage when the number of blocks is restricted. Provided that a second best direct mechanism is ID-implementable, ID schemes with a finite number of blocks can be constructed to better approximate the second best direct mechanism than AUD schemes or QF schemes with the same number of blocks do. In this sense, we say ID has the largest "approximation power".

In the rest we impose the following regularity assumptions on the virtual surplus function H, which ensure that second best direct mechanism is essential unique, and better approximating the second best direct mechanism raises profit.

Assumption 3 *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 4 *Let Q^* be a second best quantity function. Let (Q^1, T^1) and (Q^2, T^2) be two IC direct mechanisms 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 3, $\{Q^*(\theta)\} = \arg\max_{q\geq 0} H(q,\theta)$ for all $\theta \in \Theta$ except possibly $\underline{\theta}$ and $\bar{\theta}$. Apply formula (2). ■

Theorem 10 *Suppose that* $\Pi^* = \Pi_\infty^{ID}$ *and* n *is finite. (a)* $\Pi_n^{ID} \geq \Pi_n^{AUD}$*, and this inequality is strict unless some* n*-ID-optimal direct mechanism coincides with some* n*-AUD-optimal direct mechanism almost everywhere. (b)* $\Pi_n^{ID} \geq \Pi_n^{QF}$*, and this inequality is strict unless some* n*-ID-optimal direct mechanism coincides with some* n*-QF-optimal direct mechanism almost everywhere.*

The formal proof of Theorem 10 is provided in Appendix. The idea is the following. To show $\Pi_n^{ID} \geq \Pi_n^{AUD}$, it amounts to show that, given an n-AUD-optimal direct mechanism, there exists an n-ID implementable direct mechanism whose quantity function is uniformly closer to the second best than the quantity function of the n-AUD-optimal direct mechanism. In fact, we do not need to know the characteristics of the n-AUD-optimal direct mechanism except knowing that it is n-AUD-implementable. Let us fix $n = 3$ for example and start with some 3-AUD-implementable direct mechanism. The bold curve in Figure 4 illustrates the quantity function Q^{AUD} of such a 3-AUD-implementable direct mechanism, 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 direct mechanism to be ID-implementable, Q^* cannot be flatter than type-demand curves.) Then we can draw a quantity function Q^{ID} of a 3-ID-implementable direct mechanism as shown by the dashed curve in Figure 4. Notice that Q^{ID} is uniformly closer to Q^* than Q^{AUD}. Now according to Lemma 4, the 3-ID-implementable direct mechanism $\left(Q^{ID}, T^{ID}\right)$ that leaves the lowest type of consumer a zero rent would make a higher profit than the original 3-AUD-implementable direct mechanism does. Therefore $\Pi_n^{ID} \geq \Pi_n^{AUD}$. We can similarly argue that $\Pi_n^{ID} \geq \Pi_n^{QF}$ under the assumption $\Pi^* = \Pi_\infty^{ID}$.

8 Concluding remarks

In the context of nonlinear pricing (or more generally principal-agent model), we introduce and characterize the concept of menu implementability. Notably, for a direct selling mechanism to be menu implementable, the familiar monotonicity constraint in incentive theory has to be strengthened to what we call "increasing differences monotonicity" or "single crossing monotonicity".

Our theory can be comfortably used to analyze a large class of practical pricing schemes, such as incremental discounts (ID), all-units discounts (AUD) and quantity forcing (QF). So which of the three forms of pricing schemes makes the highest

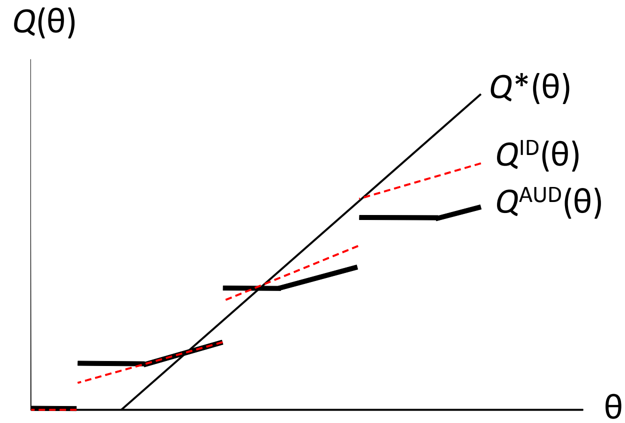

Figure 4: Improving upon AUD by ID

profit? Our comparison among them shed 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 ID, AUD and QF. QF 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. AUD is somehow in the middle: a minimum purchase tariff controls quantity through both a direct instrument, minimum purchase, and an indirect instrument, per-unit price. This is why when the number of blocks (contract complexity) is unrestricted, QF has the largest and ID the smallest implementation power (Theorem 9 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. $\Pi^* = \Pi^{ID}_{\infty}$), then this kind of flexibility must help the approximation to second best. This is why ID has the largest approximation power (Theorem 10 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

Proof of Lemma 1. We must show that if $\tau_1(q) - \tau_2(q)$ is nondecreasing in q on $\chi(\tau_1) \cup \chi(\tau_2)$, then $\tau_1^{\inf}(q) - \tau_2^{\inf}(q)$ is nondecreasing in q on $\chi\left(\tau_1^{\inf}\right) \cup \chi\left(\tau_2^{\inf}\right)$. Let $\chi_1 \equiv \chi(\tau_1)$ and $\chi_2 \equiv \chi(\tau_2)$. Notice that $\chi\left(\tau_1^{\inf}\right) = \mathrm{cl}(\chi_1)$ and $\chi\left(\tau_2^{\inf}\right) = \mathrm{cl}(\chi_2)$.[16]

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

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

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

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

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

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

[16] As usual, cl denotes closure.

Then,

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

and

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

Now, (5) implies that

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

Thus, there exist sequences $q_1^m \to q_1$ on χ_2 and $q_2^m \to q_2$ 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 Lemma 2.

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

If n is finite, then $\tau_B \cup \{\tau_{out}\}$ is a nonempty finite set. We delete any dominated tariff options in τ_B so that condition 2 holds. After the deletion, $\tau_B \cup \{\tau_{out}\}$ is still a nonempty finite set so that condition 1 holds, and the lower envelope of $\tau_B \cup \{\tau_{out}\}$ 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 $\chi \backslash \{q\}$. Then $(\Psi, \leq_q)$ is a partially ordered set. Notice that Ψ is closed since $\mathcal{C} \cup \{\tau_{out}\}$ is closed. We claim that every chain in $(\Psi, \leq_q)$ has a lower bound in Ψ. Indeed, any chain in $(\Psi, \leq_q)$, 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, $(\Psi, \leq_q)$ 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 $(\Psi, \leq_q)$.

Regard $\{\psi_q\}_{q \in \chi} \backslash \{\tau_{out}\}$ as a menu. By our construction, this menu is on $\mathcal{C}$. 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 $\tau_B \cup \{\tau_{out}\}$ share a common lower envelope $\underline{\tau}$, $\{\psi_q\}_{q \in \chi} \backslash \{\tau_{out}\}$ induces the same direct mechanism as τ_B 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. ■

In the following, we prove a stronger version of Lemma 3. Before doing it, we introduce a concept weaker than the tariff single crossing property.

Definition 15 *For any two tariff options τ_1 and τ_2, we say (τ_1, τ_2) satisfies* tariff weak single crossing property *if for any $q_1, q_2 \in \chi$ with $q_1 < q_2$, we have*

$$\tau_1(q_2) < \tau_2(q_2) \Rightarrow \tau_1(q_1) \leq \tau_2(q_1).$$

That is, $\tau_1(\cdot)$ crosses $\tau_2(\cdot)$ only once and only from below. The interpretation is that τ_1 is weakly less favorable to high quantities than τ_2.

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

Lemma 5 *If tariff options $\tau_1, \ldots, \tau_n$ do not dominate one another, and (τ_1, τ_2), (τ_2, τ_3), ..., (τ_{n-1}, τ_n) and (τ_n, τ_1) satisfy tariff weak single crossing property, then $\tau_1 = \cdots = \tau_n$.*

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

Proof of Lemma 3. Clearly, Lemma 5 implies Lemma 3. ■

The proof of Theorem 6 requires the following lemma.

Lemma 6 *The set of all two-part tariffs is closed.*

Proof. Let $\{\tau_\beta\}$ be a net of two-part tariffs, and suppose that it pointwise converges to a tariff option $\hat{\tau}$. Since each τ_β is a two-part tariff, it has the form $\tau_\beta(q) = p_\beta q + \phi_\beta$ for all $q \geq 0$. First notice that, for any $q \geq 0$, the limit $\hat{\tau}(q)$ of $\{\tau_\beta(q)\}$ cannot be $-\infty$, for otherwise $\hat{\tau}$ is not a tariff option. The limit $\hat{\tau}(q)$ cannot be ∞ either, for otherwise $\hat{\tau}$ is ∞ everywhere or is $-\infty$ somewhere (since each τ_β is linear) so that $\hat{\tau}$ is not a tariff option. Now both $\hat{\phi} \equiv \hat{\tau}(0)$ and $\hat{p} \equiv \hat{\tau}(1) - \hat{\tau}(0)$ are finite. Moreover, we have $\phi_\beta = \tau_\beta(0) \to \hat{\tau}(0) = \hat{\phi}$ and $p_\beta = \tau_\beta(1) - \tau_\beta(0) \to \hat{\tau}(1) - \hat{\tau}(0) = \hat{\phi}$. Then, for any $q \geq 0$,

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

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

Proof of Theorem 6. Let $\mathcal{C}$ be the set of all two-part tariffs.

Sufficiency. Suppose that the conditions in this theorem hold. Let τ_Θ be the menu-direct mechanism τ_Θ characterized by (A, P, Φ) as described in the last paragraph of this theorem. It suffices to check that the sufficient conditions provided in Theorem 4. 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 Theorem 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 τ_Θ satisfies condition 1 in Theorem 4. Condition 1 and the definition of A and Φ imply that τ_Θ induces (Q, T). IC of (Q, T) and Theorem 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 τ_Θ satisfies condition 2 in Theorem 4.

Necessity. Suppose that (Q, T) is $(n, \mathcal{C})$-implementable. By Lemma 6, $\mathcal{C}$ is closed, and thus $\mathcal{C} \cup \{\tau_{out}\}$ is also closed. For any two two-part tariffs τ_1 and τ_2, (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, the conditions provided by Theorem 4 are necessary for $(n, \mathcal{C})$-implementability. Then the conditions in Theorem 4 are satisfied by some menu-direct mechanism τ_Θ on $\mathcal{C}$. For any $\theta \in \Theta$ with $Q(\theta) = T(\theta) = 0$ and $\tau_\theta \neq \tau_{out}$, we redefine τ_θ as τ_{out}. Clearly the new τ_Θ still satisfies the conditions in Theorem 4. Since $\tau_\Theta \backslash \{\tau_{out}\} \subset \mathcal{C}$, τ_Θ can be characterized by some (A, P, Φ) as described in the last sentence of this theorem. Then condition 1 holds because τ_Θ induces (Q, T). Conditions 2-3 follow from conditions 1-2 in Theorem 4. ■

The proof of Theorem 7 requires the following lemma.

Lemma 7 *The closure of the set of all minimum purchase tariffs*[17] *comprises the outside option* τ_{out}*, all minimum purchase tariffs, and all "quasi-minimum purchase tariffs" of the form*

$$q \mapsto \begin{cases} pq & \text{if } q > m \\ \infty & \text{if } 0 \le q \le m \end{cases},$$

where $m \ge 0$ *and* $p \in \mathbb{R}$*.*

Proof. Let $\{\tau_\beta\}$ be a net of minimum purchase tariffs, with each τ_β characterized by (m_β, p_β). If $m_\beta = 0$ and $p_\beta \to \infty$, then τ_β pointwise converges to τ_{out}. Suppose that $\{\tau_\beta\}$ pointwise converges to a tariff option $\hat{\tau} \neq \tau_{out}$. Since $\hat{\tau}$ is a tariff option, $\hat{\tau}(q)$ is never $-\infty$, and the set $\{q \ge 0 : \hat{\tau}(q) < \infty\}$ is nonempty, and we let $\hat{m} \ge 0$ be its infimum. Then $m_\beta \to \hat{m}$. For any $q > \hat{m}$, $\hat{\tau}(q)$ is the limit of $\{p_\beta q\}$, and then $\hat{\tau}(q)$ cannot be ∞, for otherwise $p_\beta \to \infty$ and hence $\hat{\tau} = \tau_{out}$ or $\hat{\tau}$ is ∞ everywhere. Now $\hat{p} \equiv \hat{\tau}(\hat{m}+2) - \hat{\tau}(\hat{m}+1)$ is finite. Moreover, for large β,

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

Therefore, for any $q > \hat{m}$, we have $\hat{\tau}(q) = \lim \tau_\beta(q) = \lim p_\beta q = \hat{p}q$. 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 $\lim p_\beta q = \hat{p}q$ or ∞. We conclude that $\hat{\tau}$ is the minimum purchase tariff or the quasi-minimum purchase tariff with minimum purchase $\hat{m}$ and per-unit price $\hat{p}$. ■

Proof of Theorem 7. Let $\mathcal{C}$ be the set of all minimum purchase tariffs, and $\mathcal{C}^+$ be the set of all minimum purchase tariffs and all "quasi-minimum purchase tariffs" (see Lemma 7). If a direct mechanism (Q,T) is $(n,\mathcal{C})$-implementable, then it is $(n,\mathcal{C}^+)$-implementable, since $\mathcal{C} \subset \mathcal{C}^+$. The converse is also true. Indeed, if a direct mechanism (Q,T) is $(n,\mathcal{C}^+)$-implementable, then it is induced by some menu $\tau_B \equiv \{\tau_\beta\}_{\beta \in B} \subset \mathcal{C}^+$ with $|\tau_B| \le n$. Then clearly (Q,T) is also induced by $\tau_B^{\inf} \equiv \{\tau_\beta^{\inf}\}_{\beta \in B}$. But now $\tau_B^{\inf} \subset \mathcal{C}$ and $\left|\tau_B^{\inf}\right| \le n$, so that (Q,T) is $(n,\mathcal{C})$-implementable.

Sufficiency. Suppose that the conditions in this theorem hold. Let τ_Θ be the menu-direct mechanism τ_Θ characterized by (A,M,P) as described in the last paragraph of this theorem. It suffices to check that the sufficient conditions provided in Theorem 4. Notice that $\theta_1 < \theta_2$ and $Q(\theta_2) = T(\theta_2) = 0$ imply $Q(\theta_1) = T(\theta_1) = 0$,

[17] As 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 $\mathcal{C}$.

by IC of (Q,T) and Theorem 1. Hence A is an increasing subset of Θ. It, together with condition 3, implies that τ_Θ satisfies condition 1 in Theorem 4. The definition of A and conditions 1-2 imply that τ_Θ induces (Q,T). Condition 4 implies that τ_Θ satisfies condition 2 in Theorem 4. If in addition $n=\infty$, 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) is $(n,\mathcal{C})$-implementable and hence $(n,\mathcal{C}^+)$-implementable. By Lemma 7, $\mathcal{C}^+\cup\{\tau_{out}\}$ is closed. For any two tariff options τ_1 and τ_2 in $\mathcal{C}^+$, 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 per-unit 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, the conditions provided by Theorem 4 with $\mathcal{C}$ replaced by $\mathcal{C}^+$ are necessary for $(n,\mathcal{C}^+)$-implementability. Then the conditions in Theorem 4 are satisfied by some menu-direct mechanism τ_Θ on $\mathcal{C}^+$. We claim that the conditions in Theorem 4 are also satisfied by some menu-direct mechanism on $\mathcal{C}$. Indeed, the menu-direct mechanism $\tau_\Theta^{\inf}\equiv\{\tau_\theta^{\inf}\}_{\theta\in\Theta}$ does the job. Clearly, $\tau_\Theta^{\inf}$ is on $\mathcal{C}$, induces the same direct mechanism as τ_Θ does (since $S(\cdot,\theta)$ is continuous), and $|\tau_\Theta^{\inf}|\le|\tau_\Theta|\le n$. Moreover, by Lemma 1, $\tau_\Theta^{\inf}$ satisfies condition 1 in Theorem 4 as τ_Θ does.

Let us redefine τ_Θ as a menu-direct mechanism on $\mathcal{C}$ that satisfies the conditions in Theorem 4. For any $\theta\in\Theta$ with $Q(\theta)=T(\theta)=0$ and $\tau_\theta\neq\tau_{out}$, we redefine τ_θ as τ_{out}. Clearly the new τ_Θ still satisfies the conditions in Theorem 4. Since $\tau_\Theta\backslash\{\tau_{out}\}\subset\mathcal{C}$, τ_Θ can be characterized by some (A,M,P) as described in the last paragraph of this theorem. Then conditions 1-2 hold because τ_Θ induces (Q,T). Conditions 3-4 follow from conditions 1-2 in Theorem 4. ■

Proof of Theorem 9. Apply Theorems 6, 7 and 8. The first statement in this theorem is obvious.

Suppose that a direct mechanism (Q,T) is ID-implementable. Then it is IC and IR, and there exists a function $P:A\to\mathbb{R}$, where $A=\{\theta\in\Theta:(Q(\theta),T(\theta))\neq(0,0)\}$, such that conditions 1-2 in Theorem 6 hold. Suppose also that $U(\underline{\theta})=0$, where $U(\theta)\equiv S(Q(\theta),\theta)-T(\theta)$. For each $\theta\in A$, define $M(\theta)\equiv Q(\theta)$ and $P^{AUD}(\theta)\equiv T(\theta)/Q(\theta)$.

Then (A,M,P^{AUD}) satisfies condition 2 in Theorem 7. Since (Q,T) is IC, Q is nondecreasing by Theorem 1. Hence M is nondecreasing. It remains to verify that

$$Q(\theta) \in \arg\max_{q \geq Q(\theta)} \left\{S(q,\theta) - P^{AUD}(\theta)\, q\right\} \tag{6}$$

for $\theta \in A$, and that P^{AUD} is nonincreasing.

Since the menu-direct mechanism characterized by (A, P, Φ) as described in Theorem 6 is IC, type $\underline{\theta}$ has no incentive to deviate to pick the two-part tariff for any type $\theta \in A$. Thus, for any $\theta \in A$,

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

That is, Φ is nonnegative.

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

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

It, together with $Q(\theta) \in \arg\max_{q \geq 0} \{S(q,\theta) - P(\theta)\, q\}$, implies (6).

Pick any $\theta_1, \theta_2 \in A$ with $\theta_1 < \theta_2$. Since the menu-direct mechanism characterized by (A, P, Φ) is IC, type θ_2 has no incentive to deviate to pick the two-part tariff for type θ_1. Thus, $T(\theta_2) \leq P(\theta_1)\, Q(\theta_2) + \Phi(\theta_1)$. It, together with the monotonicity of Q and the nonnegativity of Φ, implies that

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

Therefore, P^{AUD} is nonincreasing. ■

Proof of Theorem 10.

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

$$Q^*(\theta) \in \arg\max_{q \geq 0} \{S(q,\theta) - P^*(\theta)\, q\} \tag{7}$$

whenever $Q^*(\theta) > 0$. By Assumption 3, $Q^*(\theta)$ 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 $\Theta \backslash \{\underline{\theta}, \bar{\theta}\}$.

Let $\left(Q_n^{AUD}, T_n^{AUD}\right)$ be an n-AUD-optimal direct mechanism with a finite n. Apply Theorem 7. $\left(Q_n^{AUD}, T_n^{AUD}\right)$ is associated with some A, M, P as described in Theorem 7. 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 $\{\Theta_1, \ldots, \Theta_{n'}\}$ is a partition of A, $n' \leq n$, each Θ_i $(i = 1, \ldots, n')$ is nonempty, and

$$\Theta_1 < \Theta_2 < \cdots < \Theta_{n'},$$

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

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

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

$$Q_n^{AUD}(\theta) \in \arg\max_{q \geq m_i} \{S(q, \theta) - p_i q\}. \tag{8}$$

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

To prove Claim 1, suppose first that both (I) and (II) are false. Then $Q_n^{AUD}(\theta_1) > Q^*(\theta_1)$ and $Q_n^{AUD}(\theta_2) < Q^*(\theta_2)$ for some $\theta_1, \theta_2 \in \Theta_i$. Notice that Q^* is continuous (by Assumption 3) and nondecreasing, and Q_n^{AUD} is nondecreasing. If $\theta_1 < \theta_2$, then there exists some $\mu_i \in (\theta_1, \theta_2)$ such that $Q_n^{AUD}(\mu_i) = Q^*(\mu_i)$. If $\theta_2 < \theta_1$, then $p_i \leq P^*(\theta_1)$ (otherwise $Q_n^{AUD}(\theta_1) \leq Q^*(\theta_1)$, a contradiction), and $p_i \geq P^*(\theta_2)$ (otherwise $Q_n^{AUD}(\theta_2) \geq Q^*(\theta_2)$, 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 $Q_n^{AUD}(\theta) = Q^*(\theta)$ for almost all $\theta \in (\theta_2, \theta_1)$ (otherwise Q_n^{AUD} is not n-AUD-optimal, a contradiction). Therefore, Claim 1 is true.

For every $i = 1, \ldots, n'$, 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 $Q_n^{AUD}(\mu_i) = Q^*(\mu_i)$ 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, $P_n^{ID}: A \to \mathbb{R}$ is nonincreasing because P^* is.

We will design a direct mechanism $\left(Q_n^{ID}, T_n^{ID}\right)$ such that

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

$$Q_n^{ID}(\theta) \in \arg\max_{q \geq 0} \{S(q,\theta) - P^*(\mu_i)\, q\} \text{ for } \theta \in \Theta_i, \tag{9}$$

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

By Theorem 6, any such $\left(Q_n^{ID}, T_n^{ID}\right)$ is n-ID-implementable.

Notice that $S\left(Q_n^{ID}(\underline{\theta}), \underline{\theta}\right) - T_n^{ID}(\underline{\theta}) = 0 \leq S\left(Q_n^{AUD}(\underline{\theta}), \underline{\theta}\right) - T_n^{AUD}(\underline{\theta})$. In order to apply Lemma 4, we will prove, for every $i = 1, \ldots, n'$, that $\inf \Theta_i < \theta < \mu_i$ implies $Q_n^{AUD}(\theta) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$, and that $\mu_i < \theta < \sup \Theta_i$ implies $Q_n^{AUD}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$.

Pick any $i = 1, \ldots, n'$ 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 (8) for type μ_i, then $p_i \leq P^*(\mu_i) \leq P^*(\theta)$. Comparing (7), (8) and (9), we can select $Q_n^{ID}(\theta)$ such that $Q_n^{AUD}(\theta) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$. If the constraint $q \geq m_i$ is binding in problem (8) for type μ_i, then $Q_n^{AUD}(\theta) = m_i = Q_n^{AUD}(\mu_i) \geq Q^*(\mu_i)$. Since $P^*(\mu_i) \leq P^*(\theta)$, comparing (7) and (8), 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 (7), (8) and (9), we can select $Q_n^{ID}(\theta)$ such that $Q_n^{AUD}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$.

Suppose that case III in Claim 1 occurs and the constraint $q \geq m_i$ is not binding in problem (8) for type μ_i. Then $p_i = S_q\left(Q_n^{AUD}(\mu_i), \mu_i\right) = S_q\left(Q^*(\mu_i), \mu_i\right) = P^*(\mu_i)$. 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 $Q_n^{AUD}(\theta) \geq Q_n^{ID}(\theta) \geq Q^*(\theta)$ whenever $Q_n^{AUD}(\theta) \geq Q^*(\theta)$. If $Q_n^{AUD}(\theta) < Q^*(\theta)$, then, for small $\varepsilon > 0$ and $x \in (\theta - \varepsilon, \theta]$, we have $Q_n^{AUD}(x) < Q^*(x)$ (since Q^* is continuous and Q_n^{AUD} is nondecreasing) and $p_i = P^*(x)$, and then Q_n^{AUD} is not

n-AUD-optimal because $Q_n^{AUD}(x)$ 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 $Q_n^{AUD}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$ whenever $Q_n^{AUD}(\theta) \leq Q^*(\theta)$. If $Q_n^{AUD}(\theta) > Q^*(\theta)$, then, for small $\varepsilon > 0$ and $x \in [\theta, \theta + \varepsilon)$, we have $Q_n^{AUD}(x) > Q^*(x)$ (again since Q^* is continuous and Q_n^{AUD} is nondecreasing) and $p_i = P^*(x)$, and then Q_n^{AUD} is not n-AUD-optimal because $Q_n^{AUD}(x)$ 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 (8) for type μ_i. Then $p_i \geq P^*(\mu_i)$ and $m_i = Q_n^{AUD}(\mu_i) = Q^*(\mu_i)$. For $\inf \Theta_i < \theta < \mu_i$, we have $Q_n^{AUD}(\theta) = m_i = Q^*(\mu_i)$. 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 $Q_n^{AUD}(\theta) \leq Q_n^{ID}(\theta) \leq Q^*(\theta)$ whenever $Q_n^{AUD}(\theta) \leq Q^*(\theta)$. Repeating our previous logic, one can show that $Q_n^{AUD}(\theta) > Q^*(\theta)$ is impossible.

Apply Lemma 4, we see that $\Pi_n^{ID} \geq \Pi_n^{AUD}$. If $\Pi_n^{ID} = \Pi_n^{AUD}$, then it must be the case that $\left(Q_n^{AUD}, T_n^{AUD}\right)$ and $\left(Q_n^{ID}, T_n^{ID}\right)$ make the same profit and $\left(Q_n^{ID}, T_n^{ID}\right)$ is n-ID-optimal. Then Lemma 4 implies that $\left(Q_n^{AUD}(\theta), T_n^{AUD}(\theta)\right) = \left(Q_n^{ID}(\theta), T_n^{ID}(\theta)\right)$ for almost all $\theta \in \Theta$. It completes the proof of part (a). ■

References

Bergemann, D., J. Shen, Y. Xu, and E. M. Yeh (2010): "Mechanism Design with Limited Information: The Case of Nonlinear Pricing," http://ssrn.com/abstract=1717904, (forthcoming, 2nd International ICST Conference on Game Theory for Networks, Shangai, 2011).

Chu, L. Y., and D. E. M. Sappington (2007): "Simple cost-sharing contracts," *American Economic Review*, pp. 419–428.

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.