Duality Approach to Nonlinear Pricing Schedules with Applications

Masahiro Watabe*

Department of Economics

Meisei University

August 16, 2013

Preliminary Draft

Abstract

The paper provides a reverse of the revelation principle and its applications in a principal-agent model with type-dependent reservation utility under asymmetric information. I show that any incentive compatible and individually rational direct revelation mechanism is voluntarily implementable by a single indirect mechanism/price schedule. The first application of my result can be used to check the optimality of posting a block tariff. I also compare the second-best price schedule to marginal cost pricing. The second application is the existence of a unique Nash equilibrium in terms of nonlinear pricing for delegated common agency games.

Keywords: nonlinear pricing · taxation principle · bunching · block tariffs · marginal cost pricing · countervailing incentives · delegated common agency

JEL Classification Numbers: D82 · D86 · L12 · L13

*Department of Economics, Meisei University, 2-1-1 Hodokubo, Hino, Tokyo, Japan 191-8506. Email: masahiro.watabe@meisei-u.ac.jp.

I thank for helpful comments to seminar participants at the 2013 Contract Theory Workshop, Fall 2010 Midwest Economic Theory Meeting, TOBB University of Economics and Technology, and Washington University in St. Louis.

Contents

1 Introduction

I consider a model of contracting by a principal and an agent under asymmetric information. There is adverse selection because the agent's characteristics/types are not observable to the principal (the distribution being known, however). The principal and the agent contract on a product characteristic such as quality and quantity, and a monetary transfer. The agent's product choice can be described as a *decision rule* that assigns a product choice for each type. The strategy space of the principal is a set of nonlinear price schedules. Each nonlinear pricing schedule is considered a catalog of products and prices. The agent chooses a product so as to maximize his utility from trade. I say that a decision rule is *implementable via price schedule* if it is consistent with the agent's type-dependence best responses for some price schedule. The principal selects a price schedule so as to maximize his expected profit from trade, subject to the implementability constraints and the participation constraints.

The standard approach to the screening problem is reformulating the principal's strategy space. If a decision rule is implementable by a nonlinear price schedule, then a transfer function indexed by agent types is obtained as the composite function of the price schedule and the decision rule. By construction, such a direct revelation mechanism is incentive compatible in the sense that it is optimal for the agent to announce the true value of his private information. This is so-called the *revelation principle*.[1] The principal's problem is rewritten as an expected profit maximization problem over a set of incentive compatible decisions and transfers satisfying the participation constraints.

The use of the revelation principle has become widespread in the optimal contracts literature. However, because of the intensive use of the revelation principle, the literature has focused on analyzing the properties of contracts indexed by agent types, such as distortions and information rents, *rather than* the properties of price schedules. The purpose of the paper is to establish the converse of the revelation principle, that is, to construct an indirect mechanism or a price schedule as a solution to the principal's optimization problem that are flexible to applications.

I present the following example in order to illustrate the basic methodology in the literature, and then I shall give my motivations in the present paper.[2]

Example 1 (Vertical Differentiation)**.** The seminar article in the literature on nonlinear pricing, Mussa and Rosen (1978), predict that the monopolist's optimal quality allocation exhibits no-distortion for the highest type and downward distortions for all other types when the participation constraints are deterministic and type-independent. Mussa and Rosen (1978) and Rochet and Stole (2002) assume $u(x, \theta) = x\theta$. Let $C(x)$ be a quadratic cost function of the principal with $C'(x) = a + bx$ for some $a \geqslant 0$ and $b > 0$. Taste θ is uniformly distributed over $\Theta = [\underline{\theta}, \overline{\theta}]$ such that $\underline{\theta} > a$, and $\theta^* = \frac{1}{2}(\overline{\theta} + a) > \underline{\theta}$. The principal selects a price schedule or nonlinear pricing $t : X \to \mathbb{R}$ to solve

$$\max_{t(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} [t(x(\theta)) - C(x(\theta))] f(\theta) d\theta \text{ subject to for every } \theta \in \Theta,$$

$$x(\theta) \in \text{argmax}[u(x, \theta) - t(x) \mid x \in X],$$

$$\max[u(x, \theta) - t(x) \mid x \in X] \geqslant 0.$$

The basic methodology in the literature reformulates the principal's problem over incentive compatible and individually rational direct revelation mechanism $\langle x(\cdot), U(\cdot)\rangle$, where $U(\theta) = \max[u(x, \theta) - t(x) \mid x \in X]$. As is well-known, such a direct revelation mechanism in this context is incentive compatible if and only is $\dot{U}(\theta) = u_\theta(x(\theta), \theta)$ and $x(\cdot)$ is non-decreasing. Using the information rent $U(\theta) = u(x(\theta), \theta) - t(x(\theta))$, the variables in the principal's problem are transformed from $t(\cdot)$ to $\langle x(\cdot), U(\cdot)\rangle$:

[1] For instance, see Laffont and Martimort (2002, Proposition 2.2).

[2] Throughout the paper, I indicate derivatives taken with respect to θ with a "dot" superscript, while derivatives with respect to x with "prime" superscript. Moreover, subscripts denote partial derivatives of the agent's utility function $u(x, \theta)$.

$$\max_{\langle x(\cdot),U(\cdot)\rangle} \int_{\underline{\theta}}^{\overline{\theta}} [u(x(\theta),\theta) - C(x(\theta)) - U(\theta)]f(\theta)d\theta \text{ subject to for every } \theta \in \Theta,$$

$$\dot{U}(\theta) = u_\theta(x(\theta),\theta), \ \ \dot{x}(\theta) \geqslant 0, \ \text{ and } \ U(\theta) \geqslant 0.$$

The literature has argued the optimal menu of tariffs $\langle x(\cdot), U(\cdot)\rangle$ consisting of a decision rule and an information rent that maximizes the principal's expected surplus. According to Mussa and Rosen (1978), the principal's optimal decision rule is given by

$$x(\theta) = \begin{cases} 0 & \text{for } \theta < \theta^* \\ \frac{1}{b}(2\theta - (\overline{\theta} + a)) & \text{for } \theta \geqslant \theta^*. \end{cases}$$

Consumers for whom $\theta < \theta^*$ are excluded from the market. The full-information decision rule is given by $x^*(\theta) \in \text{argmax}[u(x,\theta) - C(x) \mid x \in X]$, where $u(x,\theta) - C(x)$ is the social surplus function. The first-order condition yields that $x^*(\theta) = \frac{1}{b}(\theta - a)$ for all $\theta \in \Theta$. For the highest type $\overline{\theta}$, $x(\overline{\theta}) = \frac{1}{b}(\overline{\theta} - a) = x^*(\overline{\theta})$. For any $\theta < \overline{\theta}$, $x(\theta) = \frac{1}{b}(2\theta - (\overline{\theta} + a)) < \frac{1}{b}(2\theta - (\theta + a)) = \frac{1}{b}(\theta - a) = x^*(\theta)$. In other words, the decision rule $x(\theta)$ exhibits no-distortion at the top and downward distortions for all other types. Figure 1 shows the pattern of distortions and pooling at $x = 0$. The existing analysis mainly focuses on distortion patterns due to the revelation principle.

As shown, after applying the revelation principle, we have obtained an outcome $\langle x(\cdot), U(\cdot)\rangle$ or $\langle x(\cdot), p(\cdot)\rangle$. A natural question is how the optimal price schedule $t : X \to \mathbb{R}$ looks like. Denote by $X = [x(\underline{\theta}), x(\overline{\theta})]$, where $x(\underline{\theta}) = 0$, the corresponding product line. Consider the following price schedule $t : X \to \mathbb{R}$ shown in Figure 2:

$$t(x) = \tfrac{1}{4}bx^2 + \tfrac{1}{2}(\overline{\theta} + a)x.$$

The corresponding marginal price schedule becomes $t'(x) = \frac{1}{2}(bx + \overline{\theta} + a) > 0$ and so, this price schedule is increasing. Furthermore, the price schedule $t(\cdot)$ is strictly convex because $t''(x) = \frac{b}{2} > 0$.

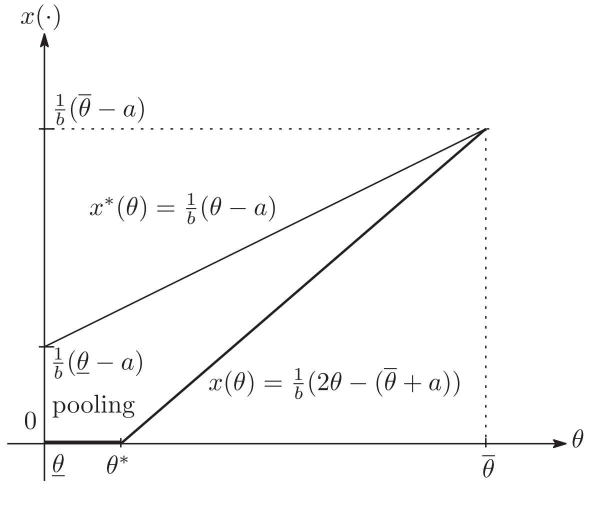


Figure 1: Decision Rules $x(\cdot)$ and $x^*(\cdot)$

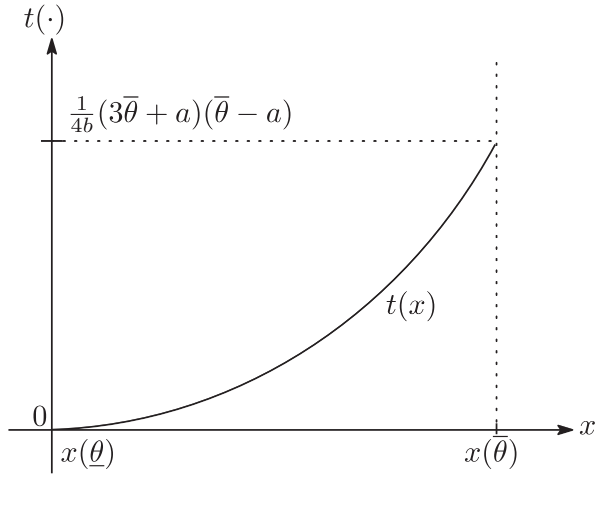


Figure 2: Price Schedule $t(\cdot)$

I can show that the price schedule $t(\cdot)$ actually implements the decision rule $x(\cdot)$, or equivalently, it satisfies the implementability constraint $x(\theta) \in \text{argmax}[u(x,\theta) - t(x) \mid x \in X]$ for the above decision rule $x(\cdot)$. Since the price schedule $t(\cdot)$ is strictly convex, it follows that the net utility $u(x,\theta) - t(x)$ is a strictly concave function in x. Consider, first, any $\theta \geqslant \theta^*$. The first-order condition with respect to x is given by $0 = u_x(x,\theta) - t'(x) = \theta - \left(\frac{1}{2}bx + \frac{1}{2}(\overline{\theta} + a)\right)$, and so $x = \frac{1}{b}(2\theta - (\overline{\theta} + a)) = x(\theta)$ is the solution to utility maximization problem of type $\theta \geqslant \theta^*$. Consider, next, any $\theta \leqslant \theta^*$. Similarly, $u_x(x,\theta) - t'(x) = \theta - \left(\frac{1}{2}bx + \frac{1}{2}(\overline{\theta} + a)\right) \leqslant \theta^* - \frac{1}{2}bx - \theta^* = -\frac{1}{2}bx \leqslant 0$ with equality at $x = 0$. Since the net utility is strictly concave in x, it follows that $x = 0$ is the only solution to utility maximization problem of type $\theta \leqslant \theta^*$. Therefore,

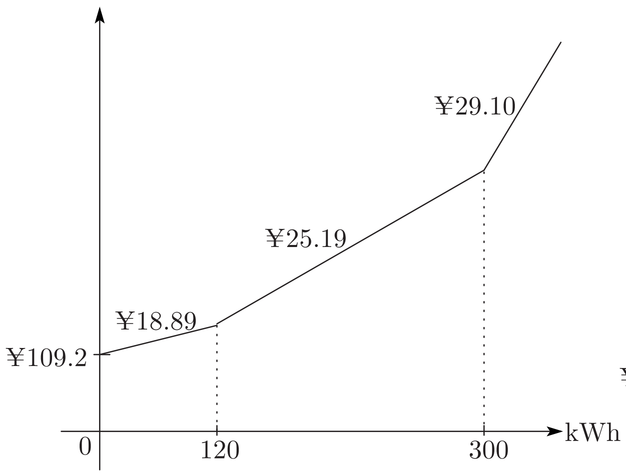
$$\text{argmax}[u(x,\theta) - t(x) \mid x \in X] = \{x(\theta)\}, \ \ \forall \theta \in \Theta.$$

To sum up, the decision rule $x(\cdot)$ is the unique solution to the utility maximization problem of the agent of type θ under the nonlinear pricing $t(x) = \frac{1}{4}bx^2 + \frac{1}{2}(\overline{\theta} + a)x$. □

It is not difficult to construct a nonlinear price schedule implementing a given decision rule when the decision rule exhibits no pooling (cf. Laffont and Martimort, 2002, Proposition 9.6). The reservation utility is fixed to zero for all of types as in the above example. The usual method obtains the optimal menu of contracts $\langle x(\cdot), U(\cdot)\rangle$ along the strictly convex price schedule in the above example. My concern in Section 3 is how to construct such a price schedule in a more general setting in which the reservation utility is type-specific. This means that it is not always exclude a possibility of pooling. If there is a procedure to obtain a nonlinear price function from any incentive compatible and individually rational direct revelation mechanism, then it means to claim that, without loss of generality, we can restrict attention to a class of direct revelation mechanisms. The character of nonlinear pricing including convexity/concavity with respect to quality/quantity remains largely unexplored because of the intensive use of direct revelation mechanisms. Theorem 2 in the paper establishes the implementability of any decision rules possibly involving bunching in a general model of nonlinear pricing. Maggi and Rodriguez-Clare (1995) consider a model where the reservation utility is deterministic but type-dependent. They argue that when the reservation utility depends on the private information, the structure of optimal contracts, in particular, the occurrence of pooling, crucially depends on the shape of the reservation utility function. I introduce the notion of *voluntary implementability*, taking into account the agent's type-dependent reservation utility. Theorem 2, together with the revelation principle, states that any pair of a decision rule and an information rent is incentive compatible and individually rational if and only if it is voluntarily implementable by some indirect mechanism (Theorem 3). The advantage of my construction is that there is no need to exclude bunching in decision rules.

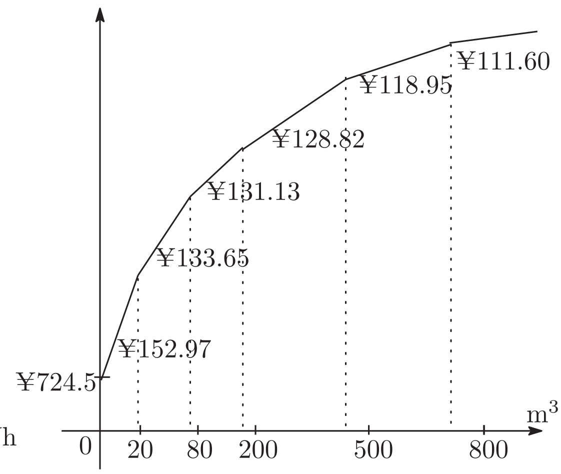
In Section 4, I explore the economic interpretations behind the constructed price schedule in the previous section. It is shown that the inverse of a decision rule plays a crucial role in constructing the optimal price schedule in my approach (Theorem 4). This is the reason that I refer to the methodology proposed in the present paper the *duality approach* to nonlinear pricing schedules.

In Section 5.1, I discuss under what conditions, the optimal price schedule belongs to the principal's strategy space consisting of *piecewise linear* price schedules. Piecewise linear tariffs are commonly used to public utilities. For instance, TEPCO (Tokyo Electric Power Company, Inc) and Tokyo Gas Co., Ltd. use the following block tariffs for the use of electricity and gas in 2012, respectively (see Figures 3 and 4). I show that a single equation is necessary and sufficient condition for that the optimal price schedule constructed in the proof of Theorem 2 takes the form of a block tariff indeed (Theorem 6). Furthermore, the expression for the second-derivative of the optimal price schedule can be used to check the optimality of quality premia or quantity discounts.



(a) TEPCO (40A)

Figure 3: Increasing Block Tariff



(b) Tokyo Gas

Figure 4: Declining Block Tariff

In Section 5.2, I will examine the deviation of the optimal price schedule deviates from marginal cost pricing schedule under countervailing incentives. Maggi and Rodriguez-Clare (1995) work on properties of optimal contracts (i.e., direct revelation mechanisms) under countervailing incentives such as (a) allocative distortion, and (b) the distribution of the agent's information rents with respect to the shape of the agent's type-dependent reservation utility. On the other hand, my primary concern is how the optimal price schedule for the principal deviates from marginal cost pricing, where both price schedules are not direct revelation mechanisms. The present paper provides a unified analysis of indirect mechanisms. I will discuss a relationship between the optimal marginal price and marginal cost, based on the steepness of the agent's type-dependent reservation utility. If the slope of the agent's reservation utility takes an intermediate value, then the decision rule involves pooling around a non-degenerate interval of types that earn zero information rents, and there is no distortion at both extremes of types and an interior type. The optimal marginal price schedule is lower than marginal cost up to the consumption level for that interior type, and higher than marginal cost from then on. A combination of two concave price schedules can arise as the optimal price schedule if the principal's cost function is linear or strictly concave over the product line, as long as the sign of the second-derivative of the optimal price schedule is constant over the product line. In contrast with the case of concave cost functions, how the marginal price deviates from marginal cost is indeterminate if the principal's cost function is convex.

In Section 5.3, I build a theoretical model of firm competition via nonlinear pricing schemes. I show that an equilibrium exists and any equilibrium outcome (specifying decision rules and who sells to which markets) is implementable by a profile of price schedules constructed in Section 3 (Theorem 8). Common agency is a mechanism design problem with multiple principals and an agent. The literature has been considering an environment with two principals for simplicity. There are two distinct environments with common agency. In the context of *intrinsic* common agency, the agent must choose between contracting with both principals or contracting with neither. On the other hand, in the context of *delegated* common agency, the agent can choose whether to contract with both, one, or none of the principals. There are several theoretical works in contract games that have largely restricted attention to intrinsic settings. However, most competitive nonlinear pricing applications assume exclusive purchasing. In other words, the agent chooses at most one of the principals. In the present paper, I consider such an extreme case of delegated common agency in which the principals are perfectly substitutes. The price schedules are modeled as a Nash equilibrium of the delegated common agency game.

There are several works on competing mechanisms under asymmetric information. Ivaldi and Martimort (1994) solve for the equilibrium of intrinsic common agency games with non-identical duopolists within the class of linear quadratic price functions. Contrary to Ivaldi and Martimort (1994), I establish the *competitive implementability* (formally discussed in Section 5.3.1) without restricting attention to such a particular class of price functions. Rochet and Stole (2002) solve for the equilibrium of delegated common agency games with identical duopolists on a horizontal market segment when transportation cost is sufficiently small. Quality distortions disappear completely, and thus the efficient equilibrium outcome is implementable by the cost-plus-fixed fee price functions. However, they do not look for a Nash equilibrium in the case that quality allocations are distorted. My concern is what sorts of nonlinear pricing schedules emerge in an equilibrium. The proof of my competitive implementability shows how to construct nonlinear pricing schedules.

In my setting, the reservation utility is the outside option determined *endogenously* as a Nash equilibrium outcome. This complicates the analysis in delegated common agency games (formally discussed in Section 5.3.2). The *client set* of each principal is defined as the set of consumer types served by that principal, which is divided in two parts: the *captive* market over which that principal is strictly dominant and the *competitive* market over which both principals are indifferent for consumers. The paper analyzes a market segmentation endogenously determined, without assuming a degenerate competitive market. The present paper also provides results that are consistent with empirical evidences. McManus (2007) conducts an empirical examination of distortions in product lines in an oligopolistic market for specialty coffee. The estimated distortions are small at the top and the bottom, with larger distortion in between. His interpretation is that the bottom of the product line is in closer competition with the outside option than the top of the product line. The intuition seems to be that the competitive pressure below makes quality allocations distorted less. Rochet and Stole (2002) mentioned above show that when competition is sufficiently intense, there is no distortion everywhere. The present paper shows that quality distortions disappear only at the top and the bottom of the client sets.

2 Principal-Agent Model in a Quasi-Linear Context

I consider a principal-agent model which can be described as follows. The principal is a Stackelberg leader of the two-player game with asymmetric information. In the first stage, the principal chooses his strategy given the optimizing behavior of the agent in the second stage. The principal and the agent contract on two types of variables: a product characteristic x such as quality or quantity, and a monetary transfer y. Both are observable to both players. Denote by $X \subseteq \mathbb{R}_+$ the product line that the principal can offer. A strategy of the principal is a nonlinear *price schedule* $t : X \to \mathbb{R}$. The principal's utility is $y - C(x)$, where $C : X \to \mathbb{R}$ is the cost function.

The agent chooses a product x sold at price y. His choice is made according to preferences represented as a quasi-linear utility function $u(x, \theta) - y$, where type θ is a one-dimensional parameter that belongs to a compact set $\Theta = [\underline{\theta}, \overline{\theta}] \subseteq \mathbb{R}_{++}$. There is adverse selection problem because this parameter is known to the agent but unobservable to the principal. I assume that $u_x(x, \theta) > 0$ and $u_{x\theta}(x, \theta) > 0$.[3] The latter condition is called the single-crossing property. Given a price schedule $t(\cdot)$, a consumer of type θ maximizes his net utility $u(x, \theta) - t(x)$ over X. Let $\pi_t(\theta) = \max[u(x, \theta) - t(x) \mid x \in X]$ be the indirect utility function of type θ. Finally, the agent of type θ may have an outside opportunity, from which he can derive a utility level $\bar{\pi}(\theta)$.

From the principal's perspective, the agent's type is continuously distributed over Θ with a density function $f(\theta) > 0$ for every $\theta \in \Theta$. The principal chooses a price schedule to solve the following profit-maximization problem:

$$\max_{t(\cdot)} \int_{\Theta} [t(x(\theta)) - C(x(\theta))] f(\theta) d\theta$$

subject to the *implementability constraints*

$$x(\theta) \in \text{argmax}[u(x, \theta) - t(x) \mid x \in X], \quad \forall \theta \in \Theta$$

[3] Throughout the paper, subscripts denote partial derivatives of the agent's utility function $u(x, \theta)$. Moreover, I indicate derivatives taken with respect to θ with a "dot" superscript, while derivatives with respect to x with "prime" superscript.

and the *participation constraints*

$$\pi_t(\theta) = \max[u(x,\theta) - t(x) \mid x \in X] \geqslant \bar{\pi}(\theta), \quad \forall \theta \in \Theta.$$

The standard approach to the screening problem is reformulating the principal's strategy space. The participation constraints can be replaced by the system of inequalities, $r(\theta) = \pi_t(\theta) - \bar{\pi}(\theta) \geqslant 0$ for every $\theta \in \Theta$. Introducing the social surplus function $v(x,\theta) = u(x,\theta) - C(x)$, I rewrite profit margin as $t(x(\theta)) - C(x(\theta)) = u(x(\theta),\theta) - \pi_t(\theta) - C(x(\theta)) = v(x(\theta),\theta) - r(\theta) - \bar{\pi}(\theta)$. Moreover, if $x(\theta) \in \text{argmax}[u(x,\theta) - t(x) \mid x \in X]$ holds under a price schedule $t(\cdot)$, then it is the case $\theta \in \text{argmax}[u(x(\hat{\theta}),\theta) - t(x(\hat{\theta})) \mid \hat{\theta} \in \Theta]$. Using the information rent, the payment $t(x(\hat{\theta}))$ in this expression is written as $t(x(\hat{\theta})) = u(x(\hat{\theta}),\hat{\theta}) - \pi_t(\hat{\theta}) = u(x(\hat{\theta}),\hat{\theta}) - r(\hat{\theta}) - \bar{\pi}(\hat{\theta})$. To sum up, a profile $\langle x(\cdot), r(\cdot)\rangle$ of a decision rule and an information rent is *incentive compatible* in the sense that $\theta \in \text{argmax}[u(x(\hat{\theta}),\theta) - u(x(\hat{\theta}),\hat{\theta}) + r(\hat{\theta}) + \bar{\pi}(\hat{\theta}) \mid \hat{\theta} \in \Theta]$ for every $\theta \in \Theta$. In what follows, such a profile $\langle x(\cdot), r(\cdot)\rangle$ is called a *direct revelation mechanism*.

The principal's problem can be written as

$$\max_{\langle x(\cdot), r(\cdot)\rangle} \int_\Theta [v(x(\theta),\theta) - r(\theta) - \bar{\pi}(\theta)] f(\theta) d\theta$$

subject to the incentive constraints

$$\theta \in \text{argmax}[u(x(\hat{\theta}),\theta) - u(x(\hat{\theta}),\hat{\theta}) + r(\hat{\theta}) + \bar{\pi}(\hat{\theta}) \mid \hat{\theta} \in \Theta], \quad \forall \theta \in \Theta,$$

and the participation constraints

$$r(\theta) \geqslant 0, \quad \forall \theta \in \Theta.$$

Finally, a direct revelation mechanism $\langle x(\cdot), r(\cdot)\rangle$ is *individually rational* if $r(\theta) \geqslant 0$ for every $\theta \in \Theta$. The monotonicity of the information rent $r(\cdot)$ is not guaranteed in general.

3 Voluntary Implementability via Price Schedule

In this section, I construct a price schedule under which any non-decreasing decision rule $x : \Theta \to X$ emerges as a solution to utility maximization problem of consumers. Without loss of generality, I may assume that $X = [x(\underline{\theta}), x(\overline{\theta})]$, given a non-decreasing decision rule $x(\cdot)$.

The literature has been focused on the following incentive compatibility through a transfer function defined over the type space.

Definition 1 (Rochet, 1987)**.** A decision rule $x(\cdot)$ is said to be *rationalizable or implementable via transfer* if there exists a *transfer function* $p : \Theta \to \mathbb{R}$ such that the direct revelation mechanism $\langle x(\cdot), p(\cdot)\rangle$ induces truthful revelation: $\theta \in \text{argmax}[u(x(\hat{\theta}),\theta) - p(\hat{\theta}) \mid \hat{\theta} \in \Theta]$ for every $\theta \in \Theta$.

On the other hand, my concern is the implementability in the following sense.

Definition 2. A decision rule $x(\cdot)$ is said to be *implementable via price schedule* if there exists a price schedule $t : X \to \mathbb{R}$ such that $x(\theta) \in \text{argmax}[u(x,\theta) - t(x) \mid x \in X]$ for every $\theta \in \Theta$.

When the reservation utility is not type-dependent, Rochet (1985, Principle 2) shows how to recover the requirement (1) in the above definition. The participation constraints are incorporated into the implementability in the following manner.

Definition 3. A direct revelation mechanism $\langle x(\cdot), r(\cdot)\rangle$ is *voluntarily implementable via price schedule* if there exists a price schedule $t : X \to \mathbb{R}$ such that for every $\theta \in \Theta$,
(1) $x(\theta) \in \text{argmax}[u(x,\theta) - t(x) \mid x \in X]$,
(2) $r(\theta) = u(x(\theta),\theta) - t(x(\theta)) - \bar{\pi}(\theta) \geqslant 0$.

The necessary conditions for the voluntary implementability are summarized as follows.

Theorem 1 (Revelation Principle)**.** *If a direct revelation mechanism $\langle x(\cdot), r(\cdot)\rangle$ is voluntarily implementable, then it is incentive compatible and individually rational.*

Proof. The revelation principle (cf. Laffont and Martimort, 2002, Proposition 2.2) states that condition (1) in Definition 3 implies the incentive compatibility. The individual rationality is trivially satisfied by condition (2) in Definition 3. This establishes the theorem. □

In what follows, I shall explore the reverse of the revelation principle. My question is whether it is possible to construct a nonlinear price schedule for *any* given incentive compatible and individually rational direct revelation mechanism for the voluntary implementability.

Assumption 1. The reservation utility function $\bar{\pi}(\cdot)$ is differentiable almost everywhere.

As is well-known, the incentive compatibility is characterized as follows.

Lemma 1. *A direct revelation mechanism $\langle x(\cdot), r(\cdot)\rangle$ is incentive compatible if and only if $x(\cdot)$ is non-decreasing and $\dot{r}(\theta) = u_\theta(x(\theta),\theta) - \dot{\bar{\pi}}(\theta)$ for every $\theta \in \Theta$.*

The purpose of the paper is to show how to construct a price schedule satisfying the voluntary implementability for any feasible direct revelation mechanisms. The following theorem states that the reverse of Theorem 1 actually holds.

Theorem 2 (Taxation Principle)**.** *If a direct revelation mechanism $\langle x(\cdot), r(\cdot)\rangle$ is incentive compatible and individually rational, then it is voluntarily implementable.*

Proof. For each $x \in X$, define

$$t(x) = \max\,[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x,\hat{\theta}) \mid \hat{\theta} \in \Theta].$$

Step 1. $x(\theta) \in$ *argmax*$[u(x,\theta) - t(x) \mid x \in X]$ *for every* $\theta \in \Theta$.

Proof of Step 1. It suffices to show that $\pi_t(\theta) = u(x(\theta),\theta) - t(x(\theta))$ for every $\theta \in \Theta$. Consider any $\theta \in \Theta$.

Claim 1. $u(x(\theta),\theta) - t(x(\theta)) = r(\theta) + \bar{\pi}(\theta)$.

Proof of Claim 1. Let $h(x,\hat{\theta}) = u(x,\hat{\theta}) - (r(\hat{\theta}) + \bar{\pi}(\hat{\theta}))$. Using the envelope condition $\dot{r}(\theta) = u_\theta(x(\theta),\theta) - \dot{\bar{\pi}}(\theta)$ in Lemma 1, I obtain $h_\theta(x,\hat{\theta}) = u_\theta(x,\hat{\theta}) - u_\theta(x(\hat{\theta}),\hat{\theta}) = \displaystyle\int_{x(\hat{\theta})}^{x} u_{x\theta}(z,\hat{\theta})dz$. The first-order condition $0 = h_\theta(x(\theta),\hat{\theta})$ yields $x(\theta) = x(\hat{\theta})$ by the single-crossing property. The second-order condition at $x = x(\theta)$ is written as $h_{\theta\theta}(x(\theta),\hat{\theta}) = u_{\theta\theta}(x(\theta),\hat{\theta}) - u_{x\theta}(x(\hat{\theta}),\hat{\theta}) \cdot \dot{x}(\hat{\theta}) - u_{\theta\theta}(x(\hat{\theta}),\hat{\theta}) = -u_{x\theta}(x(\hat{\theta}),\hat{\theta}) \cdot \dot{x}(\hat{\theta}) \leqslant 0$ almost everywhere. Therefore, $h(x(\theta),\hat{\theta})$ is maximized at $\hat{\theta} = \theta$. Thus, $t(x(\theta)) = h(x(\theta),\theta) = u(x(\theta),\theta) - (r(\theta) + \bar{\pi}(\theta))$. This establishes the claim.

Claim 2. $\pi_t(\theta) = r(\theta) + \bar{\pi}(\theta)$.

Proof of Claim 2. Consider any $x \in X$. By the definition of $t(x)$, I see that $t(x) \geqslant -(r(\theta) + \bar{\pi}(\theta)) + u(x,\theta)$, which implies that $u(x,\theta) - t(x) \leqslant r(\theta) + \bar{\pi}(\theta)$. Since x was arbitrary, it follows that $\pi_t(\theta) \leqslant r(\theta) + \bar{\pi}(\theta)$. It remains to show that $\pi_t(\theta) \geqslant r(\theta) + \bar{\pi}(\theta)$. I see that $\pi_t(\theta) - (r(\theta) + \bar{\pi}(\theta)) = \max[u(x,\theta) - t(x) \mid x \in X] - (r(\theta) + \bar{\pi}(\theta)) \geqslant u(x(\theta),\theta) - t(x(\theta)) - (r(\theta) + \bar{\pi}(\theta)) = 0$, where the last equality follows from Claim 1. Therefore, $\pi_t(\theta) \geqslant r(\theta) + \bar{\pi}(\theta)$. This establishes the claim.

By Claims 1 and 2, $\pi_t(\theta) = u(x(\theta),\theta) - t(x(\theta))$. This establishes the step.

Step 2. $r(\theta) = \pi_t(\theta) - \bar{\pi}(\theta) \geqslant 0$.

Proof of Step 2. The equality is immediate from Claim 2. By the individual rationality, $r(\theta) \geqslant 0$ for every $\theta \in \Theta$. By Step 1, $\pi_t(\theta) = u(x(\theta),\theta) - t(x(\theta)) = r(\theta) + \bar{\pi}(\theta) \geqslant \bar{\pi}(\theta)$. This establishes the inequality and the step.

Steps 1 and 2 establish the theorem. □

I have established the following characterization result.

Theorem 3. *A direct revelation mechanism* $\langle x(\cdot), r(\cdot)\rangle$ *is incentive compatible and individually rational if and only if it is voluntarily implementable.*

Proof. Immediate from Theorems 1 and 2. □

I have constructed a particular price schedule for voluntary implementation in the proof of Theorem 2. The following remark states that, without loss of generality, I can restrict attention to the price schedule $t(x) = \max[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x,\hat{\theta}) \mid \hat{\theta} \in \Theta]$, and is hereafter referred to as the *optimal* price schedule.

Proposition 1. *If a price schedule voluntarily implements a direct revelation mechanism $\langle x(\cdot), r(\cdot)\rangle$, then such a price schedule is uniquely determined over the range of the decision rule.*

Proof. Suppose that $t(\cdot)$ and $\tilde{t}(\cdot)$ voluntarily implements $x(\cdot)$. Then, for every $\theta \in \Theta$, $u(x(\theta), \theta) - t(x(\theta)) - \bar{\pi}(\theta) = r(\theta) = u(x(\theta), \theta) - \tilde{t}(x(\theta)) - \bar{\pi}(\theta)$, which yields that $t(x(\theta)) = \tilde{t}(x(\theta))$. Therefore, $t(x) = \tilde{t}(x)$ for every $x \in [x(\underline{\theta}), x(\overline{\theta})]$. $\square$

By the previous proposition, whenever I mention the *optimal* price schedule, it is of the form: for each $x \in [x(\underline{\theta}), x(\overline{\theta})]$,

$$t(x) = -(r(\psi(x)) + \bar{\pi}(\psi(x))) + u(x, \psi(x)),$$
$$\text{where } \psi(x) \in \text{argmax}[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta].$$

To end this section, I want to summarize some properties of the optimal price schedule constructed in the proof of Theorem 2.

Proposition 2. *The optimal price schedule $t(\cdot)$ is continuous and increasing.*

Proof. The optimal price schedule $t(x) = \max[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta]$ is continuous by the maximum theorem because the type space is compact. It remains to show the monotonicity of $t(\cdot)$. Let $\psi(x) \in \text{argmax}[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta]$. If $x > y$, then

$$\begin{aligned} t(x) - t(y) &\geqslant -(r(\psi(y)) + \bar{\pi}(\psi(y))) + u(x, \psi(y)) - [-(r(\psi(y)) + \bar{\pi}(\psi(y))) + u(y, \psi(y))] \\ &= u(x, \psi(y)) - u(y, \psi(y)) = \int_y^x u_x(z, \psi(y)) dz > 0 \end{aligned}$$

because $u_x(x, \theta) > 0$.[4] Hence, $t(x) > t(y)$. This establishes the proposition. $\square$

Since the price schedule $t(x) = \max[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta]$ is increasing by Proposition 2, it is differentiable almost everywhere, and it has kinks at non-differentiable points.

Theorem 4. *The optimal price schedule $t(\cdot)$ satisfies the following envelope condition:*

$$t'(x) = u_x(x, \psi(x)),$$
$$\text{where } \psi(x) \in \text{argmax}[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta].$$

Proof. Immediate from applying the envelope theorem to $t(x) = \max[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta]$. $\square$

[4] This implies that the optimal price schedule $t(\cdot)$ is a *flat rate tariff* if $u_x(x, \theta) = 0$ for every $x \in X$ and $\theta \in \Theta$.

The standard approach to the screening problems (when the reservation utility is not type-dependent) employs a non-decreasing decision rule $x : \Theta \to X$. Instead of using such a non-decreasing decision rule $x(\cdot)$, Goldman et al. (1984) consider a dual problem as choosing a non-decreasing type assignment function $\psi : X \to \Theta$ for a Ramsey pricing problem. Nöldeke and Samuelson (2007) reformulate principal-agent problems as choosing a non-decreasing function $\psi : X \to \Theta$ satisfying the envelope condition $t'(x) = u_x(x, \psi(x))$ shown in the previous proposition.

In addition, Maskin and Riley (1984, Proposition 3) also derive the optimal price schedule in an extended framework of Mussa and Rosen (1978). Their result can be obtained as a corollary of Theorem 2.

Example 2 (Bunching)**.** Just for an illustration, I shall derive the optimal price function in the monopoly case in which the reservation utility is zero for every type. Let $\overline{\theta} > 5a > 0$ and $b > 0$. Let θ be distributed on $[\underline{\theta}, \overline{\theta}]$, where $\underline{\theta} = \frac{1}{2}(\overline{\theta} + a)$. Let $\theta_a = \frac{1}{4}(3\overline{\theta} - \underline{\theta})$ and $\theta_b = \frac{1}{4}(5\overline{\theta} - 3\underline{\theta})$. Substituting $\underline{\theta} = \frac{1}{2}(\overline{\theta} + a)$, I see that $\theta_a = \frac{1}{8}(5\overline{\theta} - a)$ and $\theta_b = \frac{1}{8}(7\overline{\theta} - 3a)$. Then, $\theta_a < \theta_b$ if and only if $\overline{\theta} > a$. Consider the following decision rule depicted in Figure 5:

$$x(\theta) = \begin{cases} \frac{1}{b}(2\theta - (\overline{\theta} + a)) & \text{for } \theta \in [\underline{\theta}, \theta_a) \\ \frac{1}{2b}(\overline{\theta} - 2a - \underline{\theta}) & \text{for } \theta \in [\theta_a, \theta_b] \\ \frac{1}{2b}(4\theta - (3\overline{\theta} + a)) & \text{for } \theta \in (\theta_b, \overline{\theta}]. \end{cases}$$

There is pooling at $y = \frac{1}{2b}(\overline{\theta} - 2a - \underline{\theta})$. Figure 6 illustrates the optimal price schedule $\langle \mathbb{R}_+, t(\cdot) \rangle$ that implements $x(\cdot)$. The corresponding marginal price function is given by

$$t'(x) = \begin{cases} \frac{1}{2}(bx + \overline{\theta} + a) & \text{for } x \leqslant \frac{1}{2b}(\overline{\theta} - 2a - \underline{\theta}) \\ \frac{1}{4}(2bx + 3\overline{\theta} + a) & \text{for } x > \frac{1}{2b}(\overline{\theta} - 2a - \underline{\theta}). \end{cases}$$

There is a jump in the optimal marginal price function at $y = \frac{1}{2b}(\overline{\theta} - 2a - \underline{\theta})$: $\lim_{x \uparrow y} t'(x) = \theta_a < \theta_b = \lim_{x \downarrow y} t'(x)$ (see Figure 6). Moreover, $\lim_{x \downarrow 0} t'(x) = \underline{\theta}$.

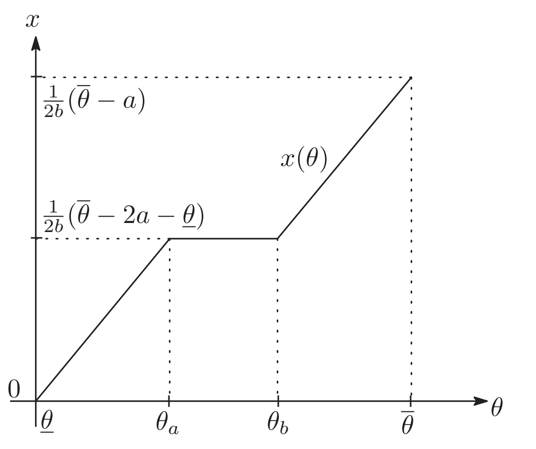


Figure 5: Pooling over $[\theta_a, \theta_b]$

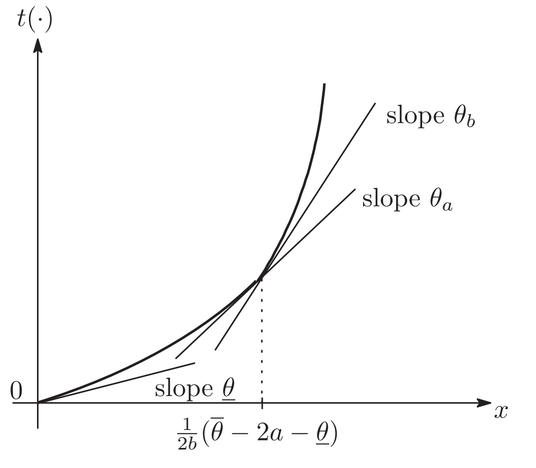


Figure 6: Kink at $\frac{1}{2b}(\overline{\theta} - 2a - \underline{\theta})$

The optimal price schedule $t : X \to \mathbb{R}$ is anonymous. It is defined over the product line X, rather than over the type space Θ. $\square$

4 Duality Approach

In this section, I shall derive further properties of the optimal price schedule $t(\cdot)$. In Theorem 2, the optimal price schedule is obtained as $t(x) = -(r(\psi(x)) + \bar{\pi}(\psi(x))) + u(x, \psi(x))$, where $\psi(x) \in \text{argmax}[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta]$. Let us call $\mathit{\Gamma}(\cdot) = \text{argmax}[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(\cdot, \hat{\theta}) \mid \hat{\theta} \in \Theta]$ the *price adjustment correspondence*. It will be shown that a selection $\psi : X \to \Theta$ from $\mathit{\Gamma} : X \twoheadrightarrow \Theta$ can be referred to as a *type assignment function*.

4.1 Price Adjustment Correspondence

Any non-decreasing decision rule $x(\cdot)$ is implementable by the optimal price schedule. Intuitively speaking, the optimal price schedule calculates the optimal marginal price or the optimal targeted willingness to pay for each $x \in X$.

Definition 4. A decision rule $x(\cdot)$ *has a bunch at* $y \in X$ if $x(\theta) = y$ over some $[\theta_1, \theta_2] \subseteq \Theta$ with $\theta_1 < \theta_2$.

Lemma 2. *Let $\langle x(\cdot), r(\cdot)\rangle$ be an incentive compatible and individually rational direct revelation mechanism. Then,*
(1) $\mathit{\Gamma}(y)$ is a compact subset of Θ for every $y \in X$.
(2) the composite $\mathit{\Gamma} \circ x : \Theta \twoheadrightarrow \Theta$ is self-belonging, that is, $\theta \in \mathit{\Gamma}(x(\theta))$ for every $\theta \in \Theta$.
(3) $\mathit{\Gamma}(y) = \text{argmax}\left[u(y, \hat{\theta}) - \int_{\theta^*}^{\hat{\theta}} u_\theta(x(s), s)ds \mid \hat{\theta} \in \Theta\right] - \bar{\pi}(\theta^*)$ where $\theta^* \in \Theta$ such that $r(\theta^*) = 0$ for every $y \in X$.

Proof. (1) Immediate from Berge's maximum theorem.
(3) Since $\dot{r}(\theta) = u_\theta(x(\theta), \theta) - \dot{\bar{\pi}}(\theta)$, it follows that $r(\theta) = \int_{\theta^*}^{\theta} [u_\theta(x(s), s) - \dot{\bar{\pi}}(s)]ds$, where $\theta^* \in \Theta$ such that $r(\theta^*) = 0$. Recall that $\mathit{\Gamma}(x) = \text{argmax}\left[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta\right]$. The expression inside square brackets can be written as

$$\begin{aligned} -(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) &= -\int_{\theta^*}^{\hat{\theta}} [u_\theta(x(s), s) - \dot{\bar{\pi}}(s)]ds - \bar{\pi}(\hat{\theta}) + u(x, \hat{\theta}) \\ &= -\int_{\theta^*}^{\hat{\theta}} u_\theta(x(s), s)ds + [\bar{\pi}(\hat{\theta}) - \bar{\pi}(\theta^*)] - \bar{\pi}(\hat{\theta}) + u(x, \hat{\theta}) \\ &= u(x, \hat{\theta}) - \int_{\theta^*}^{\hat{\theta}} u_\theta(x(s), s)ds - \bar{\pi}(\theta^*), \end{aligned}$$

and hence,

$$\mathit{\Gamma}(x) = \text{argmax}\left[u(x,\hat{\theta}) - \int_{\theta^*}^{\hat{\theta}} u_\theta(x(s),s)ds \mid \hat{\theta} \in \Theta\right] - \bar{\pi}(\theta^*).$$

(2) Suppose, by way of contradiction, that $\theta \notin \mathit{\Gamma}(x(\theta))$. By definition, $\psi(x(\theta)) \in \mathit{\Gamma}(x(\theta))$. Since $\theta \in \Theta$, it must be the case that

$$u(x(\theta),\psi(x(\theta))) - \int_{\theta^*}^{\psi(x(\theta))} u_\theta(x(s),s)ds > u(x(\theta),\theta) - \int_{\theta^*}^{\theta} u_\theta(x(s),s)ds.$$

There are two possible cases to be considered. If $\theta \geqslant \psi(x(\theta))$, then this inequality gives

$$\begin{aligned} 0 &> \int_{\psi(x(\theta))}^{\theta} u_\theta(x(\theta),s)ds - \int_{\psi(x(\theta))}^{\theta} u_\theta(x(s),s)ds \\ &\geqslant \int_{\psi(x(\theta))}^{\theta} u_\theta(x(\theta),s)ds - \int_{\psi(x(\theta))}^{\theta} u_\theta(x(\theta),s)ds = 0. \end{aligned}$$

This is a contradiction. If $\theta \leqslant \psi(x(\theta))$, then the above inequality gives

$$\begin{aligned} 0 &> \int_{\psi(x(\theta))}^{\theta} u_\theta(x(\theta),s)ds + \int_{\theta}^{\psi(x(\theta))} u_\theta(x(s),s)ds \\ &\geqslant -\int_{\theta}^{\psi(x(\theta))} u_\theta(x(\theta),s)ds + \int_{\theta}^{\psi(x(\theta))} u_\theta(x(\theta),s)ds = 0. \end{aligned}$$

This is a contradiction. Therefore, it must be the case $\theta \in \mathit{\Gamma}(x(\theta))$. □

4.2 Type-Assignment Function

The following proposition states that the construction of the optimal price schedule actually involves the inverse of decision rules.

Lemma 3. *Let $\langle x(\cdot), r(\cdot)\rangle$ be an incentive compatible and individually rational direct revelation mechanism. For every selection $\psi(\cdot) \in \mathit{\Gamma}(\cdot)$,*
(1) the composite $x \circ \psi : X \to X$ is the identity.
(2) the composite $\psi \circ x : \Theta \to \Theta$ is the identity at which $x(\cdot)$ has no bunch.

Proof. (1) Let $y \in X$. Any interior optimum $\psi(y)$ yields the following first-order condition:

$$\begin{aligned} 0 &= -(\dot{r}(\psi(y)) + \dot{\bar{\pi}}(\psi(y))) + u_\theta(y,\psi(y)) \\ &= -u_\theta(x(\psi(y)),\psi(y)) + u_\theta(y,\psi(y)) \\ &= \int_{x(\psi(y))}^{y} u_{x\theta}(z,\psi(y))dz, \end{aligned}$$

and hence $y = x(\psi(y)) = (x \circ \psi)(y)$. This establishes the assertion.

(2) Without loss of generality, I may choose $\psi(x) = \min \Gamma(x)$. Since $\theta \in \Gamma(x(\theta))$ and $\psi(x(\theta)) \in \Gamma(x(\theta))$, it follows that $\psi(x(\theta)) \leqslant \theta$. It remains to show that this holds with equality. Suppose, by way of contradiction, that $\theta > \psi(x(\theta))$. If $x(\cdot)$ has no bunch, then it is strictly increasing. This yields that $x(\theta) > x(\psi(x(\theta))) = (x \circ \psi)(x(\theta)) = x(\theta)$, a contradiction. Therefore, it must be the case $\psi(x(\theta)) = \theta$. This establishes the assertion. □

The above observation holds regardless of the differentiability of $t(\cdot)$ at y.[5]

Theorem 5. *Let $\langle x(\cdot), r(\cdot)\rangle$ be an incentive compatible and individually rational direct revelation mechanism. Then, $\Gamma(y) = \{\theta \in \Theta \mid x(\theta) = y\}$ for every $y \in X$.*

Proof. It suffices to consider $y \in X$ at which $x(\cdot)$ has a bunch. Since $x(\cdot)$ is non-decreasing by Proposition 1, it must be the case that $\{\theta \in \Theta \mid x(\theta) = y\} = [\theta_1, \theta_2]$ for some $\theta_1, \theta_2 \in \Theta$ with $\theta_1 < \theta_2$. Since $x(\theta) = y$ for every $\theta \in [\theta_1, \theta_2]$, it follows from Lemma 2(2) that $\theta \in \Gamma(x(\theta)) = \Gamma(y)$, which implies that $[\theta_1, \theta_2] \subseteq \Gamma(y)$. It remains to show that $\Gamma(y) \subseteq [\theta_1, \theta_2]$. Let $\theta \in \Gamma(y)$. Then, $\theta = \psi(y)$ for some selection $\psi(y)$ from $\Gamma(y)$. Then, $x(\theta) = x(\psi(y)) = (x \circ \psi)(y) = y$, where the last equality follows from Lemma 3(1). This implies that $\theta \in [\theta_1, \theta_2]$. Therefore, $\Gamma(y) \subseteq [\theta_1, \theta_2]$. This establishes the proposition. □

Corollary 1. *Let $\langle x(\cdot), r(\cdot)\rangle$ be an incentive compatible and individually rational direct revelation mechanism. Then, $\Gamma(y) = \{x^{-1}(y)\}$ for every $y \in X$ at which $x(\cdot)$ has no bunch.*

5 Applications of the Analysis

5.1 Block Tariffs, Quality Premia, and Quantity Discounts

Block Tariffs are simple, but it is trivial that the principal could increase his expected profit beyond that achievable with an optimal block tariff by choosing a more complex nonlinear pricing scheme. A *two-part tariff* $t(x) = px + q$ can be considered as a special case of a block tariff. The current model corresponds to second-degree price discrimination due to Pigou (1932). In the context of first-degree price discrimination, the use of two-part tariffs has been discussed. However, the optimality of a two-part tariff or a block tariff in general has not been explored in the context of second-degree price discrimination because of the intensive use of direct revelation mechanisms.

Definition 5. A *block tariff* is a piecewise linear price schedule.

[5] The second assertion can be proved, provided that the optimal price schedule $t(\cdot)$ is differentiable. Recall that the implementability constraint is satisfied under the price schedule $t(x) = \max[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x, \hat{\theta}) \mid \hat{\theta} \in \Theta]$, that is, $x(\theta) \in \text{argmax}[u(x, \theta) - t(x) \mid x \in X]$. The first-order condition is given by $0 = u_x(x(\theta), \theta) - t'(x(\theta))$. By Theorem 4, $t'(x) = u_x(x, \psi(x))$. Thus, $u_x(x(\theta), \theta) = u_x(x(\theta), \psi(x(\theta)))$. By the strict single-crossing property $u_{x\theta}(x, \theta) > 0$, it must be the case $\theta = \psi(x(\theta)) = (\psi \circ x)(\theta)$.

The following theorem provides a necessary and sufficient condition for that the optimal price schedule obtained in Theorem 2 takes the form of a block tariff indeed.

Theorem 6 (Block Tariff). *Let $\langle x(\cdot), r(\cdot)\rangle$ be an incentive compatible and individually rational direct revelation mechanism. Then, the following statements are equivalent:*
(1) the optimal price schedule $t(\cdot)$ takes the form of a block tariff.
(2) $\frac{d}{d\theta}u_x(x(\theta),\theta) = 0$ except at $\theta \in \Theta$ where $x(\cdot)$ has no bunch at $x(\theta)$.

Proof. Recall the envelope condition $t'(x) = u_x(x, \psi(x))$, where $\psi(x) \in \text{argmax}[-(r(\hat{\theta}) + \bar{\pi}(\hat{\theta})) + u(x,\hat{\theta}) \mid \hat{\theta} \in \Theta]$. In addition, $\psi(y) = x^{-1}(y)$ if and only if $x(\cdot)$ has no bunch at $y \in X$ by Lemma 3. Differentiating $t'(\cdot)$ to obtain

$$\begin{aligned} t''(y) &= u_{xx}(y,\psi(y)) + u_{x\theta}(y,\psi(y)) \cdot \psi'(y) \\ &= u_{xx}(x(\theta),\theta) + u_{x\theta}(x(\theta),\theta) \cdot \tfrac{1}{\dot{x}(\theta)} \\ &= \tfrac{1}{\dot{x}(\theta)} \left(u_{xx}(x(\theta),\theta) \cdot \dot{x}(\theta) + u_{x\theta}(x(\theta),\theta) \right) \\ &= \tfrac{1}{\dot{x}(\theta)} \cdot \tfrac{d}{d\theta} u_x(x(\theta),\theta). \end{aligned}$$

This implies that for every $x \in X$ where the price schedule is twice differentiable, it must be the case that $t''(x) = 0$ is equivalent to saying that $\frac{d}{d\theta}u_x(x(\theta),\theta) = 0$, using the duality between $\theta = \psi(x)$ and $x = x(\theta)$. This establishes the theorem. □

As a conclusion of the previous theorem, the emergence of a block tariff as a solution to principal-agent problems depends on the form of the utility function of the agent. In the following, I consider a commonly used utility function of the agent to check whether a block tariff can realize a given decision rule.

Proposition 3. *Let $\langle x(\cdot), r(\cdot)\rangle$ be an incentive compatible and individually rational direct revelation mechanism. If the utiliy function of the agent takes of the form $u(x,\theta) = \theta v(x)$, where $v : X \to \mathbb{R}$ is strictly increasing and concave, then the optimal price schedule $t(\cdot)$ takes the form of a block tariff if and only if $\frac{d}{dy}(\psi(y)v'(y)) = 0$, where $\psi(y) = x^{-1}(y)$ except at bunching points of $x(\cdot)$.*

Proof. The expression of $t''(x)$ obtained in the proof of Theorem 6 can be written as

$$\begin{aligned} t''(y) &= u_{xx}(y,\psi(y)) + u_{x\theta}(y,\psi(y)) \cdot \psi'(y) \\ &= \psi(y) \cdot v''(y) + v'(y) \cdot \psi'(y) \\ &= \tfrac{d}{dy}(\psi(y)v'(y)) \end{aligned}$$

for every $y \in X$ except at bunching points of $x(\cdot)$. The equivalence result is immediate from Theorem 6. □

Recall that the following expression for the second-derivative of the optimal price schedule shown in the proof of Theorem 6:

$$\text{sign } t''(y) = \text{sign } [u_{xx}(x(\theta),\theta)\cdot\dot{x}(\theta) + u_{x\theta}(x(\theta),\theta)], \text{ whenever } x(\theta) \text{ has no bunch at } y.$$

This yields that the sign of $t''(y)$ is determined by the relationship between $\mid u_{xx}(y,\theta)\cdot\dot{x}(\theta)\mid$ and $\mid u_{x\theta}(y,\theta)\mid$, where $x(\theta) = y$. A few findings based on this observation can be stated as follows.

Proposition 4 (Quality Premia or Quantity Discounts)**.** *Let $\langle x(\cdot), r(\cdot)\rangle$ be an incentive compatible and individually rational direct revelation mechanism. Then,*
(1) if the optimal price schedule $t(\cdot)$ is of the form of $t(x) = px + q$, then it must be the case that $u_{xx}(x,\theta) < 0$ for every x and θ,
(2) if the optimal price schedule $t(\cdot)$ is strictly concave, then it must be the case that $u_{xx}(x,\theta) < 0$ for every x and θ,
(3) if $u_{xx}(x,\theta) = 0$ for every x and θ, then the optimal price schedule $t(\cdot)$ must be strictly convex.

To end this subsection, I will show the optimality of a quantity discounts in a commonly used framework.

Example 3 (Constant Elasticity)**.** In what follows, for illustrative purposes, I assume that $\bar{\pi}(\theta) = \bar{\pi}$ (constant) for all $\theta \in \Theta$, and $f(\cdot)$ is the uniform distribution over a unit interval $[\underline{\theta}, \overline{\theta}]$. When the reservation utility is not type-dependent, it is well-known that the principal's problem can be written as

$$\max_{x(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} \tilde{v}(x(\theta),\theta) f(\theta) d\theta - \bar{\pi} \quad \text{subject to} \quad x(\cdot) \text{ is non-decreasing},$$

where $\tilde{v}(x,\theta) = u(x,\theta) - C(x) - \frac{1-F(\theta)}{f(\theta)} u_\theta(x,\theta)$ is the *virtual surplus* function. Under the standard regularity conditions for the virtual surplus function (including $u_\theta(x,\theta) > 0$), the decision rule $x(\cdot)$ satisfying $\tilde{v}_x(x(\theta),\theta) = 0$ solves the relaxed problem, and the resulting transfer function is given as $p(\theta) = t(x(\theta)) = u(x(\theta),\theta) - \left(\int_{\underline{\theta}}^{\theta} u_\theta(x(s),s)ds + \bar{\pi}\right)$.

Now, suppose $u(x,\theta) = \theta x^{1-\eta}$, where $0 < \eta < 1$, and $C'(x) = a > 0$. Assume $\underline{\theta} > 1$.[6] Using distribution uniformity,

$$\begin{aligned}\tilde{v}_x(x,\theta) &= \theta\varphi'(x) - a - (1 - (\theta - \underline{\theta}))\varphi'(x)\\ &= (2\theta - 1 - \underline{\theta})\varphi'(x) - a = (2\theta - 1 - \underline{\theta})(1-\eta)x^{-\eta} - a,\end{aligned}$$

and thus, $0 = \tilde{v}_x(x(\theta),\theta)$ yields $x(\theta) = \big((2\theta - 1 - \underline{\theta})(1-\eta)/a\big)^{1/\eta}$. The corresponding type assignment function is given by $\psi(x) = \frac{1}{2}\big(ax^\eta/(1-\eta) + 1 + \underline{\theta}\big)$. The expression for the

[6] The virtual type $\theta - \frac{1-F(\theta)}{f(\theta)} = \theta - (1 - (\theta - \underline{\theta})) = 2\theta - 1 - \underline{\theta}$ is positive for all $\theta \in \Theta$ if $\underline{\theta} > 1$.

second-derivative of the optimal price schedule in the proof of Proposition 3 can be written as

$$\begin{aligned} t''(x) &= \tfrac{d}{dx}\left(\tfrac{1}{2}\left(ax^{\eta}/(1-\eta)+1+\underline{\theta}\right)\cdot(1-\eta)x^{-\eta}\right) \\ &= \tfrac{d}{dx}\left(\tfrac{1}{2}ax^{0}+\tfrac{1}{2}(1+\underline{\theta})(1-\eta)x^{-\eta}\right)=\tfrac{1}{2}(1+\underline{\theta})(1-\eta)(-\eta)/x^{1+\eta}<0. \end{aligned}$$

Therefore, it is optimal for the principal to offer a quantity discount price schedule rather than a block tariff or a two-part tariff. □

5.2 Deviation from Marginal Cost Pricing under Countervailing Incentives

I examine how the optimal price schedule for the principal deviates from marginal cost pricing schedule. When the social surplus $u(x,\theta)-C(x)$ is strictly concave in x for all θ, the pointwise maximization $\max\,[u(x,\theta)-C(x)\mid x\in X]$ ensures that the welfare-optimal decision rule $x^{*}(\theta)$ is uniquely determined. The strict concavity of $u(x,\theta)-C(x)$ in x, together with the single-crossing property, shows that $\dot{x}^{*}(\theta)=-[u_{xx}(x^{*}(\theta),\theta)-C''(x^{*}(\theta))]^{-1}u_{x\theta}(x^{*}(\theta))>0$ for every $\theta\in\Theta$. It will be discussed that the second-best marginal price schedules depend crucially on the shape of the type-dependent reservation $\bar{\pi}(\cdot)$. It will be discussed how the departure of the optimal marginal price schedule from marginal cost crucially depends on the steepness of the agent's information rents.

An interesting situation arises when $u_{\theta}(x,\theta)$ and $\dot{\bar{\pi}}(\cdot)$ intersect at some point $(x,\theta)\in X\times\Theta$. It will be discussed that the decision rule will be distorted except at both extremes.[7] Consider a simple case, in which $\bar{\pi}(\theta)$ is linear in θ. Assume that $u_{\theta\theta}(x,\theta)=0$ and $f(\cdot)$ is uniform.

In the principal's optimization problem, the Hamiltonian is

$$\mathcal{H}(r,x,\lambda,\theta)=[u(x,\theta)-C(x)-r-\bar{\pi}(\theta)]f(\theta)+\lambda[u_{\theta}(x,\theta)-\dot{\bar{\pi}}(\theta)],$$

and the Lagrangian is

$$\mathcal{L}(r,x,\lambda,\tau,\theta)=\mathcal{H}(r,x,\lambda,\theta)+\tau r.$$

To solve this problem, I make use of a set of sufficient conditions for optimality in the optimal control problem with pure state constraints:
(a) there exists a function $\lambda(\cdot)$, which is piecewise continuous and piecewise continuously differentiable with jump discontinuities at $\theta_1,\cdots,\theta_m$ with $\underline{\theta}_1<\theta_1<\cdots<\theta_m\leqslant\overline{\theta}$, and
(b) there exist a piecewise continuous function $\tau(\cdot)$ and non-negative numbers $\beta_1,\cdots,\beta_m$ such that the following conditions are satisfied:
(1) $x(\theta)\in$ argmax$[\mathcal{H}(r(\theta),x,\lambda(\theta),\theta)\mid x\in X]$,
(2) $\tau(\theta)\geqslant 0$ and $\tau(\theta)r(\theta)=0$,

[7] There are two extreme cases in which there is no bunching. When the slope of the reservation utility function $\bar{\pi}(\cdot)$ is sufficiently small so that $u_{\theta}(x,\theta)>\dot{\pi}(\theta)$ for every $x\in X$ and $\theta\in\Theta$, the standard "no-distortion-at-the-top" result holds, whereas when $\dot{\bar{\pi}}(\cdot)$ is sufficiently high such that $u_{\theta}(x,\theta)<\dot{\bar{\pi}}(\theta)$ for every $x\in X$ and $\theta\in\Theta$, the "no-distortion-at-the-bottom" result holds.

(3) $\dot{\lambda}(\theta) = -\frac{\partial \mathcal{L}}{\partial r}(r(\theta), x(\theta), \lambda(\theta), \tau(\theta), \theta)$, (the costate equation)
(4) $\lambda(\theta_k^-) - \lambda(\theta_k^+) = \beta_k$ (Put $\lambda(\theta_k^+) = \lambda(\overline{\theta})$ if $\theta_m = \overline{\theta}$),
(5) $\beta_k = 0$ if either $r(\theta_k) > 0$ or $[\theta_k \in \text{int } \Theta$, $r(\theta_k) = 0$ and $x(\theta)$ is discontinuous at $\theta_k]$,
(6) $\lambda(\underline{\theta})r(\underline{\theta}) = 0$ and $\lambda(\underline{\theta}) \leqslant 0$, (the initial transversality condition)
(7) $\lambda(\overline{\theta})r(\overline{\theta}) = 0$ and $\lambda(\overline{\theta}) \geqslant 0$, (the terminal transversality cond)
(8) $\hat{\mathcal{H}}(r, \lambda(\theta), \theta) = \max\left[\mathcal{H}(r, x, \lambda(\theta), \theta) \mid x \in X\right]$ is concave in r,
(9) the pure state constraint $r \geqslant 0$ is quasi-concave in r.

For each $x \in \mathbb{R}$ and $\lambda \in \mathbb{R}$, define

$$\sigma(x, \lambda, \theta) = u(x, \theta) - C(x) + \frac{\lambda}{f(\theta)} u_\theta(x, \theta).$$

To formulate the decision rule $x(\cdot)$, define the following decision rules commonly used in the literature:

$$x_u^\sigma(\theta) \in \text{argmax}\left[\sigma(x, F(\theta), \theta) \mid x \in X\right],$$
$$x_d^\sigma(\theta) \in \text{argmax}\left[\sigma(x, F(\theta) - 1, \theta) \mid x \in X\right].$$

Both decision rules $x_u^\sigma(\theta)$ and $x_d^\sigma(\theta)$ are uniquely determined by the strict concavity of $u(x, \theta) - C(x)$ in x. Moreover, these are strictly increasing, $\dot{x}_u^\sigma(\cdot) > 0$ and $\dot{x}_d^\sigma(\cdot) > 0$, because the uniform distribution satisfies the monotone hazard rate property. Since $\dot{\bar{\pi}}(\cdot)$ takes intermediate values, I may suppose that there exists $\theta_0 \in \text{int } \Theta$ such that $u_\theta(x^*(\theta), \theta) = \dot{\bar{\pi}}(\theta)$ at $\theta = \theta_0$. Also, there exist θ_1 and θ_2 with $\theta_1 \leqslant \theta_0 < \theta_2$ such that $x_u^\sigma(\theta_1) = x^*(\theta_0) = x_d^\sigma(\theta_2)$ (see Figure 7).

The critical part of the proof is to construct the right solution for $x(\cdot)$, $\lambda(\cdot)$, and $\tau(\cdot)$. Consider the following decision rule:

$$x(\theta) = \begin{cases} x_u^\sigma(\theta) & \text{for } \theta < \theta_1 \\ x^*(\theta_0) & \text{for } \theta_1 \leqslant \theta < \theta_2 \\ x_d^\sigma(\theta) & \text{for } \theta_2 \leqslant \theta. \end{cases}$$

The corresponding type-assignment function is shown in Figure 8. Because of the presence of bunching at $x = x^*(\theta_0)$, the jump from θ_1 to θ_2 occurs in the type-assignment function. I shall verify the discontinuity of $\psi(\cdot)$ at $x = x^*(\theta_0)$ after defining the information rent $r(\cdot)$.

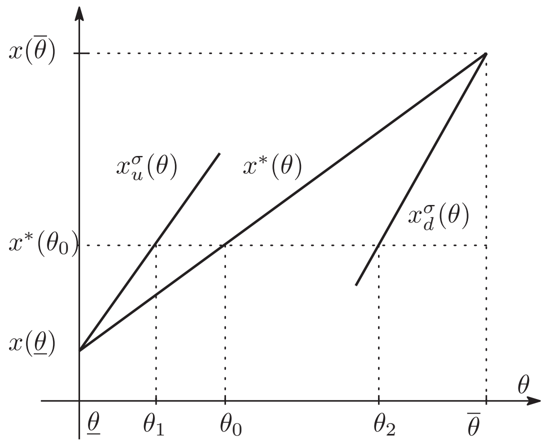


Figure 7: Decision Rules

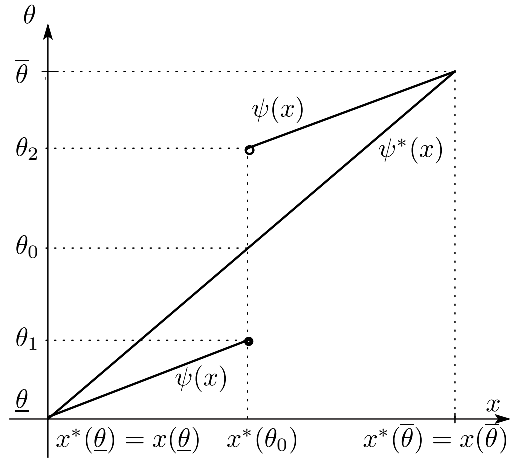


Figure 8: Type-Assignment Functions

It can be shown that the costate equation $\dot{\lambda}(\theta) = -\frac{\partial \mathcal{L}}{\partial r} = f(\theta) - \tau(\theta)$ is consistent with the following pair $\lambda(\cdot)$ and $\tau(\cdot)$:

$$\lambda(\theta) = \begin{cases} F(\theta) & \text{for } \theta < \theta_1 \\ \{\lambda \in \mathbb{R} \mid \sigma_x(x^*(\theta_0), \lambda, \theta) = 0\} & \text{for } \theta_1 \leqslant \theta < \theta_2 \\ F(\theta) - 1 & \text{for } \theta_2 \leqslant \theta, \end{cases}$$

$$\tau(\theta) = \begin{cases} 0 & \text{for } \theta < \theta_1 \\ 2f(\theta) & \text{for } \theta_1 \leqslant \theta < \theta_2 \\ 0 & \text{for } \theta_2 \leqslant \theta. \end{cases}$$

The optimal contract $\langle x(\cdot), r(\cdot)\rangle$ involves a bunch at $x(\theta) = x^*(\theta_0)$ in $[\theta_1, \theta_2]$.[8] Since the multiplier $\tau(\cdot)$ of the participation constraints $r(\cdot) \geqslant 0$ is positive in $[\theta_1, \theta_2]$,all types in $[\theta_1, \theta_2]$ earn no information rents (see Figure 10).

[8] This is the phenomenon discussed in Lewis and Sappington (1989a), Lewis and Sappington (1989b), and Maggi and Rodriguez-Clare (1995).

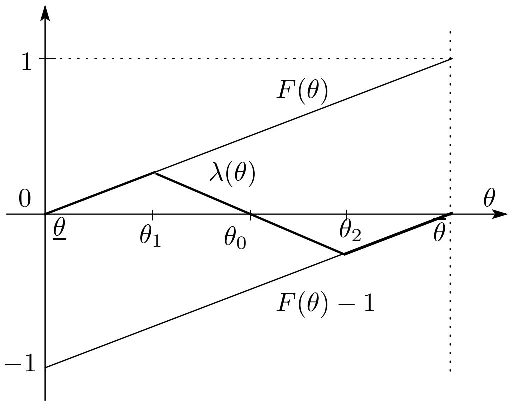


Figure 9: Multiplier associated with the Envelope Condition

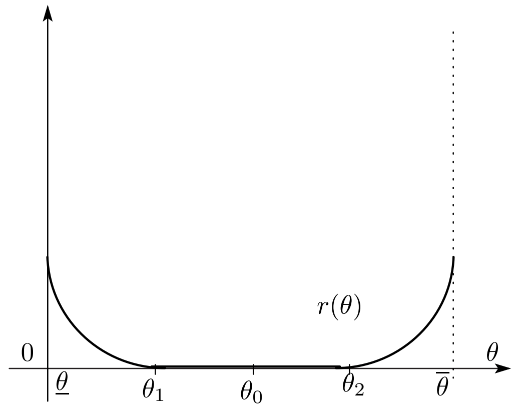


Figure 10: Information Rents

By construction, it is obvious that $\dot{\lambda}(\theta) = f(\theta) = f(\theta) - \tau(\theta)$ for every $\theta \in [\underline{\theta}, \theta_1) \cup [\theta_2, \overline{\theta}]$. Let $\theta \in [\theta_1, \theta_2)$. In order for $\lambda(\theta)$ to satisfy the costate equation, I need to check that $\dot{\lambda}(\theta) = -f(\theta)$. Solving the equation $\sigma_x(x^*(\theta_0), \lambda(\theta), \theta) = 0$ for $\lambda(\theta)$ to obtain $\lambda(\theta) = -[u_{x\theta}(x^*(\theta_0), \theta)]^{-1} v_x(x^*(\theta_0), \theta) f(\theta)$, where $v(x, \theta) = u(x, \theta) - C(x)$ is the surplus function. This gives the expression for $\dot{\lambda}(\theta)$ over the interval (θ_1, θ_2):

$$\begin{aligned}\dot{\lambda}(\theta) &= -\frac{\{v_{x\theta}(x^*(\theta_0), \theta) f(\theta) + v_x(x^*(\theta_0), \theta) \dot{f}(\theta)\} u_{x\theta}(x^*(\theta_0), \theta) - v_x(x^*(\theta_0), \theta) u_{x\theta\theta}(x^*(\theta_0), \theta) f(\theta)}{[u_{x\theta}(x^*(\theta_0), \theta)]^2} \\ &= -\frac{[u_{x\theta}(x^*(\theta_0), \theta)]^2 f(\theta)}{[u_{x\theta}(x^*(\theta_0), \theta)]^2} = -f(\theta),\end{aligned}$$

since I have assumed that $\dot{f}(\theta) = 0$ and $u_{\theta\theta}(x, \theta) = 0$.

I have shown that the sufficient conditions for optimality are satisfied. Given the optimal contract $\langle x(\cdot), r(\cdot) \rangle$, the corresponding type-assignment function is not single-valued at $x = x^*(\theta_0)$. Since the type-assignment function maximizes $-[r(\hat{\theta} + \bar{\pi}(\theta))] + u(x, \hat{\theta})$ with respect to $\hat{\theta}$, the first-order condition at $x = x^*(\theta_0)$ becomes $0 = -[\dot{r}(\hat{\theta}) + \dot{\bar{\pi}}(\hat{\theta})] + u_\theta(x^*(\theta_0), \hat{\theta}) = -u_\theta(x(\hat{\theta}), \hat{\theta}) + u_\theta(x^*(\theta_0), \hat{\theta})$ by using the envelope condition $\dot{r}(\theta) = u_\theta(x, \theta) - \dot{\bar{\pi}}(\theta)$. By the single-crossing property, an optimal choice of $\hat{\theta}$ satisfies $x(\hat{\theta}) = x^*(\theta_0)$. The second-order condition is met at $\hat{\theta} \in [\theta_1, \theta_2]$: $-[\ddot{r}(\hat{\theta}) + \ddot{\bar{\pi}}(\hat{\theta})] + u_{\theta\theta}(x^*(\theta_0), \hat{\theta}) = u_{\theta\theta}(x^*(\theta_0), \hat{\theta}) = 0$ because $\ddot{r}(\hat{\theta}) = 0$ in (θ_1, θ_2) by construction, and $\ddot{r}(\hat{\theta}) = 0 = u_{\theta\theta}(x^*(\theta_0), \hat{\theta})$. To sum up, any type $\hat{\theta} \in [\theta_1, \theta_2]$ is optimal at $x = x^*(\theta_0)$: $\underset{[}{\text{argmax}} -[r(\hat{\theta} + \bar{\pi}(\theta))] + u(x^*(\theta_0), \hat{\theta}) \mid \hat{\theta} \in \Theta] = [\theta_1, \theta_2]$. In this case, set $\psi(x^*(\theta_0)) = \theta_1$ so that the type-assignment function is continuous from the left, as shown in Figure 14.

The following propositions summarize the structure of marginal price schedules. The optimal marginal price schedule $t'(\cdot)$ intersects $C'(\cdot)$ only for the boundary points $\underline{\theta}$ and $\overline{\theta}$.

Proposition 5. *Suppose $u_\theta(x^*(\theta_0), \theta_0) = \dot{\bar{\pi}}(\theta_0)$ for some interior type θ_0. If (a) $\bar{\pi}(\cdot)$ is linear, (b) $u_{\theta\theta}(x, \theta) = 0$ for every $x \in X$ and $\theta \in \Theta$, and (c) $f(\theta)$ is uniform, then:*
(1) the optimal price schedule has a kink at the interior point $x(\theta_0)$.
(2) the optimal marginal price schedule deviates downward from marginal cost pricing schedule up to $x = x^*(\theta_0)$, whereas deviates upward from marginal cost pricing schedule beyond $x = x^*(\theta_0)$, except at the top and the bottom:

$$t'(x) \begin{cases} \leqslant C'(x) & \text{for } x \leqslant x^*(\theta_0) \text{ with equality only at } x = x^*(\underline{\theta}) = x(\underline{\theta}) \\ \geqslant C'(x) & \text{for } x \geqslant x^*(\theta_0) \text{ with equality only at } x = x^*(\overline{\theta}) = x(\overline{\theta}). \end{cases}$$

(3) Moreover, assume the sign of $t''(\cdot)$ is constant. Then the optimal price schedule is a combination of two concave price schedules for any concave cost function. On the other hand, if the cost function is convex, then the optimal price schedule is a combination of either two concave functions or two convex functions.

The shape of the optimal price schedule is indeterminate except that its marginal price jumps upward at $x(\theta_0)$ when the cost function is convex. Notice that if the sign of $t''(\cdot)$ is constant, then the convexities of the optimal price schedule and the cost function are inconsistent.

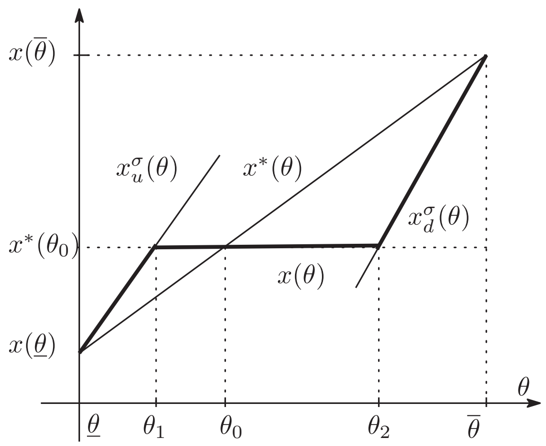


Figure 11: Marginal Price Schedules

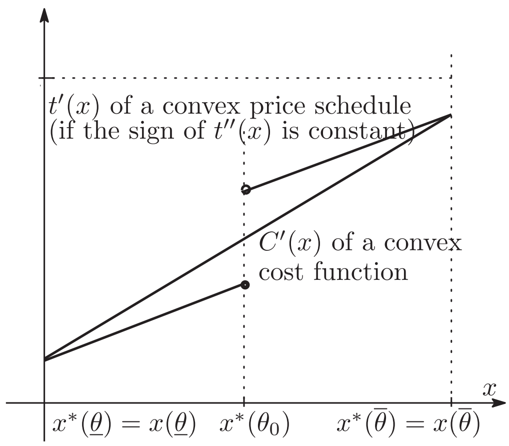


Figure 12: Marginal Price Schedules

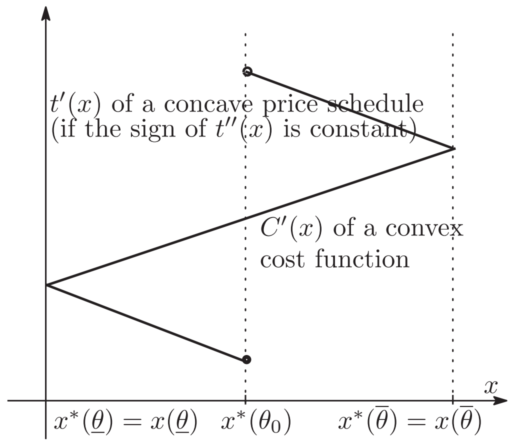


Figure 13: Decision Rules

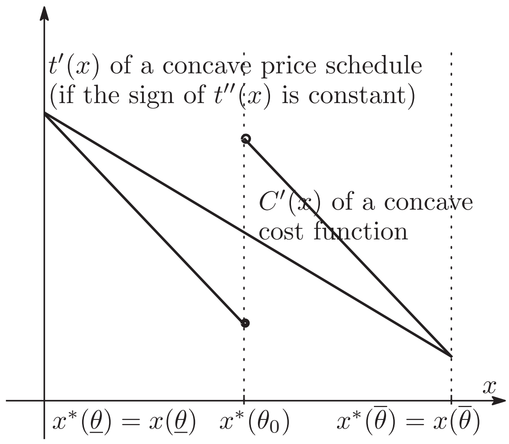


Figure 14: Marginal Price Schedules

5.3 Delegated Common Agency Games

In this section, I consider economies with two principals and one agent. Two sellers compete for an agent through nonlinear pricing schemes. As in the previous sections, one commodity x is referred to as product quality and another commodity y is a monetary transfer. Denote by $(x, y) \in \mathbb{R}^2$ the agent's net trade. In what follows, for the sake of simplicity, I assume that $u(x, \theta) = x\theta$. Throughout this section, I assume that taste for quality θ is uniformly distributed over $[\underline{\theta}, \overline{\theta}]$.[9] Let $X \subseteq \mathbb{R}_+$ be a non-empty set, interpreted as the whole range of qualities that principals can offer. A *nonlinear price schedule for principal* i is a pair $\langle X_i, t_i(\cdot)\rangle$ consisting of a *product line* $X_i \subseteq X$ and a price function $t_i : X_i \to \mathbb{R}_+$. For each $x_i \in X_i$, principal i's profit margin is given by $t_i(x_i) - C_i(x_i)$ for some cost function $C_i : X_i \to \mathbb{R}_+$. I do not allow for common values. I consider environments in which the cost functions are linear-quadratic in quality, that is, $C_i(x_i) = a_i x_i + \frac{b_i}{2} x_i^2$ for each principal i. I assume that the cost functions of the principals are not identical, but the asymmetry is small: $0 < a_1 < a_2$ and $0 < b_1 = b_2 = b$.

Assumption 2. $0 < C_1'(x) < C_2'(x)$ and $0 < C_1''(x) = C_2''(x)$ for every $x \in X$.

Consider any principal i. For each type θ, define his demand correspondence and indirect utility function from net trades by, respectively,

$$X_{t_i}(\theta) = \{x_i \in X_i \mid u(x_i, \theta) - t_i(x_i) \geqslant u(\tilde{x}_i, \theta) - t_i(\tilde{x}_i) \text{ for every } \tilde{x}_i \in X_i\}$$

and

$$\pi_{t_i}(\theta) = u(x_i(\theta), \theta) - t_i(x_i(\theta))$$

for each selection $x_i(\theta) \in X_{t_i}(\theta)$. The demand correspondence $X_{t_i}(\theta)$ might be set-valued under nonlinear pricing schedules.

[9] The uniformity is used only in the proof of Lemma 18.

I introduce a principal-specific component $\langle\alpha_1, \alpha_2\rangle$. The net utility of type θ derived from purchasing a product from principal i is assumed to be $\pi_{t_i}(\theta) + \alpha_i$. I assume that α_i is observable and independent of quality choice. The difference $\Delta\alpha_i = \alpha_j - \alpha_i$ is referred to as the *comparative advantage of principal j over principal i.*

The *social surplus* $v_i(x_i, \theta) + \alpha_i = u(x_i, \theta) + \alpha_i - C_i(x_i)$ is adjusted by α_i. Since the social surplus $v_i(x_i, \theta)$ is strictly concave in x_i under Assumption 2, it follows that the *full-information decision scheme* $x_i^*(\theta) \in \text{argmax}[v_i(x_i, \theta) \mid x_i \in X_i]$ is uniquely determined. Going back to the linear-quadratic environment, I obtain the following full-information outcome: for each principal i,

$$x_i^*(\theta) = \tfrac{1}{b}(\theta - a_i) \text{ and } v_i(x_i^*(\theta), \theta) = \tfrac{1}{2b}(\theta - a_i)^2.$$

Since the social surplus functions are linear in the comparative advantage parameters $\langle\alpha_1, \alpha_2\rangle$, the full-information decision schemes $\langle x_1^*(\cdot), x_2^*(\cdot)\rangle$ won't be affected by $\langle\alpha_1, \alpha_2\rangle$. Finally, I assume that the production of the lowest quality is socially valuable when there is no asymmetry of information.

Assumption 3. $v_i(x_i^*(\underline{\theta}), \underline{\theta}) + \alpha_i \geqslant 0.$

This assumption is rewritten as $\frac{1}{2b}(\underline{\theta} - a_i)^2 + \alpha_i \geqslant 0$. To end this subsection, I summarize properties of the full-information schemes.

Lemma 4. $x_1^*(\theta) \geqslant x_2^*(\theta)$ *and* $\dot{x}_1^*(\theta) = \dot{x}_2^*(\theta) > 0$ *for every* $\theta \in \Theta$.

5.3.1 Competitive Implementability

I model strategic competition among the principals in a market with asymmetric information as a non-cooperative game. The principals move first, choosing their price schedules simultaneously, given the distribution of types and the optimizing behavior of the agent. The agent moves second, contracting with at most one principal. This is a situation of the delegated common agency. Principal i's payoff is determined by the price schedules offered by all the principals and the induced demand correspondence. Denote the *client set of principal* i by

$$\Theta_i = \{\theta \in \Theta \mid \pi_{t_i}(\theta) + \alpha_i \geqslant \max\{0, \pi_{t_j}(\theta) + \alpha_j\}\}.$$

Each principal i faces his own *captive* market $\Theta_i \setminus \Theta_j$ and the *competitive* market $\Theta_i \cap \Theta_j$.[10] Principal i is strictly dominant in his captive market, while he is weakly dominant in the competitive market. In what follows, without loss of generality, I may assume that $\pi_{t_j}(\cdot) + \alpha_j \geqslant 0$, and thus $\pi_{t_i}(\cdot) + \alpha_i \geqslant \max\{0, \pi_{t_j}(\cdot) + \alpha_j\} = \pi_{t_j}(\cdot) + \alpha_j$.

Given the price schedule $\langle X_j, t_j(\cdot)\rangle$ posted by his competitor, the problem of principal i is to design his price schedule $\langle X_i, t_i(\cdot)\rangle$, so as to maximize his expected payoff defined by

$$P_i(t_i) = \int_{\Theta_i} [t_i(x_i(\theta)) - C_i(x_i(\theta))] f(\theta) d\theta$$

[10] Notations on set operations: $A \setminus B = \{a \mid a \in A \text{ and } a \notin B\}$, $A \cap B = \{a \mid a \in A \text{ and } a \in B\}$.

subject to the incentive constraints

$$x_i(\theta) \in \text{argmax}[u(x_i, \theta) - t_i(x_i) \mid x_i \in X_i], \quad \forall \theta \in \Theta_i.$$

The participation constraints are incorporated into the client set Θ_i. I assume that the agent of type θ obediently follows principal i's instruction on his decision $x_i(\cdot)$ if the demand correspondence is set-valued. Denote by S_i the finite set of all possible price schedules for principal i, and by $s_i = \langle X_i, t_i(\cdot)\rangle$ a generic element of S_i. Denote the set of best responses for principal i to the opponent's strategy s_j by $\mathcal{N}_i(s_j) \subseteq S_i$. By construction, any strategy profile $(s_1, s_2) \in S_1 \times S_2$ such that $(s_1, s_2) \in \mathcal{N}_1(s_2) \times \mathcal{N}_2(s_1)$ is a Nash equilibrium. The price schedule $\langle X_i, t_i(\cdot)\rangle$ is the instrument chosen by principal i. Given the opponent's strategy $\langle X_j, t_j(\cdot)\rangle$, I shall transform principal i's problem that consists in using a pair of an *decision rule* $x_i(\cdot)$ and an *information rent* $r_i(\cdot) = \pi_{t_i}(\cdot) + \alpha_i - \max\{0, \pi_{t_j}(\cdot) + \alpha_j\}$, rather than the payment $p_i(\cdot) = t_i(x_i(\cdot))$, as the instruments.

Principal i is weakly dominant over his client set Θ_i. The original participation constraints depends on the opponent's strategy $\langle X_j, t_j(\cdot)\rangle$ in a complex fashion. I replace the participation constraints with the system of inequalities, $r_i(\theta) = \pi_{t_i}(\theta) - \pi_{t_j}(\theta) - \Delta\alpha_i \geqslant 0$ for every $\theta \in \Theta_i$. Then, profit margin is written as $t_i(x_i(\theta)) - C_i(x_i(\theta)) = u(x_i(\theta), \theta) - \pi_{t_i}(\theta) - C_i(x_i(\theta)) = v_i(x_i(\theta), \theta) - r_i(\theta) - \pi_{t_j}(\theta) - \Delta\alpha_i$. If $x_i(\theta) \in \text{argmax}[u(x_i, \theta) - t_i(x_i) \mid x_i \in X_i]$ holds for every $\theta \in \Theta_i$ under the price schedule $\langle X_i, t_i(\cdot)\rangle$, then it must be the case $\theta \in \text{argmax}[u(x_i(\hat{\theta}), \theta) - t_i(x_i(\hat{\theta})) \mid \hat{\theta} \in \Theta_i]$ for every $\theta \in \Theta_i$. Using the information rent $r_i(\theta) = \pi_{t_i}(\theta) - \pi_{t_j}(\theta) - \Delta\alpha_i$, the payment $t_i(x_i(\hat{\theta}))$ in this expression is written as $t_i(x_i(\hat{\theta})) = u(x_i(\hat{\theta}), \hat{\theta}) - \pi_{t_i}(\hat{\theta}) = u(x_i(\hat{\theta}), \hat{\theta}) - (r_i(\hat{\theta}) + \pi_{t_j}(\hat{\theta}) + \Delta\alpha_i)$. Therefore, the pair $\{x_i(\theta), r_i(\theta)\}_{\theta \in \Theta_i}$ is *incentive compatible* in the sense that $\theta \in \text{argmax}[u(x_i(\hat{\theta}), \theta) - u(x_i(\hat{\theta}), \hat{\theta}) + r_i(\hat{\theta}) + \pi_{t_j}(\hat{\theta}) + \Delta\alpha_i \mid \hat{\theta} \in \Theta_i]$ for every $\theta \in \Theta_i$. Given the opponent's strategy $\langle X_j, t_j(\cdot)\rangle$, principal i's problem can be written as

$$\max_{\langle x_i(\cdot), r_i(\cdot), \Theta_i\rangle} \int_{\Theta_i} [v_i(x_i(\theta), \theta) - r_i(\theta) - \pi_{t_j}(\theta) - \Delta\alpha_i] f(\theta) d\theta$$

subject to the incentive constraints

$$\theta \in \text{argmax}[u(x_i(\hat{\theta}), \theta) - u(x_i(\hat{\theta}), \hat{\theta}) + r_i(\hat{\theta}) + \pi_{t_j}(\hat{\theta}) + \Delta\alpha_i \mid \hat{\theta} \in \Theta_i], \quad \forall \theta \in \Theta_i,$$

and the participation constraints

$$r_i(\theta) \geqslant 0, \quad \forall \theta \in \Theta_i.$$

The monotonicity of the information rent $r_i(\cdot)$ is not guaranteed in general. I need to identify the pattern of the information rent $r_i(\cdot)$ over the client set Θ_i.

The following lemma shows the equivalence between the incentive constraints in terms of $\langle x_i(\cdot), r_i(\cdot)\rangle$ and the monotonicity of $x_i(\cdot)$ together with the envelope condition of $r_i(\cdot)$.

Lemma 5. *A direct revelation mechanism $\langle x_i(\cdot), r_i(\cdot)\rangle$ is incentive compatible if and only if $x_i(\cdot)$ is non-decreasing and $\dot{r}_i(\theta) = u_\theta(x_i(\theta),\theta) - \dot{\pi}_{t_j}(\theta)$ for every $\theta \in \Theta_i$.*

Proof. Similar to that of Lemma 1. □

Since utility function $u(x,\theta)$ is linear in θ, the slope of the indirect utility function, at which it is differentiable, is given the corresponding quality choice by the envelope theorem. The envelope condition becomes $\dot{r}_i(\theta) = u_\theta(x_i(\theta),\theta) - u_\theta(x_j(\theta),\theta)$ for some $x_j(\theta) \in \text{argmax}[u(x_j,\theta) - t_j(x_j) \mid x_j \in X_j]$. I have deleted the indirect utility function $\pi_{t_i}(\cdot)$ from principal i's problem. In Section 5.3.3, I am going to guess and verity a Nash equilibrium outcome. I provide a pair $\langle x_i(\cdot), x_j(\cdot)\rangle$ of decision rules including off-equilibrium components explicitly. Using the sufficient conditions for an optimal control with pure state constraints, I verify that the decision rule $x_i(\cdot)$ is a solution to the problem of principal i against $x_j(\cdot)$. I identify the pattern of the information rents $\langle r_1(\cdot), r_2(\cdot)\rangle$ describing $\langle\Theta_1,\Theta_2\rangle$ as well. Once a Nash equilibrium outcome $\langle\Theta_i, x_i(\cdot)\rangle$ is obtained, my concern is whether I can associate nonlinear price schedules that achieves the same outcome in the following definition. Not only the allocations but also the market segmentation are preserved under a profile of nonlinear pricing schemes.

Definition 6. A profile $\langle\Theta_i, x_i(\cdot)\rangle$ of client sets and decision rules is *competitively implementable* if there exists a profile $\langle X_i, t_i(\cdot)\rangle$ of price schedules such that for each principal i,
(1) $x_i(\theta) \in \text{argmax}[u(x_i,\theta) - t_i(x_i) \mid x_i \in X_i]$ for every $\theta \in \Theta_i$,
(2) $\pi_{t_i}(\theta) + \alpha_i \geqslant \pi_{t_j}(\theta) + \alpha_j$ for every $\theta \in \Theta_i$.

The competitive implementability tells that given any profile of price schedules, there exists a profile of price schedules which generates the same decision rules and the same market segmentation. There is no need to exclude bunching in decision rules.

Theorem 7. *Let $\langle\Theta_i, x_i(\cdot)\rangle$ be any Nash equilibrium outcome. If $x_i(\cdot)$ is non-decreasing, then the outcome is competitively implementable by some profile $\langle X_i, t_i(\cdot)\rangle$, where the price schedule $t_i(\cdot)$ is of the form $t_i(x_i) = a_i(x_i) + b_i(x_i)x_i$ for each $x_i \in X_i$.*

Proof. Since $\langle\Theta_i, x_i(\cdot)\rangle$ is a Nash equilibrium outcome, it is voluntarily implemented by a profile $\langle X_i, \tilde{t}_i(\cdot)\rangle$ of price schedules. For each $x_i \in X_i$, define

$$t_i(x_i) = \max\left[-(r_i(\hat{\theta}) + \pi_{\tilde{t}_j}(\hat{\theta}) + \Delta\alpha_i) + u(x_i,\hat{\theta}) \mid \hat{\theta} \in \Theta\right].$$

For each $\hat{\theta} \in \Theta$, denote $U_i(\hat{\theta}) = r_i(\hat{\theta}) + \pi_{\tilde{t}_j}(\hat{\theta}) + \Delta\alpha_i$. The convexity follows the fact that $t_i(\cdot)$ is the maximum of a collection of affine functions because of the linearity of $u(x_i,\theta)$ in θ. Furthermore, $t_i(x_i) = -(r_i(\psi_i(x_i)) + \pi_{\tilde{t}_j}(\psi_i(x_i)) + \Delta\alpha_i) + u(x_i,\psi_i(x_i)) = a_i(x_i) + b_i(x_i)x_i$, where $\psi_i(x_i) \in \text{argmax}[-(r_i(\hat{\theta}) + \pi_{\tilde{t}_j}(\hat{\theta}) + \Delta\alpha_i) + u(x_i,\hat{\theta}) \mid \hat{\theta} \in \Theta]$. The fact that the price schedule $t_i(\cdot)$ implements $x_i(\cdot)$ is immediate from the proof of Step 1 in Theorem 2. It remains to show that the market segmentation is actually preserved under $\langle t_1(\cdot), t_2(\cdot)\rangle$. It suffices to show that $\Theta_i = \Theta_i^*$, where $\Theta_i^* = \{\theta \in \Theta \mid \pi_{\tilde{t}_i}(\theta) + \alpha_i \geqslant \pi_{\tilde{t}_j}(\theta) + \alpha_j\}$. Note that for every $\theta \in \Theta_i$, $\pi_{t_i}(\theta) = U_i(\theta) = r_i(\theta) + \pi_{\tilde{t}_j}(\theta) + \Delta\alpha_i = \pi_{\tilde{t}_i}(\theta)$. This implies that for every $\theta \in \Theta_i$, $\pi_{\tilde{t}_i}(\theta) + \alpha_i = \pi_{t_i}(\theta) + \alpha_i \geqslant \pi_{t_j}(\theta) + \alpha_j = \pi_{\tilde{t}_j}(\theta) + \alpha_j$, which implies that $\theta \in \Theta_i^*$. Therefore, $\Theta_i \subseteq \Theta_i^*$. The proof of the converse inclusion $\Theta_i^* \subseteq \Theta_i$ is similar. This establishes the theorem. □

5.3.2 Endogenous Market Segmentation

In this subsection, I argue how the participation constraints are determined endogenously. In order to examine the participation constraints, I summarize properties of the indirect utility function in this setting.

Lemma 6. *The indirect utility function* $\pi_{t_i}(\theta) = \max[u(x_i, \theta) - t_i(x_i) \mid x_i \in X_i]$ *is convex, non-decreasing and continuous.*

Proof. See Appendix 1. □

If the difference $\pi_{t_i}(\cdot)+\alpha_i-(\pi_{t_j}(\cdot)+\alpha_j)$ is strictly increasing, there exists at most one $\hat{\theta} \in \Theta$ such that the set of consumers splits in two closed and convex intervals: $\pi_{t_j}(\theta)+\alpha_j \geqslant \pi_{t_i}(\theta)+\alpha_i$ for $\theta \in [\underline{\theta}, \hat{\theta}]$ and $\pi_{t_i}(\theta) + \alpha_i \geqslant \pi_{t_j}(\theta) + \alpha_j$ for $\theta \in [\hat{\theta}, \overline{\theta}]$, where equalities hold only at taste $\hat{\theta}$ of the unique indifferent consumer. That is, $\Theta_i \cap \Theta_j = \{\hat{\theta}\}$. However, there is no reason for neither that the client sets are closed and convex intervals nor that the competitive market in which agents are indifferent between the two principals is degenerate. It is not appropriate to assume a particular market segmentation for the analysis because the participation constraints are determined *endogenously* as a Nash equilibrium outcome. More precisely, the slopes of $\pi_{t_i}(\cdot)$ and $\pi_{t_j}(\cdot)$ are given by $x_i(\cdot) \in X_{t_i}(\cdot)$ and $x_j(\cdot) \in X_{t_j}(\cdot)$, respectively. The relative steepness between indirect utility functions are determined by the profile $\langle t_1(\cdot), t_2(\cdot)\rangle$ of price functions. The following figure depicts the situation in which there is a re-switching of domination.

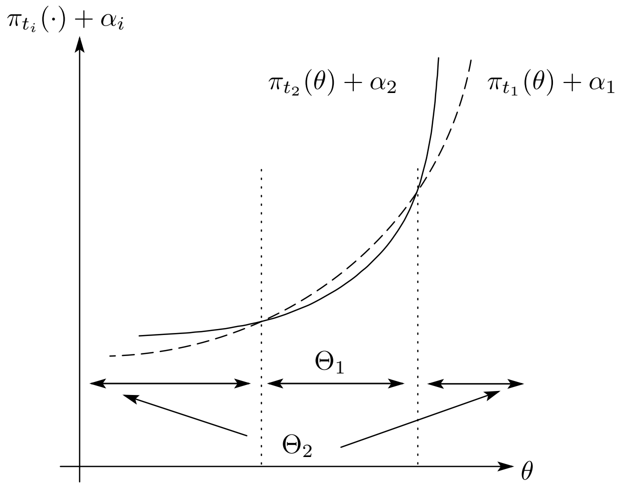


Figure 15: Re-switching of domination

It is easy to guess that if there is no asymmetry between the two principals, then both client sets $\langle \Theta_1, \Theta_2 \rangle$ are equal to the entire type space Θ. The following lemma shows that the market is fully covered under Assumption 3 in any Nash equilibrium.

Lemma 7. *There is no exclusion in equilibrium:* $\Theta_1 \cup \Theta_2 = \Theta$.

Proof. See Appendix 2. □

The following lemma claims that a non-degenerate competitive market is not excluded. Because of the presence of asymmetric cost structures, principal 1 gets high types, whereas principal 2 gets low types.

Lemma 8. $\Theta_1 = [\,\underline{\theta}_1, \overline{\theta}\,]$ and $\Theta_2 = [\,\underline{\theta}, \overline{\theta}_2\,]$ with $\underline{\theta}_1 \leqslant \overline{\theta}_2$.

Proof. See Appendix 3. □

Lemma 8 in Section 5.3.2 states that the market segmentation $\langle \Theta_1, \Theta_2 \rangle$ is parameterized by two parameters, $\underline{\theta}_1$ and $\overline{\theta}_2$. Lemma 10 below states that how both boundaries $\underline{\theta}_1$ and $\overline{\theta}_2$ are determined in a particular Nash equilibrium outcome.

5.3.3 Oligopolistic Competition

In this subsection, I shall construct a particular Nash equilibrium outcome $\langle \Theta_i, x_i(\cdot) \rangle$. For each $x_i \in \mathbb{R}$ and $\lambda_i \in \mathbb{R}$, define

$$\sigma_i(x_i, \lambda_i, \theta) = v_i(x_i, \theta) + \frac{\lambda_i}{f(\theta)} u_\theta(x_i, \theta).$$

Denote by $x_1^\sigma(\theta)$ is the solution to

$$0 = D_x v_1(x_1, \theta) - \frac{1 - F(\theta)}{f(\theta)} u_{x\theta}(x_1, \theta).$$

Similarly, $x_2^\sigma(\theta)$ is the solution to

$$0 = D_x v_2(x_2, \theta) + \frac{F(\theta)}{f(\theta)} u_{x\theta}(x_2, \theta).$$

Note that both $x_1^\sigma(\cdot)$ and $x_2(\cdot)$ are determined uniquely by the strict concavity of $v_i(x_i, \theta)$ in x_i.

I need to specify a rule for determining which principal wins the tie.[11]

Assumption 4. Tie-breaking is determined by an efficiency rule: the principal that makes higher profit wins.

[11] Monteiro and Page (2008) argue that a choice of tie-breaking rules is crucial to guarantee the existence of a Nash equilibrium in mixed strategies over *catalogs* which are closely related nonlinear prices.

The following lemma characterizes the bunching allocation in terms of the comparative advantage $\Delta\alpha_1$ of principal 2.

Lemma 9. *If the competitive market,* $\Theta_1 \cap \Theta_2 = [\underline{\theta}_1, \overline{\theta}_2]$*, is a non-degenerate interval, then it must be the case that the product lines are overlapping and there is pooling:* $x_1(\theta) = z(\Delta\alpha_1) = x_2(\theta)$ *for every* $\theta \in \Theta_1 \cap \Theta_2$*, where* $C_1(z(\Delta\alpha_1)) + \Delta\alpha_1 = C_2(z(\Delta\alpha_1))$ *and* $z(\alpha_1) = \frac{\Delta\alpha_1}{a_2 - a_1}$*. Moreover,* $\dot{z}(\Delta\alpha_1) > 0$*.*

Proof. See Appendix 4. □

The parameter $\Delta\alpha_1 = \alpha_2 - \alpha_1$ measures the comparative advantage of principal 2 from view point of the agent. I can incorporate the comparative advantage in the cost structure as well. Decompose the cost function of principal i as $C_i(x_i) = \gamma_i + C_i^v(x_i)$, where $C_i^v(x_i)$ is the variable cost function and γ_i is the fixed cost. Denote $\Delta\gamma_1 = \gamma_2 - \gamma_1$. The quality allocation in the pooling region is implicitly determined by $C_1^v(z(\Delta\alpha_1 - \Delta\gamma_1)) + (\Delta\alpha_1 - \Delta\gamma_1) = C_2^v(z(\Delta\alpha_1 - \Delta\gamma_1))$. Not ambiguously, the higher $\Delta\alpha_1 - \Delta\gamma_1$, principal 2 has the comparative advantage. I can replace $\Delta\alpha_1$ by $\Delta\alpha_1 - \Delta\gamma_1$. In particular, if cost functions are of the form $C_i(x_i) = \gamma_i + a_i x_i + \frac{b_i}{2} x_i^2$ for each principal i, then it must be the case that $\Delta\alpha_1 - \Delta\gamma_1 > 0$.[12] For simplicity, I say that $\Delta\alpha_1 > 0$ summarizes all of components for the comparative advantage of principal 2.

In order to identify a Nash equilibrium outcome, I make use of sufficient conditions in Seierstad and Sydsæter (1987, Chapter 5, Theorem 1) for an optimal control problem with pure state constraints. Principal 1 is weakly dominant over his client set $\Theta_1 = [\underline{\theta}_1, \overline{\theta}]$. It was shown that, given the client set $\Theta_1 = [\underline{\theta}_1, \overline{\theta}]$, principal 1's problem can be written as

$$\max_{\langle x_1(\cdot), r_1(\cdot)\rangle} \int_{\underline{\theta}_1}^{\overline{\theta}} [v_1(x_1(\theta), \theta) - r_1(\theta) - \pi_{t_2}(\theta) - \Delta\alpha_1] f(\theta) d\theta$$

subject to the incentive constraints $\dot{x}_1(\theta) \geqslant 0$ and $\dot{r}_1(\theta) = u_\theta(x_1(\theta), \theta) - \dot{\pi}_{t_2}(\theta)$ for every $\theta \in \Theta_i$, and the participation constraints $r_1(\theta) \geqslant 0$ for every $\theta \in \Theta_i$. Define the Hamiltonian and the Lagrangian by, respectively:

$$\mathcal{H}_1(r_1, x_1, \lambda_1, \theta) = [v_1(x_1, \theta) - r_1 - \pi_{t_2}(\theta) - \Delta\alpha_1] f(\theta) + \lambda_1 [u_\theta(x_1, \theta) - \dot{\pi}_{t_2}(\theta)]$$

and

$$\mathcal{L}_1(r_1, x_1, \lambda_1, \tau_1, \theta) = \mathcal{H}_1(r_1, x_1, \lambda_1, \theta) + \tau_1 r_1.$$

I have one state variable $r_1(\cdot)$, one control variable $x_1(\cdot)$ and one pure state constraint $r_1(\cdot) \geqslant 0$. The multipliers associated with the envelope condition and the participation constraint are $\lambda_1(\cdot)$ and $\tau_1(\cdot)$, respectively. I shall show that

[12] By Lemma 9, this is immediate from the fact that $\Delta\alpha_1 - \Delta\gamma_1 = (a_2 - a_1) z(\Delta\alpha_1 - \Delta\gamma_1) > 0$ because $a_2 > a_1 > 0$ and $b_1 = b_2$.

(a) there exists a function $\lambda_1(\cdot)$, which is piecewise continuous and piecewise continuously differentiable with jump discontinuities at $\theta_1, \cdots, \theta_m$ with $\underline{\theta}_1 < \theta_1 < \cdots < \theta_m \leqslant \overline{\theta}$, and
(b) there exist a piecewise continuous function $\tau_1(\cdot)$ and non-negative numbers $\beta_1, \cdots, \beta_m$ such that the following conditions are satisfied:
(1) $x_1(\theta) \in$ argmax$[\mathcal{H}_1(r_1(\theta), x_1, \lambda_1(\theta), \theta) \mid x_1 \in X_1]$,
(2) $\tau_1(\theta) \geqslant 0$ and $\tau_1(\theta) r_1(\theta) = 0$,
(3) $\dot{\lambda}_1(\theta) = -\frac{\partial \mathcal{L}_1}{\partial r_1}(r_1(\theta), x_1(\theta), \lambda_1(\theta), \tau_1(\theta), \theta)$, (the costate equation)
(4) $\lambda_1(\theta_k^-) - \lambda_1(\theta_k^+) = \beta_k$ (Put $\lambda(\theta_k^+) = \lambda(\overline{\theta})$ if $\theta_m = \overline{\theta}$),
(5) $\beta_k = 0$ if either $r_1(\theta_k) > 0$ or [$\theta_k \in$ int Θ_1, $r_1(\theta_k) = 0$ and $x_1(\theta)$ is discontinuous at θ_k],
(6) $\lambda_1(\underline{\theta}_1) r_1(\underline{\theta}_1) = 0$ and $\lambda_1(\underline{\theta}_1) \leqslant 0$, (the initial transversality condition)
(7) $\lambda_1(\overline{\theta}) r_1(\overline{\theta}) = 0$ and $\lambda_1(\overline{\theta}) \geqslant 0$, (the terminal transversality condition)
(8) $\hat{\mathcal{H}}_1(r_1, \lambda_1(\theta), \theta) = \max[\mathcal{H}_1(r_1, x_1, \lambda_1(\theta), \theta) \mid x_1 \in X_1]$ is concave in r_1,
(9) the pure state constraint $r_1 \geqslant 0$ is quasi-concave in r_1.

I shall construct a profile $\langle x_1(\cdot), x_2(\cdot) \rangle$ of decision rules including *off-equilibrium* components, not only realized consumptions $x_i(\cdot)$ over Θ_i. This is because the slope of $r_i(\cdot)$ is given by $\dot{r}_i(\theta) = \frac{d}{d\theta}\{\pi_{t_i}(\theta) - \pi_{t_j}(\theta) - \Delta\alpha_i\} = u_\theta(x_i(\theta), \theta) - u_\theta(x_j(\theta), \theta)$ for some $x_k(\theta) \in$ argmax$[u(x_k, \theta) - t_k(x_k) \mid x_k \in X_k]$, $k = i, j$. For each type $\theta \in \Theta_i \setminus \Theta_j$ in the captive market for principal i, $x_j(\theta)$ was *not realized.* My conjecture is as follows. For each decision rule $x_2(\cdot)$, define

$$x_1(\theta) = \begin{cases} x_1^*(\theta) \text{ for } \theta < \underline{\theta}_1 \\ x_2(\theta) \text{ for } \theta \in [\underline{\theta}_1, \hat{\theta}_1) \\ x_1^\sigma(\theta) \text{ for } \theta \geqslant \hat{\theta}_1 \end{cases}$$

and

$$\lambda_1(\theta) = \begin{cases} \{\lambda_1 \in \mathbb{R} \mid D_x\sigma_1(x_2(\theta), \lambda_1, \theta) = 0\} & \text{for } \theta \in [\underline{\theta}_1, \hat{\theta}_1) \\ -(1 - F(\theta)) & \text{for } \theta \geqslant \hat{\theta}_1 \end{cases}$$

where $\hat{\theta}_1 = \max\{\underline{\theta}_1, \{\theta \in \Theta \mid -D_x v_1(x_2(\theta), \theta) f(\theta) = -(1 - F(\theta))\}\}$. I see that for every $\theta \in [\underline{\theta}_1, \hat{\theta}_1)$, $\lambda_1(\theta) = -D_x v_1(x_2(\theta), \theta) f(\theta)$, which is uniquely determined. Both $x_1(\cdot)$ and $\lambda_1(\cdot)$ are single-valued over Θ_1. Notice that $0 = D_x\sigma_1(x_1(\theta), \lambda_1(\theta), \theta)$ for every $\theta \in \Theta_1$.

The costate equation becomes $\dot{\lambda}_1(\theta) = f(\theta) - \tau_1(\theta)$. I need to define $\lambda_1(\cdot)$ and $\tau_1(\cdot)$ jointly. The crucial step is to construct the right functions to $\lambda_1(\cdot)$ and $\tau_1(\cdot)$. Define

$$\tau_1(\theta) = \begin{cases} 2f(\theta) & \text{for } \theta \in [\underline{\theta}_1, \overline{\theta}_2) \\ f(\theta) & \text{for } \theta \in [\overline{\theta}_2, \hat{\theta}_1) \\ 0 & \text{for } \theta \in [\hat{\theta}_1, \overline{\theta}_1]. \end{cases}$$

As mentioned above, the envelope condition $\dot{r}_1(\theta) = u_\theta(x_1(\theta), \theta) - \dot{\pi}_{t_2}(\theta)$ contains $\dot{\pi}_{t_2}(\theta) = u_\theta(x_2(\theta), \theta)$ for some $x_2(\theta) \in$ argmax$[u(x_2, \theta) - t_2(x_2) \mid x_2 \in X_2]$. I provide my conjecture on $x_2(\cdot)$. It will be shown that both $x_1(\cdot)$ and $x_2(\cdot)$ are non-decreasing, continuous and differentiable almost everywhere. Define

$$x_2(\theta) = \begin{cases} x_2^\sigma(\theta) & \text{for } \theta < \hat{\theta}_2 \\ x_1^*(\theta) & \text{for } \theta \in [\hat{\theta}_2, \underline{\theta}_1) \\ z(\Delta\alpha_1) & \text{for } \theta \in [\underline{\theta}_1, \overline{\theta}_2) \\ x_2^*(\theta) & \text{for } \theta \geqslant \overline{\theta}_2 \end{cases}$$

where $\hat{\theta}_2 = \min\{\overline{\theta}_2, \{\theta \in \Theta \mid -D_x v_2(x_1(\theta), \theta) f(\theta) = F(\theta)\}\}$. Then, $x_2(\cdot)$ is single-valued. The decision rule $x_1(\cdot)$ is continuous at $\hat{\theta}_1$ because $x_1^\sigma(\hat{\theta}_1) = x_2(\hat{\theta}_1) = x_2^*(\hat{\theta}_1)$. Similarly, $x_2(\cdot)$ is continuous at $\hat{\theta}_2$. On the other hand, the continuity of $x_1(\cdot)$ at $\underline{\theta}_1$ and $\overline{\theta}_2$ is not straightforward. It suffices to show that $x_1^*(\underline{\theta}_1) = z(\Delta\alpha_1) = x_2^*(\overline{\theta}_2)$.

Lemma 10. *If the competitive market, $\Theta_1 \cap \Theta_2 = [\underline{\theta}_1, \overline{\theta}_2]$, is a non-degenerate interval, then it must be the case that $x_1^*(\underline{\theta}_1) = z(\Delta\alpha_1) = x_2^*(\overline{\theta}_2)$. Both $\underline{\theta}_1$ and $\overline{\theta}_2$ are uniquely determined with respect to $\Delta\alpha_1$:*

$$\underline{\theta}_1 = \frac{b(a_2 - a_1)}{\Delta\alpha_1} + a_1 \quad \text{and} \quad \overline{\theta}_2 = \frac{b(a_2 - a_1)}{\Delta\alpha_1} + a_2.$$

Proof. See Appendix 5. □

Lemma 11. *If the competitive market, $\Theta_1 \cap \Theta_2 = [\underline{\theta}_1, \overline{\theta}_2]$, is a non-degenerate interval, then it must be $\overline{\theta}_2 \leqslant \hat{\theta}_1$ and $\hat{\theta}_2 \leqslant \underline{\theta}_1$ under the outcome $\langle x_1(\cdot), x_2(\cdot)\rangle$, where*

$$\hat{\theta}_1 = \overline{\theta} - (a_2 - a_1) \quad \text{and} \quad \hat{\theta}_2 = \underline{\theta} + (a_2 - a_1).$$

Proof. See Appendix 6. □

Proposition 6. *The sufficient conditions for optimality in the optimal control problem of principal 1 are satisfied. The decision rule $x_1(\cdot)$ exhibits pooling in the competitive market $[\underline{\theta}_1, \overline{\theta}_2]$ and downward distortions except at the top and bottom over his client set (see Figure 16):*

$$x_1(\theta) = \begin{cases} \frac{1}{b}(\theta - a_1) & \text{for } \theta < \underline{\theta}_1 \\ \frac{\Delta\alpha_1}{a_2 - a_1} & \text{for } \theta \in [\underline{\theta}_1, \overline{\theta}_2) \\ \frac{1}{b}(\theta - a_2) & \text{for } \theta \in [\overline{\theta}_2, \hat{\theta}_1) \\ \frac{1}{b}(2\theta - \overline{\theta} - a_1) & \text{for } \theta > \hat{\theta}_1. \end{cases}$$

Proof. Immediate from Lemmas 12 to 20 in Appendix 7. □

The best response of principal 2 can be obtained in a similar way.

Proposition 7. *The sufficient conditions for optimality in the optimal control problem of principal 2 are satisfied. The decision rule $x_2(\cdot)$ exhibits pooling in the competitive market $[\underline{\theta}_1, \overline{\theta}_2]$ and upward distortions except at the top and bottom over his client set (see Figure 17):*

$$x_2(\theta) = \begin{cases} \frac{1}{b}(2\theta - \underline{\theta} - a_2) & \text{for } \theta < \hat{\theta}_2 \\ \frac{1}{b}(\theta - a_1) & \text{for } \theta \in [\hat{\theta}_2, \underline{\theta}_1) \\ \frac{\Delta\alpha_1}{a_2 - a_1} & \text{for } \theta \in [\overline{\theta}_2, \hat{\theta}_1) \\ \frac{1}{b}(\theta - a_2) & \text{for } \theta > \hat{\theta}_1. \end{cases}$$

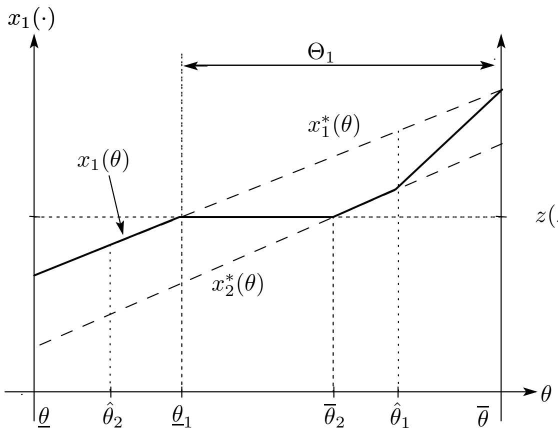


Figure 16: No distortion at $\underline{\theta}_1$ and $\overline{\theta}$

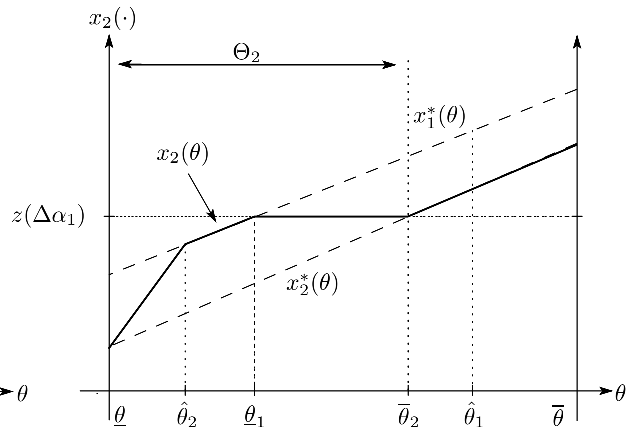


Figure 17: No distortion at $\underline{\theta}$ and $\overline{\theta}_2$

A Nash equilibrium outcome with the competitive market of positive measure is summarized below.

Theorem 8. *There exists a Nash equilibrium outcome in which the competitive market is non-degenerate. The corresponding quality allocation of each principal is distorted everywhere, except at the top and the bottom of his client set. There is pooling in the competitive market. Moreover, such equilibrium outcome can be competitively implementable by convex price schedules.*

To end this subsection, I want to summarize the resulting information rents in the Nash equilibrium in Theorem 8 (Figures 18 and 19). Integrating the envelope condition in Lemma 1, the information rents are given as

$$r_1(\theta) = \int_{\underline{\theta}_1}^{\theta} [u_\theta(x_1(s), s) - u_\theta(x_2(s), s)]ds = \begin{cases} 0 & \text{for } \theta < \hat{\theta}_1 \\ \frac{1}{2b}(\overline{\theta} - \theta - (a_2 - a_1))^2 & \text{for } \theta \geqslant \hat{\theta}_1, \end{cases}$$

$$r_2(\theta) = \int_{\theta}^{\overline{\theta}_2} [u_\theta(x_2(s), s) - u_\theta(x_1(s), s)]ds = \begin{cases} \frac{1}{2b}(\theta - \underline{\theta} - (a_2 - a_1))^2 & \text{for } \theta \leqslant \hat{\theta}_2 \\ 0 & \text{for } \theta > \hat{\theta}_2. \end{cases}$$

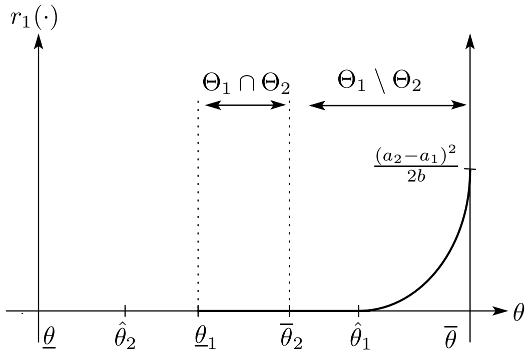


Figure 18: Information rent $r_1(\cdot)$

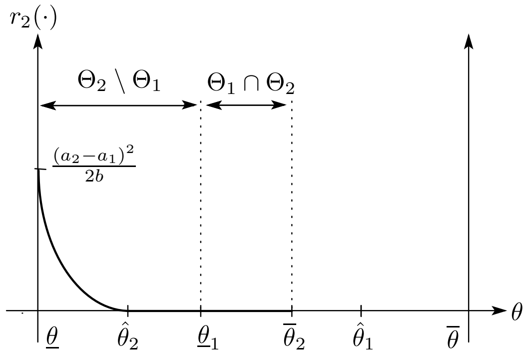


Figure 19: Information rent $r_2(\cdot)$

6 Discussion: Beyond Quasi-Linear Context

This section deals with the derivation of the price schedule as an indirect mechanism in a context with income effects. Roberts (1979) considers the optimal pricing schedule of a public utility subject to a profit constraint. In the economy, there are two commodities; a good x subject to nonlinear pricing $t(\cdot)$, and the composite good y. The agent with income θ maximizes $u(x, y)$ with $\nabla u(x, y) \gg 0$ subject to the budget constraint $x + t(x) \leqslant y$.

The method developed in the present paper can be applied to obtain the optimal price schedule for the government. Let $\langle x(\cdot), p(\cdot)\rangle$ be incentive compatible direct revelation mechanism consisting of a decision rule and a transfer function, and $U(\theta) = u(x(\theta), \theta - p(\theta))$. By the strict monotonicity of $u(x, y)$ with respect to y, there is a unique $\phi(x, \hat{\theta})$ such that $u(x, \theta - \phi(x, \hat{\theta})) = U(\hat{\theta})$ for each $x \in X$ and $\hat{\theta} \in \Theta$. Define $t(x) = \max[\phi(x, \hat{\theta}) \mid \hat{\theta} \in \Theta]$ for each $x \in X$. A similar procedure in the proof of Theorem 2 will show that $x(\theta) \in \text{argmax}[u(x, \theta - t(x)) \mid x \in X]$ for every $\theta \in \Theta$ indeed.

7 Appendices

Appendix 1: Proof of Lemma 6

Proof. The convexity follows the fact that $\pi_{t_i}(\theta)$ is the maximum of a collection of affine functions. Consider any pair $(\theta, \tilde{\theta}) \in \Theta \times \Theta$ with $\theta < \tilde{\theta}$. For every selection $x_i(\theta) \in X_{t_i}(\theta)$, I see that $\pi_{t_i}(\tilde{\theta}) - \pi_{t_i}(\theta) \geqslant u(x_i(\theta), \tilde{\theta}) - t_i(x_i(\theta)) - \pi_{t_i}(\theta) = u(x_i(\theta), \tilde{\theta}) - u(x_i(\theta), \theta) \geqslant 0$ since $u_\theta(x, \theta) \geqslant 0$. Therefore, $\pi_{t_i}(\theta)$ is non-decreasing.[13] Regarding the continuity, suppose, by way of contradiction, that $\pi_{t_i}(\theta)$ is not continuous at θ. There exists some $\varepsilon_1 > 0$ such that no matter what $\delta_1 > 0$, there is some $\tilde{\theta} \in \Theta$ with $\mid \theta - \tilde{\theta} \mid < \delta_1$ such that

[13] Notice that $\pi_{t_i}(\cdot)$ is strictly increasing if $x_i(\cdot)$ is positive.

$\mid \pi_{t_i}(\theta) - \pi_{t_i}(\tilde{\theta}) \mid \geqslant \varepsilon_1$. Since $u(x,\theta)$ is continuous in θ, it follows that for every $\varepsilon_2 > 0$, there exists $\delta_2 > 0$ such that $\mid u(x,\theta) - u(x,\hat{\theta}) \mid < \frac{\varepsilon_2}{2}$ whenever $\mid \theta - \hat{\theta} \mid < \delta_2$. There are two possibilities to be considered. Suppose first that $\pi_{t_i}(\theta) - \pi_{t_i}(\tilde{\theta}) \leqslant -\varepsilon_1$. Then, $u(x_i(\tilde{\theta}),\theta) - t_i(x_i(\tilde{\theta})) > u(x_i(\tilde{\theta}),\tilde{\theta}) - t_i(x_i(\tilde{\theta})) - \frac{\varepsilon_2}{2} = \pi_{t_i}(\tilde{\theta}) - \frac{\varepsilon_2}{2}$, where $\mid \theta - \tilde{\theta} \mid < \delta_2$ and $\mid \pi_{t_i}(\theta) - \pi_{t_i}(\tilde{\theta}) \mid \geqslant \varepsilon_1$. Since ε_2 was arbitrary, letting $\varepsilon_2 < 2\varepsilon_1$ to obtain $u(x_i(\tilde{\theta}),\theta) - t_i(x_i(\tilde{\theta})) > \pi_{t_i}(\theta) + \varepsilon_1 - \frac{\varepsilon_2}{2} > \pi_{t_i}(\theta)$, a contradiction. Suppose next that $\pi_{t_i}(\theta) - \pi_{t_i}(\tilde{\theta}) \geqslant \varepsilon_1$. Then, $u(x_i(\theta),\tilde{\theta}) - t_i(x_i(\theta)) > u(x_i(\theta),\theta) - t_i(x_i(\theta)) - \frac{\varepsilon_2}{2} = \pi_{t_i}(\theta) - \frac{\varepsilon_2}{2} \geqslant \pi_{t_i}(\tilde{\theta}) + \varepsilon_1 - \frac{\varepsilon_2}{2} > \pi_{t_i}(\tilde{\theta})$, a contradiction. Thus, $\pi_{t_i}(\theta)$ is continuous. This establishes the lemma. □

Appendix 2: Proof of Lemma 7

Proof. Suppose, by way of contradiction, that there exists $\theta \in \Theta \backslash (\Theta_1 \cup \Theta_2)$. Then, $\max\{\pi_{t_1}(\theta) + \alpha_1, \pi_{t_2}(\theta) + \alpha_2\} < 0$. Each principal i makes zero profit, so it is weakly optimal to offer $x_i(\theta)$ such that $t_i(x_i(\theta),\theta) - C_i(x_i(\theta)) = 0$. The agent of type θ gets $u(x_i(\theta),\theta) + \alpha_i - t_i(x_i(\theta)) = v_i(x_i(\theta),\theta) + \alpha_i$, which is maximized at the full-information scheme $x_i^*(\theta)$. Without loss of generality, I may assume that both principals could not attract type θ even they offer the full-information surplus $v_i(x_i^*(\theta),\theta) + \alpha_i$. It follows that $\pi_{t_i}(\theta) + \alpha_i = v_i(x_i^*(\theta),\theta) + \alpha_i < 0$, which contradicts that $0 \leqslant v_i(x_i^*(\underline{\theta}),\underline{\theta}) + \alpha_i \leqslant v_i(x_i^*(\theta),\theta) + \alpha_i$ because the full-information surplus is non-decreasing: $\frac{d}{d\theta} v_i(x_i^*(\theta),\theta) = u_\theta(x_i^*(\theta),\theta) \geqslant 0$ by the envelope theorem. This establishes the lemma. □

Appendix 3: Proof of Lemma 8

Proof. Firstly, I shall show that both client sets are convex: $\Theta_1 = [(\underline{\theta}_1, \overline{\theta}\,]$ and $\Theta_2 = [\,\underline{\theta}, \overline{\theta}_2)]$ with $\underline{\theta}_1 \leqslant \overline{\theta}_2$, where the boundaries $\underline{\theta}_1$ and $\overline{\theta}_2$ may or may not belong to the intervals.

Recall that, by Lemma 7, the market is fully covered in the sense that $\Theta_1 \cup \Theta_2 = \Theta$. I first consider the case that the client sets of the two principals are convex subsets of Θ, and then I verify this conjecture later.[14] Let $\Theta_1 = [(\underline{\theta}_1, \overline{\theta}_1)]$ and $\Theta_2 = [(\underline{\theta}_2, \overline{\theta}_2)]$, where the notation $[(a,b)]$ indicates that the end points may or may not belong to the interval. By the convexity of the client sets, there is no re-switching of domination in the sense that there is no triplet $\{\theta_1, \theta_2, \theta_3\} \subseteq \Theta$ with $\theta_1 < \theta_2 < \theta_3$ such that one principal weakly dominates at θ_1 and θ_3 and the other principal strictly dominates at θ_2. Denote by $x_i(\theta^+)$ and $x_i(\theta^-)$ the right-hand limit of $x_i(\cdot)$ at θ and the left-hand limit of $x_i(\cdot)$ at θ, respectively. Let $\{x_i(\theta), \pi_{t_i}(\theta) + \alpha_i\}_{\theta \in \Theta}$ be an equilibrium outcome, including off-equilibrium components.

[14] The following proof is similar to the proofs of Lemmas 4 and 5 in Biglaiser and Mezzetti (1993). They analyze competition in a labor market model. As far as I know, Biglaiser and Mezzetti (1993) is the only work for a delegated common agency with non-identical principals. They analyze competition between two non-identical firms for a worker in a labor market model. Contracts are compensation schemes associated with output targets.

Step 1. $\Theta_1 = [(\underline{\theta}_1, \overline{\theta}]$ and $\Theta_2 = [\underline{\theta}, \overline{\theta}_2)]$. In other words, principal 1 gets high types, whereas principal 2 gets low types.

Proof of Step 1. I first show that $\underline{\theta}_2 = \underline{\theta}$. Suppose, by way of contradiction, that $\underline{\theta} < \underline{\theta}_2$. Since there is no re-switching of domination, it must be the case that principal 1 captures low types and principal 1 is strictly dominant up to $\underline{\theta}_2$ (See Figure 20). Then, $x_1(\underline{\theta}_2^-) \leqslant x_2(\underline{\theta}_2^+)$ (see Figure 21). Recall that $\pi_{t_i}(\theta)$ is continuous, convex and non-decreasing in θ. Since $\dot{\pi}_{t_i}(\theta) = x_i(\theta)$ for some $x_i(\theta) \in \operatorname{argmax}[u(x_i, \theta) - t_i(x_i) \mid x_i \in X_i]$ at which $\pi_{t_i}(\theta)$ is differentiable, it follows that $x_i(\theta^-) \leqslant x_i(\theta) \leqslant x_i(\theta^+)$.

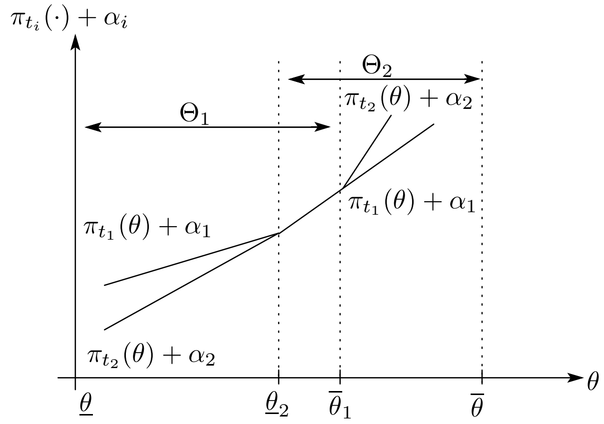


Figure 20: Market segmentation

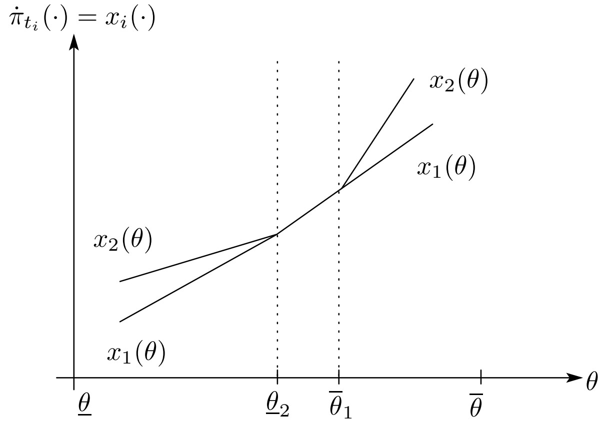


Figure 21: $\dot{\pi}_{t_i}(\theta)$ around $[\underline{\theta}_2, \overline{\theta}_1]$

Consider the following menu of principal 2:

$$(\hat{x}_2(\theta), \hat{\pi}_{t_2}(\theta) + \alpha_2) = \begin{cases} (x_1(\theta), \pi_{t_1}(\theta) + \alpha_1) \text{ for } \theta < \underline{\theta}_2 \\ (x_2(\theta), \pi_{t_2}(\theta) + \alpha_2) \text{ for } \theta \geqslant \underline{\theta}_2. \end{cases}$$

Obviously, the decision rule $\hat{x}_2(\theta)$ is non-decreasing. Let $\hat{r}_2(\theta) = \hat{\pi}_{t_2}(\theta) + \alpha_2 - (\pi_{t_1}(\theta) + \alpha_1) = \hat{\pi}_{t_2}(\theta) - \pi_{t_1}(\theta) - \Delta\alpha_2$ be the corresponding information rent. By construction, $\hat{r}_2(\theta) \geqslant 0$ for every $\theta \in \Theta$, which yields that the corresponding client set becomes $\hat{\Theta}_2 = \Theta$. Now, any type $\theta < \underline{\theta}_2$ is indifferent between the two principals. Then, $\hat{r}_2(\theta)$ satisfies the envelope condition: For $\theta < \underline{\theta}_2$, $u_\theta(\hat{x}_2(\theta), \theta) - \dot{\pi}_{t_1}(\theta) - \frac{d}{d\theta}\hat{r}_2(\theta) = u_\theta(x_1(\theta), \theta) - \dot{\pi}_{t_1}(\theta) - 0 = x_1(\theta) - x_1(\theta) = 0$. For $\theta \geqslant \underline{\theta}_2$, $u_\theta(\hat{x}_2(\theta), \theta) - \dot{\pi}_{t_1}(\theta) - \frac{d}{d\theta}\hat{r}_2(\theta) = u_\theta(x_2(\theta), \theta) - \dot{\pi}_{t_1}(\theta) - \dot{r}_2(\theta) = 0$. Therefore, $\frac{d}{d\theta}\hat{r}_2(\theta) = u_\theta(\hat{x}_2(\theta), \theta) - \dot{\pi}_{t_1}(\theta)$ for every $\theta \in \hat{\Theta}_2$. By Lemma 5, the pair $\langle \hat{x}_2(\cdot), \hat{\pi}_{t_2}(\cdot) + \alpha_2 \rangle$ is incentive compatible. Therefore, principal 2 can imitate principal 1 over $\Theta_1 \setminus \Theta_2$ by offering the menu $\langle \hat{x}_2(\cdot), \hat{\pi}_{t_2}(\cdot) + \alpha_2 \rangle$, without violating the incentive constraints. In equilibrium, this feasible deviation cannot be profitable for principal 2. Under any efficient tie-breaking rule, it must be the case that for every $\varepsilon > 0$ sufficiently small, $v_1(x_1(\underline{\theta}_2 - \varepsilon), \underline{\theta}_2 - \varepsilon) - \pi_{t_1}(\underline{\theta}_2 - \varepsilon) > v_2(x_1(\underline{\theta}_2 - \varepsilon), \underline{\theta}_2 - \varepsilon) - (\pi_{t_1}(\underline{\theta}_2 - \varepsilon) - \Delta\alpha_1)$. This implies that $C_2(x_1(\underline{\theta}_2 - \varepsilon)) > C_1(x_1(\underline{\theta}_2 -$

$\varepsilon)) + \Delta\alpha_1$. Similarly, principal 1 can imitate principal 2 over $\Theta_2 \setminus \Theta_1$ by offering the following menu, without violating the incentive constraints:

$$(\hat{x}_1(\theta), \hat{\pi}_{t_1}(\theta) + \alpha_1) = \begin{cases} (x_1(\theta), \pi_{t_1}(\theta) + \alpha_1) \text{ for } \theta < \underline{\theta}_2 \\ (x_2(\theta), \pi_{t_2}(\theta) + \alpha_2) \text{ for } \theta \geqslant \underline{\theta}_2. \end{cases}$$

The corresponding client set becomes $\hat{\Theta}_1 = \Theta$. In equilibrium, it must be the case that for every $\varepsilon > 0$ sufficiently small, $v_2(x_2(\underline{\theta}_2 + \varepsilon), \underline{\theta}_2 + \varepsilon) - \pi_{t_2}(\underline{\theta}_2 + \varepsilon) \geqslant v_1(x_2(\underline{\theta}_2 + \varepsilon), \underline{\theta}_2 + \varepsilon) - (\pi_{t_2}(\underline{\theta}_2 + \varepsilon) + \Delta\alpha_1)$. This implies that $C_1(x_2(\underline{\theta}_2 + \varepsilon)) + \Delta\alpha_1 \geqslant C_2(x_2(\underline{\theta}_2 + \varepsilon))$. Therefore, $C_2(x_1(\underline{\theta}_2 - \varepsilon)) - C_1(x_1(\underline{\theta}_2 - \varepsilon)) > \Delta\alpha_1 \geqslant C_2(x_2(\underline{\theta}_2 + \varepsilon)) - C_1(x_2(\underline{\theta}_2 + \varepsilon))$. I conclude that $C_1(x_2(\underline{\theta}_2 + \varepsilon)) - C_1(x_1(\underline{\theta}_2 - \varepsilon)) > C_2(x_2(\underline{\theta}_2 + \varepsilon)) - C_2(x_1(\underline{\theta}_2 - \varepsilon))$. However, under Assumption 2, it follows from $x_1(\underline{\theta}_2 - \varepsilon) \leqslant x_2(\underline{\theta}_2 + \varepsilon)$ that $C_2(x_2(\underline{\theta}_2 + \varepsilon)) - C_2(x_1(\underline{\theta}_2 - \varepsilon)) = \int_{x_1(\underline{\theta}_2 - \varepsilon)}^{x_2(\underline{\theta}_2 + \varepsilon)} C_2'(y)dy \geqslant \int_{x_1(\underline{\theta}_2 - \varepsilon)}^{x_2(\underline{\theta}_2 + \varepsilon)} C_1'(y)dy = C_1(x_2(\underline{\theta}_2 + \varepsilon)) - C_1(x_1(\underline{\theta}_2 - \varepsilon))$, a contradiction. Hence, it must be $\underline{\theta}_2 = \underline{\theta}$. Similarly, $\overline{\theta}_1 = \overline{\theta}$. This establishes the step.

It remains to verify the hypothesis.

Step 2. $\Theta_1 = [(\underline{\theta}_1, \overline{\theta}_1)]$ and $\Theta_2 = [(\underline{\theta}_2, \overline{\theta}_2)]$. In other words, both client sets are convex.

Proof of Step 2. Suppose, by way of contradiction, that there is a triplet $\{\theta_1, \theta_2, \theta_3\} \subseteq \Theta$ with $\theta_1 < \theta_2 < \theta_3$ such that principal i weakly dominates at θ_1 and θ_3 and principal j strictly dominates at θ_2. Let $A_i = \{\theta \in [\theta_1, \theta_3] \mid \pi_{t_i}(\theta) + \alpha_i \geqslant \pi_{t_j}(\theta) + \alpha_j\}$ and $A_j = \{\theta \in [\theta_1, \theta_3] \mid \pi_{t_j}(\theta) + \alpha_j > \pi_{t_i}(\theta) + \alpha_i\}$. Then, A_i is non-empty because $\{\theta_1, \theta_3\} \subseteq A_i$ and A_j is non-empty because $\theta_2 \in A_j$. If there exists some $\theta \in A_i \cap A_j$, it follows from the fact $\theta \in A_i$ that $\pi_{t_i}(\theta) + \alpha_i \geqslant \pi_{t_j}(\theta) + \alpha_j$, which contradicts the fact that $\theta \in A_j$. Therefore, $A_i \cap A_j = \emptyset$. It is obvious that $A_i \cup A_j = [\theta_1, \theta_3]$. Therefore, $\langle A_i, A_j \rangle$ is a partition of $[\theta_1, \theta_3]$. Let $B_i = A_i \cap [\theta_1, \theta_2]$ and $B_j = A_j \cap [\sup B_i, \theta_2]$. Then, B_i is non-empty because $\theta_1 \in B_i$ and B_j is non-empty because $\theta_2 \in B_j$. Moreover, B_i and B_j are disjoint because $A_i \cap A_j = \emptyset$. In other words, $\sup B_i$ is the boundary of the two disjoint and adjacent subsets of $[\theta_1, \theta_2]$. Let $\hat{\theta}_i = \sup B_i$. By the convexity of the indirect utility functions, $x_i(\hat{\theta}_i^-) \leqslant x_j(\hat{\theta}_i^+)$ (see Figure 22). Repeating a similar argument, I consider a boundary of two disjoint and adjacent subsets of $[\theta_2, \theta_3]$. Let $D_i = A_i \cap [\theta_2, \theta_3]$ and $D_j = A_j \cap [\theta_2, \inf D_i]$. Let $\tilde{\theta}_i = \inf D_i$. Then, $x_j(\tilde{\theta}_i^-) \leqslant x_i(\tilde{\theta}_i^+)$ (see Figure 23).

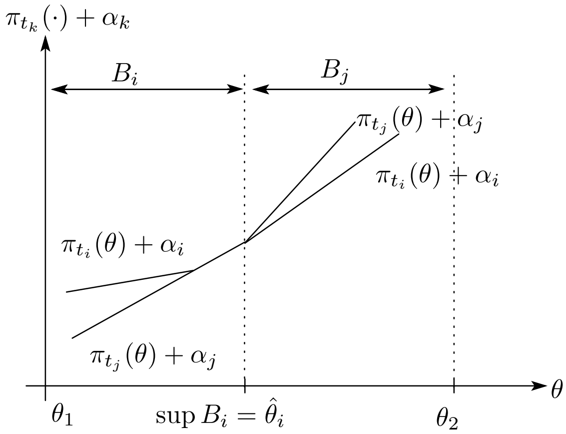


Figure 22: $x_i(\hat{\theta}_i^-) \leqslant x_j(\hat{\theta}_i^+)$

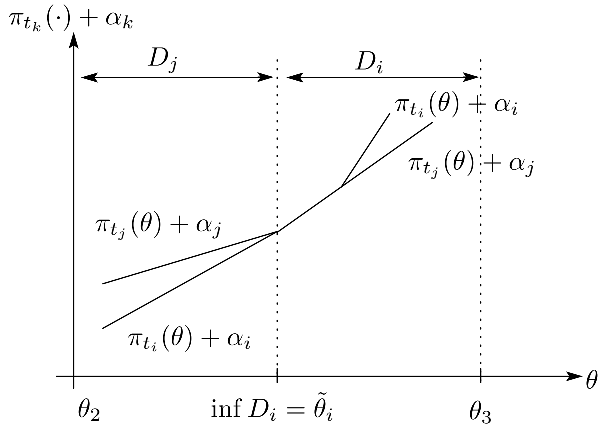


Figure 23: $x_j(\tilde{\theta}_i^-) \leqslant x_i(\tilde{\theta}_i^+)$

Principal i can imitate principal j in $B_j \cup D_j$ by offering the menu $\langle x_j(\cdot), \pi_{t_j}(\cdot) + \Delta\alpha_i \rangle$, without violating the incentive constraints. Analogously, principal j can imitate principal i in $B_i \cup D_i$ by offering the menu $\langle x_i(\cdot), \pi_{t_i}(\cdot) + \Delta\alpha_j \rangle$, without violating the incentive constraints. In equilibrium, principal i must not generate a profit greater than principal j in B_j so that for every $\varepsilon > 0$ sufficiently small, $v_j(x_j(\hat{\theta}_i + \varepsilon), \hat{\theta}_i + \varepsilon) - \pi_{t_j}(\hat{\theta}_i + \varepsilon) > v_i(x_j(\hat{\theta}_i + \varepsilon), \hat{\theta}_i + \varepsilon) - (\pi_{t_j}(\hat{\theta}_i + \varepsilon) + \Delta\alpha_i)$. On the other hand, since principal i is weakly dominant in B_i, it follows that $v_i(x_i(\hat{\theta}_i - \varepsilon), \hat{\theta}_i - \varepsilon) - \pi_{t_i}(\hat{\theta}_i - \varepsilon) \geqslant v_j(x_i(\hat{\theta}_i - \varepsilon), \hat{\theta}_i - \varepsilon) - (\pi_{t_i}(\hat{\theta}_i - \varepsilon) + \Delta\alpha_j)$. Then, I see that $C_i(x_j(\hat{\theta}_i + \varepsilon)) - C_j(x_j(\hat{\theta}_i + \varepsilon)) > -\Delta\alpha_i = \Delta\alpha_j \geqslant C_i(x_i(\hat{\theta}_i - \varepsilon)) - C_j(x_i(\hat{\theta}_i - \varepsilon))$, and hence $C_i(x_j(\hat{\theta}_i + \varepsilon)) - C_i(x_i(\hat{\theta}_i - \varepsilon)) > C_j(x_j(\hat{\theta}_i + \varepsilon)) - C_j(x_i(\hat{\theta}_i - \varepsilon))$. This yields that $0 < \int_{x_i(\hat{\theta}_i - \varepsilon)}^{x_j(\hat{\theta}_i + \varepsilon)} [C_i'(y) - C_j'(y)]dy$. It must be that $C_i'(y) \geqslant C_j'(y)$ for every $y \in [x_i(\hat{\theta}_i - \varepsilon), x_j(\hat{\theta}_i + \varepsilon)]$ with strict inequality for some $y \in [x_i(\hat{\theta}_i - \varepsilon), x_j(\hat{\theta}_i + \varepsilon)]$. Similarly, using the facts that principal j is strictly dominant in D_j and principal i is weakly dominant in D_i, I obtain $v_j(x_j(\tilde{\theta}_i - \varepsilon), \tilde{\theta}_i - \varepsilon) - \pi_{t_j}(\tilde{\theta}_i - \varepsilon) > v_i(x_j(\tilde{\theta}_i - \varepsilon), \tilde{\theta}_i - \varepsilon) - (\pi_{t_j}(\tilde{\theta}_i - \varepsilon) + \Delta\alpha_i)$ and $v_i(x_i(\tilde{\theta}_i + \varepsilon), \tilde{\theta}_i + \varepsilon) - \pi_{t_i}(\tilde{\theta}_i + \varepsilon) \geqslant v_j(x_i(\tilde{\theta}_i + \varepsilon), \tilde{\theta}_i + \varepsilon) - (\pi_{t_i}(\tilde{\theta}_i + \varepsilon) + \Delta\alpha_j)$. These yield that $C_j(x_i(\tilde{\theta}_i + \varepsilon)) - C_i(x_i(\tilde{\theta}_i + \varepsilon)) \geqslant -\Delta\alpha_j = \Delta\alpha_i > C_j(x_j(\tilde{\theta}_i - \varepsilon)) - C_i(x_j(\tilde{\theta}_i - \varepsilon))$, and hence $0 < \int_{x_j(\tilde{\theta}_i - \varepsilon)}^{x_i(\tilde{\theta}_i + \varepsilon)} [C_j'(z) - C_i'(z)]dz$. Therefore, $C_j'(z) \geqslant C_i'(z)$ for every $z \in [x_j(\tilde{\theta}_i - \varepsilon), x_i(\tilde{\theta}_i + \varepsilon)]$ with strict inequality for some $z \in [x_j(\tilde{\theta}_i - \varepsilon), x_i(\tilde{\theta}_i + \varepsilon)]$. Therefore, $C_i'(y) > C_j'(y)$ and $C_j'(z) > C_i'(z)$ for some $\{y, z\} \subseteq X$. But this contradicts Assumption 2.

It remains to show that the boundaries $\underline{\theta}_1$ and $\overline{\theta}_2$ belong to the client sets Θ_1 and Θ_2, respectively. The proof is similar to that of Lemma 7 in Biglaiser and Mezzetti (1993). This establishes the step and the lemma. □

Appendix 4: Proof of Lemma 9

Proof. Recall that client sets are given by $\Theta_1 = \{\theta \in \Theta \mid \pi_{t_1}(\theta) \geqslant \pi_{t_2}(\theta) + \Delta\alpha_1\}$ and $\Theta_2 = \{\theta \in \Theta \mid \pi_{t_2}(\theta) \geqslant \pi_{t_1}(\theta) + \Delta\alpha_2\}$, where $\Delta\alpha_1 + \Delta\alpha_2 = 0$. Consider any $\theta \in \Theta_1 \cap \Theta_2$. Then, $\pi_{t_1}(\theta) \geqslant \pi_{t_2}(\theta)+\Delta\alpha_1 \geqslant \pi_{t_1}(\theta)+\Delta\alpha_2+\Delta\alpha_1 = \pi_{t_1}(\theta)$, and hence $\pi_{t_1}(\theta) = \pi_{t_2}(\theta)+\Delta\alpha_1$. In other words, agent of type θ gets the same utility from both principals. This implies that $r_1(\theta) = \pi_{t_1}(\theta)-\pi_{t_2}(\theta)-\Delta\alpha_1 = 0$ and $r_2(\theta) = \pi_{t_2}(\theta)-\pi_{t_1}(\theta)-\Delta\alpha_2 = 0$. Since θ was arbitrary, it follows that $r_1(\theta) = r_2(\theta)$ over the competitive market. If $\Theta_1 \cap \Theta_2$ is an interval, I conclude that $\dot{r}_{t_1}(\theta) = \dot{r}_{t_2}(\theta)$ in the competitive market, where $\dot{r}_{t_1}(\theta) = u_\theta(x_1(\theta), \theta) - u_\theta(x_2(\theta), \theta)$ and $\dot{r}_{t_2}(\theta) = u_\theta(x_2(\theta), \theta) - u_\theta(x_1(\theta), \theta)$. Therefore, $x_1(\theta) = x_2(\theta)$. The first assertion is obvious, that is, it must be the case that $X_1 \cap X_2$ is non-empty. Denote by $z(\theta)$ the common value.

It remains to show that $z(\theta)$ is independent of θ. It is necessary that both principals make the same profit under any efficient tie-breaking rules. Thus, $v_1(x_1(\theta), \theta) - r_1(\theta) - \pi_{t_2}(\theta) - \Delta\alpha_1 = v_2(x_2(\theta), \theta) - r_2(\theta) - \pi_{t_1}(\theta) - \Delta\alpha_2$. Substituting $x_1(\theta) = z(\theta) = x_2(\theta)$ and $\pi_{t_1}(\theta) = \pi_{t_2}(\theta) + \Delta\alpha_1$ to obtain $u(z(\theta), \theta) - C_1(z(\theta)) - (\pi_{t_2}(\theta) + \Delta\alpha_1) = u(z(\theta), \theta) - C_2(z(\theta)) - \pi_{t_2}(\theta)$, which implies that $C_1(z(\theta)) + \Delta\alpha_1 = C_2(z(\theta))$. It is obvious that $z(\theta)$ is uniquely determined by the implicit function theorem and the value of $z(\theta)$ is independent of θ: $\dot{z}(\Delta\alpha_1) = -[C_1'(z(\Delta\alpha_1)) - C_2'(z(\Delta\alpha_1))]^{-1} > 0$ where the inequality follows from Assumption 2 (i). Denote the unique value by $z(\Delta\alpha_1)$ rather than $z(\theta)$. This establishes the lemma. □

Appendix 5: Proof of Lemma 10

Proof. I shall show that in any equilibrium associated with non-decreasing information rents, the optimal boundaries $\underline{\theta}_1$ and $\overline{\theta}_2$ must satisfy $x_1^*(\underline{\theta}_1) = z(\Delta\alpha_1) = x_2^*(\overline{\theta}_2)$. Notice that if $x_1^*(\underline{\theta}_1) = z(\Delta\alpha_1) = x_2^*(\overline{\theta}_2)$ holds, then $x_1(\theta) \geqslant x_2(\theta)$ for every $\theta \in \Theta_1$ by construction, and then the information rent $r_1(\theta)$ is actually non-decreasing.[15] Substituting the information rent $r_1(\theta)$ in Lemma 17 below, the expected profit for principal 1 becomes

$$\int_{\underline{\theta}_1}^{\overline{\theta}} \left(v_1(x_1(\theta), \theta) - \pi_{t_2}(\theta) - \Delta\alpha_1 - \int_{\underline{\theta}_1}^{\theta} [u_\theta(x_1(s), s) - u_\theta(x_2(s), s)] ds \right) f(\theta) d\theta.$$

[15] There is no need to assume that both information rents are monotone. For any decision rule $x_2(\theta) \leqslant x_1^*(\theta)$ for every $\theta \in \Theta_1$, if $x_1(\theta)$ is a solution to the optimal control problem for principal 1, then it must be the case that $x_1(\theta) \geqslant x_2(\theta)$ for every $\theta \in \Theta_1$. This yields the monotonicity of the information rent $r_1(\theta)$.

By an integration by parts,

$$\int_{\underline{\theta}_1}^{\overline{\theta}} \left(\int_{\underline{\theta}_1}^{\theta} [u_\theta(x_1(s), s) - u_\theta(x_2(s), s)] ds \right) f(\theta) d\theta$$

$$= \left(\int_{\underline{\theta}_1}^{\theta} [u_\theta(x_1(s), s) - u_\theta(x_2(s), s)] ds \right) F(\theta) \Big|_{\underline{\theta}_1}^{\overline{\theta}} - \int_{\underline{\theta}_1}^{\overline{\theta}} F(\theta)[u_\theta(x_1(\theta), \theta) - u_\theta(x_2(\theta), \theta)] d\theta$$

$$= \int_{\underline{\theta}_1}^{\overline{\theta}} [u_\theta(x_1(s), s) - u_\theta(x_2(s), s)] ds - \int_{\underline{\theta}_1}^{\overline{\theta}} F(\theta)[u_\theta(x_1(\theta), \theta) - u_\theta(x_2(\theta), \theta)] d\theta$$

$$= \int_{\underline{\theta}_1}^{\overline{\theta}} \left(\frac{1 - F(\theta)}{f(\theta)} [u_\theta(x_1(\theta), \theta) - u_\theta(x_2(\theta), \theta)] \right) f(\theta) d\theta.$$

Therefore, the objective for principal 1 is written as

$$\int_{\underline{\theta}_1}^{\overline{\theta}} \left(v_1(x_1(\theta), \theta) - \pi_{t_2}(\theta) - \Delta\alpha_1 - \frac{1 - F(\theta)}{f(\theta)} [u_\theta(x_1(\theta), \theta) - u_\theta(x_2(\theta), \theta)] \right) f(\theta) d\theta.$$

Differentiating the objective for principal 1 with respect to $\underline{\theta}_1$, (assuming that the second-order condition with respect to $\underline{\theta}_1$ is satisfied) to obtain:

$$v_1(x_1(\underline{\theta}_1), \underline{\theta}_1) - \pi_{t_2}(\underline{\theta}_1) - \Delta\alpha_1 = \frac{1 - F(\underline{\theta}_1)}{f(\underline{\theta}_1)} [u_\theta(x_1(\underline{\theta}_1), \underline{\theta}_1) - u_\theta(x_2(\underline{\theta}_1), \underline{\theta}_1)] = 0,$$

where the last equality holds because $x_1(\underline{\theta}_1) = z(\Delta\alpha_1) = x_2(\underline{\theta}_1)$ by Lemma 6. Using $\pi_{t_1}(\underline{\theta}_1) = \pi_{t_2}(\underline{\theta}_1) + \Delta\alpha_1$, I obtain $v_1(z(\Delta\alpha_1), \underline{\theta}_1) = \pi_{t_1}(\underline{\theta}_1)$. Recall that $\pi_{t_1}(\underline{\theta}_1) = v_1(x_1^*(\underline{\theta}_1), \underline{\theta}_1)$. By the strict concavity of $v_1(x_1, \underline{\theta}_1)$, it is maximized only at $x_1^*(\underline{\theta}_1)$. Therefore, it must be $z(\Delta\alpha_1) = x_1^*(\underline{\theta}_1)$. Since the full-information scheme $x_1^*(\theta)$ is strictly increasing by Lemma 4, it follows that the value of $\underline{\theta}_1$ is uniquely determined with respect to $z(\Delta\alpha_1)$, and so is with respect to $\Delta\alpha_1$ because $\dot{z}(\Delta\alpha_1) > 0$.

Similarly, the expected profit for principal 2 is written as

$$\int_{\underline{\theta}}^{\overline{\theta}_2} \left(v_2(x_2(\theta), \theta) - \pi_{t_1}(\theta) - \Delta\alpha_2 + \frac{F(\theta)}{f(\theta)} [u_\theta(x_2(\theta), \theta) - u_\theta(x_1(\theta), \theta)] \right) f(\theta) d\theta.$$

For principal 2, the information rent $r_2(\theta)$ is non-increasing over Θ_2 by construction. The first-order condition respect to $\overline{\theta}_2$ becomes:

$$v_2(x_2(\overline{\theta}_2), \overline{\theta}_2) - \pi_{t_1}(\overline{\theta}_2) - \Delta\alpha_2 = -\frac{F(\overline{\theta}_2)}{f(\overline{\theta}_2)} [u_\theta(x_2(\overline{\theta}_2), \overline{\theta}_2) - u_\theta(x_1(\overline{\theta}_2), \overline{\theta}_2)].$$

Similarly, this yields that $x_2^*(\overline{\theta}_2) = z(\Delta\alpha_1)$. This establishes the lemma. □

Appendix 6: Proof of Lemma 11

Proof. Suppose, by way of contradiction, that $\hat{\theta}_1 < \overline{\theta}_2$. In the region $[\hat{\theta}_1, \overline{\theta}_2]$, $x_1(\theta) = x_1^\sigma(\theta)$, which is strictly increasing (see the proof of Lemma 9). However, Lemma 6 tells that $x_1(\theta) = z(\Delta\alpha_1)$ in the region $[\hat{\theta}_1, \overline{\theta}_2]$, a contradiction. Similarly, $\hat{\theta}_2 \leqslant \underline{\theta}_1$. This establishes the lemma. □

Appendix 7: Statements and Proofs of Lemmas 12 – 20

Lemma 12. *$x_1(\cdot)$ is non-decreasing over Θ_1.*

Proof. By construction, $\dot{x}_1(\theta) \geqslant 0$ for every $\theta < \hat{\theta}_1$ at which $x_1(\theta)$ is differentiable. It remains to show that $x_1(\tilde{\theta}) \geqslant x_1(\theta)$ for every $\tilde{\theta} > \theta \geqslant \hat{\theta}_1$. By the log-concavity of $1 - F(\theta)$, $\frac{d}{d\theta}\frac{f(\theta)}{1-F(\theta)} \geqslant 0$. I see that $D_x\sigma_1(x_1(\tilde{\theta}), -(1-F(\tilde{\theta})), \tilde{\theta}) = 0 = D_x\sigma_1(x_1(\theta), -(1-F(\theta)), \theta)$, which implies that $D_x v_1(x_1(\tilde{\theta}), \tilde{\theta}) = \frac{1-F(\tilde{\theta})}{f(\tilde{\theta})} \leqslant \frac{1-F(\theta)}{f(\theta)} = D_x v_1(x_1(\theta), \theta)$. Then, $\tilde{\theta} - C_1'(x_1(\tilde{\theta})) \leqslant \theta - C_1'(x_1(\theta))$. This yields that $0 < \tilde{\theta} - \theta \leqslant C_1'(x_1(\tilde{\theta})) - C_1'(x_1(\theta))$, and hence $C_1'(x_1(\theta)) < C_1'(x_1(\tilde{\theta}))$. By the strict convexity of $C_1(x)$, it must be $x_1(\theta) < x_1(\tilde{\theta})$. This establishes the lemma. □

Lemma 13. *Condition (1) is satisfied.*

Proof. I see that $D_x\mathcal{H}_1(r_1, x_1, \lambda_1, \theta) = D_x v_1(x_1, \theta) f(\theta) + \lambda_1 u_{x\theta}(x_1, \theta) = D_x\sigma_1(x_1, \lambda_1, \theta) f(\theta)$. The Hamiltonian is strictly concave in x_1 since $D_{xx}\mathcal{H}_1(r_1, x_1, \lambda_1, \theta) = D_{xx}v_1(x_1, \theta) f(\theta) < 0$. By construction, $0 = D_x\sigma_1(x_1(\theta), \lambda_1(\theta), \theta)$ for every $\theta \in \Theta_1$. Therefore, the first-order condition $0 = D_x\mathcal{H}_1(r_1(\theta), x_1(\theta), \lambda_1(\theta), \theta)$ is sufficient for condition (1). This establishes the lemma. □

Lemma 14. *Conditions (8) and (9) is satisfied.*

Proof. By Lemma 10, $x_1(\theta)$ maximizes $\mathcal{H}_1(r_1(\theta), x_1, \lambda_1(\theta), \theta)$ over $x_1 \in X_1$. Consider any $\theta \in \Theta_1$. The first-order condition of $\mathcal{H}_1(r_1, x_1, \lambda_1(\theta), \theta)$ with respect to x_1 is independent of r_1, and so that I can replace $r_1(\theta)$ in $\mathcal{H}_1(r_1(\theta), x_1, \lambda_1(\theta), \theta)$ by any $r_1 \in \mathbb{R}$. Thus, $x_1(\theta)$ maximizes $\mathcal{H}_1(r_1, x_1, \lambda_1(\theta), \theta)$ over $x_1 \in X_1$ as well, and $\hat{\mathcal{H}}_1(r_1, \lambda_1(\theta), \theta) = \mathcal{H}_1(r_1, x_1(\theta), \lambda_1(\theta), \theta)$, which is linear in r_1. Therefore, condition (8) is satisfied. Condition (9) is trivial because the pure state constraint is linear in r_1. This establishes the lemma. □

Lemma 15. *The value of $\hat{\theta}_1$ is well-defined:*

$$\{\theta \in \Theta \mid -D_x v_1(x_2(\theta), \theta) f(\theta) = -(1 - F(\theta))\} = \{\hat{\theta}_1\}$$

whenever it is non-empty.

Proof. Consider any $\theta \in \Theta$ such that $-D_x v_1(x_2(\theta), \theta) f(\theta) = -(1-F(\theta))$. Suppose, by way of contradiction, that there exists a distinct element $\tilde{\theta} \in \Theta$ such that $-D_x v_1(x_2(\tilde{\theta}), \tilde{\theta}) f(\tilde{\theta}) = -(1-F(\tilde{\theta}))$. Without loss of generality, I may assume that $\theta > \tilde{\theta}$. Let $\eta(\theta) = -D_x v_1(x_2(\theta), \theta) f(\theta)$ and $\eta(\tilde{\theta}) = -D_x v_1(x_2(\tilde{\theta}), \tilde{\theta}) f(\tilde{\theta})$. For every $\theta \in [\underline{\theta}_1, \hat{\theta}_1)$, I have $0 = D_x v_1(x_1(\theta), \theta) + \frac{\eta(\theta)}{f(\theta)} u_{x\theta}(x_1(\theta), \theta)$ by the definition of $x_1(\theta)$, and hence $\frac{\eta(\theta)}{f(\theta)} = -D_x v_1(x_1(\theta), \theta)$. Since $1 - F(\theta)$ is log-concave, it follows that $-\frac{\eta(\theta)}{f(\theta)} = \frac{1-F(\theta)}{f(\theta)} \leqslant \frac{1-F(\tilde{\theta})}{f(\tilde{\theta})} = -\frac{\eta(\tilde{\theta})}{f(\tilde{\theta})}$. Then, $0 \leqslant \frac{\eta(\theta)}{f(\theta)} - \frac{\eta(\tilde{\theta})}{f(\tilde{\theta})} = D_x v_1(x_1(\tilde{\theta}), \tilde{\theta}) - D_x v_1(x_1(\theta), \theta)$, and hence $D_x v_1(x_1(\theta), \theta) \leqslant D_x v_1(x_1(\tilde{\theta}), \tilde{\theta})$. On the other hand, by construction, $x_2(\theta) \leqslant x_1^*(\theta)$. Since $x_1(\tilde{\theta}) \leqslant x_1(\theta) = x_2(\theta) \leqslant x_1^*(\theta)$, it follows that

$D_x v_1(x_1(\theta),\theta) - D_x v_1(x_1(\tilde\theta),\tilde\theta) \geqslant D_x v_1(x_1(\tilde\theta),\theta) - D_x v_1(x_1(\tilde\theta),\tilde\theta) = \theta - C_1'(x_1(\tilde\theta)) - (\tilde\theta - C_1'(x_1(\tilde\theta))) = \theta - \tilde\theta > 0$, where the first inequality follows the strict concavity of $v_1(x,\theta)$ in x, and hence $D_x v_1(x_1(\theta),\theta) > D_x v_1(x_1(\tilde\theta),\tilde\theta)$, a contradiction. Therefore, the value of $\hat\theta_1$ is well-defined. This establishes the lemma. □

Lemma 16. *The value of $\hat\theta_2$ is well-defined:*

$$\{\theta \in \Theta \mid -D_x v_2(x_1(\theta),\theta) f(\theta) = F(\theta)\} = \{\hat\theta_2\}$$

whenever it is non-empty.

Proof. Similar to that of Lemma 15. □

Lemma 17. *Conditions (4) and (5) are satisfied.*

Proof. Immediate from the continuity of $\lambda_1(\cdot)$. □

Lemma 18. *Condition (3) is satisfied. Moreover, $\lambda_1(\theta) \leqslant 0$ for every $\theta \in \Theta_1$.*

Proof. It is obvious that $\dot\lambda_1(\theta) = f(\theta)$ and $\lambda_1(\theta) \leqslant 0$ for every $\theta \geqslant \hat\theta_1$. Consider any $\theta < \hat\theta_1$. By Lemma 8, it was shown that $\underline{\theta}_1 \leqslant \overline{\theta}_2$. There are two possibilities to be considered. If $\underline{\theta}_1 = \overline{\theta}_2$, then $x_1(\theta) = x_2^*(\theta)$. If $\underline{\theta}_1 < \overline{\theta}_2$ and $\overline{\theta}_2 = \hat\theta_1$, then $x_1(\theta) = z(\Delta\alpha)$. Finally, if $\underline{\theta}_1 < \overline{\theta}_2$ and $\overline{\theta}_2 < \hat\theta_1$, then $x_1(\theta) = z(\Delta\alpha)$ for $\theta < \overline{\theta}_2$ and $x_1(\theta) = x_2^*(\theta)$ for $\theta \in [\overline{\theta}_2, \hat\theta_1)$.

Case 1. If $x_1(\theta) = x_2^*(\theta)$ then $\lambda_1(\theta) \leqslant 0$ and $\dot\lambda_1(\theta) = 0$.

Proof of Case 1. Under Assumption 2, $C_1'(x_2^*(\theta)) < C_2'(x_2^*(\theta)) = \theta$, and hence $\theta > C_1'(x_2^*(\theta))$. This implies that $\lambda_1(\theta) = -D_x v_1(x_2^*(\theta),\theta) f(\theta) = -[\theta - C_1'(x_2^*(\theta))] f(\theta) < 0$. Differentiating $\lambda_1(\theta) = -[\theta - C_1'(x_2^*(\theta))] f(\theta)$ with respect to θ to obtain

$$\begin{aligned}\dot\lambda_1(\theta) &= -[f(\theta) + \theta f'(\theta) - C_1''(x_2^*(\theta))\dot x_2^*(\theta) f(\theta) - C_1'(x_2^*(\theta)) f'(\theta)] \\ &= -[f(\theta) + \theta f'(\theta)] + C_1''(x_2^*(\theta))\dot x_2^*(\theta) f(\theta) + C_1'(x_2^*(\theta)) f'(\theta).\end{aligned}$$

Since it was shown that $\dot x_2^*(\theta) = [C_2''(x_2^*(\theta))]^{-1}$ in the proof of Lemma 4 and $C_1''(x_2^*(\theta)) = C_2''(x_2^*(\theta))$, it follows that $C_1''(x_2^*(\theta))\dot x_2^*(\theta) f(\theta) = f(\theta)$. Using distribution uniformity, $\dot\lambda_1(\theta) = -f(\theta) + f(\theta) = 0$. Therefore, $\lambda_1(\theta) \leqslant 0$ and $\dot\lambda_1(\theta) = 0$ for every $\theta < \hat\theta_1$. This establishes the case.

Case 2. If $x_1(\theta) = z(\Delta\alpha_1)$ then $\lambda_1(\theta) \leqslant 0$ and $\dot\lambda_1(\theta) = -f(\theta)$.

Proof of Case 2. By the definition, $\lambda_1(\theta) = -D_x v_1(x_1(\theta),\theta) f(\theta) = -[\theta - C_1'(x_1(\theta))] f(\theta) = -[\theta - C_1'(z(\Delta\alpha_1))] f(\theta)$. Recall that $x_1(\theta) \leqslant x_1^*(\theta)$ by construction. Then, $\theta = C_1'(x_1^*(\theta)) \geqslant C_1'(x_1(\theta)) = C_1'(z(\Delta\alpha_1))$. Therefore, $\lambda_1(\theta) \leqslant 0$. Moreover, using distribution uniformity,

$$\begin{aligned}\dot\lambda_1(\theta) &= -[(1 - C_1''(z(\Delta\alpha_1))\frac{d}{d\theta} z(\Delta\alpha_1)) f(\theta) + (\theta - C_1'(z(\Delta\alpha_1))) f'(\theta)] \\ &= -[f(\theta) + \theta f'(\theta) - C_1'(z(\Delta\alpha_1)) f'(\theta)] = -f(\theta).\end{aligned}$$

Therefore, $\lambda_1(\theta) \leqslant 0$ and $\dot{\lambda}_1(\theta) = -f(\theta)$ for every $\theta < \hat{\theta}_1$. This establishes the case.

To sum up,

$$\dot{\lambda}_1(\theta) = \begin{cases} 0 & \text{if } \underline{\theta}_1 = \overline{\theta}_2 \text{ or } \theta \in [\overline{\theta}_2, \hat{\theta}_1) \\ -f(\theta) & \text{if } \underline{\theta}_1 < \overline{\theta}_2 = \hat{\theta}_1 \text{ or } \theta \notin [\overline{\theta}_2, \hat{\theta}_1). \end{cases}$$

On the other hand, by the definition,

$$\tau_1(\theta) = \begin{cases} f(\theta) & \text{if } \underline{\theta}_1 = \overline{\theta}_2 \text{ or } \theta \in [\overline{\theta}_2, \hat{\theta}_1) \\ 2f(\theta) & \text{if } \underline{\theta}_1 < \overline{\theta}_2 = \hat{\theta}_1 \text{ or } \theta \notin [\overline{\theta}_2, \hat{\theta}_1). \end{cases}$$

I am ready to verify condition (3). Recall that the costate condition is written as $\dot{\lambda}_1(\theta) = -[-f(\theta) + \tau_1(\theta)] = f(\theta) - \tau_1(\theta)$. Therefore, I conclude that $\dot{\lambda}_1(\theta) + \tau_1(\theta) = f(\theta)$. This establishes the lemma. □

Lemma 19. *Conditions (6) and (7) are satisfied.*

Proof. By construction, $\dot{r}_1(\theta) = u_\theta(x_1(\theta), \theta) - u_\theta(x_2(\theta), \theta) = x_1(\theta) - x_2(\theta) \geqslant 0$, and hence the information rent $r_1(\theta)$ must be non-decreasing. Then, $r_1(\underline{\theta}_1) = 0$ for sure, and the pure state constraint $r_1(\theta) \geqslant 0$ will be binding in the left part of the distribution. That is, $r_1(\underline{\theta}_1) = 0$ is fixed and $r_1(\overline{\theta})$ is free. By Lemma 18, $\lambda_1(\theta) \leqslant 0$ for every $\theta \in \Theta_1$. Since $\lambda_1(\underline{\theta}_1) \leqslant 0$ and $r_1(\underline{\theta}_1) = 0$, it follows that $\lambda_1(\underline{\theta}_1) r_1(\underline{\theta}_1) = 0$. Therefore, the initial transversality condition is satisfied. Since $\lambda_1(\overline{\theta}) = 0$ by construction and $r_1(\overline{\theta}) \geqslant 0$, it follows that $\lambda_1(\overline{\theta}_1) r_1(\overline{\theta}_1) = \lambda_1(\overline{\theta}) r_1(\overline{\theta}) = 0$. Therefore, the terminal transversality condition is satisfied. This establishes the lemma. □

Lemma 20. *Condition (2) is satisfied. Moreover,* $r_1(\theta) = \displaystyle\int_{\underline{\theta}_1}^{\theta} [u_\theta(x_1(s), s) - u_\theta(x_2(s), s)] ds$ *for every* $\theta \geqslant \underline{\theta}_1$.

Proof. Recall the multiplier $\tau_1(\theta) \geqslant 0$: $\tau_1(\theta) > 0$ for every $\theta < \hat{\theta}_1$ and $\tau_1(\theta) = 0$ for every $\theta \geqslant \hat{\theta}_1$ by construction. In the proof of Lemma 19, it was shown that $r_1(\underline{\theta}_1) = 0$ and $\dot{r}_1(\theta) = 0$ for every $\theta < \hat{\theta}_1$, which implies that $r_1(\theta) = 0$ for every $\theta < \hat{\theta}_1$. Moreover, since $r_1(\theta)$ is non-decreasing, it follows that $r_1(\theta) \geqslant 0$ for every $\theta \geqslant \hat{\theta}_1$. Therefore, $\tau_1(\theta) r_1(\theta) = 0$ for every $\theta \in \Theta_1$. Moreover, the envelope condition $\dot{r}_1(\theta)$ and the boundary condition $r_1(\underline{\theta}_1) = 0$ yields the integrability condition. This establishes the lemma. □

References

BIGLAISER, G. AND C. MEZZETTI (1993): "Principals Competing for an Agent in the Presence of Adverse Selection and Moral Hazard," *Journal of Economic Theory*, 61, 302–330.

GOLDMAN, M. B., H. E. LELAND, AND D. S. SIBLEY (1984): "Optimal Nonuniform Prices," *Review of Economic Studies*, 51, 305–19.

IVALDI, M. AND D. MARTIMORT (1994): "Competition under Nonlinear Pricing," *Annales d'Donomie et de Statistique*, 34, 71–114.

LAFFONT, J.-J. AND D. MARTIMORT (2002): *The Theory of Incentives : the Principal-Agent Model*, Princeton University Press, Princeton, N.J.

LEWIS, T. R. AND D. E. M. SAPPINGTON (1989a): "Countervailing incentives in agency problems," *Journal of Economic Theory*, 49, 294–313.

——— (1989b): "Inflexible Rules in Incentive Problems," *The American Economic Review*, 79, 69–84.

MAGGI, G. AND A. RODRIGUEZ-CLARE (1995): "On Countervailing Incentives," *Journal of Economic Theory*, 66, 238–263.

MASKIN, E. AND J. RILEY (1984): "Monopoly with Incomplete Information," *The RAND Journal of Economics*, 15, 171–196.

MCMANUS, B. (2007): "Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee," *The RAND Journal of Economics*, 38, 512–532.

MONTEIRO, P. K. AND F. PAGE (2008): "Catalog Competition and Nash Equilibrium in Nonlinear Pricing Games," *Economic Theory*, 34, 503–524.

MUSSA, M. AND S. ROSEN (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

NÖLDEKE, G. AND L. SAMUELSON (2007): "Optimal Bunching without Optimal Control," *Journal of Economic Theory*, 134, 405 – 420.

PIGOU, A. C. (1932): *The Economics of Welfare*, London: Macmillan, fourth ed.

ROBERTS, K. (1979): "Welfare Considerations of Nonlinear Pricing," *Economic Journal*, 89, 66–83.

ROCHET, J.-C. (1985): "The Taxation Principle and Multi-Time Hamilton-Jacobi Equations," *Journal of Mathematical Economics*, 14, 113–128.

——— (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191–200.

ROCHET, J.-C. AND L. A. STOLE (2002): "Nonlinear Pricing with Random Participation," *Review of Economic Studies*, 69, 277–311.

SEIERSTAD, A. AND K. SYDSÆETER (1987): *Optimal control theory with economic applications.*, Amsterdam; New York; Oxford and Tokyo.