Two-Part Tariffs with Quality Degradation*

Sissel Jensen†

March 23, 2006

Abstract

There is a gap between the recommendations of the theory of second degree price discrimination and the practices of firms that target consumer segments with varying willingness to pay with two or more distinct tariffs. We present a model where consumers' private information is single dimensional and the allocation rule is two-dimensional. In contrast to the established result in nonlinear pricing, we find that the per-unit price may be non-monotonic: low-demand consumers face a two-part tariff with a per-unit price possibly below marginal cost, and even zero, whereas high-demand consumers face tariffs with per-unit charges above marginal cost. On the other hand, all consumers but the one on the top of the distribution, are faced with a quality restriction, quality being monotonically increasing in type. Finally, we show that this practice increases welfare due to increased consumption efficiency.

JEL. CLASS. NO: D42, D82, L96
KEYWORDS: Price discrimination, two-part tariffs, quantity discounts, telecommunications

---

*I am grateful for comments from Petter Osmundsen, Fred Schroyen, Lars Sørgard, Jon Vislie, and two anonymous referees. I also thank seminar participants at the Norwegian School of Economics and Business Administration, participants at the 3rd Nordic Workshop in Industrial Organization (NORIO III) in Helsinki. Financial support from Telenor is gratefully acknowledged.

†Norwegian School of Economics and Business Administration. Mailing address: Helleveien 30, 5045 Bergen, Norway. Phone: 55 95 94 57, Internet: sissel.jensen@nhh.no.

1 Introduction

There is a wide gap between the recommendations one can draw from theoretical models on second degree price discrimination and the actual practices of firms that aim to target consumer segments with different willingness to pay with two or more distinct tariffs. Attempting to bridge this gap, we present a model where consumers' private information is single dimensional and the allocation rule is two-dimensional. Hence, we explore an extension of a simple two-part pricing arrangement by assuming that the firm can observe a customer's usage of its service along more than one dimension.[1] The firm offers a menu of two part tariffs, where each tariff is characterized by a fixed fee, a uniform per unit charge, and some quality restriction (a usage restriction, e.g., line speed or calling circle).[2] The quality restriction is intended to separate consumers with different willingness to pay for the firm's service, similar to the standard model in Mussa and Rosen (1978) and Deneckere and McAfee (1996). However, rather than assuming inelastic demand, it is assumed that consumers have elastic demand, as in Maskin and Riley (1984) and Oi (1971).

To illustrate the general idea, consider the strategy used in broadband pricing. Users of broadband services have very diverse needs when it comes to Internet surfing, e-mailing, music and video downloads, and high-quality video and audio streaming, and this reflects their demand for speed and their intensity of usage with respect to download/upload. While surfing the Internet is slightly faster on a high-speed connection, high-quality video streaming will perform badly on a low-speed connection. If low demand types are served with a higher line speed this is valuable first of all because it frees up time to explore further content on the Internet, and to some extent consume new high speed services. However, the propensity to use free time to explore further content is diminishing and increased line speed will eventually be of no use. For high demand consumers a higher line speed is valuable primarily because it gives access to a new series of services and contents that the consumer values highly. Hence, consumers' willingness to pay for access speed depends partly on which

[1] Matthews and Moore (1987) generalize one extension of the bench-mark model by assuming that consumers' private information is single dimensional, while the firm offers contracts with two or more attributes (quality and warranty) in addition to the monetary payment from consumer to the firm. García (2005) extends the results in Matthews and Moore (1987).

[2] Two part tariffs where each tariff consists of only a fixed fee and a per usage charge have been studied in Oi (1971); Faulhaber and Panzar (1977); Goldman, Leland and Sibley (1984); Sharkey and Sibley (1993), and Wilson (1993, chapter 6). Mirman and Sibley (1980) study the problem in Goldman et al. (1984) in a multiproduct firm.

services they use and partly on their usage intensity.

Table 1 shows an example. The broadband company Tiscali charges high-demand consumers £24.99 for broadband at 512 kbps, while low-demand consumers pay £10 less for the same speed, but on different terms, since they must stay online less than 50 hours, or download less than 1 GB, per month. Another example that is consistent with the framework in this paper is the widespread practice of various kinds of calling circle tariffs, for instance "Friends and Family" tariffs.[3] Table 2 lists some additional examples from telecommunications.

Table 1: *Pricing of broadband, Tiscali (UK, June 2004)*

Product	*Downstream speed*	*Cost per month (Pounds)*
Broadbandx3	150 kbps	15.99, Free usage
Broadbandx5	256 kbps	19.99, Free usage
Broadbandx10 50 Hours	512 kbps	19.99, After 50 hours 2p per minute
Broadbandx10 1 GB	512 kbps	19.99, After 1 GB 2p per Mb
Broadbandx10 Unlimited	512 kbps	24.99, Free usage

The firm's task is to design a menu of two part tariffs with appropriate fixed fees, per unit charges, and usage restriction, in such a way that all consumers find it individual rational to select the tariff that is in fact intended for his/her type, given that consumers have private information about a one dimensional characteristic. This is another way of saying that the solution to the problem must obey the incentive compatibility constraint and the participation constraint. In the early days when ISDN was the sole access technology, price per dial-up online minute was the sole instrument in addition to the fixed fee in the firm's pricing decision.[4] The optimization problem was simplified by imposing the "downward adjacent" incentive compatibility constraint, together with the participation constraint for the lowest consumer type buying the product. The simplified problem is the solution to the full problem under the condition that the per unit charge is decreasing over the type space (monotonicity condition).

[3]Subscribers are billed according to aggregate minutes of calling to a restricted set of network subscribers. Firms' use of calling circle tariffs has received some attention in other areas of economics literature as well. Wang and Wen (1998) consider a duopoly model with demand side heterogeneity, where such pricing behavior enables a new firm to enter the market despite the presence of consumer switching costs. Laffont, Rey and Tirole (1998) examine the effects of discriminatory pricing on the negotiated interconnection agreements between rival network operators. In a recent publication written independently of this study, Shi (2003) study the use of calling circle tariffs from a social network theory perspective.

[4]This is not the full story since many telecom firms offered internet surfing (dialling the ISP) at different rates contingent on the consumers choice of "Internet calling plans", i.e., calling plans with a single number.

Table 2: *Examples of telecommunications services pricing (June 2004)*

Company/Product	*Service restriction*	*Pricing arrangement*
Vodafone (UK) Perfect Fit	Lower rates anytime Lower rates daytime Lower rates evening/weekend	Two-part tariffs with inclusive minutes
Orange (UK) Your Plan	Lower rates any network anytime Lower rates Orange-Orange anytime Lower rates Orange-Orange off-peak	Two-part tariffs with inclusive minutes
O_2 (UK)	Lower rates anytime Lower rates daytime/evening time	Two-part tariffs with inclusive minutes
BT (UK) Together 1,2,3	Low rate evening/weekend Free calls evening/weekend Free calls anytime	Two-part tariffs
Telenor (Norway) Friends & Family	Lower rates on calls to mobile Lower rates on national calls Lower rates on international calls	Two-part tariffs
Tiscali (UK) Dial-up internet access	Unlimited surfing anytime Unlimited surfing daytime, weekdays Unlimited surfing daytime all week	Flat rate Per minute outside hours
BT Broadband 512 kbps	Less than 15 GB monthly download Upgrade to free download	Flat rate
BT Broadband 1Mb	Less than 30 GB monthly download Upgrade to free download	Flat rate

In our model, the firm faces a slightly different problem. While holding on to the assumption that the private information is single dimensional (for instance, willingness to pay for viewing content on the internet), and that the outside option for the consumer is type-independent, we assume that consumers' willingness to pay is correlated, not only with one, but with two variables that are observed by the firm.[5] One is a quantity variable, and the other one is some variable related to the quality of the service (e.g., line speed). We show that introducing an additional instrument might change the incentive constraint

[5]Models where the firm aims to screen consumers according to multiple dimensions of uncertainty soon become difficult to solve, partly because the incentive compatibility conditions are frequently not only binding between adjacent types. Discrete models with fewer incentive compatibility constraints can be tractable, Jensen (2001) model a discrete version of the present model. Rochet and Stole (2003) gives a comprehensive survey of the literature related to multidimensional screening. Rochet and Stole (2002) and Armstrong and Vickers (2001) relax the assumption that the reservation utility is perfectly known by the firm and introduce stochastic participation. The principal must induce both truthful information revelation and voluntary participation. Both find that efficient two part tariffs may emerge as an equilibrium. Other extensions of the bench-mark models are to introduce more than one instrument or more than one observable variable (see Matthews and Moore (1987); Sappington (1983); Caillaud, Guesnerie, Rey and Tirole (1988); García (2005)).

and that we have to be more careful in using the simplified approach to profit maximization. If consumers have a systematic incentive to understate their private information, informational rent increases over the type space, and the firm only has to be concerned about participation at the very lowest end of the type space. While we show that the need to secure incentive compatibility does not conflict with the need to ensure participation and that complete separation between consumer types is reached, we show that the firm's allocation rule may not be monotonic along both dimensions. However, the allocation has to satisfy a "weighed monotonicity constraint". Especially, and in sharp contrast to the existing literature, we show that the per unit charge may be below marginal cost in the lower end of the type space, and that it may increase in some subinterval. In this respect, our paper is close to Matthews and Moore (1987), which also shows that the optimal contracts need not be monotonic in type. However, the allocation in their model depend on consumers' attitude towards risk. García (2005) extend the non-monotonicity properties in Matthews and Moore (1987) to a setup with quasi-linear preferences.

If we change the interpretation a little, the model can be used to analyze nonlinear pricing and bundling in a multiproduct monopoly setting. Assuming that the firm sells a very large number of products, the firm can bundle a subset of the products and charge units within this product bundle according to a distinct two-part tariff. In a model with unit demand, Bakos and Brynjolfsson (1999) study the strategy of bundling a large number of information goods (goods with zero or very low marginal costs of production) and selling them for a fixed price. One of their findings is that the firm should offer a menu of different bundles aimed at each market segment and practice price discrimination when consumers' tastes are positively correlated.[6] Armstrong (1999) studies optimal multiproduct nonlinear pricing when the firm offers a very large number of products, applicable to telecommunications.[7] When consumers' tastes are correlated across products, he find that a menu of two part tariffs, each of which have prices proportional to marginal costs, can extract

[6]Since the literature on bundling to a large extent deals with a setting with only two products and linear pricing, most is not relevant to our model.

[7]Multiproduct nonlinear pricing is also studied elsewhere. Mirman and Sibley (1980) consider a multiproduct monopoly facing consumers who are differentiated by a single characteristic, where the firm offers a menu of commodity bundles together with the price for the bundle. Hence, Mirman and Sibley (1980) has similarities with our paper. Sibley and Srinagesh (1997) explore the difference between screening the different dimensions of consumer types independently by means of two-part tariffs and the alternative of bundling all taste parameters to design a single two-part tariff. Miravete (2001) studies multidimensional screening where different type components distinguish quality dimensions of products that can be aggregated.

almost all available profits. However, Armstrong (1999) covers only the case where all products are sold in all segments.

Section 2 presents the framework with usage pattern heterogeneity used in this paper. Section 3 presents the results we obtain within this setting. Section 4 offers some concluding remarks.

2 A model with usage pattern heterogeneity

The market is served by a monopoly, and resale opportunities are absent.[8] The cost function is assumed to be linear, and the fixed cost is excluded from the profit measure. There is a continuum of consumers on the demand side, having heterogeneous and unobserved willingness to pay for the service in question. Consumers also have heterogeneous usage patterns, and this can be observed by the firm. Section 2.1 describes the details of the demand side of the model. If the firm, say, for some exogenously given reason chooses not to restrict consumers' mode of usage, the qualitative results are that the per unit price is set above marginal cost for every consumer but the one with the highest willingness to pay. The fixed fee increases and the per unit price decreases over the type space. Hence, if the heterogeneity is very large the model can result in a situation where some consumer segments face a high price cost margin, while other segments are excluded from purchasing.

As the introduction suggests, demand side heterogeneity may come about because different consumers use the service very differently. While some broadband subscribers only surf the Internet and read e-mails, services that perform well on low line speed, others may use the connection to watch live video, which requires high line speed to perform reasonably well. Just as call minutes to one network node (your boyfriend, for example) is a bad substitute for call minutes to a different network node (say, your mother), a low speed connection is a bad substitute for a high speed connection if one wants to watch live video. If the firm offers a tariff with a restricted line speed it restricts the consumers' mode of usage since they will not have access to all available content on the internet.[9]

[8]Although telecommunications is subject to competition almost all over the world, we do not add imperfect competition to the framework. The reason for this is simply that it adds too much complexity (see Rochet and Stole (2003) and Stole (2005)).

[9]Vodafone has recently introduced a tariff option under the name "At Home". On this tariff the usage charge is lower when the consumer make calls from the home zone (some radio coverage area around the home or the office). Hence, consumers' mode of usage is restricted since full mobility is possible otherwise.

In our framework then, consumers' willingness to pay for the service depends partly on their usage patterns and partly on their usage intensity.

We assume that consumers' usage pattern is captured by a conditional distribution function $H(s|\theta)$ over a continuous variable $s \in [0, 1]$, with a unimodal probability density function $h(s|\theta)$. The usage pattern then is defined as each consumer type's intensity over the various modes of usage s (e.g. line speed).[10] Consumers' usage intensity depends on a single dimensional and unobserved demand parameter θ. The conditional distribution function is derived from a cumulative bivariate distribution $H(s, \theta)$ on $s \in [0, 1]$ and $\theta \in [1, 2]$, with a marginal cumulative distribution $F(\theta)$. Both distributions are prior knowledge for the firm. The assumption that high demand types have a more dispersed usage pattern implies the further assumption that the distribution $H(s|\theta')$ first-order stochastically dominates the distribution $H(s|\theta'')$ if $\theta'' > \theta'$. That is, consumers with higher quantity preferences do also have higher preferences for "quality", and usage mode is an intrinsic part of consumers' preferences. One implication of this assumption is that a price increase will not change the cumulative distribution across usage modes, even though a price increase will change individual consumption levels across all usage modes. This implication may be questioned. However, there is no obvious alternative assumption – i.e., that the usage pattern will be more concentrated or more dispersed when the per unit charge increases.[11] Appendix C describes the family of distribution function that our conclusions are derived from.

2.1 Utility

The subutility of a consumer of type θ from consuming q units of the service at some given usage mode s is given by the following subutility function

$$u(q, \theta; s) = \theta q - \frac{1}{2h(s|\theta)} q^2, \tag{1}$$

[10] For instance, a telecom firm keeps records of each subscriber's dispersion of calls in the network, i.e., number of call minutes to all available network nodes, and a mobile company can observe the location a call is made from (mobile stations). The firm could also learn about consumers' usage patterns from market research and market surveys. A rationale behind the difference in information held by the firm can also be that regulations prohibit the firm from giving exclusive offers so that consumers must self select tariffs. However, it can still be a legal pricing strategy to offer tariffs with restrictions on the mode of usage. Accordingly, we assume that the usage pattern is specific to each individual consumer type θ, and that there is a correlation between θ and the usage pattern.

[11] The same assumption is made in Bousquet and Ivaldi (1997).

where $h(s|\theta)$ is the conditional probability distribution function of s on $[0,1]$ given θ, with a conditional cumulative distribution function $H(s|\theta)$, $h'_s(s|\theta) \leq 0$ and $H'_\theta(s|\theta) \leq 0$.[12] The subutility function takes into account that modes of usage with higher usage intensity contributes more to aggregate utility than modes of usage with lower usage intensity, and that different consumers have different usage patterns.

Each consumer is billed according to a two part tariff $T = \{E, p, s\}$, where $E \geq 0$ is a fixed entry fee, $p \geq 0$ is a charge per unit of usage, and $s \in [0,1]$ is a restriction on the mode of usage on this particular tariff. The case with no restriction on the mode of usage is normalized to $s = 1$. Hence if $s = 1$, the mode of usage is not restricted at all , and if $s = 0$ the consumer is de facto prevented from using the service. If $0 < s < 1$ the consumer can enjoy consumption on every mode of usage up to s. If a consumer of type θ finds it individual rational to pay the fixed fee E, the price is the same across all usage modes and equal to p per unit of usage. The volume at each mode of usage maximizes the quasilinear subutility function $u(q, s, \theta) - pq$. Hence, expected quantity demand at some given s is

$$q(p, s, \theta) = (\theta - p)h(s|t) \equiv x(p, \theta)h(s|\theta). \tag{2}$$

Aggregate consumption on all usage modes up to s is given by

$$Q(p, s, \theta) = \int_0^s (\theta - p)h(z|\theta)dz = x(p, \theta)H(s|\theta), \tag{3}$$

$q(\cdot)$ and $Q(\cdot)$ are both nonincreasing in p, while $Q(\cdot)$ is also nondecreasing in s. The signs of the other derivatives of $q(\cdot)$ and $Q(\cdot)$ depends on the schedule $\{p(\theta), s(\theta)\}$. When each subutility function is quasilinear, the aggregate demand function appears to maximize aggregate consumer surplus, and a consumer's gross surplus measured in monetary terms is represented by the area under the demand function.[13] We can write the indirect utility for a consumer

[12] We simplify the notation in the following manner: If we have a function, say, $f(x, z)$ we use the notation f_x for the derivative of $f(x, z)$ with respect to x. If there is no ambiguity about the arguments of a function, these will be omitted.

[13] We abstract from the fact that some consumers may have positive utility even in the case when consumption is zero. In the case of broadband usage, this is not at all problematic. In the case of fixed or mobile telephony, the case is different since a consumer may want a network connection in order to receive calls only, or to be able to make emergency calls. Oren, Smith and Wilson (1982) and Bousquet and Ivaldi (1997) study nonlinear pricing under the presence of demand externalities, for instance when the benefit a consumer receives in a communication network depends on his or her access to communication partners and increases with the size of the network.

type that is charged according to two part tariff $\{E, p, s\}$ as

$$\begin{aligned} V(E, p, s, \theta) &= \int_p^\theta x(z, \theta) H(s|\theta) dz - E \\ &\equiv \omega(p, \theta) H(s|\theta) - E \equiv v(p, s, \theta) - E. \end{aligned} \tag{4}$$

By Roy's identity we have

$$V_p(\cdot) = v_p(p, s, \theta) = \omega_p(p, \theta) H(s|\theta) = -x(p, \theta) H(s|\theta).$$

Consumers will buy if there exists a tariff $\{E, p, s\}$ such that $V(E, p, s, \theta) \geq 0$. If not, they are better off not buying. Furthermore, we will assume that the outside option is the same for all consumers and normalize this to zero. The individual rationality constraint (participation constraint) is given by the constraint

$$\int_p^\theta x(z, \theta) H(s|\theta) dz - E \geq 0. \tag{5}$$

Figure 1 illustrates the individual rationality constraint for two different types θ_1 and θ_2. A reduction in s has an adverse effect on consumers' participation constraint, and the effect is more severe for high demand types compared to low demand types.

The indirect utility is convex in (E, p) and the marginal rate of substitution between the per unit price and the fixed fee, MRS_{pE}, is given by

$$\frac{dE}{dp} = -xH \leq 0. \tag{6}$$

Hence, the slope of $V(E, p, s, \theta)$ is negative and the consumer is willing to pay a higher fixed fee against a reduction in the per unit charge.

The marginal rate of substitution varies with s and θ.

$$\frac{d^2E}{dpds} = -xh \leq 0, \quad \text{and} \quad \frac{d^2E}{dpd\theta} = -\big(x_\theta H + xH_\theta\big) \gtrless 0.$$

The slope of a consumer type's indifference curve is steeper the higher is s and a restriction on s causes a negative shift in $V(E, p, s, \theta)$. Since $H_\theta(\cdot)$ is negative, we cannot be certain that the marginal rate of substitution increases with θ for any profile $s(\theta)$.

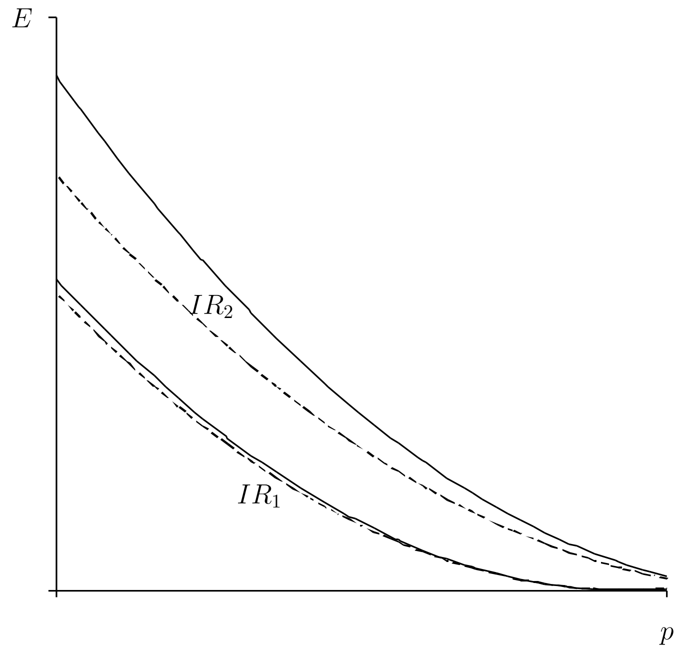


Figure 1: *Participation constraints with (dashed lines) and without (solid lines) a usage restriction for two different consumer types (IR_1 for type θ_1 and IR_2 for θ_2).*

2.2 Welfare maximization

With constant returns to scale technology the first best solution to the problem is obtained by maximizing social welfare as the sum of consumer and produce surplus with respect to p and s for each θ.

$$\max_{p(\theta)\geq 0\; s(\theta)\in[0,1]} \int_{p(\theta)}^{\theta} x(z,\theta)H(s(\theta)|\theta)dz + (p(\theta)-c)x(p(\theta),\theta)H(s(\theta)|\theta),$$

which yields first order conditions

$$\big(p(\theta)-c\big)x_p H(s(\theta)|\theta) = 0 \quad\text{and}\quad \Big(\omega(p(\theta),\theta)+(p(\theta)-c)x(p(\theta),\theta)\Big)h(s(\theta)|\theta) = 0.$$

The two above conditions can only hold simultaneously if $s(\theta) = 1$ and $p(\theta) = c$.

2.3 Profit maximization

The firm maximizes profit under two constraints. The individual rationality constraint states that consumers must receive at least the utility they can obtain from spending their money on other goods or services, hence

$$V(\theta) = V(E(\theta), p(\theta), s(\theta), \theta) \geq 0. \tag{7}$$

The other constraint on the firm's maximization problem is the incentive compatibility constraint

$$V(E(\theta), p(\theta), s(\theta), \theta) \geq V(E(\theta), p(\theta'), s(\theta'), \theta), \qquad (8)$$
$$V(\theta, \theta) \geq V(\theta, \theta').$$

For continuous profiles $p(\theta)$ and $s(\theta)$, the incentive compatibility constraint is found by solving

$$\theta \in \arg\max_{\theta'} \Big(\omega(p(\theta'), \theta) H(s(\theta')|\theta) - E(\theta')\Big).$$

Hence, the firm may increase the fixed fee if the per unit price $p(\theta)$ is reduced, or if the allowance $s(\theta)$ is increased.

$$-x(p, \theta) H(s|\theta) p'(\theta) + \omega(p, \theta) h(s|\theta) s'(\theta) = E'(\theta). \qquad (9)$$

The second order condition for incentive compatibility requires the following condition (differentiating condition (9))

$$\frac{d^2 V}{(d\theta')^2}(\theta, \theta) \leq 0 \qquad \Rightarrow \qquad \frac{d^2 V}{d\theta' d\theta}(\theta, \theta) \geq 0. \qquad (10)$$

The last condition in (10) can be stated as (dropping all functions arguments)[14]

$$-\Big\{x_\theta H + x H_\theta\Big\} \frac{dp}{d\theta} + \Big\{\omega_\theta h + \omega h_\theta\Big\} \frac{ds}{d\theta} \geq 0, \qquad (11)$$

which can also be written as

$$\Big(\frac{V_{\theta p}}{V_{\theta p} + V_{\theta s}}\Big) \frac{dp}{d\theta} + \Big(\frac{V_{\theta s}}{V_{\theta p} + V_{\theta s}}\Big) \frac{ds}{d\theta} \geq 0.$$

In this case, incentive compatibility does not longer require that $p(\theta)$ is monotonically decreasing. Instead the second order condition requires a weighted monotonicity constraint to hold (see García (2005)). However, one of the allocations must be monotonically decreasing (p) or increasing (s) in type. If $V_{\theta p} > 0$ and $V_{\theta s} > 0$, sufficient conditions for global incentive compatibility are that $p(\theta)$ is monotonically nonincreasing in θ, and $s(\theta)$ is monotonically nondecreasing in θ. However, we can allow other profiles for $p(\theta)$ and $s(\theta)$ as long as the positive term outweighs the negative. A consumer will not choose

[14] If the per unit price is the single instrument with tariffs $\{E(\theta), p(\theta)\}$, global incentive compatibility can be replaced by the local downward incentive compatibility constraint given that $p(\theta)$ is nonincreasing together with the single crossing condition $\omega_{p\theta} \leq 0$.

a tariff with a lower per unit price if the restriction in the mode of usage is sufficiently severe. Since H_θ and h_θ can both be negative it is neither sufficient nor necessary that $p(\theta)$ being nonincreasing and $s(\theta)$ being nondecreasing in θ.

If we ignore the second order condition for global incentive compatibility (11) for the moment, letting local incentive compatibility be the only binding constraint, we can apply the envelope theorem and write

$$\frac{\partial V(\theta)}{\partial \theta} = v_\theta(p(\theta), s(\theta), \theta).$$

Hence, the informational rent can be expressed as

$$V(\theta) = \int_{\underline{\theta}}^{\theta} v_\theta(p(u), s(u), u)du. \tag{12}$$

If the informational rent in (12) is increasing we know that consumers always obtain positive consumer surplus if they choose a contract that gives nonnegative surplus for some lower type. Therefore, the only binding individual rationality constraint will be for the lowest type. Otherwise, the firm sacrifices profit if it leaves consumers with higher utility than necessary. Hence, if $V'(\theta) > 0, \forall\theta$ the firm maximizes profit subject to (12) and the individual rationality constraint for the very lowest type

$$V(\underline{\theta}) = V(E(\underline{\theta}), p(\underline{\theta}), s(\underline{\theta}), \underline{\theta}) = 0. \tag{13}$$

On the other hand, since $H_\theta(\cdot)$ is negative, the informational rent is not unambiguously increasing in θ, and we cannot rule out the possibility of countervailing incentives. Under countervailing incentives the individual rationality constraint can bind for other types than $\underline{\theta}$.[15] In our case, $V'(\theta)$ is given by

$$V_\theta = \omega_\theta H + \omega H_\theta \gtreqless 0.$$

The first term is the marginal valuation for consumption up to s, which is increasing in θ. The second part takes into account that higher types have higher probability weight on higher s (the assumption about first-order stochastic dominance).

The monopoly maximizes the sum of fixed fees and variable profits, subject to individual rationality and incentive compatibility. With respect to incentive

[15] Countervailing incentives can arise if the individual rationality constraint is type dependent, or if the sign of the informational rent is ambiguous. See Lewis and Sappington (1989); Maggi and Rodriguez-Clare (1995); Jullien (2000)

compatibility, we will assume that the solution satisfies $V'(\theta) > 0$, and that the second order condition (11) is satisfied. Hence, we maximize the profit subject to (12) and (13). After we have obtained a solution it is necessary to check that the second order condition for global incentive compatibility as well as the assumption $V'(\theta) > 0$ are in fact satisfied.

The firm's profit is given by

$$\max_{E(\theta),p(\theta),s(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \big\{E(\theta) + (p(\theta) - c)x(p(\theta),\theta)H(s(\theta)|\theta)\big\} f(\theta)d\theta \tag{14}$$

s.t.

$$V(\underline{\theta}) = 0 \qquad \text{and} \qquad V(\theta) = \int_{\underline{\theta}}^{\theta} v_\theta(p(u), s(u), u)du$$

$$E(\theta) \in [0,\infty), \quad p(\theta) \in [0,\infty), \quad s(\theta) \in [0,1].$$

Substituting for $E(\theta)$ from the participation constraint, and integrating by part gives the profit as (see Appendix A)

$$\begin{aligned}\max_{p(\theta)\geq 0, s(\theta)\in[0,1]} \int_{\underline{\theta}}^{\bar{\theta}} \Big\{\omega\big(p(\theta),\theta\big)H(s(\theta)|\theta) + \big(p(\theta) - c\big)x(p(\theta),\theta)H(s(\theta)|\theta) \\ - \frac{(1-F(\theta))}{f(\theta)}\Big(\omega_\theta(p(\theta),\theta)H(s(\theta)|\theta) + \\ \omega(p(\theta),\theta)H_\theta(s(\theta)|\theta)\Big)\Big\} f(\theta)d\theta.\end{aligned} \tag{15}$$

The maximization of profit with respect to $p(\cdot)$ and $s(\cdot)$ requires that the integral in (15) is maximized with respect to $p(\theta)$ and $s(\theta)$ for all θ, subject to the constraints $p(\theta) > 0$ and $s(\theta) \in [0,1]$. The optimality conditions for this Kuhn-Tucker problem are in the Appendix B.[16]

3 Optimal pricing policy

Marginal profit at θ with respect to price is given by

$$(p - c)x_p H - (1 - F)\Big(\omega_{\theta p} H + \omega_p H_\theta\Big) = \tfrac{\partial W}{\partial p} - (1 - F)\tfrac{d}{dp}V'(\theta).$$

Social surplus increases in p as long as $p > c$. However, due to private information the monopolist is not able to appropriate the entire surplus, but has to

[16] Since the constraints are linear it is sufficient that $\Pi(p, s; \theta, c)$ is concave in (p, s) for an interior solution to solve the problem. However, the profit expression is not in general concave and we have to check that the solution to the first order conditions is indeed a maximum.

leave consumers with an information rent. If $(1-F)\frac{d}{dp}V'(\theta) < 0$ the marginal information rent decreases with p and the monopolist will increase the per unit charge above the first best level $p = c$ for all types but the very highest one $(\bar{\theta})$. The opposite will be true if $(1-F)\frac{d}{dp}V'(\theta) > 0$, and the marginal information rent increases as p is increased. If $\frac{d}{dp}V'(\theta) > 0$ it can even be the case that $p = 0$ is optimal in some parts of the type space.

Marginal profit with respect to p can also be expressed as

$$\Big\{-(p-c) + (1-F) + (1-F)x\frac{H_\theta}{H}\Big\}H.$$

Since $\frac{\partial \Pi}{\partial p}\big|_{s=1} = -(p-c) + (1-F)$ it is clear that the price will always be below the monopoly price with p as the single instrument. The isolated welfare effect from a usage restriction on the firm's per unit price is positive.

Marginal profit at θ with respect to the usage restriction s is given by

$$(p-c)xh + wh - (1-F)\left(\omega_\theta h + \omega h_\theta\right) = \frac{\partial W}{\partial s} - (1-F)\frac{d}{ds}V'(\theta).$$

Again, the sum of the first two terms evaluates the effect on social surplus at θ from an increase in s. In addition, setting s below the first best level will increase or decrease the monopolist's ability to appropriate social surplus at a given p because it affects the marginal information rent. Notice that marginal profit is zero at $s = 0$ and $s = 1$. If $\frac{d}{ds}V'(\theta) < 0\ \forall\ \theta, p$, the marginal information rent increases as s is decreased and $s = 1$ is certainly optimal. In the opposite case, when $\frac{d}{ds}V'(\theta) > 0\ \forall\ \theta, p$ it will be optimal to restrict s for all types but the very highest one.

If an interior solution with $p(\theta) > 0$ and $0 < s(\theta) < 1$, $p(\theta)$ exists, this must satisfy

$$p - c = (1-F)\Big(1 + x\frac{H_\theta}{H}\Big). \tag{16}$$

According to (16) the per unit price can be above or below marginal cost depending on the sign of the term $(1 + x\frac{H_\theta}{H})$, but the price-cost margin will never exceed $1-F$ for those consumer types being served. In the case with p as the only instrument, the price-cost margin is given by the term $1-F$. Hence, introduction of a second instrument reduces the price-cost distortion for those consumer types being served. The following propositions are derived under the assumption that the conditional probability function $H(s|\theta)$ is derived from a bivariate Beta distribution with a joint probability function $g(s, \theta)$. The shape of $h(s|\theta)$ depends on the demand parameter θ, the lower is θ the larger is the mass for low s. The distribution is defined in the Appendix C.

Proposition 1 (Efficiency at the top) *The consumer type with the very highest willingness to pay* ($\theta = \bar{\theta}$) *is offered a two-part tariff with* $p = c$ *together with* $s = 1$. *Every other consumer type is offered a two part tariff with a price-cost distortion together with a mode of usage restriction.*

Proposition (1) is proved in Appendix D.1. We recognize the "no distortion at the top" result not only with respect to the usage charge, but also with respect to the mode of usage restriction. Also, if consumers with lower demand do not have a significantly different usage pattern, the firm will sort consumers via the usage charge rather than restricting consumers' mode of usage, and we will have pooling along s. Notice that the combination of $s = 1$ and $p - c = (1 - F)$ will only happen at the very highest end of the distribution of θ. Every other consumer will face a distortion, either via the usage charge, via a restriction on usage, or both.

Proposition 2 (Free usage) *If the demand side heterogeneity is sufficiently large, together with* $c \leq \min\left\{\frac{\widetilde{\theta}}{2}, 2 - \widetilde{\theta}\right\}$, *consumers in the interval* $[1 + \frac{c}{2}, \widetilde{\theta}]$ *will be offered a tariff with a mode of usage restriction* ($s < 1$) *together with a zero usage charge* ($p = 0$). *The larger is the heterogeneity in consumers' mode of usage, the larger is* $\widetilde{\theta}$. *The larger is the marginal cost, the smaller is* $\widetilde{\theta}$. *However, the tariff is incentive compatible only if* $c \leq \frac{\theta}{2}$. *Further, it satisfies the participation constraint only if* $c \leq 2 - \theta$.

Proposition 2 is proved in the Appendix D.2. This result shows that it might be an optimal strategy to sort consumers solely via the usage restriction, and we will have pooling along p in the lower end of the type space. If the heterogeneity in consumers' mode of usage is large, the cost of restricting low demand types mode of usage (in terms of the effect on their willingness to pay) is low compared with the gain that can be achieved by the reduction in the information rent paid to higher types. Therefore, it might even be profitable to offer tariffs with free usage in low demand segments, and restore incentive compatibility via higher usage restrictions instead.

Proposition 3 (Market coverage) *Every consumer with demand parameter* $\theta > 1 + \frac{c}{2}$ *will consume a strictly positive quantity. The entire market is covered if* $c = 0$.

(i) In the case that $p > 0$, *the marginal consumer that finds it just individual rational to pay the fixed fee is given by* $\theta = 1 + \frac{c}{2}$.

(ii) In the case that $p = 0$, the firm achieves nonnegative profit by serving consumer types $\theta \geq 1 + \frac{c}{2}$

Proposition (3) is proved in the Appendix D.3. If the per unit price is the only available instrument it is easy to verify that the firm will serve consumer types in $[1 + \frac{c}{2}, \bar{\theta}]$. It may seem a little surprising that the monopolist is not inclined to serve more consumers when it gets control over an additional instrument. Restricting s enables the firm to reveal information about θ, but revealing this information has a cost side. For a given per unit charge a low s restricts low demand types' willingness to pay. Because the firm can compensate low demand types for this restriction by reducing the per unit price, the firm finds it profitable to reduce s. In designing the optimal use of the two instruments, the firm finds it unprofitable to increase market coverage.

Proposition 4 (Interior solution) *Consumers with demand parameter $\theta \in [\widetilde{\theta}, \bar{\theta}\rangle$ is confronted with a per unit price $p = c + (1 - F)\big(1 + x\frac{H_\theta}{H}\big) < 1 - F$, together with a usage restriction $0 < s < 1$.*

Proposition (4) is proved in the Appendix D.4, except for $\widetilde{\theta}$ which is defined in Proposition 2. In the case with p as the only instrument the price-cost margin is given by the term $1 - F$. Hence, introduction of a second instrument reduces the price-cost distortion.

Finally, turning attention to welfare considerations, it is clear that the monopoly solution departs from the full information solution, except for the very highest type $\bar{\theta}$, and that the monopolist serves too few consumers relative to the full information solution. However, the relevant standard of comparison is not welfare maximization under full information, but a second best solution where welfare is maximized subject to informational asymmetry. Another standard of comparison is the bench-mark solution where a monopoly does not restrict usage via s at all, but only via the usage charge p. There are two potential sources of welfare gains due to quality degradation in this framework. First, consumers gain if overall efficiency in consumption increases. Second, if introducing quality degradation induces the firm to serve consumers it would otherwise exclude, these consumers' surpluses will increase. Proposition 5 and 6 below summarizes the welfare effects of the firm's use of mode of usage restrictions by comparing the outcome under profit maximization to the outcome in a second-best welfare optimum.

The maximum second-best welfare in our context is found by maximizing the unweighed social welfare under asymmetric information subject to a break-even constraint $\Pi \geq 0$. This gives the Lagrangian (dependent on θ)

$$L = (1+\lambda)\Big[\omega\big(p,\theta\big)H(s|\theta) + \big(p-c\big)x(p,\theta)H(s|\theta)\Big] \\ - \lambda\,(1-F)\,\Big[\omega_\theta(p,\theta)H(s|\theta) + \omega(p,\theta)H_\theta(s|\theta)\Big]. \tag{17}$$

The first order conditions remains the same as with profit maximization as the objective, except that $(1-F)$ is replaced with $\frac{\lambda}{1+\lambda}(1-F)$, λ being the shadow price for public funding. First we state the second-best allocation in Proposition 5.

Proposition 5 (Second best allocation) *The following properties characterize a second best allocation where a social planner maximizes the sum of consumer surplus and profit, under the restriction that the firm breaks even:*

(i) No distortion at the top: $p(\bar{\theta}) = c$ and $s(\bar{\theta}) = 1$

(ii) Free usage is optimal only if λ is very large, given that the heterogeneity in consumers mode of usage is large as well. Then $p = 0$ is optimal for $\theta < \widetilde{\theta}_W$ where $\widetilde{\theta}_W \lll \widetilde{\theta}$.

(iii) Market coverage: Consumers with demand parameter $\theta \in [1 + \frac{c}{2} - \frac{2-c}{2(2\lambda+1)}, 2]$ consume a strictly positive quantity. If $c < \frac{1}{2}$ it is optimal to cover the entire market. If $c = \frac{2}{3}$ it is optimal to cover the entire market given that $\lambda \leq \frac{1}{2}$.

(iv) Consumers with demand parameter $\theta \in [\widetilde{\theta}_W, 2\rangle$ is confronted with a per unit price $p = c + \frac{\lambda}{1+\lambda}(1-F)\big(1 + x\frac{H_\theta}{H}\big) < 1 - F$, together with a usage restriction $s < 1$.

See the Appendix D.5 for a proof. From Proposition 5 it is clear that two part tariffs with mode of usage restrictions welfare dominates two part tariffs that sort consumers solely via the usage charge. It is also evident that the monopoly serves too few consumers relative to the second best allocation. Further, if restrictions in consumers' mode of usage is an effective means for rent extraction, the monopoly sets s and p too low relative to the second best. Since market coverage remains unchanged, welfare gains arise only if overall consumption efficiency increases. Each consumer's aggregate consumption level increases in s and p, hence, welfare increases if the net effect on aggregate consumption is positive.

Proposition 6 (Welfare gains) *Introducing mode of usage restriction in a monopoly with two part tariffs increases the welfare of every consumer being served. That is* $Q(p(\theta), s(\theta), \theta) > Q(c + 2 - \theta, 1, \theta)\ \forall\ \theta \in [1 + \frac{c}{2}, 2)$. *Since the market coverage is* $[1 + \frac{c}{2}, 2]$ *in any case, increased consumption efficiency is the only source of welfare gains.*

The Appendix D.5 provide a sketch for the proofs of Propositions 5 and 6. The Propositions show that although the monopoly will exaggerate the magnitude of the distortions, the direction of the distortions is in line with the second best. Especially, a pricing policy with mode of usage restriction and below cost pricing is preferable. By degrading quality the firm becomes better informed about consumers' privately known demand parameter θ. This enables the firm to capture a larger fraction of the social surplus and leads to a reduction in the price-cost margin for all consumers. The other side of this is that consumers are served with insufficient quality, and will therefore reduce their consumption. However, by further reductions in the price-cost margin the firm can to some extent compensate low demand types for this. If quality degradation is sufficiently effective in this respect, the price-cost margin might even be negative in the lower end of the type space. As to the second source of welfare gains, Proposition 3 shows that market coverage might increase, and that market coverage never decreases. Figure 2 show the welfare and profit maximizing choice $s(\theta)$ for $c = 0.2,\ \lambda = 0.3$ and $b = \{4,\ 5.5,\ 7\}$.

4 Conclusion

This paper examines a firm's incentive to degrade it's service along a vertical quality dimension when the firm offers a continuum of two part tariffs, and shows how the two forms of usage restrictions interact in the screening analysis. Hence, we combine the insights from Mussa and Rosen (1978) and Maskin and Riley (1984). We show that the "no distortion at the top" result is preserved in both instruments. Since the intention behind a distortion in per-unit charges and quality levels is to restrict the informational rent to higher types, by restricting the number of units the can claim informational rent on, the main insights are not new. However, the results contradict one of the most established insights in nonlinear pricing, that the per-unit charge should be monotonically decreasing over the type space. We find that allocation of quality is monotonic in type, while per usage charge might be non-monotonic. What happens is that mode of usage restrictions are used to separate consumers, partly in combina-

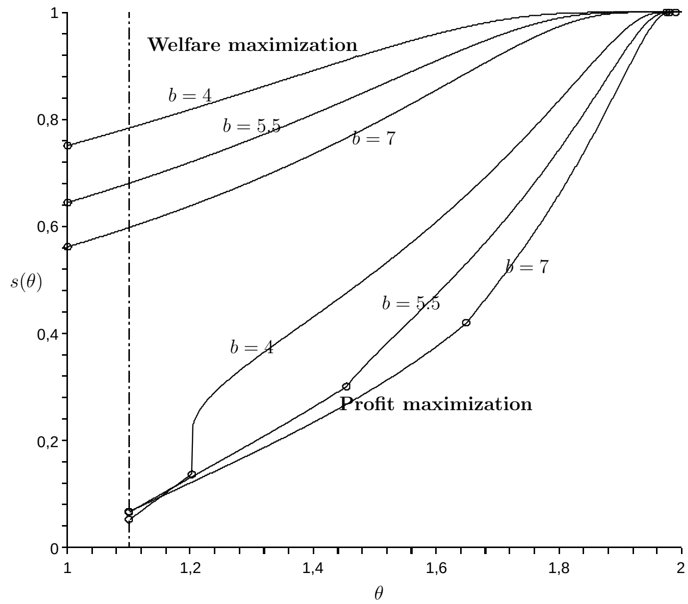


Figure 2: *The profit and welfare maximizing choice of* $s(\theta)$ *given three different assumptions about the demand side heterogeneity with respect to mode of usage* $(c = 0.2,\ \lambda = 0.3)$.

tion with distortions in per unit charges. However, if imposing the mode of usage restriction is a very efficient instrument, the monopoly might prefer to use only this instrument heavily in some segments, and rather compensate low demand types for large restrictions by offering tariffs with free usage. In the case that both instruments are used in combination to achieve sorting, they are both monotonic. If the firm relies on mode of usage restriction alone, per unit charges are typically non-monotonic, while the usage restriction is monotonically increasing in type.

In comparison, a social planner maximizing second best welfare, defined as maximizing the unweighed social welfare under asymmetric information subject to a break-even constraint, will indeed find it optimal to distort both allocation rules, but not to the same extent as the profit maximizing monopoly. The monopoly is likely to use mode of usage restriction alone, and to an excessive degree, and to balance this by setting lower per usage charges in low demand segments. The two instruments aim at the same objective, which is to restrict the consumption level in low demand segments below the efficient level. As in all screening models, this is done, not really to restrict consumption in low demand segments, but to hurt high demand consumers if they choose a tariff with a low price.

Appendix

A Derivation of the profit expression

The firm's profit is given by

$$\max_{E(\theta),p(\theta),s(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \big\{E(\theta) + (p(\theta) - c)x(p(\theta), \theta)H(s(\theta)|\theta)\big\} f(\theta)d\theta$$

s.t.

$$V(\underline{\theta}) = 0 \text{ and } V(\theta) = \int_{\underline{\theta}}^{\theta} v_\theta(p(u), s(u), u)du$$

$$E(\theta) \in [0, \infty), \; p(\theta) \in [0, \infty), \; s(\theta) \in [0, 1]$$

Substituting for $E(\theta)$ from the participation constraint gives

$$\max_{p(\theta)\geq 0, s(\theta)\in[0,1]} \int_{\underline{\theta}}^{\bar{\theta}} \big\{\omega(p(\theta), \theta)H(s(\theta)|\theta) - \int_{\underline{\theta}}^{\theta} v_\theta(p(u), s(u), u)du$$
$$+ (p(\theta) - c)x(p(\theta), \theta)H(s(\theta)|\theta)\big\} f(\theta)d\theta$$

Next, after integrating by parts we obtain the firm's profit as

$$\max_{p(\theta)\geq 0, s(\theta)\in[0,1]} \int_{\underline{\theta}}^{\bar{\theta}} \big\{\omega(p(\theta), \theta)H(s(\theta)|\theta) - (1 - F(\theta))(v_\theta(p(\theta), s(\theta), \theta))$$
$$+ (p(\theta) - c)x(p(\theta), \theta)H(s(\theta)|\theta)\big\} f(\theta)d\theta$$

and we can now write

$$\max_{p(\theta)\geq 0, s(\theta)\in[0,1]} \int_{\underline{\theta}}^{\bar{\theta}} \Big\{\omega\big(p(\theta), \theta\big)H(s(\theta)|\theta) + \big(p(\theta) - c\big)x(p(\theta), \theta)H(s(\theta)|\theta)$$
$$- \tfrac{(1-F(\theta))}{f(\theta)} \Big(\omega_\theta(p(\theta), \theta)H(s(\theta)|\theta) + \tag{A.1}$$
$$\omega(p(\theta), \theta)H_\theta(s(\theta)|\theta)\Big)\Big\} f(\theta)d\theta$$

B Optimality conditions

Maximizing the term under the integral in $\Pi(p, s; \theta, c)$ in (A.1) subject to the conditions on $p(\theta)$ and $s(\theta)$ yields the following complementary slackness conditions for the Kuhn-Tucker problem

$$\frac{\partial \Pi}{\partial p} \leq 0, \qquad p \geq 0, \qquad p\frac{\partial \Pi}{\partial p} = 0, \tag{B.1}$$

$$\frac{\partial \Pi}{\partial s} - \mu \leq 0, \qquad s \geq 0, \qquad s(\frac{\partial \Pi}{\partial s} - \mu) = 0, \tag{B.2}$$

$$s \leq 1, \qquad \mu \geq 0, \qquad \mu(1 - s) = 0, \tag{B.3}$$

where

$$\frac{\partial \Pi}{\partial p} = -(p-c)\omega_{pp}H - (1-F)(\omega_{\theta p}H + \omega_p H_\theta), \tag{B.4}$$

$$\frac{\partial \Pi}{\partial s} = (p-c)xh - (1-F)(\omega_\theta h + \omega h_\theta) + \omega h. \tag{B.5}$$

and μ is the multiplier for the constraint $s \leq 1$.

If an interior solution exists this is given by a pricing policy with $0 < p^* < 1-F$ together with $0 < s^* < 1$, and $\frac{\partial \Pi}{\partial p}(p^*, s^*; c) = \frac{\partial \Pi}{\partial s}(p^*, s^*; c) = 0$. Notice that $\frac{\partial \Pi}{\partial s} = 0$ if $s = 0$ or $s = 1$, independent of p.

- *Second order conditions*

The signs of the second order derivatives of the profit function cannot be determined in general. Hence, sufficient conditions for profit maximization must be evaluated in each case.

C Bivariate distribution

Let a bivariate distribution be defined by the standard Beta distribution with parameters α and β on the support $[0,1]$. Let $\alpha = 2$ and $\beta(\theta) \geq 2$. The bivariate probability function is then

$$g(s,\theta) = \begin{cases} \frac{s(1-s)^{(\beta(\theta)-1)}}{B(2,\beta(\theta))} & \text{if } 0 \leq s \leq 1, \\ 0 & \text{otherwise}, \end{cases}$$

and the bivariate cumulative density function is

$$G(s,\theta) = \begin{cases} \frac{\int_0^s t(1-t)^{(\beta(\theta)-1)}dt}{B(2,\beta(\theta))} & \text{if } 0 \leq s \leq 1, \\ 0 & \text{otherwise}. \end{cases}$$

assuming $\beta'(\theta) < 0$, $\beta''(\theta) \leq 0$, $B(\alpha, \beta)$ is the Beta function. $\beta(\theta)$ determines the shape of the distribution, the higher is $\beta(\theta)$ (lower is θ) the larger is the mass for low s. For $w = 2$ the distribution is symmetric around the expectation $s = \frac{1}{2}$. Otherwise, the distribution is skewed with expectation $s < \frac{1}{2}$.

The marginal pdf over θ is

$$f(\theta) = \int_0^1 g(s,\theta)ds = 1.$$

Hence, the conditional probability distribution is given by

$$h(s|\theta) = \frac{g(s,\theta)}{f(\theta)} = g(s,\theta)$$

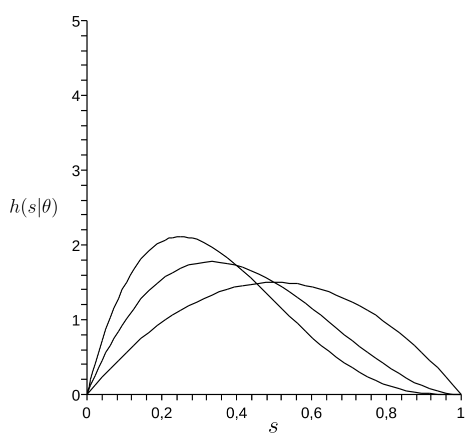
Consumers' taste parameter θ is uniform on a unit interval $[1, 2]$. As to the shape of the conditional probability function we assume that this is given by $\beta(\theta) = 2 + b(2 - \theta)$, thus $\beta(2) = 2$, and $\beta(1) = 2 + b$. The greater is b the larger is the difference in consumers' usage patterns.

The conditional cdf $H(s|\theta)$ is continuous, and the conditional pdf $h(s|\theta)$ is unimodal, positive and integrable on the support $[0, 1]$, $h(s|\theta)$, $\lim_{s\to 0} h(s|\theta) = \lim_{s\to 1} h(s|\theta) = 0$.

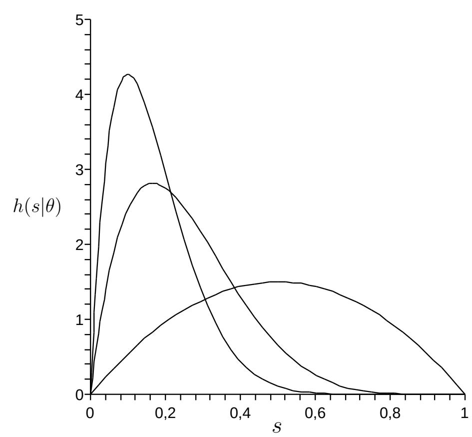
The conditional pdf and the conditional cdf will also satisfy the following

$$H_\theta = \tfrac{\partial H(s|\theta)}{\partial\theta} < 0, \quad \tfrac{\partial}{\partial\theta}\left(\tfrac{H_\theta}{H}\right) \le 0, \quad \tfrac{\partial}{\partial\theta}\left(\tfrac{h_\theta}{h}\right) \le 0,$$

$$H_s = h(s|\theta) > 0, \quad \tfrac{\partial}{\partial s}\left(\tfrac{H_\theta}{H}\right) \ge 0, \quad \tfrac{\partial}{\partial s}\left(\tfrac{h_\theta}{h}\right) \ge 0.$$



(a) $b = 2$.



(b) $b = 8$.

Figure 3: *Conditional probability distributions for $\theta = 1,\ 1.5,\ 2$.*

D Proof of Propositions 1–4

D.1 Proof of Proposition 1

- *First order conditions for $\theta = \bar{\theta}$*

The first order condition with respect to p is given by $-(p - c)H = 0$. Hence, we must have $p = c$ at $\theta = \bar{\theta}$. The first order condition with respect to s reduces to wh. This is positive whenever $s \in (0,1)$. Hence, $s=1$ is the only possible choice that satisfies the first order conditions at $\bar{\theta}$.

- *When will the constraint $s=1$ be binding?*

If $s=1$ is binding we must have that $\frac{\partial \Pi}{\partial s} \geq 0$ evaluated at $p-c=1-F=2-\theta$ and $s=1$. In addition it must also be the case that $\frac{\partial^2 \Pi}{\partial s^2} \leq 0$ for $p-c=2-\theta$ and $s=1$. However, since both $\frac{\partial \Pi}{\partial s} = \frac{\partial^2 \Pi}{\partial s^2} = 0$ we do not know whether $s=1$ is a local maximum or a local minimum. Evaluating marginal profit with respect to s at $p-c=2-\theta$ gives

$$\frac{\partial \Pi}{\partial s} = \tfrac{1}{2}x^2 h\Big(1-(2-\theta)\tfrac{h_\theta}{h}\Big).$$

Here, it is the case that $\frac{h_\theta}{h}$ is negative as s approaches zero, and infinitely positive as s approaches 1. For $s=0$ and $s=1$ we have $\frac{\partial \Pi}{\partial s}=0$. For $\theta \neq 2$, $\frac{\partial \Pi}{\partial s}=0$ for some $0<s<1$. Since $\Pi(0,c+2-\theta)=0$ and $\Pi(1,c+2-\theta)>0$, $s=1$ is optimal for $\theta=2$. For every other θ there exists a stationary point $s \neq 0,1$. If profit is concave at this point (say $\hat{s}(\theta)$) we know that $\Pi(\hat{s},c+2-\theta) > \Pi(1,c+2-\theta)$, and $s=1$ is not optimal. The second order derivative is given by

$$\frac{\partial^2 \Pi}{\partial s^2} = \tfrac{1}{2}x^2 h_s\Big(1-(2-\theta)\tfrac{h_\theta}{h}\Big) - \tfrac{1}{2}x^2 h(2-\theta)\Big[\tfrac{h_\theta}{h}\Big]_s'.$$

Since the last term is positive, it is easy to confirm that profit is concave at least close to $\hat{s}(\theta)$.

The second order condition for global incentive compatibility is given by

$$x_\theta \frac{dp}{d\theta} \geq 0,$$

which is satisfied. Since $x>0$ we also know that $V'(\theta)>0$. This completes the proof of Proposition 1.■

D.2 Proof of Proposition 2

- *When will the constraint $p=0$ be binding*

For $p=0$ to be a binding constraint it must be the case that $\frac{\partial \Pi}{\partial p}<0$. Marginal profit with respect to p evaluated at $p=0$ is given by

$$\left.\frac{\partial \Pi}{\partial p}\right|_{p=0} = cH + (1-F)(H+\theta H_\theta).$$

By inspection, it is clear that the constraint cannot be binding for $s=1$ and $\theta=2$.

We know that s is determined according to (B.5), say that this defines a function $s'(\theta)$, which is increasing in θ. Hence, if both the condition that $\frac{\partial \Pi}{\partial p}(s'(\theta), p)$ is negative when p approaches 0, and that $\frac{\partial \Pi}{\partial s}(s, 0)$ is negative for $s > s'(\theta)$, setting $p = 0$ is indeed optimal. Otherwise, the monopolist will increase s and it is less likely $p = 0$ is optimal. The two conditions states

$$\frac{-H_\theta}{H} \geq \frac{1}{\theta}\Big(1 + \frac{c}{2-\theta}\Big) \tag{D.1}$$

$$\frac{-h_\theta}{h} \leq \frac{1}{\theta}\Big(2 - \frac{\theta - 2c}{2-\theta}\Big) = \frac{1}{\theta}\Big(1 + \frac{c}{2-\theta} - \frac{2\theta - 2 - c}{2-\theta}\Big) \tag{D.2}$$

Let (D.1) when it is binding define a function $s_1(\theta)$ and (D.2) define a function $s_2(\theta)$. If $s_1(\theta) \geq s_2(\theta)$, then $p = 0$ is the optimal choice.

The left hand sides in (D.1) and (D.2) are identical, positive and finite, as s approaches zero.[17] Further, $\left[\frac{-H_\theta}{H}\right]$ and $\left[\frac{-h_\theta}{h}\right]$ are both lower for lower demand side heterogeneity with respect to mode of usage, and are also from our assumptions decreasing in θ. Hence, if demand side heterogeneity in consumers' mode of usage is sufficiently large, we expect that both conditions are satisfied at least in some interval in the lower end of the type space, and that $s_1 > s_2$ in this interval. On the other hand, if θ is close to 2 the right hand side in (D.1) approaches ∞, while the right hand side in (D.2) approaches $-\infty$. Since $\lim_{s\to 1}\left[\frac{-H_\theta}{H}\right] = 0$, while $\lim_{s\to 1}\left[\frac{-h_\theta}{h}\right] = -\infty$, we can conclude that $\lim_{\theta\to 2} s_1(\theta) < 1$ and $\lim_{\theta\to 2} s_2(\theta) = 1$ and that $p = 0$ cannot bind at $\bar{\theta}$.

Since both $\frac{d}{ds}\left[\frac{-H_\theta}{H}\right] \leq 0$ and $\frac{d}{ds}\left[\frac{-h_\theta}{h}\right] \leq 0$, while the right hand sides are are unchanged, if a solution to (D.1) and (D.2) exists $s_1(\theta)$ and $s_2(\theta)$ are unique. However, if c is sufficiently large, a nonnegative solution to (D.2) might fail to exist for low values of θ. The firm will serve these consumers with $s = 0$ (de facto exclusion). If a nonnegative solution to (D.1) fail to exist, the constraint cannot be binding at all.

The slopes of s_1 and s_2 are given by

$$\frac{ds_1}{d\theta} = \frac{\left[\frac{H_\theta}{H}\right]'_\theta - \frac{1}{\theta^2}\left[1 - \frac{2c(\theta-1)}{(2-\theta)^2}\right]}{-\left[\frac{H_\theta}{H}\right]'_s}, \tag{D.3}$$

$$\frac{ds_2}{d\theta} = \frac{\left[\frac{h_\theta}{h}\right]'_\theta - \frac{1}{\theta^2}\left[2\left(1 - \frac{2c(\theta-1)}{(2-\theta)^2}\right) + \frac{\theta^2}{(2-\theta)^2}\right]}{-\left[\frac{h_\theta}{h}\right]'_s}. \tag{D.4}$$

[17] Using L'Hôptal's rule we get $\lim_{s\to 0}\left[\frac{-H_\theta}{H}\right] = \lim_{s\to 0}\left[\frac{-h_\theta}{h}\right] = \lim_{s\to 0}\left[\frac{-h_{\theta s}}{h_s}\right] > 0$.

While $\frac{ds_2}{d\theta}$ is undoubtedly positive, $\frac{ds_1}{d\theta}$ might be negative for large c and large θ. We know that $s_1(2) < s_2(2) = 1$. If $\lim_{s\to 0}\left[\frac{h_{\theta s}}{h_s}\right] > \frac{2}{2-c}$ we know that $s_1(1) > s_2(1)$. Thus, if $s_1(\theta)$ and $s_2(\theta)$ crosses this is at most once and the constraint $p = 0$ is binding in the interval $\theta \in [1, \widetilde{\theta}]$, where $\widetilde{\theta}$ is the solution to $s_1(\theta) = s_2(\theta)$. $\widetilde{\theta}$ increases as the heterogeneity in mode of usage increases, and $\widetilde{\theta}$ decreases with c.

Profit is concave in s given that

$$\frac{\theta^2}{2}(2-\theta)h_s\Big[\frac{-h_\theta}{h} - \frac{1}{\theta}\Big(2 - \frac{\theta-2c}{2-\theta}\Big)\Big] - \frac{\theta^2}{2}(2-\theta)h\Big[\frac{h_\theta}{h}\Big]'_s. \leq 0$$

The last term is positive, hence profit is concave for s close to $s_2(\theta)$. Since the only other points satisfying $\frac{\partial\Pi}{\partial s} = 0$ is $s = 0$ and $s = 1$, $s_2(\theta)$ constitutes a maximum for profit.

At the same time, since $p < c$ it is necessary to check that the firm obtains positive profit on each type that it serves. Profit evaluated at $p = 0$ must be nonnegative

$$\tfrac{1}{2}\theta^2 H - cxH - (2-\theta)(\theta H + \tfrac{1}{2}\theta^2 H_\theta) \geq 0.$$

Rewriting this, and combining it with condition (D.2) when this is binding enable us to formulate the following implicit condition on (s, θ)

$$\Big[\frac{-H_\theta}{H}\Big] - \Big[\frac{-h_\theta}{h}\Big] \geq 0.$$

Since we know that (D.1) is met while (D.2) is binding, the difference above is given by

$$\Big[\frac{-H_\theta}{H}\Big] - \Big[\frac{-h_\theta}{h}\Big] \geq \frac{1}{\theta}\Big(\frac{\theta-c}{2-\theta} - 1\Big) \geq 0.$$

Hence, a sufficient condition for some type θ to be served is that

$$\frac{1}{\theta}\Big(\frac{\theta-c}{2-\theta} - 1\Big) \geq 0 \qquad \Rightarrow \qquad \theta \geq 1 + \frac{c}{2}.$$

Finally, we need to show that $V'(\theta) > 0$ and that the solution is incentive compatible. Knowing that s is increasing in θ, the second order condition for global incentive compatibility is given by

$$\theta h + \tfrac{1}{2}\theta^2 h_\theta \geq 0 \quad \Rightarrow \quad \frac{-h_\theta}{h} \leq \frac{2}{\theta}.$$

Since we know that condition (D.2) is met this reduces to the condition that $2c - \theta \leq 0$. If $c \leq \frac{2}{3}$ the tariff is implementable.

The marginal information rent is positive under the condition that

$$\frac{-H_\theta}{H} \leq \frac{2}{\theta}.$$

Since (D.2) is binding we can subtract each side in the inequality above and rewrite the condition as

$$\left[\frac{-H_\theta}{H}\right] - \left[\frac{-h_\theta}{h}\right] \leq \frac{1}{\theta}\left[\frac{\theta - 2c}{2-\theta}\right]. \tag{D.5}$$

Again, using the fact that (D.1) is met, while (D.2) is binding, we know that $V'(\theta) \geq 0$ if $\frac{1}{\theta}\left[\frac{\theta-c}{2-\theta} - 1\right] \leq \frac{1}{\theta}\left[\frac{\theta-2c}{2-\theta}\right]$. The condition reduces to $2 - c \geq \theta$. Altogether then, we must have that $c \leq \min\left\{\frac{\theta}{2}, 2-\theta\right\}$. Or, if we assume that $c \leq \frac{2}{3}$ we have $V'(\theta) > 0$ with global incentive compatibility satisfied if $\theta \geq 1 + \frac{c}{2}$.

This completes the proof of Proposition 2.■

Figure 4 illustrates the conditions for the same θ with high (dashed lines) and low (solid lines) heterogeneity in consumers' mode of usage. With low heterogeneity $\partial\Pi(s,p)/\partial p|_{p=0}$ is decreasing for $s \in (0, s_1)$ ((D.1) is satisfied). However, $\partial\Pi(s,p)/\partial s|_{p=0}$ is increasing for $s \in (0, s_2)$ ((D.2) is satisfied). Thus, with low heterogeneity, both constraints cannot be satisfied simultaneously. With increased heterogeneity the figure shows that $p = 0$ is optimal for $\theta = 1.2$.

D.3 Proof of Proposition 3

- *Market coverage with $p > 0$.*

The firm will serve a given consumer type θ if it is possible to satisfy the first order condition for $p \leq \theta$. A necessary condition is that

$$\lim_{p\to\theta}\Big[(p-c)x_p H - (2-\theta)(\omega_{\theta p} H + \omega_p H_\theta)\Big] \leq 0 \quad \Rightarrow \quad \theta \geq 1 + \tfrac{c}{2}.$$

Notice that this is independent of the level of s.

- *Can the constraint $s = 0$ be binding with $p > 0$?*

If profit increases for s close to zero when p approaches $1 + \frac{c}{2}$, setting $s = 0$ can never be an optimal choice.

$$\lim_{p\to 1+\frac{c}{2}}\left[\frac{\partial\Pi}{\partial s}\right] = h\big(\theta - 1 - \tfrac{c}{2}\big)^2\Big(\tfrac{3}{2} - \tfrac{1}{2}(2-\theta)\tfrac{h_\theta}{h}\Big).$$

Because $h_\theta \leq 0$ for low values of s, the limit value is positive for s close to 0. Hence, the constraint $s = 0$ will not bind if the per unit charge is positive.

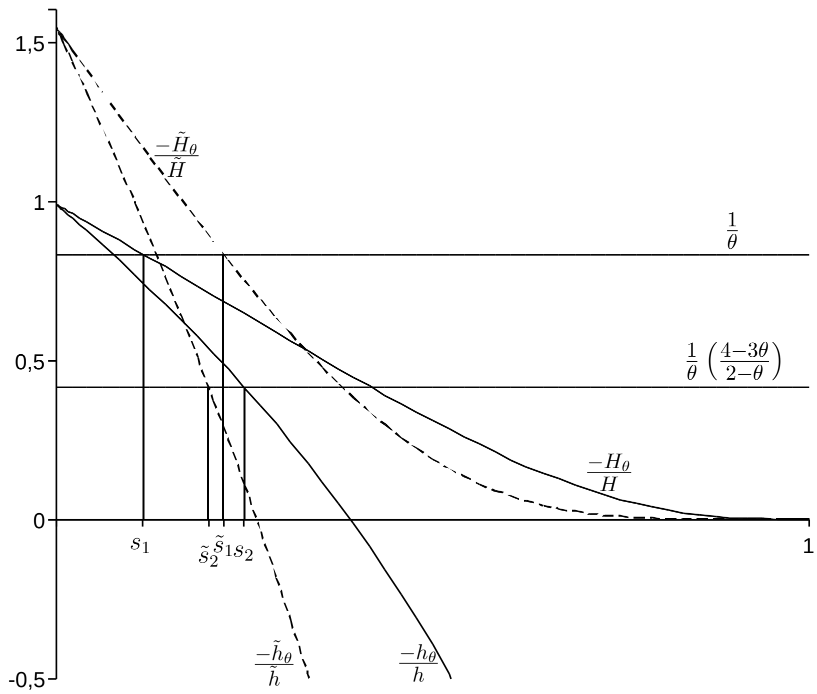


Figure 4: *The constraints in (D.1) and (D.2) given that $c = 0$ and $\theta = 1.2$, with high ($\widetilde{H}$) and low (H) heterogeneity in consumers' usage patterns.*

This proves part (i).

- *Market coverage with $p = 0$?*

An alternative strategy to exclude consumers in the low demand segment can be to set $s = 0$ in the case that the constraint $p = 0$ is binding (i.e., $s = 0$ is binding together with $p = 0$). This defines market coverage in the case that $p = 0$. If $s = 0$ together with $p = 0$ conditions (D.1) and (D.2) are both satisfied as s approaches zero, and we must have that

$$\frac{1}{\theta}\left(1 + \frac{c}{2-\theta}\right) \leq \lim_{s\to 0}\left[\frac{-h_\theta}{h}\right] \leq \frac{1}{\theta}\left(2 - \frac{\theta - 2c}{2-\theta}\right). \tag{D.6}$$

The largest possible θ must satisfy

$$1 + \frac{c}{2-\theta} \leq 2 - \frac{\theta - 2c}{2-\theta} \quad \Rightarrow \quad \theta \leq 1 + \frac{c}{2}. \tag{D.7}$$

We have already proved that profit is nonnegative given that $\theta > 1 + \frac{c}{2}$ if $p = 0$ is the optimal per unit charge. Hence, we can conclude that market coverage in this case is also given by $\theta \geq 1 + \frac{c}{2}$.

This proves part (ii), and the proof of Proposition 3 is compete.■

D.4 Proof of Proposition 4

- *Interior solution*

If the per unit charge is determined by $\frac{\partial \Pi}{\partial p} = 0$, this is given by

$$p - c = (1 - F)\Big(1 + \frac{xH_\theta}{H}\Big). \tag{D.8}$$

For $x > 0$ and $s > 0$ the left hand side will belong to the interval $[0, 1 - F]$ (the term $\Big(1 + \frac{xH_\theta}{H}\Big)$ can never exceed 1 under the assumption about first order stochastic dominance). Hence, $p - c \leq 1 - F$, which is the distortion a monopoly would apply if $s = 1$.

In an interior solution, for $\theta \in (\widetilde{\theta}, \theta')$, the per unit charge p^* and the usage restriction s^* is determined according to the two conditions

$$\frac{-H_\theta}{H} = \frac{1}{\theta - p}\Big(1 - \frac{p - c}{2 - \theta}\Big), \tag{D.9}$$

$$\frac{-h_\theta}{h} = \frac{1}{\theta - p}\Big(2 - \frac{\theta + p - 2c}{2 - \theta}\Big) = \frac{1}{\theta - p}\Big(1 - \frac{p - c}{2 - \theta} - \frac{2\theta - 2 - c}{2 - \theta}\Big). \tag{D.10}$$

Let (D.9) define a function $s_1(p;\theta)$ and (D.10) define a function $s_2(p;\theta)$. Holding θ fixed, we can determine ds/dp along the two conditions by differentiating (D.9) and (D.10) with respect to p and s. We find

$$\frac{ds_1}{dp} = \frac{1}{\big[\frac{H_\theta}{H}\big]'_s}\left[\frac{2\theta - 2 - c}{(2 - \theta)(\theta - p)^2}\right] \geq 0, \qquad \frac{ds_2}{dp} = \frac{1}{\big[\frac{h_\theta}{h}\big]'_s}\left[\frac{2(2\theta - 2 - c)}{(2 - \theta)(\theta - p)^2}\right] \geq 0.$$

Next, for $\theta \geq \widetilde{\theta}$ we have that $s_2 \geq s_1$ as p approaches zero. On the other hand, if s approaches 1 condition (D.9) are satisfied if $p = c + 2 - \theta$, whereas (D.10) only can be satisfied if p is infinitively positive. Given the slopes above $s_1(\theta)$ and $s_2(\theta)$ crosses exactly once for $p \in [0, c + 2 - \theta]$.

Since it is clear that we only have one stationary point it is sufficient to check that the second order conditions for profit maximization is satisfied close to the optimum. Under the assumption that we can cancel all first order conditions, we can write the second order conditions as

$$\frac{\partial^2 \Pi}{\partial p^2} = -H\Big(1 + (2 - \theta)\frac{H_\theta}{H}\Big) \leq 0, \qquad \frac{\partial^2 \Pi}{\partial s^2} = -\tfrac{1}{2}x^2 h(2 - \theta)\Big[\frac{h_\theta}{h}\Big]'_s \leq 0.$$

Given (D.9), a sufficient condition for profit to be concave in p is that $1/(2-\theta) \geq 1/(\theta-p)[1-(p-c)/(2-\theta)]$. This is met whenever $\theta \geq 1+c/2$. Profit is concave in s everywhere since all terms are nonnegative. The last condition is that

$$\frac{\partial^2 \Pi}{\partial p^2}\frac{\partial^2 \Pi}{\partial s^2} - \Big(\frac{\partial^2 \Pi}{\partial s \partial p}\Big)^2$$
$$= \tfrac{1}{2}x^2 h\Big\{H(2-\theta)\Big(1+(2-\theta)\tfrac{H_\theta}{H}\Big)\Big[\tfrac{h_\theta}{h}\Big]'_s - \tfrac{1}{2}h\Big(1+(2-\theta)\tfrac{h_\theta}{h}\Big)^2\Big\} \geq 0.$$

Given (D.9) and (D.10) this can be simplified further

$$\tfrac{1}{2}h^2(\theta-p)(2\theta-2-c)\Big(\tfrac{H}{h}(2-\theta)\big[\tfrac{h_\theta}{h}\big]'_s - 2\Big(\frac{2\theta-2-c}{\theta-p}\Big)\Big) \geq 0.$$

By inspection, it is easy to confirm that this is positive as long as s is above some threshold $\bar{s}(\theta)$, where $\bar{s}(\theta)$ is increasing in θ.[18] $\bar{s}$ approaches 1 when θ approaches 2, and $s(1) > 0$. Note that both terms inside the bracket parenthesis are nonnegative for $1+\frac{c}{2} < \theta < 2$, and $0 < s < 1$, it is 0 for $s = 0$ and ∞ for $s = 1$. The first term is increasing in s, while the second term is constant. The sign of the determinant of the Hessian cannot be determined in general. However, we have not been able to construct a numerical example where it is not positive at (p^*, s^*).

Next, in order to rule out countervailing incentives, the information rent must be increasing in θ. $V'_\theta \geq 0$ if $xH\big(1+\frac{1}{2}x\frac{H_\theta}{H}\big) \geq 0$. This implies that $\frac{-H_\theta}{H} \leq \frac{2}{\theta-p}$. When (D.9) is binding it is sufficient that $\frac{2}{\theta-p} \geq \frac{1}{\theta-p}\big(1-\frac{p-c}{2-\theta}\big)$. This is satisfied if $p \geq c-(2-\theta)$. When p is determined by (D.9) this is true for $\theta \in [1+\frac{c}{2}, 2]$.[19]

Finally, the second order conditions for global incentive compatibility must be satisfied. Differentiating (D.9) and (D.10) yields

$$\frac{ds^*}{d\theta} = \frac{\big[\frac{h_\theta}{h}\big]'_\theta - 2\big[\frac{H_\theta}{H}\big]'_\theta - \frac{1}{(2-\theta)^2}}{2\big[\frac{H_\theta}{H}\big]'_s - \big[\frac{h_\theta}{h}\big]'_s} \geq 0,$$

[18] By differentiating the condition $\frac{1}{2}h^2(\theta-p)(2\theta-2-c)\Big(\frac{H}{h}(2-\theta)\big[\frac{h_\theta}{h}\big]'_s - 2\Big(\frac{2\theta-2-c}{\theta-p}\Big)\Big) = 0$ we find that $ds/d\theta > 0$.

[19] Solving the first order condition for p we find that $p = \frac{c+(2-\theta)\left(1+\theta\frac{H_\theta}{H}\right)}{1+\left(2-\theta\right)\frac{H_\theta}{H}}$.

$$\frac{dp^*}{d\theta} = \frac{1}{\left(2\left[\frac{H_\theta}{H}\right]'_s - \left[\frac{h_\theta}{h}\right]'_s\right)(2-\theta)(2\theta-2-c)} \times$$
$$\left\{ \left(2\left[\frac{H_\theta}{H}\right]'_s - \left[\frac{h_\theta}{h}\right]'_s\right)\left((p-2\theta+2)(p-c)-(2-\theta)^2\right) \right.$$
$$\left. -(2-\theta)^2(\theta-p)^2\left(\left[\frac{H_\theta}{H}\right]'_\theta\left[\frac{h_\theta}{h}\right]'_s - \left[\frac{H_\theta}{H}\right]'_s\left[\frac{h_\theta}{h}\right]'_\theta\right) - \left[\frac{H_\theta}{H}\right]'_s(\theta-p)^2 \right\} \gtreqless 0.$$

To prove that the first sign above is correct, let us differentiate the first order conditions in (D.9) and (D.10), holding s fixed. This gives

$$\frac{dp_1}{d\theta} = \left[\frac{H_\theta}{H}\right]'_\theta\left[\frac{(2-\theta)(\theta-p)^2}{2\theta-2-c}\right] \leq 0, \quad \frac{dp_2}{d\theta} = \left[\frac{h_\theta}{h}\right]'_\theta\left[\frac{(2-\theta)(\theta-p)^2}{2(2\theta-2-c)}\right] \leq 0.$$

For a solution to exist it must be the case that $s_1(p)$ crosses $s_2(p)$ from below, and that $p_1(\theta)$ crosses $p_2(\theta)$ from below. Hence $\frac{ds_1}{dp} \geq \frac{ds_2}{dp}$ and $\frac{dp_1}{d\theta} \geq \frac{dp_2}{d\theta}$. These two conditions are met if

$$2\left[\frac{H_\theta}{H}\right]'_s - \left[\frac{h_\theta}{h}\right]'_s \leq 0, \qquad \left[\frac{h_\theta}{h}\right]'_\theta - 2\left[\frac{H_\theta}{H}\right]'_\theta \leq 0.$$

This proves that s^* is increasing in θ. The term $V_{\theta s}$ is positive if $p > 2c-\theta$, which is always the case when $c < \frac{2}{3}$. Hence, the second term in the sufficient condition for global incentive compatibility is positive.

If the first term is positive as well, the tariffs are implementable. The term $V_{\theta p} > 0$ if $p > c$ ($V_{\theta p} < 0$ if $p < c$). If p is close to c the term is close to zero and can be ignored. Because p can be both above and below c and because the sign of $p'(\theta)$ cannot be determined in general, it is more ambiguous whether the first term is positive. However, as θ increases above $\widetilde{\theta}$, the derivative can change sign at most once, and then from positive to negative, $p'(\theta) < 0$ for θ very close to 2. If $p < c$ it must be the case that $p(\theta)$ is increasing, and the first term is positive. On the other hand, since $V_{\theta s}\frac{ds^*}{d\theta}$ is strictly positive, a solution where $V_{\theta p}\frac{dp^*}{d\theta} < 0$ can be implementable. We have not been able to solve a numerical example with the given distributions where the second order condition is not satisfied.

This completes the proof of Proposition 4.■

D.5 Proof of Proposition 5 and 6

Proposition 5 can be verified by going through the proofs for Propositions 1 to 4, replacing the first order conditions for the monopoly problem with the appropriate first order conditions for welfare maximization. The proofs are omitted here. For convenience we state the conditions that must be met in the case that $p = 0$ is binding, and the first order conditions in an interior solution.

For $p = 0$ to be a binding condition we must have that

$$\frac{-H_\theta}{H} \geq \frac{1}{\theta}\Big(1 + \Big(\frac{1+\lambda}{\lambda}\Big)\frac{c}{2-\theta}\Big), \tag{D.11}$$

$$\frac{-h_\theta}{h} \leq \frac{1}{\theta}\Big(2 - \Big(\frac{1+\lambda}{\lambda}\Big)\frac{\theta-2c}{2-\theta}\Big). \tag{D.12}$$

For $\lambda > 0$ the left hand side in (D.11) is shifted upwards relative to equation (D.1), while the left hand side in (D.12) is shifted downwards relative to equation (D.2). It is easy to confirm that both constraints above cannot be satisfied for $\theta = \widetilde{\theta}$ (as defined in Proposition 2), except for when $\lambda \to \infty$. Hence, $p_w > 0$ for $\theta = \widetilde{\theta}$. Let $\widetilde{\theta}_W$ be defined the same way as $\widetilde{\theta}$. Then it is easy to confirm that $\widetilde{\theta}_W \lll \widetilde{\theta}$, and that $\widetilde{\theta}_W \to \widetilde{\theta}$ when $\lambda \to \infty$. In fact, λ must be large for $p = 0$ to bind in any interval at all, so that $\widetilde{\theta}_W < 1$ in all relevant cases.

In an interior solution, we must have that

$$\frac{-H_\theta}{H} = \frac{1}{\theta-p}\Big(1 - \Big(\frac{1+\lambda}{\lambda}\Big)\frac{p-c}{2-\theta}\Big), \tag{D.13}$$

$$\frac{-h_\theta}{h} = \frac{1}{\theta-p}\Big(2 - \Big(\frac{1+\lambda}{\lambda}\Big)\frac{\theta+p-2c}{2-\theta}\Big). \tag{D.14}$$

It is easy to confirm that both conditions cannot be satisfied at (p^*, s^*) solving profit maximization. Consumption efficiency increases if $Q(p(\theta), s(\theta), \theta) > Q(c + 2 - \theta, 1, \theta)$. A necessary condition for this is that

$$H(s(\theta)|\theta) > \frac{2\theta - c - 2}{\theta - p(\theta)}.$$

If θ is close to $1 + \frac{c}{2}$ it is sufficient that $s(\theta) > 0$. The condition will also hold if s is close to 1 since $p < c + 2 - \theta$. In the case that $p = 0$ binds, the condition defines a lower bound for $s(\theta)$, $\theta \in [1 + \frac{c}{2}, \widetilde{\theta}\rangle$. Otherwise, the condition states a restriction on $(p(\theta), s(\theta))$ for $\theta \in [\widetilde{\theta}, 2\rangle$. Numerical simulation confirms that the condition is satisfied everywhere.

References

Armstrong, M. (1999). "Price discrimination by a many-product firm." *Review of Economic Studies*, **66**(1), 151–168.

Armstrong, M. and J. Vickers (2001). "Competitive price discrimination." *RAND Journal of Economics*, **32**, 579–605.

Bakos, Y. and E. Brynjolfsson (1999). "Bundling information goods: Pricing, profits and efficiency." *Management Science*, **45**(12), 1613–1630.

Bousquet, A. and M. Ivaldi (1997). "Optimal pricing of telephone usage: An econometric implementation." *Information Economics and Policy*, **9**, 219–239.

Caillaud, B., R. Guesnerie, P. Rey and J. Tirole (1988). "Government intervention in production and incentives theory: A review of recent contributions." *RAND Journal of Economics*, **19**(1), 1–26.

Deneckere, R. J. and R. P. McAfee (1996). "Damaged goods." *Journal of Economics and Management Strategy*, **5**(2), 149–174.

Faulhaber, G. R. and J. C. Panzar (1977). "Optimal two-part tariffs with self-selection." Bell Laboratories Economic Discussion Paper No. 74.

García, D. (2005). "Monotonicity in direct revelation mechanisms." *Economic Letters*, **88**(1), 21–26.

Goldman, M. B., H. E. Leland and D. S. Sibley (1984). "Optimal nonuniform prices." *Review of Economic Studies*, **51**, 304–319.

Jensen, S. (2001). "Two-part tariffs with partial unbundling." Discussion paper 19/2001. Department of Economics, Norwegian School of Economics and Business Administration, Norway.

Jullien, B. (2000). "Participation constraints in adverse selection models." *Journal of Economic Theory*, **93**, 1–47.

Laffont, J.-J., P. Rey and J. Tirole (1998). "Network competition: II. Price discrimination." *RAND Journal of Economics*, **29**(1), 38–56.

Lewis, T. and D. Sappington (1989). "Countervailing incentives in agency problems." *Journal of Economic Theory*, **49**, 294–313.

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." *RAND Journal of Economics*, **15**(2), 171–196.

Matthews, S. and J. Moore (1987). "Monopoly provision of quality and warranties: An exploration in the theory of multidimensional screening." *Econometrica*, **55**(2), 441–467.

Miravete, E. J. (2001). "Screening through bundling." CARESS Discussion Paper No. 01-01.

Mirman, L. M. and D. Sibley (1980). "Optimal nonlinear prices for multiproduct monopolies." *The Bell Journal of Economics*, **11**, 659–670.

Mussa, M. and S. Rosen (1978). "Monopoly and product quality." *Journal of Economic Theory*, **18**, 301–317.

Oi, W. Y. (1971). "A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly." *Quarterly Journal of Economics*, **85**, 77–96.

Oren, S. S., S. A. Smith and R. B. Wilson (1982). "Nonlinear pricing in markets with interdependent demand." *Marketing Science*, **1**(3), 287–313.

Rochet, J.-C. and L. A. Stole (2002). "Nonlinear pricing with random participation constraints." *Review of Economic Studies*, **69**(1), 277–311.

Rochet, J.-C. and L. A. Stole (2003). "The economics of multidimensional screening." In M. Dewatripont, L. P. Hansen and S. J. Turnovsky, eds., *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*. Cambridge University Press.

Sappington, D. (1983). "Optimal regulation of a multiproduct monopoly with unknown technological capabilities." *Bell Journal of Economics*, **14**, 453–463.

Sharkey, W. W. and D. S. Sibley (1993). "Optimal nonlinear pricing with regulatory preference over customer type." *Journal of Public Economics*, **50**(2), 197–229.

Shi, M. (2003). "Social network-based discriminatory pricing strategy." *Marketing Letters*, **14**(4), 239–256.

Sibley, D. S. and P. Srinagesh (1997). "Multiproduct nonlinear pricing with multiple taste characteristics." *RAND Journal of Economics*, **28**(4), 684–707.

Stole, L. A. (2005). "Price Discrimination in Competitive Environments." In M. Armstrong and R. H. Porter, eds., *Handbook of Industrial Organization. Vol 3.* North-Holland.

Wang, R. and Q. Wen (1998). "Strategic invasion in markets with switching costs." *Journal of Economics and Management Strategy*, **7**(4), 521–549.

Wilson, R. (1993). *Nonlinear Pricing.* Oxford University Press, Oxford.