

# Selling Service Plans to Differentially Informed Customers\*

Roman Inderst<sup>†</sup>

University of Frankfurt and Imperial College London

Martin Peitz<sup>‡</sup>

University of Mannheim

November 2011

## Abstract

We analyze a model of monopolistic price discrimination where only some customers are originally informed about their preferences. We find that uninformed customers benefit from the presence of informed customers, even though information helps to predict only a customer's own demand. In stark contrast to the implications of models with only uninformed customers, we find that when some customers are already informed, reducing the costs of information acquisition can increase welfare. As this is not in the interest of the monopolistic seller, we thus provide conditions for when policy intervention may be warranted.

**Keywords:** Nonlinear pricing; price discrimination; information acquisition.

---

\*Martin Peitz gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft (SFB TR 15). We thank the Editor Yongmin Chen, two anonymous reviewers, Benjamin Hermalin, David Myatt, and various seminar audiences for helpful comments.

<sup>†</sup>Johann Wolfgang Goethe-University, IMFS, Mertonstrasse 17, 60054 Frankfurt am Main, Germany. E-mail: [inderst@finance.uni-frankfurt.de](mailto:inderst@finance.uni-frankfurt.de).

<sup>‡</sup>Department of Economics, University of Mannheim, D-68131 Mannheim, Germany. Email: [martin.peitz@googlegmail.com](mailto:martin.peitz@googlegmail.com). Also affiliated with CEPR, CESifo, ENCORE, and ZEW.

# 1 Introduction

We consider subscribers' choice between different service plans when at least some of them do not yet know their future demand or can only learn it at a cost. Customers may be choosing between different fixed or mobile telephone call plans or between contracts for the supply of electricity or other utilities. At the time that customers choose a particular contract, some may be well informed about their future usage. For other customers, future demand may be less predictable; for example, they may have little recollection of their past usage of the service, or they may be subscribing for the first time.

Our first step in the analysis is to solve for the optimal contracts offered by a monopolistic provider. Here, it is again useful to consider two benchmarks. If all customers were *ex-ante* uninformed about their future demand, a monopolist's optimal menu would specify first-best consumption levels and would allow the firm to extract all consumer rent. As a second benchmark, if initially all customers knew their demand type, the consumption level of customers with low demand would be distorted downwards, provided that low-type customers were served at all—this is the standard screening benchmark. When both informed and uninformed customers are present, as in our model, we find that contracts for all low-demand customers are (weakly) more distorted compared to these two benchmark cases: Contracts for uninformed low-demand customers are distorted compared to the efficient outcome where all customers are uninformed; and contracts for informed low-demand customers are more distorted compared to the already distorted outcome in the standard screening benchmark.

The fact that some customers are perfectly informed about their consumption, while others are not, gives rise to different contractual forms that are signed in equilibrium. Uninformed customers still face uncertainty about their future demand, and the respective tariff is more flexible in this sense: Precisely, in our model with only two demand types it allows for two possible quantity realizations. Such flexibility is, in our model, only valued by those customers who a priori have insufficient information about their future consumption.<sup>1</sup> Our simple setup with only two (low- and high-valuation) demand types, allows us to undertake various comparative statics analyses. In particular, we find that uninformed customers benefit from the presence of informed customers, even though information helps to predict only a customer's own demand.<sup>2</sup> This has interesting policy implications.

Policies intended to assist customers in making better-informed decisions may benefit *all* customers, including those who stay uninformed. Note again that in our model, such information does not increase transparency in a competitive market place, but it relates only to a consumer's own demand schedule. With the deregulation of many utilities, including fixed-line telephony, electricity, or gas, public agencies have set up internet services to assist households with their decision making. For instance, they may provide "calcula-

---

<sup>1</sup>For evidence that firms offer a variety of subscription plans, offering different degrees of flexibility, see, for instance, Lambrecht et al. (2007) or Narayanan et al. (2007).

<sup>2</sup>The fact that some customers are perfectly informed about their consumption allows us also to distinguish between "simple tariffs" and "flexible tariffs," where the latter allow customers to adjust the consumed quantity. As evidence from the marketing literature shows (e.g., Lambrecht et al. 2007; Narayanan et al. 2007), for different subscription services, firms' range of offers seems to indeed take into account that only some customers sufficiently value such flexibility.

tors” that force households to key in an expected demand profile and, thereby, calculate their expected bill for a given tariff.<sup>3</sup> Recently, such a policy of assisting consumers in predicting their own demand (type) has also been prominently advocated for by Thaler and Sunstein (2008). However, we show as well that, at least in our context of a model of monopolistic screening, more information and, thereby, a lower fraction of uninformed customers may lead to lower welfare, though it benefits consumers.

Our main analysis proceeds in a model where consumers’ information is exogenous, though it can be influenced by policy. Uninformed customers may, however, be in a position to learn about their future demand (type) already at the contracting stage, albeit only after incurring costs. In an extension of our analysis we solve such a model of endogenous information acquisition.

Uninformed customers’ costs of information acquisition could include the time and effort to go through past bills or to think about future consumption needs. If these costs are low, this additionally constrains the firm’s offers. Intuitively, as costs become smaller, contracts designed for informed and uninformed customers become more similar. Our analysis with the possibility of endogenous information acquisition follows the seminal paper by Crémer and Khalil (1992),<sup>4</sup> though there all consumers are ex-ante homogeneous. As we show, the simultaneous presence of both informed and uninformed customers leads to strikingly different welfare results. To see this, suppose first that there were only uninformed customers, as in Crémer and Khalil (1992). Then, welfare is strictly lower as costs of information acquisition decrease. While results are generally ambiguous in our model, where consumers differ in their information, for a (standard) linear-quadratic functional specification, which allows us to obtain explicit solutions, we show that the opposite holds: Welfare is higher as costs of information acquisition decrease. Hence, even though the (potentially "subsidized") option of information acquisition may not be used in equilibrium, the constraint that it imposes on the firm’s price discriminating offer may enhance welfare.

Taken together, our model thus reveals two channels through which public policy that aims to increase consumer knowledge can enhance consumer surplus or welfare. If such a policy turns uninformed into informed customers, then this increases the "information rent" that these customers can realize (in expectation) and, as we show, it can also raise the utility of those customers who are still uninformed. In addition, even if (some) customers do not become informed but could do so at lower cost after policy intervention, this can make contracts more efficient.

This paper contributes to the growing literature on sequential screening models. Perhaps the most distinctive feature of our setting compared to recent approaches such as Courty and Li (2000) or Miravete (2005) is the specification of three different ex-ante

---

<sup>3</sup>In addition, these websites often offer price comparison services as well.

<sup>4</sup>See, also, Matthews and Persico (2007), where, through its refund policy, the firm can affect customers’ incentives to ex-ante acquire information. Also compare, more generally, the literature on mechanism design with endogenous information acquisition, as surveyed in Bergemann and Välimäki (2006). Finally, there is also a strand of the literature in which the *principal* (i.e., the firm in our model) has information or can, at a cost, acquire information about the characteristics of the good and must decide whether to share this with the agents (i.e., the consumers in our model). See, in particular, Lewis and Sappington (1994) and Johnson and Myatt (2006).

types of customers: one group of uninformed customers (though they may become informed at a cost) and two informed groups of customers, one with high and the other with low valuation. This allows us, in particular, to conduct a comparative analysis of the fraction of informed customers.<sup>5</sup>

While in Lewis and Sappington (1997) there are also both informed and uninformed agents, the focus there is on how to get the informed agent to exert (more) effort into acquiring information that is valuable to the principal. Somewhat more closely related, in Dai, Lewis, and Lopomo (2006), agents differ initially in the precision with which they can forecast their costs of production. In our setting, however, the better information that some agents have *ex-ante* also creates *ex-ante* heterogeneity in a second dimension: low- and high-demand types. Consequently, in our model, offers to both informed and uninformed agents will be distorted, while in their model, only the menu offered to the less knowledgeable agent is inefficient. The novel focus of the paper is on the comparative statics properties of the model with respect to the share of informed customers.

Finally, there is also a small but growing literature that combines demand uncertainty with non-standard utilities incorporating overconfidence, procrastination, or projection bias. In particular, in Grubb (2009) customers systematically underestimate the variability of their future demand.<sup>6</sup> His main objective is to explain the three-part tariff structure that seems to prevail, in particular, in some contracts for mobile telephone network usage. In contrast, the *ex-ante* heterogeneity of consumers in our model implies that informed and uninformed users subscribe to different tariffs, where essentially only the latter consumers' tariff must account for (additional) flexibility in their unpredicted usage. In this respect, our model is admittedly too stylized to account for many features of real-world contracts, though this comes at the advantage of generating a tractable analysis with rational consumers that allows for an evaluation of information provision policies.

The rest of this paper is organized as follows. In Section 2, we set up the model. Section 3 contains the analysis with informed and uninformed customers who may have low or high demand. Section 4 provides some results on comparative statics, while Section 5 extends the analysis by allowing uninformed customers to acquire information, albeit at a cost, before choosing from the offered contracts. Section 6 concludes. All proofs are relegated to the Appendix.

---

<sup>5</sup>Though Courty and Li (2000) consider only single-unit purchases, Rochet and Stole (2003) have shown how their results can be re-interpreted with multi-unit purchases, as well as with decreasing marginal values. Earlier papers include Baron and Besanko (1984), Riordan and Sappington (1987), and Miravete (1996, 2005). Miravete (1996) is of particular interest, as this paper also considers non-linear pricing: Consumers have *ex-ante* knowledge about their demand function, which together with some additional "shock" generates their willingness to pay at the time of consumption. Also Clay et al. (1992) introduce *ex-ante* uncertainty into a model of price discrimination, though their analysis is restricted to a comparison of two particular tariff structures.

<sup>6</sup>In Eliaz and Spiegel (2008) and Uthemann (2005), customers have biased priors about having low or high demand later.

## 2 The Model

Consider a monopolistic firm offering a long-term service contract to customers. Though our model applies to many different settings, as discussed in the Introduction, it may be convenient in what follows to have in mind an application to mobile telephone call plans.

The firm has constant marginal cost  $\tilde{c}$ . A customer of (real-valued) demand type  $\theta$ , which can be low or high with  $0 < \theta_l < \theta_h$ , derives gross utility  $\theta\tilde{u}(q)$  from consuming  $q$  "units" (e.g., minutes) of the particular service. Here, the continuously differentiable function  $\tilde{u}(q)$  is assumed to be strictly increasing and concave with  $\tilde{u}(0) = 0$ . It is convenient to additionally invoke the (standard) boundary conditions  $\lim_{q \downarrow 0} \tilde{u}'(q) = \infty$  and  $\lim_{q \rightarrow \infty} \tilde{u}'(q) = 0$ , which together imply that the first-best level of service will be both finite and strictly positive for any choice  $\theta > 0$  and  $\tilde{c}$ . We also suppose that  $\tilde{u}$  is twice continuously differentiable.

Our choice of specification derives from the main motivating examples in the Introduction, where the assumption of constant marginal cost is probably realistic. Before describing the model, however, it is useful to rephrase the customer's choice problem. Instead of choosing quantity  $q$ , we suppose that the customer selects a certain level of gross "base utility"  $u = \tilde{u}(q)$ . Since  $\tilde{u}$  can be inverted, we define  $C(u) := \tilde{c}\tilde{u}^{-1}(u) = \tilde{c}q$ . That is, to generate customer utility of  $\theta u$ , the firm must incur the cost  $C(u)$ . The invoked properties of  $\tilde{u}(q)$  imply that  $C(u)$  is strictly increasing and strictly convex with  $C'(0) = \lim_{u \downarrow 0} C'(u) = 0$ . Denote total surplus by  $s(u; \theta) := \theta u - C(u)$ , which for  $\theta_i$  is uniquely maximized by some bounded and strictly positive value  $u_i^{FB}$ ,  $i = h, l$ . Note that  $0 < u_l^{FB} < u_h^{FB}$ .

Suppose that there is mass one of customers. The *ex-ante* probability with which an individual customer ultimately has high demand is given by  $\mu \in (0, 1)$ . However, only the fraction  $\pi$  of customers initially know their type. Later, at the stage of consumption, all customers become informed about their type. (In Section 5, an uninformed customer may also learn his type early, albeit only at a cost.) A customer's knowledge is his private information.

Without loss of generality, we can restrict consideration to the following set of offers. For *ex-Ante* informed customers, the firm designates, at most, two different consumption profiles  $u_{A,i}$  and respective transfers  $t_{A,i}$ , where  $i = l, h$ . For only *ex-Post* informed customers, the firm specifies, instead, a menu of two options:  $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$ . Each customer decides which, if any, contract to sign. Note that while contracts  $(u_{A,i}, t_{A,i})$  specify a fixed allowance, the menu  $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$  still allows for flexibility: The customer pays  $t_{P,l}$  for an allowance up to  $u_{P,l}$ , while if she wants to consume more, she can purchase the additional allowance  $u_{P,h} - u_{P,l}$  at an incremental price of  $t_{P,h} - t_{P,l}$ .<sup>7</sup>

To summarize the contractual game, we thus have the following sequence of moves. Initially, at some period  $\tau = 1$ , nature chooses the type of each customer, but only a fraction  $\pi$  is informed about this. At  $\tau = 2$ , the firm offers a menu of three tariffs, comprising  $(u_{A,i}, t_{A,i})$  with  $i = l$  and  $i = h$  and  $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$ . A customer can then choose, in  $\tau = 3$ , the preferred tariff or not to sign a contract at all. When a contract was signed, in  $\tau = 4$  also the previously uninformed customers perfectly learn their demand

---

<sup>7</sup>This contract is, thus, similar to a "three-part tariff" contract.

type. Finally, in  $\tau = 5$  all customers who have signed a contract realize the respective quantities and make the respective payments.

### 3 Analysis of the Optimal Contract

#### 3.1 Benchmarks

**Only Informed Customers.** With only informed customers, the firm would face a standard screening problem to choose contracts  $(u_{A,i}, t_{A,i})$ . With net utility levels  $V_{A,i} := \theta_i u_{A,i} - t_{A,i}$ , the incentive constraint of the high-demand type,  $IC_{A,h}$ , becomes  $V_{A,h} \geq \theta_h u_{A,l} - t_{A,l}$ ; the individual rationality constraint of the low-demand type,  $IR_{A,l}$ , becomes  $V_{A,l} \geq 0$ . It is well known that both constraints bind at the optimal offer. Moreover, all other constraints can be ignored, while the high type consumes the first-best service level,  $u_{A,h} = u_h^{FB}$ . High-demand customers realize a rent equal to  $u_{A,l}(\theta_h - \theta_l)$ , where the optimal service level for low-demand customers,  $u_{A,l} = u_l^S$ , solves

$$s'(u_l^S; \theta_l) = \frac{\mu}{1 - \mu}(\theta_h - \theta_l) \quad (1)$$

whenever this is positive, while otherwise  $u_l^S = 0$ . Substituting  $C'(0) = 0$  such that  $s'(0; \theta_l) = \theta_l$ , we have from (1) that  $u_l^S > 0$  holds strictly if and only if  $\mu < \theta_l/\theta_h$ .

**Only Uninformed Customers.** As a second benchmark, suppose that all customers are uninformed ( $\pi = 0$ ). In this case, customers' individual rationality constraint need only be satisfied in expectation:  $IR_P$  with  $\mu V_{P,h} + (1 - \mu)V_{P,l} \geq 0$ , where  $V_{P,i} := \theta_i u_{P,i} - t_{P,i}$ . By optimality for the firm,  $IR_P$  binds and both consumption profiles are efficient,  $u_{P,i} = u_i^{FB}$ . Finally, with discrete types, there is some freedom in specifying the optimal transfers, which have to satisfy  $IR_P$  as well as both *ex-post* incentive-compatibility constraints:  $IC_{P,i}$  with  $V_{P,i} \geq \theta_i u_{P,j} - t_{P,j}$ . For instance, one possibility is to choose transfers that reflect incremental costs:  $t_{P,h} - t_{P,l} = C(u_h^{FB}) - C(u_l^{FB})$ .

#### 3.2 The Firm's Program with Informed and Uninformed Customers

**Constraints.** With both informed and uninformed customers present, the firm faces in addition to the constraints spelled out in the two benchmarks additional incentive-compatibility constraints *across* the respective offers, which we denoted by subscripts  $A$  (for the *ex-ante* informed customers) and  $P$  (for the only *ex-post* informed customers). We will show that we can ignore the possibility of informed low-demand customers mimicking uninformed customers, while we refer to the respective ("cross") constraint of informed high-demand customers,  $V_{A,h} \geq V_{P,h}$ , as  $ICC_{A,h}$ . Next, for an uninformed customer, who is supposed to pick the menu  $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$ , we suppose first that only the option to mimic the informed low-demand type may become attractive. We refer to this ("cross") constraint as  $ICC_P$ :

$$\mu V_{P,h} + (1 - \mu)V_{P,l} \geq \mu(\theta_h u_{A,l} - t_{A,l}) + (1 - \mu)V_{A,l}.$$

As is usual, we will later show that the solution to the "relaxed program" satisfies all neglected constraints, including the constraint that the uninformed customer may want to mimic the informed high-demand customer.

**Firm's Program.** Summing up, with both informed and uninformed customers present, the firm faces the following (relaxed) program. The firm chooses contracts to maximize expected profits

$$\begin{aligned} & \pi \{ \mu [t_{A,h} - C(u_{A,h})] + (1 - \mu) [t_{A,l} - C(u_{A,l})] \} \\ & + (1 - \pi) \{ \mu [t_{P,h} - C(u_{P,h})] + (1 - \mu) [t_{P,l} - C(u_{P,l})] \} \end{aligned}$$

subject to the following set of constraints: (i) The downward incentive-compatibility constraints for both informed and uninformed customers  $IC_{A,h}$  and  $IC_{P,h}$  (as introduced in the benchmarks of Section 3.1); (ii) the individual rationality constraints for the informed low-type customer  $IR_{A,l}$  and the uninformed customer  $IR_P$  (as also introduced in Section 3.1); and (iii) the two "cross" incentive-compatibility constraints, namely for the uninformed customer  $ICC_P$  and the informed high-type customer  $ICC_{A,h}$ . In addition, note that all  $u$  must be non-negative.<sup>8</sup>

**Solution Procedure.** We characterize the optimal contract in several steps. We first solve the firm's program under the assumption that all customers purchase a positive level of services so that  $u_{,i} > 0$ . Here, we encounter two cases, which we refer to as Cases 1 and 2. Subsequently, we show that there are two more cases possible, Cases 3 and 4, in which not all customers are served. Finally, we derive conditions for when Cases 1-4 apply.

To characterize the outcome, we will have to establish which constraints bind (and when). Though in our model, consumers differ along two dimensions—i.e., whether they have high or low demand and whether they are initially informed or uninformed—from an *ex-ante* perspective, there are only three distinct types, which may be ordered in one dimension, namely through their prior beliefs of whether they have high or low valuation.<sup>9</sup> As we argue below, this allows us to reinterpret some of our results in the language of a more standard, one-dimensional screening problem. Still, the procedure that we adopt will be particularly helpful in extending our results to the case where uninformed customers can acquire information, albeit at a cost.

### 3.3 Preliminary Results

**No Exclusion.** Suppose first that all customers purchase a strictly positive level of services. We are particularly interested in the *allocation* of consumption levels  $u_{A,l}$ ,  $u_{A,h}$ ,

---

<sup>8</sup>To save space, we have chosen not to first write out the full program explicitly. Note, however, that in the relaxed program, the following constraints are ignored: the upward incentive-compatibility constraints; the individual rationality constraint for the high type; the constraint that an informed low type does not want to mimic an uninformed low type; the constraint that an uninformed consumer does not want to mimic the informed high type; the constraint that an informed low type does not want to mimic an uninformed high type; and the constraint that an informed high type does not want to mimic an uninformed low type.

<sup>9</sup>Instead, see for multi-dimensional screening problems, e.g. Armstrong and Rochet (1999).

$u_{P,l}, u_{P,h}$ . They are characterized as follows:

**Proposition 1** *Suppose that according to the optimal offer all customers purchase a positive level of services. This offer has the following properties:*

*Case 1) If  $\pi \geq \frac{1}{2-\mu}$ , informed and uninformed customers obtain the same allocations:  $u_{.,h} = u_h^{FB}$  and  $u_{.,l} = u_l^S$ .*

*Case 2) If  $\pi < \frac{1}{2-\mu}$ , only high-demand customers obtain the same allocation regardless of whether or not they are informed—i.e.,  $u_{P,h} = u_{A,h} = u_h^{FB}$ . The allocation of the informed low type is more distorted than that of the uninformed low type—i.e.,  $u_{A,l} < u_l^S < u_{P,l} < u_l^{FB}$ .*

Recall, for Case 2, that  $u_l^S$  denotes the distorted consumption level for low-demand types under a "standard screening" contract (cf. in Section 3.1, the case with  $\pi = 1$ ). The key for Proposition 1 are the two "cross" constraints. To see this, we first compare the characterization in Proposition 1 with the outcome of the two benchmark cases with only uninformed customers ( $\pi = 0$ ) or only informed customers ( $\pi = 1$ ). For  $\pi = 0$  the first-best allocation results, but in the presence of both informed and uninformed customers, the distortion of  $u_{P,l}$  follows from the constraint  $ICC_{A,h}$ , which requires that an informed high-demand customer does not want to mimic an uninformed customer. Compared to the case with  $\pi = 1$ , where there are only informed customers,  $u_{A,l}$  becomes even further distorted in Case 2 of Proposition 1, in which the two incentive constraints across informed and uninformed customers,  $ICC_{A,h}$  and  $ICC_P$ , bind.

Next, we provide more details. If the uninformed customer picks the informed low-type customer's contract, thus violating  $ICC_P$ , she realizes a rent of  $(\theta_h - \theta_l)u_{A,l}$  in case she ultimately turns out to have high demand. The level of  $u_{A,l}$  also determines the rent of an informed high-type customer. In Case 2, however, this does not follow from the ("standard") constraint  $IC_{A,h}$ , which remains slack, but instead more indirectly, as both  $ICC_{A,h}$  and  $ICC_P$ , bind.

Put differently, in this case it is particularly helpful to think of the three initial types of consumers as being ordered according to their ex-ante expected valuation, that is with the informed low-demand customer at the bottom and the informed high-demand customer at the top. In Case 2, all downwards incentive compatibility constraints are binding and the derived contracts are globally incentive compatible (cf. also the discussion below after Corollary 3). Each tariff, i.e.,  $(u_{A,i}, t_{A,i})$  for the informed customers  $i = l$  and  $i = h$  and the menu  $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$  for the uninformed customers are then derived from the same principle, namely to trade off the objective to minimize the information rent for all "higher-type" customers with the objective to maximize the respective surplus. For the "highest type", the informed  $\theta = h$  customers, this obtains efficiency, while there is inefficiency for all "lower types". (That being said, for the "middle type", namely the ex-ante uninformed customer, the efficiency shows up only when they turn out to have low demand, while  $u_{P,h} = u_{A,h} = u_h^{FB}$ .) In particular, for the type at the bottom, the informed low-demand customer, the proof of Proposition 1 shows that  $u_{A,l}$  solves

$$s'(u_{A,l}; \theta_l) = \frac{\mu}{1-\mu} \frac{1-\pi+\pi\mu}{\pi} (\theta_h - \theta_l). \quad (2)$$

Comparing this to (1) confirms  $u_{A,l} < u_l^S$  for Case 2, where  $\pi < \frac{1}{2-\mu}$ . The respective trade-off for the "middle type" is next as follows. Recall that it is through the binding constraint  $ICC_P$  that a higher level of  $u_{P,l}$  allows informed high-type customers to extract a higher rent when an uninformed customer realizes a higher utility. Taking this into account,  $u_{P,l}$  optimally trades off surplus maximization with rent extraction if

$$s'(u_{P,l}; \theta_l) = \mu \frac{\pi}{1-\pi} (\theta_h - \theta_l). \quad (3)$$

Comparing this to (1) confirms  $u_{P,l} > u_l^S$  for Case 2.

**Exclusion.** We suppose next that not all customers are served. Here, we have to distinguish between two cases: In Case 3, *all* low-type customers are excluded, whereas in Case 4, only those who are also informed are excluded.

**Proposition 2** *If not all customers are served under the optimal offer, all high-demand customers purchase the first-best level of services,  $u_{.,h} = u_h^{FB}$ . For low-demand customers, the following cases are possible:*

*Case 3) All low-demand customers are excluded.*

*Case 4) Only informed low-demand customers are excluded, while uninformed low-demand customers purchase an inefficiently low level of services,  $0 < u_{P,l} < u_l^{FB}$ .*

Of particular interest is Case 4 in Proposition 2. There, by no longer serving informed low-type customers, the firm can extract all consumer surplus from uninformed customers. The optimal choice of  $u_{P,l}$  for low-demand uninformed customer trades off surplus maximization with rent extraction from informed high-demand customers. This is the same trade-off as in Case 2 of Proposition 1, which is why  $u_{P,l}$  is again determined from the first-order condition (3).

### 3.4 Solution to the Firm's Problem

Which of the different cases in Propositions 1 and 2 applies depends on the fractions of the different types of customers.

**Proposition 3** *Which of the characterized four cases applies depends, as follows, on the fractions of the different customer types:*

*i) Suppose that the fraction of high-demand customers is low with  $\mu < \theta_l/\theta_h$ : In this case, uninformed low-type customers always purchase a positive quantity. If, for given  $\mu$ , the fraction of informed customers  $\pi$  is low, then informed low-demand customers are excluded (Case 4). Otherwise, all customers are served, with Case 1 applying for high values of  $\pi$  and Case 2 for intermediate values.*

*ii) Suppose, instead, that  $\mu \geq \theta_l/\theta_h$ : Then, for given  $\mu$ , all low-type customers are excluded if  $\pi$  is sufficiently high (Case 3). For lower values of  $\pi$ , however, only informed low-demand customers are excluded (Case 4).*

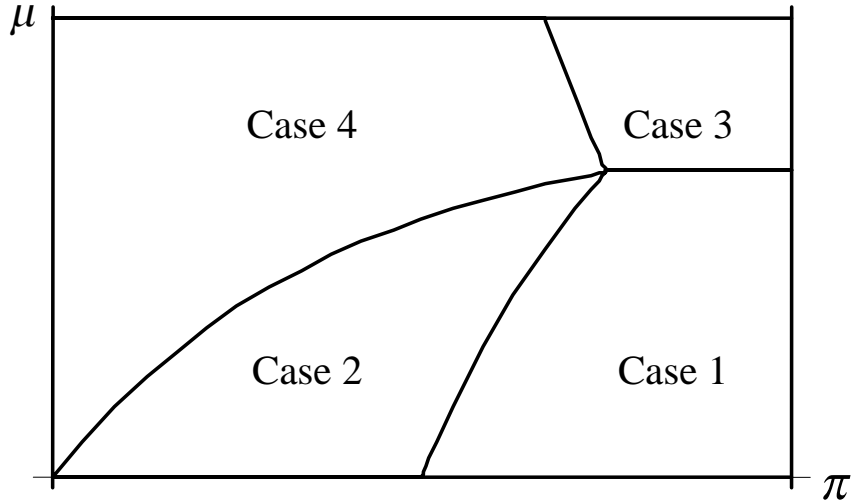


Figure 1: Optimal contracts for  $\theta_h = 3/4$  and  $\theta_l = 1/2$ .

We illustrate this in Figure 1 (which is drawn for the particular values  $\theta_h = 3/4$  and  $\theta_l = 1/2$ ). Furthermore, the respective thresholds for  $\mu$  and  $\pi$  that determine which of the four cases apply are given explicitly in the proof of Proposition 3. We now provide more intuition for the case distinction in Proposition 3.

The role of the fraction  $\mu$  of high-demand customers is intuitive and standard: As there are more customers with high demand, it becomes more likely that low-demand types are excluded so that the firm can extract more rents from high-demand types. Hence, when moving upwards in Figure 1, we move from Cases 1 and 2 to Cases 3 and 4, respectively. Next, if  $\mu$  is high but also  $\pi$  low, implying that there altogether few informed customers, it is intuitive that only informed low types, but not uninformed low types, are excluded. Hence, as we move to the left in the upper part of Figure 1 we move from Case 3 to Case 4, implying that fewer types are excluded. Interestingly, the opposite holds in the lower part of Figure 1—i.e., for relatively low values of  $\mu$ . There, with few high-type customers, as there are more uninformed customers (lower  $\pi$ ), the offer made to informed low-demand customers becomes increasingly distorted in an attempt to extract more rents from uninformed customers. Ultimately, as  $\pi$  becomes too low, informed low-demand customers no longer purchase a positive quantity—i.e., we move from Case 2 to Case 4. At this point, there is a discrete jump up in the number of excluded customers. As  $\pi$  decreases further, the number of excluded consumers declines.

For simplicity, Proposition 3 was phrased mainly in terms of high or low values of  $\pi$ . To see how the respective boundaries that separate the four cases change in both parameters  $\pi$  and  $\mu$ , it is again instructive to consult Figure 1: As is shown in the proof of Proposition 3, the respective monotonicity of the boundaries displayed in Figure 1 holds generally. Before further discussing the solution to the firm’s program in the following section, the rest of this section offers a more formal discussion of how the boundaries of the cases behave.

Recall first that in the standard screening problem, a horizontal line with  $\mu = \theta_l/\theta_h$  separates the case where all customers are served from that where only high-demand customers are served. This line separates Cases 1 and 3 in Figure 1. From Proposition 1, we have next that Cases 1 and 2 are separated by a function that we denote by  $\tilde{\pi}_{12} = 1/(2-\mu)$ . Applying a similar notation, we have that  $\tilde{\pi}_{24}$  separates Cases 2 and 4, while  $\tilde{\pi}_{34}$  separates Cases 3 and 4. Note that  $\tilde{\pi}_{24}$  is determined from the requirement that  $u_{A,l} = 0$  holds in Case 2, where  $u_{A,l}$  is strictly decreasing in  $\mu$  but strictly increasing in  $\pi$ . This implies that  $\tilde{\pi}_{24}$  is indeed upward-sloping as a function of  $\mu$ , as depicted in Figure 1. Finally, the boundary between Cases 3 and 4,  $\tilde{\pi}_{34}$ , is obtained from setting  $u_{P,l} = 0$  in Case 4. As  $u_{P,l}$  becomes more distorted as there are more informed customers and more high-type customers,  $\tilde{\pi}_{34}$  is strictly decreasing in  $\mu$ .

### 3.5 Further Discussion

Serving informed customers with low demand comes with high opportunity costs to the firm in terms of lost profits with *both* informed high-demand customers and uninformed customers. Optimally, the firm makes the corresponding “basic” contract  $(u_{A,l}, t_{A,l})$  relatively unattractive, in particular, if the fraction of uninformed customers or that of high-demand customers is relatively high.

**Corollary 1** *Suppose that Cases 2 or 4 apply. Then, as the fraction of uninformed customers or of high-demand customers increases (lower  $\pi$  or higher  $\mu$ , respectively), the more unattractive becomes the “basic” tariff, which is offered to informed low-demand customers. This results first in a lower level of  $u_{A,l}$  and, ultimately, in the exclusion of these customers (corresponding to  $u_{A,l} = 0$ ).*

Recall now our previous discussion of reinterpreting our results in terms of a more standard one-dimensional screening problem. From an *ex-ante* perspective, informed low-demand types then represent the “bottom type,” while uninformed customers and high-demand informed customers represent the respective “adjacent higher” types. As all “adjacent downwards” incentive-compatibility constraints bind in the presently considered cases, it is intuitive that the distortion “at the bottom” increases as the probabilities of the “higher types” increase (specifically, through an increase in  $\mu$  or  $\pi$ ).

For what follows, note further that an uninformed customer generates (weakly) higher revenues for the firm compared to an informed customer.

**Corollary 2** *Suppose that Cases 2 or 4 apply. Then, the firm always realizes strictly higher revenues from an uninformed customer than from an informed customer, both in expectation (over high- and low-demand types) and when considering only low-demand customers.*

From the customer’s side, it is immediate that an informed customer is better off (strictly for Cases 2 and 4). In particular, note that uninformed low-demand customers end up realizing strictly negative utility:  $V_{P,l} < 0$ .<sup>10</sup> Clearly, from an *ex-post* perspective,

---

<sup>10</sup>Cf. also Miravete (1996).

uninformed low-demand customers would have been better off purchasing the “basic” tariff  $(u_{A,l}, t_{A,l})$  instead. However, given their own initial demand uncertainty, the offer designed for uninformed customers was equally attractive, as it also contained the option to make use of an additional allowance  $u_h^{FB} - u_{P,l} > 0$  at an incremental price  $t_{P,h} - t_{P,l}$  that is smaller than the respective utility increment  $\theta_h(u_h^{FB} - u_{P,l})$ .

## 4 Comparative Analysis

Given the explicitly characterized solution to the firm’s program in Propositions 1, 2, and 3, the present model lends itself to comparative analysis. This section conducts such an analysis in terms of an increase in the share of informed customers  $\pi$ .

This analysis seems interesting for the following reasons. First, in light of the results from other models, which we review below, it is interesting to analyze how the presence of (more) informed customers affects the utility of those who are less informed (though they do not suffer from any other, exploitable behavioral biases). Second, as noted in the Introduction, public policy in some recently deregulated industries aims to encourage customers to become more knowledgeable, including about their own demand profile.<sup>11</sup> In the extreme, firms may be forced to apply the contract that turns out to be, ex post, the most favorable for any given customer—the introduction of such a policy corresponds to a discrete jump to  $\pi = 1$ . The comparative analysis in  $\pi$  may help to shed more light on the implications of such policies.

**Impact of  $\pi$  on Uninformed Customers’ Surplus.** From implicitly differentiating (3) for  $u_{P,l}$  and using  $u_{P,h} = u_h^{FB}$ , it follows that the expected service level of uninformed customers decreases as there are more informed customers (higher  $\pi$ ). Still, it turns out that uninformed customers are better off. This holds strictly for Case 2, where it follows immediately from the proof of Proposition 1. Precisely, this holds as the lowest service level  $u_{A,l}$  increases in  $\pi$ , which, through the binding (“cross”) constraint  $ICC_P$ , leads to a higher rent for uninformed customers. (For all other cases uninformed customers’ utility is constant in  $\pi$ .<sup>12</sup>)

**Corollary 3** *As the fraction of informed customers increases (higher  $\pi$ ), uninformed customers’ utility increases.*

Note that a customer’s information is only with respect to her *own* demand type. Still, as more of the uninformed customers become informed, this increases not only their own utility but also that of customers who still remain uninformed. The mechanism through

---

<sup>11</sup>In our monopolistic setting, we can abstract from other, more well-known implications of such policies, which serve to induce more effective competition by reducing search (and/or switching) costs (cf. the literature discussed below).

<sup>12</sup>Recall that Case 2 applies for intermediate values of  $\pi$ , provided that  $\mu$  is not too high. (Cf. also Figure 1.) For low values of  $\pi$ , where Case 4 applies instead, uninformed customers realize zero rent, while for high  $\pi$ , where Case 1 applies and uninformed customers realize the highest rent, we know that a further increase in  $\pi$  does not affect contracts and, thus, utilities. Finally, for high  $\mu$ , where only Cases 3 and 4 apply as  $\pi$  changes, uninformed customers always realize zero utility.

which the presence of informed customers benefits those who are uninformed differs from the mechanism that is at work in models with search and shopping costs, where the presence of customers who are better informed about rivals' offers brings down expected prices (cf. Varian 1980; or, more recently, Janssen and Moraga-González 2004).<sup>13</sup>

Formally, the result from Corollary 3 can be also restated in the following way. Recall that once the way incentive constraints in Case 2 bind has been worked out, uninformed customers represent the "intermediate type" in a standard screening model, with informed low- and high-demand types representing the "top" and "bottom types", respectively. The respective *ex-ante* probabilities of these "types" are  $(1 - \mu)\pi$  ("bottom"),  $1 - \pi$  ("intermediate"), and  $\mu\pi$  ("top"). As  $\pi$  increases, the rent of uninformed customers is affected in the following way through the chosen service level "at the bottom,"  $u_{A,l}$ . First, as the probability of the "bottom type" increases, the distortion "at the bottom" is optimally decreased, resulting in more rent for all "higher types." Second, as the probability of the "intermediate type" decreases, this reduces the firm's benefits from extracting rent, which further pushes up  $u_{A,l}$ . The third effect, however, goes in the opposite direction: As a higher  $\pi$  increases the probability of the "top type," this increases the benefits from rent extraction "at the top," which pushes down  $u_{A,l}$  (next to  $u_{P,l}$ ). Corollary 3 shows that the first two effects together dominate.

Corollary 3 may also be interesting in light of claims that more-informed or sophisticated customers are cross-subsidized at the expense of less-informed customers. For instance, in Gabaix and Laibson (2006), this holds, albeit under competition, if only some customers are knowledgeable about their future demand of an "add-on service," while other customers are "naively" unaware of this. In a monopolistic context and with perfectly rational customers, Corollary 3 provides a different benchmark, where the presence of informed customers benefits uninformed customers.<sup>14</sup>

**Impact of  $\pi$  on Informed Customers' Surplus.** The present comparative analysis still ignores the impact of a change in  $\pi$  on all (other) informed customers. This is generally ambiguous. To explore this, note first that in Cases 1 and 3,  $\pi$  has no effect on contracts and, thus, utilities. Next, in Case 4, in which only informed low-demand customers are excluded, the impact of a higher  $\pi$  on informed high-demand customers is unambiguous: As their fraction increases (higher  $\pi$ ), the firm optimally extracts more rent (at the cost of a lower surplus realized with uninformed customers). Finally, in the Case 2, all customers are served. Here, the utility of informed high-demand customers,  $V_{A,h}$ , depends (positively)

---

<sup>13</sup>Interestingly, in Anderson and Renault (2000), where customers may lack information about "match value," which is again specific, a greater share of informed customers has a *negative* externality through increasing the prevailing price.

<sup>14</sup>With perfect competition, all contracts would be undistorted in our model, while high- and low-demand types would realize the same surplus irrespective of whether they are initially informed or not. From the arguments in Armstrong and Vickers (2001) and Rochet and Stole (2002), it could be conjectured that as long as the market remains fully covered and as long as horizontal differentiation is "type-independent" (i.e., additive), price differences only reflect cost differences. However, if these two conditions do not jointly hold, then under imperfect competition, there remains scope for profitable price discrimination (cf. also Stole 1995 and Inderst 2004.) An analysis of imperfect competition in this sense is beyond the scope of the paper.

on both  $u_{A,l}$  and  $u_{P,l}$  through the binding constraints  $ICCP$  (the "cross" constraint to the uninformed customers's menu),  $IC_{P,l}$  (the constraint in the menu), and  $ICC_{A,h}$  (the "cross" constraint for uninformed customers). Precisely, from the proof of Proposition 1, we have

$$V_{A,h} = \mu(\theta_h - \theta_l)u_{A,l} + (1 - \mu)(\theta_h - \theta_l)u_{P,l}. \quad (4)$$

Results are now ambiguous, as we already know that a change in  $\pi$  has the opposite effect on  $u_{A,l}$  and  $u_{P,l}$ . In fact, the proof of Corollary 4 shows that the set of parameters for which the overall effect on  $V_{A,h}$  is positive and where it is negative are both generally non-empty.

We can make further progress by requiring that  $C'''$  is zero.<sup>15</sup> With this specification, contracts in Case 2 stipulate

$$\begin{aligned} u_{A,l} &= \frac{1}{2c} \left[ \theta_l - \left( \frac{\mu}{1 - \mu} \right) \left( \frac{1 - \pi + \pi\mu}{\pi} \right) (\theta_h - \theta_l) \right], \\ u_{P,l} &= \frac{1}{2c} \left[ \theta_l - \mu \left( \frac{\pi}{1 - \pi} \right) (\theta_h - \theta_l) \right]. \end{aligned}$$

Substituting these contracts into (4),  $dV_{A,h}/d\pi > 0$  holds if and only if

$$\pi < \hat{\pi} := \frac{\sqrt{\mu}}{1 + \sqrt{\mu} - \mu}. \quad (5)$$

This threshold may or may not fall into the area that is covered by Case 2—i.e., the interval  $(\pi_{24}, \pi_{12})$ , where both boundaries also depend on  $\mu$  (cf. Figure 1). As shown in the proof of Corollary 4,  $\hat{\pi} \in (\pi_{24}, \pi_{12})$  holds only if  $\mu$  is not too large.

**Corollary 4** *As the fraction of informed customers increases (higher  $\pi$ ), the impact on the utility of informed customers is generally ambiguous: It is strictly negative in Case 4 and may be positive or negative in Case 2. With  $C''' = 0$ , the impact is strictly positive if and only if both  $\pi$  and  $\mu$  are sufficiently low.*

For the case with  $C''' = 0$ , Corollary 4 and its proof explicitly delineate a parameter region for which, together with Corollary 3, *all* customers are better off as  $\pi$  increases. From Corollary 4, this is the case if the probability of the "top type,"  $\pi\mu$ , is not too high.

**Impact on Welfare.** As there are more informed customers, the impact on welfare is also generally ambiguous. While in Cases 1 and 3, contracts are not affected, in Case 4, where only informed low-demand customers are excluded, an increase in  $\pi$  leads to a reduction in welfare on two counts: First, it reduces the service level  $u_{P,l}$ ; second, it increases the fraction of excluded customers. For Case 2, instead, one can show that there is still always a non-empty set of parameters for which total welfare strictly increases with

---

<sup>15</sup>Recall that  $C$  specifies costs as a function of the delivered "base utility,"  $\tilde{u}(q)$ , where  $q$  denotes quantity and where the ultimate utility is given by  $\theta\tilde{u}(q)$  for a customer of type  $\theta$ . In terms of the model's primitives, stipulating that  $C''' = 0$  is then equivalent to specifying some utility function  $\tilde{u}(q) = \sqrt{q}/\gamma$  (together with marginal cost  $\tilde{c}$ ), where  $\gamma > 0$ . (We use here, also, that  $C(0) = 0$  and  $C'(0) = 0$ . and note that  $C'''(u) = \tilde{c}\gamma$ .)

$\pi$  (see Appendix 2). In this case, the positive effect that this has on  $u_{A,l}$  is sufficiently strong, compensating for the two negative effects on welfare, which arise from a reduction in  $u_{P,l}$  and from the fact that the newly informed low-demand customers now consume a strictly lower service level  $u_{A,l}$  instead of  $u_{P,l}$ .

## 5 Information Acquisition

Customers who are initially uninformed about their future demand (type) may be able to acquire additional information before signing a contract. For instance, a customer may be able to go through the records of her past consumption of the respective service—e.g., her past phone bills—to better estimate her future demand. To allow for this possibility, we stipulate in what follows that at the contracting stage, uninformed customers also can observe their demand type, albeit only after incurring private disutility  $k > 0$ .

The game between the firm and customers can then be described as follows: At stage 1, the firm proposes a set of contracts. At stage 2, uninformed customers decide whether to spend  $k$  to learn their type. At stage 3, customers decide which, if any, contract to sign. At stage 4, every customer observes his type. Customers who have chosen the contract that is targeted at uninformed customers decide which option in the contract to pick.

In terms of the firm’s program, the possibility of information acquisition requires that the incentive-compatibility constraint for an uninformed customer be modified. Her alternatives, in addition to accepting the designated offer  $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$ , are now threefold: first, to reject all offers, as captured by the individual rationality constraint  $IR_P$ ; second, to stay uninformed and pick a contract designed for an informed customer; and third to become informed and subsequently make the *best* choice among all possible options, namely to either reject all contracts on offer or to accept one of them.

In what follows, for brevity’s sake, we restrict consideration to the case where the firm’s offer is acceptable to all types. Moreover, while the full program is solved in the proof of the subsequent Proposition 4, in the main text, we address only the most salient issues.

### 5.1 Analysis

If the uninformed customers did acquire information, the firm would face only informed customers and thus a standard screening problem. The resulting optimal offer would then clearly deprive customers of the incentives to acquire information.<sup>16</sup> Recall from our analysis that, without the option of information acquisition, for high  $\pi$  the standard screening solution was still optimal (Case 1 in Proposition 1). Intuitively, in this case, the option to acquire information has no impact. The remaining case is that of Case 2, where  $\pi$  is sufficiently low and where previously (cf. Proposition 1) the offer designated for informed low-demand customers was more distorted:  $u_{A,l} < u_{P,l}$ .

---

<sup>16</sup>As the consumer’s indifference can be broken by a marginal adjustment of contracts, it is straightforward to also rule out the case where the uninformed consumer would mix between acquiring information or not.

Given the additional option to acquire information, the uninformed customers' incentive-compatibility constraint becomes now (cf. the proof of Proposition 4)

$$\mu V_h + (1 - \mu)V_{P,l} \geq \max \{(\theta_h - \theta_l)u_{A,l}, \mu V_h - k\}, \quad (6)$$

where it has already been used that  $V_{A,h} = V_{P,h} = V_h$ . The first term on the right-hand side of (6) arises again from the option to mimic informed low-demand customers. As  $V_{A,l} = 0$ , the uninformed customer would then only realize a positive rent, namely of  $(\theta_h - \theta_l)u_{A,l}$ , if she turns out to have high demand. The second term on the right-hand side of (6) captures the new option to become informed at cost  $k$ . In this case, the customer will instead realize utility  $V_h$  when being of the high-demand type.

Now, take the values for  $u_{A,l}$  and  $u_{P,l}$  as obtained in Proposition 1. Once we substitute for  $V_h$ , we can show that under the previously derived offer, the option to acquire information does *not* become sufficiently attractive for uninformed customers whenever

$$k \geq \mu(1 - \mu)(\theta_h - \theta_l)(u_{P,l} - u_{A,l}) \quad (7)$$

holds. Note that this is trivially the case if  $\pi \geq \frac{1}{2-\mu}$ , where  $u_{P,l} = u_{A,l} = u_l^S$  (Case 1), which confirms our previous observation. On the other hand, if  $\pi < \frac{1}{2-\mu}$  holds, then (7) defines an upper boundary on  $k$  such that we can ignore the new constraint that arises from the possibility of information acquisition if the respective costs  $k$  are sufficiently high. Otherwise, the firm has to adjust its offer. (Note that, as in the analysis of Crémer and Khalil (1992), where there are only uninformed types, in equilibrium there is no information acquisition.)

**Proposition 4** *Suppose that uninformed customers can become informed at cost  $k > 0$ . If, under the firm's optimal offer, all customers purchase a positive level of services, the following characterization applies:*

*Case 1) If  $\pi \geq \frac{1}{2-\mu}$ , Case 1 of Proposition 1 applies.*

*Case 2a) If  $\pi < \frac{1}{2-\mu}$  and  $k$  is sufficiently large such that it satisfies (7), then the contract specified in Case 2 is optimal.*

*Case 2b) If  $\pi < \frac{1}{2-\mu}$  and  $k$  is small such that it violates (7), then the optimal offer still has the property  $u_{A,l} < u_l^S < u_{P,l} < u_l^{FB}$  as in Case 2 of Proposition 1, albeit  $u_{P,l}$  is now smaller and  $u_{A,l}$  larger compared to the characterization there.*

## 5.2 Comparative Analysis

In Case 2b of Proposition 4, it is optimal for the firm to distort the informed low-type contract *less* and the uninformed low-type contract *more* compared to the characterization in Case 2 of Proposition 1. In fact, the difference between the respective values  $u_{P,l} > u_{A,l}$  is then pinned down by the binding condition (7), which becomes

$$u_{P,l} - u_{A,l} = \frac{k}{\mu(1 - \mu)(\theta_h - \theta_l)}. \quad (8)$$

This implies, in particular, that for  $k \rightarrow 0$ , both offers become the same. Intuitively, as uninformed customers can become informed at (almost) zero costs, the firm's problem

reduces to a standard screening problem:  $u_{.,l} \rightarrow u_l^S$ . More generally speaking, as  $k$  becomes smaller, the firm's ability to price discriminate between informed and uninformed customers shrinks, which undermines a key reason why the firm previously made the (most basic) offer  $u_{A,l}$  so unattractively low. This leads us to the following Corollary.

**Corollary 5** *Suppose that  $\pi < \frac{1}{2-\mu}$  and  $k$  small such that it violates (7). As the costs of information acquisition  $k$  decrease, the difference in the allocation for informed and uninformed customers,  $u_{A,l} - u_{P,l} > 0$ , decreases according to (8). For  $k \rightarrow 0$  we have that  $u_{.,l} \rightarrow u_l^S$ .*

From Corollary 5, contracts for informed customers become more efficient and contracts for uninformed customers less efficient as  $k$  decreases. To conclude this section, we ask—in analogy to the comparative analysis in  $\pi$  from Section 4—how changes in  $k$  affect consumers and welfare. As with changes in  $\pi$ , this may also be of interest for policy, provided that it aims to aid consumers in becoming more informed.

Intuitively, even though in the present model uninformed customers do not become more or less informed as  $k$  decreases (see, however, the concluding remarks in Section 6), their expected utility increases. (This holds strictly in Case 2b—i.e., whenever both  $\pi$  and  $k$  are not too high.) The effect that this has on informed consumers, however, is generally ambiguous—as is the effect on welfare. To make further progress, as in Section 6, suppose that  $C''' = 0$ .

**Corollary 6** *Uninformed customers benefit from a reduction in their own costs of information acquisition,  $k$ , while the impact on informed customers and welfare is generally ambiguous. With  $C''' = 0$ , we have that (i) informed customers benefit if and only if  $\pi$  is sufficiently small and that (ii) welfare always increases.*

Thus, for  $C''' = 0$ , a reduction in the costs of information acquisition benefits all customers if the fraction of informed customers is not too large. Otherwise, in case this is the result of a policy measure, this measure has distributional consequences: Uninformed customers benefit, while those customers who are already informed are hurt.

Interestingly, however, for  $C''' = 0$ , we find that the impact on welfare is always strictly positive. As noted in the Introduction, this provides a striking contrast to the result in Crémer and Khalil (1992), where in the presence of *only* uninformed customers, a reduction in  $k$  always decreases welfare. This difference follows from the fact that, in their case, the positive impact (in terms of efficiency) on informed customers' offers is absent.

## 6 Conclusion

For many subscription services, tariff choice and consumption decisions are made at different points in time. Thus, when signing a contract, a share of customers are still uncertain about their future level of demand. This paper considers the contract-design problem of a monopolist facing both uninformed customers and customers who, at the contracting stage, are already informed about their demand (type). In an extension, we also allow for the possibility that uninformed customers can acquire information at a cost.

Initially, the firm faces both informed and uninformed customers, as well as informed customers with high or low demand. The respective share of informed vs. uninformed and high- vs. low-demand types determines the prevailing distortion of contracts and whether all customers are served in the first place. The restrictions to only two types and fully revealing signals allow us to make this case distinction explicit and simple. In the comparative analysis, we found that the presence of informed customers benefits uninformed customers, even though information is only about a customer's own demand. In particular, in the present model, there is thus no "free-riding" of informed customers (through being "cross-subsidized") on uninformed customers. We used this result to argue how policy intervention that helps consumers to obtain more knowledge about their own demand (type) could be justified, also given the resulting externality. In an extension we analyze the situation where policy intervention may reduce the costs of information acquisition, albeit in equilibrium the firm adjusts the offer so that this option is actually not used. For a tractable specification we found that policy intervention increases welfare, in striking contrast to the case where all consumers are ex-ante uninformed.

The tractable framework that this paper introduces may allow for several extensions and applications in future work. It could be interesting to endogenize the differential information that customers possess at the contracting stage. If customers have different costs of acquiring information, those with low costs should become informed, while those with higher costs should stay uninformed. The firm's design of the price-discriminating offer would then determine also the fraction of customers who are informed, while public policy could more generally affect the costs of information acquisition.<sup>17</sup>

Our analysis is restrictive as we only consider two demand types (low and high) and as consumers are ex-ante either informed about their type or uninformed. Clearly, a more general model could allow for more types, including a continuous variation, in either dimension: Demand could take more than two values and consumers could obtain an ex-ante signal about their future demand that varies in its precision. Our main contribution does not lie in the derivation of the optimal tariff structure, but rather in the comparative analysis that allows us to obtain normative implications. We must leave it to future work how, with a still tractable characterization of optimal contracts, our comparative analysis and normative implications formally survive.

---

<sup>17</sup>In related work, Bar-Isaac et al. (2010) analyze the decision of the firm to facilitate information acquisition for consumers with heterogeneous preferences.

## Appendix: Proofs

**Proof of Proposition 1.** The proof proceeds in several steps.

**Step 1:** If the optimal contract menu has the property that all customers take contracts with strictly positive quantities, then we show next the following: (i)  $u_{A,h} = u_h^{FB}$ ; (ii)  $u_{P,h} = u_h^{FB}$ ; (iii)  $V_{A,l} = 0$ ; and (iv) that the constraint  $ICC_P$  is binding.

To see this, note first that if (i) does not hold, then we can adjust  $u_{A,h}$  and  $t_{A,h}$  so as to keep  $V_{A,h}$  constant, while increasing the surplus and thus the firm's profits. This is possible since in the relaxed program there are no constraints to mimic the informed high type. If (ii) does not hold, we can adjust  $u_{P,h}$  and  $t_{P,h}$  to increase the surplus, while leaving  $V_{P,h}$  constant and thus also  $ICC_{A,h}$  satisfied. Regarding assertion (iii), we only have to note that in the relaxed program there is no incentive constraint for the informed low type. Finally, assertion (iv) trivially holds in the relaxed program as, otherwise, one can adjust  $t_{P,i}$  downwards, while still satisfying all remaining constraints.

**Step 2:** Next, if all customers are served, we show that  $ICC_{A,h}$  must be binding, i.e., that  $V_{A,h} = V_{P,h}$ . To show this, suppose by contradiction that  $ICC_{A,h}$  is not binding. Then the firm optimally raises  $t_{A,h}$  until  $IC_{A,h}$  binds. It is trivial that in this case  $u_{A,l} > 0$  must hold (so that  $V_{A,h} > 0$ ). Note next that if  $ICC_{A,h}$  does not bind, then  $u_{P,l}$  is optimally chosen so as to maximize surplus:  $u_{P,l} = u_l^{FB}$ . Substituting next the binding  $IC_{A,h}$  into the binding  $ICC_P$ , we have the requirement that  $\mu V_{P,h} + (1 - \mu)V_{P,l} = \mu(\theta_h u_{A,l} - t_{A,l}) = \mu V_{A,h}$ . As from  $ICC_{A,h}$  we have  $V_{A,h} \geq V_{P,h}$ , this requires that  $V_{P,l} \geq 0$ . If we substitute all of this into the firm's program, then the remaining consumption profile to specify is  $u_{A,l}$ . For this note that the expected surplus with this type of customers is  $\pi(1 - \mu)s(u_{A,l}; \theta_l)$ , while the information rent for the informed high type is  $\pi\mu(\theta_h - \theta_l)u_{A,l}$ . Moreover, from the binding constraints  $IC_{A,h}$  and  $IR_{A,l}$  it follows that the utility of an uninformed high type equals  $(\theta_h - \theta_l)u_{A,l}$ , which in expected terms (for the firm) equals  $(1 - \pi)\mu(\theta_h - \theta_l)u_{A,l}$ . As a consequence, we must clearly have that  $u_{A,l} < u_l^{FB}$ .

We argue now that, contrary to the assumption,  $ICC_{A,h}$  is violated as the derived contract implies, in fact, that  $V_{A,h} < V_{P,h}$ . This follows from two observations: (i)  $V_{P,l} \geq 0$  and  $u_{P,l} = u_l^{FB}$  imply together with  $IC_{P,h}$  that  $V_{P,h} \geq (\theta_h - \theta_l)u_l^{FB}$ ; and (ii)  $V_{A,l} = 0$  and the binding  $IC_{A,h}$  imply  $V_{A,h} = (\theta_h - \theta_l)u_{A,l}$ . Since  $u_{A,l} < u_l^{FB}$ , this gives  $V_{A,h} = (\theta_h - \theta_l)u_{A,l} < (\theta_h - \theta_l)u_l^{FB} \leq V_{P,h}$ , which is a contradiction.

**Step 3:** We now first solve the remaining program under the hypothesis that  $IC_{A,h}$  does *not* bind. Hence, because of Step 2 we consider the situation in which  $ICC_{A,h}$  binds but not  $IC_{A,h}$ . It is then immediate that  $IC_{P,h}$  must bind. Together with  $V_{A,h} = V_{P,h} \equiv V_h$  we have  $V_h = \theta_h u_{p,l} - t_{p,l} = V_{P,l} + (\theta_h - \theta_l)u_{p,l}$ . The firm then obtains all expected surplus minus "rents" that are obtained by all uninformed and informed high-type customers. The former group obtains in expectation  $\mu(\theta_h - \theta_l)u_{A,l}$ , the latter simply  $V_h$ .

To determine  $V_h$ , we proceed as follows. From  $ICC_P$  the expected surplus of an uninformed customer is  $\mu V_h + (1 - \mu)V_{P,l} = \mu(\theta_h - \theta_l)u_{A,l}$ . As from  $IC_{P,h}$  we have  $V_{P,l} = V_h - (\theta_h - \theta_l)u_{p,l}$ , it also holds that  $\mu V_h + (1 - \mu)(V_h - (\theta_h - \theta_l)u_{p,l}) = \mu(\theta_h - \theta_l)u_{A,l}$ ,

such that jointly this implies that

$$V_h = \mu(\theta_h - \theta_l)u_{A,l} + (1 - \mu)(\theta_h - \theta_l)u_{P,l}. \quad (9)$$

Therefore, the total expected rent that goes to customers is

$$(1 - \pi)\mu(\theta_h - \theta_l)u_{A,l} + \pi\mu[\mu(\theta_h - \theta_l)u_{A,l} + (1 - \mu)(\theta_h - \theta_l)u_{P,l}],$$

implying that

$$u_{P,l} = \arg \max_u (1 - \mu)[(1 - \pi)s(u; \theta_l) - \pi\mu(\theta_h - \theta_l)u], \quad (10)$$

while

$$u_{A,l} = \arg \max_u \pi(1 - \mu)s(u; \theta_l) - (1 - \pi)\mu(\theta_h - \theta_l)u - \pi\mu\mu(\theta_h - \theta_l)u. \quad (11)$$

**Step 4:** We now establish a condition that the constraint  $IC_{A,h}$  is indeed slack, as claimed in the previous step. For  $IC_{A,h}$  slack—i.e.,  $V_h > (\theta_h - \theta_l)u_{A,l}$ —equation (9) holds. Thus the condition for slackness is

$$\mu(\theta_h - \theta_l)u_{A,l} + (1 - \mu)(\theta_h - \theta_l)u_{P,l} > (\theta_h - \theta_l)u_{A,l},$$

which is equivalent to  $u_{A,l} < u_{P,l}$ . Whenever  $IC_{A,h}$  is slack, it follows from step 3 that we obtain  $u_{P,l}$  and  $u_{A,l}$  as (10) and (11), respectively. After setting up the first-order conditions and rearranging terms, we have that  $u_{A,l} < u_{P,l}$  holds if and only if

$$\frac{\pi\mu}{1 - \pi} < \frac{(1 - \pi)\mu + \pi\mu\mu}{\pi(1 - \mu)}$$

which is equivalent to  $\pi < \frac{1}{2-\mu}$ . Note that, at  $\pi = \frac{1}{2-\mu}$ , we have  $u_{A,l} = u_{P,l} = u_l^S$ . Hence, we have established that  $IC_{A,h}$  can indeed be ignored when solving the program if and only if condition  $\pi \leq \frac{1}{2-\mu}$  holds (at least weakly), while otherwise the constraint must bind.

**Step 5:** Suppose now that the converse of this condition holds, so that  $\pi > \frac{1}{2-\mu}$ . The argument in Step 4 implies that then  $IC_{A,h}$  must bind. Note also that by Step 2 also condition  $ICC_{A,h}$  must bind by optimality. Taken together, it then holds that  $\mu V_{P,h} + (1 - \mu)V_{P,l} = \mu(\theta_h u_{A,l} - t_{A,l}) = \mu V_{A,h}$  and thus that  $V_{P,l} = 0$ . Note first that it is not feasible to have  $u_{A,l} < u_{P,l}$ , given  $IC_{P,h}$ ,  $V_{A,h} = V_{P,h}$ , and as by assumption  $IC_{A,h}$  binds. While it could be feasible that  $u_{P,l} < u_{A,l}$ , it is easily shown from  $u_{A,l} < u_l^{FB}$  that this is not optimal. With  $u_{A,l} = u_{P,l}$  we then have the standard screening program and thus  $u_{A,l} = u_{P,l} = u_l^S$ .

**Step 6:** Finally, note that the solution to the relaxed program satisfies the neglected constraints. In fact, the only case where this is not immediately obvious is that were the informed low type would want to mimic an uninformed customer. Since  $IC_{P,h}$  is binding, we only have to exclude the option to ultimately select  $(u_{P,l}, t_{P,l})$ . This is, however, strictly unprofitable as we obtain

$$V_{P,l} = \mu(\theta_h - \theta_l)(u_{A,l} - u_{P,l}) < 0.$$

**Q.E.D. (of Proposition 1)**

**Proof of Proposition 2.** To see first that it cannot be the case that *only* uninformed low-type customers have a zero level of services, implying  $u_{A,l} > u_{P,l} = 0$ , recall from the proof of Proposition 1, which solves the relaxed program, that in fact  $u_{A,l} \leq u_{P,l}$ . Next, if the firm only serves high-demand customers, then it is immediate that  $u_{.,h} = u_h^{FB}$  and  $t_{.,h} = \theta_h u_h^{FB}$  (Case 4). This leaves us with only one remaining case: Case 3, where only informed low-type customers are excluded.

Note next that  $ICC_P$  becomes irrelevant, but that now the *ex ante* individual rationality constraint  $IR_P$  becomes binding:  $\mu V_{P,h} + (1 - \mu)V_{P,l} = 0$ . We show that  $ICC_{A,h}$  is binding. Suppose otherwise. Then the firm would propose a contract with  $V_{A,h} = 0$ . In order not to violate  $ICC_{A,h}$  we must have that  $V_{P,h} \leq 0$ . Because of individual rationality this requires that  $V_{P,l} \geq 0$ . But in this case the uninformed high-type customer would profit from (later) choosing  $(u_{P,l}, t_{P,l})$  such that  $IC_{P,h}$  would be violated. This establishes that  $ICC_{A,h}$  is indeed binding such that  $V_{A,h} = V_{P,h}$ . For Case 3 note finally that  $IC_{P,h}$  is always binding. Otherwise, the firm could increase  $t_{P,h}$ , while simultaneously decreasing  $t_{P,l}$  so as to still satisfy  $IR_P$ , which would be profitable as it allows also to increase  $t_{A,h}$ .

Having thus established which constraints must be binding in Case 3, note that the rent of the informed high type is given by  $V_{P,l} + (\theta_h - \theta_l)u_{P,l}$ , which after substituting  $V_{P,l} = -\mu(\theta_h - \theta_l)u_{P,l}$  from  $IR_P$  becomes  $(1 - \mu)(\theta_h - \theta_l)u_{P,l}$ . This shows finally that  $u_{P,l}$  maximizes again (10). **Q.E.D. (of Proposition 2)**

**Proof of Proposition 3.** Using Proposition 1, define the function  $\tilde{\pi}_{12} := \frac{1}{2-\mu}$  to separate Case 1 from Case 2. Recall next that if informed and uninformed types obtain the same ("standard screening") contract, then only high-demand customers are served if  $\mu > \theta_l/\theta_h$ , which separates Cases 1 and 3. To separate Cases 2 and 4, we use  $C'(0) = 0$  together with  $u_{A,l} = 0$  to solve from (2) for a function

$$\tilde{\pi}_{24}(\mu) := \frac{(\theta_h - \theta_l)\mu}{(1 - \mu)[\theta_l + (\theta_h - \theta_l)\mu]}$$

such that Case 2 only applies if  $\pi \geq \tilde{\pi}_{24}(\mu)$ . Note here that  $\tilde{\pi}_{24}(0) = 0$ ,  $\tilde{\pi}'_{24}(0) = (\theta_h - \theta_l)/\theta_l$ , and  $\tilde{\pi}'_{24}(\mu) > 0$ . Separating Cases 3 and 4, we proceed likewise and use  $C'(0) = 0$  next to  $u_{P,l} = 0$  to obtain from (3) that

$$\tilde{\pi}_{34}(\mu) := \frac{\theta_l}{\theta_l + (\theta_h - \theta_l)\mu},$$

which is strictly decreasing in  $\mu$ .

The assertions in Proposition 3 follow then immediately from applying the derived boundaries for the different cases. Note here, in particular, that all three boundaries ( $\tilde{\pi}_{12}$ ,  $\tilde{\pi}_{24}$ , and  $\tilde{\pi}_{34}$ ) together with the horizontal line  $\mu = \theta_l/\theta_h$  intersect at a single point:  $\mu = \theta_l/\theta_h$  and  $\pi = \frac{\theta_h}{2\theta_h - \theta_l}$ . **Q.E.D. (of Proposition 3)**

**Proof of Corollary 4.** Implicit differentiation of (2) and (3) yields

$$\begin{aligned}\frac{du_{A,l}}{d\pi} &= -\frac{\theta_h - \theta_l}{s''(u_{A,l}; \theta_l)} \frac{\mu}{1 - \mu} \frac{1}{\pi^2}, \\ \frac{du_{P,l}}{d\pi} &= \frac{\theta_h - \theta_l}{s''(u_{P,l}; \theta_l)} \mu \frac{1}{(1 - \pi)^2},\end{aligned}$$

such that  $dV_{A,h}/d\pi > 0$  holds if and only if

$$\frac{s''(u_{A,l}; \theta_l)}{s''(u_{P,l}; \theta_l)} < \frac{\mu}{(1 - \mu)^2} \frac{(1 - \pi)^2}{\pi^2}. \quad (12)$$

To show that generally the set of parameters for which (12) holds is non-empty, as well as the set for which the converse holds strictly, we consider parameters at the boundaries of Case 2. First, for  $\pi$  close to  $1/(2 - \mu)$ , (12) does not hold, given that in this case the right-hand side of (12) exceeds one, while as  $u_{A,l}$  and  $u_{P,l}$  become close to  $u_{S,l}$  the left-hand side of (12) converges to one. Second, when  $\pi$  is close to  $\tilde{\pi}_{24}$ , (12) holds at least for sufficiently small  $\mu$ . More formally, at the boundary  $\pi = \tilde{\pi}_{24}(\mu)$  the right-hand side of (12) becomes  $\frac{1}{\mu} \left( \frac{\theta_l}{\theta_h - \theta_l} + \frac{\mu(1-2\mu)}{1-\mu} \right)^2$  and thus tends to infinity as  $\mu \rightarrow 0$ .

Next, for  $C''' = 0$  condition (12) transforms to (5). Note next that  $\hat{\pi} < \tilde{\pi}_{12}$  holds if

$$\frac{\sqrt{\mu}}{1 + \sqrt{\mu} - \mu} < \frac{1}{2 - \mu},$$

which is always satisfied. After substitution into  $\hat{\pi} > \tilde{\pi}_{24}$  and some transformations, this condition becomes

$$\frac{\theta_l}{\theta_h - \theta_l} > \frac{\sqrt{\mu} - \mu\sqrt{\mu} + \mu^2}{1 - \mu},$$

which holds for  $\mu = 0$  but is violated close to the upper boundary  $\mu \rightarrow \theta_l/\theta_h$ . As the numerator on the right-hand side is increasing in  $\mu$  and the denominator is decreasing, this implies existence of an interior threshold such that the condition holds if and only if  $\mu$  is sufficiently small. **Q.E.D. (of Corollary 4)**

**Proof of Proposition 4.** The firm's offer must satisfy the new incentive compatibility constraint

$$\mu V_{P,h} + (1 - \mu)V_{P,l} \geq \mu V_{A,h} + (1 - \mu) \max\{V_{A,l}, V_{P,l}\} - k, \quad (13)$$

where we already used that  $V_{A,h} \geq V_{P,h}$  from  $ICC_{A,h}$  as well as  $V_{A,h} \geq 0$  from  $IR_{A,h}$ . We refer to (13) as  $ICC'_P$ . We characterize now stepwise the solution to the firm's new program.

**Step 1:** We first show that we can ignore the additional constraint  $ICC'_P$  in case the solution to the relaxed program (see Proposition 1) satisfies (7). Take thus the solution to the relaxed program (i.e., with  $k = \infty$ ). Recall from the proof of Proposition 1 that in this case  $ICC_P$  binds such that  $\mu V_{P,h} + (1 - \mu)V_{P,l}$  equals  $(\theta_h - \theta_l)u_{A,l}$ , while also

$V_{A,l} = 0$ ,  $V_{P,l} \leq 0$ , and  $V_{A,h} = V_{P,h} = V_h$  satisfies  $V_h = \mu(\theta_h - \theta_l)u_{A,l} + (1 - \mu)(\theta_h - \theta_l)u_{P,l}$ . Substituting these expressions into (13), we obtain

$$k \geq \mu(1 - \mu)(\theta_h - \theta_l)(u_{P,l} - u_{A,l}) - (1 - \mu)(\theta_h - \theta_l)u_{A,l},$$

implying that  $ICC'_P$  holds from (7).

In what follows we can thus focus on the case where condition (7) does not hold such that  $ICC'_P$  must bind.<sup>18</sup>

**Step 2:** Note that from the same arguments as in the proof of Proposition 1 it holds that  $u_{.,h} = u_h^{FB}$  and  $V_{A,l} = 0$ .

**Step 3:** We claim that if  $ICC'_P$  binds, then also  $ICC_{A,h}$  must bind such that  $V_{A,h} = V_{P,h} = V_h$ . We prove this by contradiction and suppose that  $V_{A,h} > V_{P,h}$ . Clearly, as the firm optimally increases  $t_{A,h}$  as much as possible and as  $ICC_{A,h}$  does not bind by assumption, the constraint  $IC_{A,h}$  must bind such that  $V_{A,h} = (\theta_h - \theta_l)u_{A,l}$ .

We next determine  $u_{A,l}$  and  $u_{P,l}$ . As  $ICC_{A,h}$  is supposed not to bind, it is immediate that optimally  $u_{P,l} = u_l^{FB}$ . To determine  $u_{A,l}$  note that from  $V_{A,h} > V_{P,h}$  and from the binding  $ICC'_P$  a reduction  $dt_{A,h} < 0$  increases the utility of the uninformed customer by  $-\mu dt_{A,h}$ . Recall also that  $IC_{A,h}$  is binding. Consequently, the choice of  $u_{A,l}$  optimally trades off the maximization of the surplus  $s(u_{A,l}; \theta_l)$  with the reduction of the rent  $(\theta_h - \theta_l)u_{A,l}(1 - \pi + \pi\mu)$ . As we assume that the firm serves all customers, we thus have that  $u_{A,l}$  solves

$$s'(u_{A,l}; \theta_l) = \frac{1 - \pi + \pi\mu}{\pi(1 - \mu)}(\theta_h - \theta_l). \quad (14)$$

Note that the resulting value of  $u_{A,l}$  is thus strictly *smaller* than that determined for Case 2 in (2). As also  $u_{P,l} = u_l^{FB}$  is strictly larger than the respective value in Case 2, we thus have that the difference  $u_{P,l} - u_{A,l}$  is strictly larger than the respective difference for the solution in Case 2. Consequently, as by assumption (7) was not satisfied for the solution to the original program (Case 2), where  $u_{P,l} - u_{A,l}$  was smaller, it must hold *a fortiori* that now

$$k < \mu(1 - \mu)(\theta_h - \theta_l)(u_{P,l} - u_{A,l}). \quad (15)$$

Note next that  $V_{A,h} = (\theta_h - \theta_l)u_{A,l}$ , while from  $IC_{P,h}$  we have that  $V_{P,h} \geq V_{P,l} + (\theta_h - \theta_l)u_{P,l}$ . Substituting into  $V_{A,h} > V_{P,h}$ , which holds by assumption, we have that  $(\theta_h - \theta_l)u_{A,l} > V_{P,h} \geq V_{P,l} + (\theta_h - \theta_l)u_{P,l}$ . It follows that

$$(\theta_h - \theta_l)(u_{P,l} - u_{A,l}) < -V_{P,l}. \quad (16)$$

As we have from the binding  $ICC'_P$  in (13) that  $-(1 - \mu)V_{P,l} = k - \mu(V_{A,h} - V_{P,h})$ , together with (16) this yields the requirement

$$k > (1 - \mu)(\theta_h - \theta_l)(u_{P,l} - u_{A,l}) + \mu(V_{A,h} - V_{P,h}), \quad (17)$$

---

<sup>18</sup>Note that we can rule out the case where  $ICC'_P$  constrains the firm's choice *without* being binding at the optimal offer. If (7) does not hold, then the firm's expected profits are strictly decreasing in  $\mu V_{P,h} + (1 - \mu)V_{P,l}$  and thus in the "slack" that would arise when  $ICC'_P$  did not hold with equality.

contradicting (15).

**Step 4:** Substituting  $V_{.,h} = V_h$  into the binding  $ICC'_P$ , we have that (13) becomes

$$(1 - \mu)V_{P,l} = (1 - \mu) \max \{V_{A,l}, V_{P,l}\} - k,$$

which clearly requires  $V_{P,l} < 0$  and which from  $V_{A,l} = 0$  thus yields that

$$V_{P,l} = -\frac{k}{1 - \mu}. \quad (18)$$

**Step 5:** We next claim that if  $ICC'_P$  binds, then also  $ICC_P$  must bind. Substituting for  $\mu V_{P,h} + (1 - \mu)V_{P,l}$  from the binding  $ICC'_P$  and from (18) (together also with  $V_{.,h} = V_h$ ), note that  $ICC_P$  becomes

$$\mu V_h - k \geq \mu(\theta_h - \theta_l)u_{A,l}. \quad (19)$$

Suppose, by contradiction, that  $ICC_P$  does not bind. Then in the optimal contract it must clearly hold that  $IC_{A,h}$  or  $IC_{P,h}$  (possibly both) must bind. We argue now that  $IC_{A,h}$  must bind. If only  $IC_{P,h}$  binds, then note first that  $V_h = -k/(1 - \mu) + (\theta_h - \theta_l)u_{P,l}$ , while an uninformed customer realizes  $\mu V_h - k$ . It is immediate that the optimal offer must satisfy  $u_{P,l} < u_{A,l} = u_l^{FB}$ . As then  $V_{A,h} \geq (\theta_h - \theta_l)u_l^{FB}$ , while  $V_{P,h} = V_{P,l} + (\theta_h - \theta_l)u_{P,l}$  with  $V_{P,l} < 0$  and  $u_{P,l} < u_A^{FB}$ , we have  $V_{A,l} > V_{P,l}$ . This contradicts that  $ICC_{A,h}$  must be binding (as proved in step 3).

As  $IC_{A,h}$  must thus bind, we have that  $V_h = (\theta_h - \theta_l)u_{A,l}$ . Substituting this into  $ICC_P$  in (19) yields then the requirement  $\mu(\theta_h - \theta_l)u_{A,l} - k \geq \mu(\theta_h - \theta_l)u_{A,l}$ , which clearly can not hold.

**Step 6:** We claim that if  $ICC'_P$  binds, then  $IC_{P,h}$  is binding but not  $IC_{A,h}$ . To prove this claim, we first argue that we can ignore the constraint  $IC_{A,h}$ . This follows immediately as by combining the binding constraints  $ICC_P$  and  $ICC'_P$  (using step 5) we have that

$$V_h = (\theta_h - \theta_l)u_{A,l} + \frac{k}{\mu}. \quad (20)$$

If  $IC_{P,h}$  was *also* not binding, then the firm could benefit from simply reducing  $V_h = V_{.,h}$  (by increasing the transfer). Consequently,  $IC_{P,h}$  must bind.

**Step 7:** Note next that, as in the proof of Proposition 1, we have from the binding constraints  $IC_{P,h}$  and  $ICC_P$  that  $V_h$  is given by (9). Together with the binding constraint  $ICC'_P$  this implies condition (8) for the difference  $u_{P,l} - u_{A,l}$ .

We turn now to the determination of  $u_{A,l}$  and  $u_{P,l}$ . Note that, expressed solely as a function of  $u_{A,l}$ , we have for the informed high type  $V_h = \frac{k}{\mu} + (\theta_h - \theta_l)u_{A,l}$  and for the uninformed customer the expected utility  $\mu(\theta_h - \theta_l)u_{A,l}$ . Hence, trading off surplus maximization with customer rent extraction, the optimal choice of  $u_{A,l}$  maximizes

$$\pi(1 - \mu)s(u_{A,l}; \theta_l) + (1 - \pi)(1 - \mu)s(u_{P,l}; \theta_l) - \mu(\theta_h - \theta_l)u_{A,l},$$

where  $u_{P,l}$  depends on  $u_{A,l}$  according to (8) (i.e.,  $du_{P,l}/du_{A,l} = 1$ ). Given that we focus on the case where it is optimal for the firm that all customers purchase a positive level  $u_{.,i} > 0$ , this yields the first-order condition

$$\pi(1 - \mu)s'(u_{A,l}; \theta_l) + (1 - \pi)(1 - \mu)s'(u_{P,l}; \theta_l) = \mu(\theta_h - \theta_l). \quad (21)$$

**Step 8:** We claim the following: If the solution in Case 2 of Proposition 1 does not satisfy (7), then equations (21) and (8) pin down a unique solution  $u_{A,l} < u_{P,l} < u_l^{FB}$  such that  $u_{A,l}$  is larger and  $u_{P,l}$  smaller than in the offer of Case 2.

To prove this claim it is convenient to consider  $u_{P,l}$  as the remaining variable, with  $u_{A,l}$  determined by (8). We argue first that when setting  $u_{P,l} = u_l^{FB}$  and the corresponding value  $u_{A,l} = u_{P,l} - y$  with  $y := \frac{k}{\mu(1-\mu)(\theta_h - \theta_l)}$ , then the left-hand side of (21) is strictly lower than the right-hand side. To see this, note that the left-hand side of (21) then becomes  $\pi(1 - \mu)s'(u_{A,l}; \theta_l)$ . Take now as a comparison the solution  $u_{A,l}$  in Case 2 as given by (14), which as we know must clearly be strictly lower. The assertion follows then as at this lower value of  $u_{A,l}$  we have that  $\pi(1 - \mu)s'(u_{A,l}; \theta_l)$  equals  $(1 - \pi + \pi\mu)(\theta_h - \theta_l)$ , which is in turn strictly lower than the right-hand side  $\mu(\theta_h - \theta_l)$  of (21). As a final note, observe that using  $du_{A,l}/du_{P,l} = 1$  uniqueness of  $u_{A,l}$  follows from strict concavity of the surplus function, implying that the left-hand side of (21) is strictly monotonic.

**Step 9:** For a comparison with Case 2 at the upper boundary for  $k$ , recall first that from the characterization in Case 1 we have that  $\pi(1 - \mu)s'(u_{A,l}; \theta_l) = (\mu(1 - \pi + \pi\mu)(\theta_h - \theta_l)$  and that  $(1 - \pi)(1 - \mu)s'(u_{P,l}; \theta_l) = \pi\mu(1 - \mu)(\theta_h - \theta_l)$ . Adding up the right-hand sides yields exactly  $(\theta_h - \theta_l)\mu$ . Hence, the solutions for  $u_{A,l}$  and  $u_{P,l}$  satisfy (21). Moreover, by definition we have that at the upper boundary of  $k$ , where  $ICCP'$  just starts to bind, (7) is satisfied with equality, yielding condition (8). **Q.E.D (of Proposition 4.)**

**Proof of Corollary 5.** We claim that  $u_{A,l}$  is strictly decreasing and  $u_{P,l}$  strictly increasing in  $k$ , where also  $u_{.,l} \rightarrow u_l^S$  for  $k \rightarrow 0$  and where at  $k$  satisfying (7) there is continuity with respect to the offers of Case 2. Using the derivations in the proof of Proposition 4, monotonicity in  $k$  follows immediately from implicit differentiation of (21), which establishes that

$$\frac{du_{P,l}}{dy} = \frac{\pi s''(u_{A,l}; \theta_l)}{\pi s''(u_{A,l}; \theta_l) + (1 - \pi)s''(u_{P,l}; \theta_l)} > 0$$

and

$$\frac{du_{A,l}}{dy} = -\frac{(1 - \pi)s''(u_{P,l}; \theta_l)}{\pi s''(u_{A,l}; \theta_l) + (1 - \pi)s''(u_{P,l}; \theta_l)} < 0. \quad (22)$$

For the convergence (and continuity) results, note that we can substitute  $y = 0$  for the case of  $k = 0$ . **Q.E.D (of Corollary 5.)**

**Proof of Corollary 6.** To show that uninformed customers benefit from a reduction in information acquisition costs, recall first that condition (7) just binds in Case 2b. There, where offers satisfy (8), an uninformed customer becomes indifferent between her two options for a deviation: the option of acquiring information and mimicking the respective,

preferred informed type and the option of mimicking an informed low-type customer without acquiring information. From the latter option, and as the incentive constraint binds, an uninformed customer realizes  $\mu(\theta_h - \theta_l)u_{A,h}$  (cf. also equation (6)). As, from Corollary 5,  $u_{A,l}$  increases in response to a decrease in  $k$ , the uninformed customer's expected utility thus indeed strictly increases.

Next, for Case 2b we know from Proposition 4 that informed customers obtain the utility  $V_h = \frac{k}{\mu} + (\theta_h - \theta_l)u_{A,l}$ . Hence, using (22) we have

$$\frac{dV_h}{dk} = \frac{1}{\mu} - \frac{1}{\mu(1-\mu)} \frac{(1-\pi)s''(u_{P,l}; \theta_l)}{\pi s''(u_{A,l}; \theta_l) + (1-\pi)s''(u_{P,l}; \theta_l)}.$$

This derivative is, in general, of ambiguous sign. In the special case that  $C'''$  is zero, this reduces to

$$\frac{dV_h}{dk} = \frac{1}{\mu} \left( 1 - \frac{1-\pi}{1-\mu} \right),$$

which is negative if and only if  $\mu > \pi$ .

Finally, using the derivations in the proof of Proposition 4, the impact on welfare from a change in  $k$  can be determined from

$$\pi(1-\mu)s'(u_{A,l}; \theta_l) \frac{du_{A,l}}{dy} + (1-\pi)(1-\mu)s'(u_{P,l}; \theta_l) \frac{du_{P,l}}{dy},$$

where we use  $y := \frac{k}{\mu(1-\mu)(\theta_h - \theta_l)}$ . Substituting for  $\frac{du_{P,l}}{dy}$  and  $\frac{du_{A,l}}{dy}$ , the term is strictly negative whenever

$$\pi(1-\pi)(1-\mu)^2 s'(u_{P,l}; \theta_l) |s''(u_{A,l}; \theta_l)| < \pi(1-\pi)(1-\mu)^2 s'(u_{A,l}; \theta_l) (1-\mu) |s''(u_{P,l}; \theta_l)|,$$

which reduces to

$$\frac{s'(u_{P,l}; \theta_l)}{|s''(u_{P,l}; \theta_l)|} < \frac{s'(u_{A,l}; \theta_l)}{|s''(u_{A,l}; \theta_l)|}.$$

This is always satisfied if  $C'''$  is zero since  $u_{P,l} > u_{A,l}$ . **Q.E.D. (of Corollary 6)**

## Appendix 2: Omitted Material from Welfare Analysis

We first show that  $dW/d\pi$  is sometimes positive and sometimes negative (depending on the concrete specification) at  $\pi = \tilde{\pi}_{24} + \varepsilon$  for  $\varepsilon > 0$  sufficiently small and  $\mu > 0$  sufficiently small. Note that as  $\pi \downarrow \tilde{\pi}_{24}$ , we have  $u_{A,l} \rightarrow 0$  and thus  $s(u_{A,l}; \theta_l) \rightarrow 0$ . Concerning the second term in  $dW/d\pi$ , recall that  $\frac{du_{P,l}}{d\pi} = \frac{\theta_h - \theta_l}{s''(u_{P,l}; \theta_l)} \frac{\mu}{(1-\pi)^2}$ . Using that for  $\pi \downarrow \tilde{\pi}_{24}$  we can substitute  $\pi$  by  $\frac{(\theta_h - \theta_l)\mu}{(1-\mu)[\theta_l + (\theta_h - \theta_l)\mu]}$ , where

$$\begin{aligned} & \lim_{\mu \downarrow 0} \lim_{\pi \downarrow \tilde{\pi}_{24}(\mu)} (1-\pi) s'(u_{P,l}; \theta_l) \frac{du_{P,l}}{d\pi} \\ &= \left[ \lim_{\mu \downarrow 0} \lim_{\pi \downarrow \tilde{\pi}_{24}(\mu)} s'(u_{P,l}; \theta_l) \right] \left[ \lim_{\mu \downarrow 0} \frac{\theta_h - \theta_l}{-C''(u_{P,l})} \frac{\mu(1-\mu)[\theta_l + (\theta_h - \theta_l)\mu]}{(1-\mu)\theta_l - \mu^2(\theta_h - \theta_l)} \right] \\ &= 0 \end{aligned}$$

since the second term is zero in the limit (using that  $C$  is strictly convex everywhere and that  $u_{P,l}$  falls into a bounded interval). Concerning the third term in  $dW/d\pi$ , recall that  $\frac{du_{P,l}}{d\pi} = \frac{\theta_h - \theta_l}{s''(u_{P,l}; \theta_l)} \mu \frac{1}{(1-\pi)^2}$ . Note also that  $s'(u_{A,l}; \theta_l) = \theta_l$  for  $\pi \rightarrow \tilde{\pi}_{24}$ . Then

$$\lim_{\mu \downarrow 0} \lim_{\pi \downarrow \tilde{\pi}_{24}(\mu)} \pi s'(u_{A,l}; \theta_l) \frac{du_{A,l}}{d\pi} = -\frac{\theta_l}{C''(0)} \lim_{\mu \downarrow 0} [\theta_l + (\theta_h - \theta_l)\mu] = -\frac{\theta_l^2}{C''(0)},$$

which is a finite number as  $C$  is everywhere strictly convex. We thus have that

$$\lim_{\mu \downarrow 0} \lim_{\pi \downarrow \tilde{\pi}_{24}(\mu)} \frac{dW}{d\pi} = -[\lim_{\mu \downarrow 0} \lim_{\pi \downarrow \tilde{\pi}_{24}(\mu)} s(u_{P,l}; \theta_l) - \frac{\theta_l^2}{C''(0)}],$$

which may be positive or negative.

We next show that  $dW/d\pi > 0$  holds at  $\pi = \tilde{\pi}_{12} - \varepsilon$  for  $\varepsilon > 0$  sufficiently small. To see this, note that as  $\pi \uparrow \tilde{\pi}_{12}$  we have  $u_{P,l} \rightarrow u_l^S$  and  $u_{A,l} \rightarrow u_l^S$ . Hence, we have that

$$\lim_{\pi \uparrow \tilde{\pi}_{12}} \frac{dW}{d\pi} = (\theta_h - \theta_l) \frac{s'(u_l^S; \theta_l)}{s''(u_l^S; \theta_l)} \lim_{\pi \uparrow \tilde{\pi}_{12}} \left( \frac{\mu(1-\mu)}{1-\pi} - \frac{\mu}{\pi} \right) = 0,$$

implying that  $W$  is indeed locally increasing in  $\pi$  at  $\tilde{\pi}_{12} - \varepsilon$ , for  $\varepsilon$  sufficiently small, if we can show that  $W$  is locally concave in a neighborhood to the left of  $\tilde{\pi}_{12}$ . Using continuity, it thus remains to be shown that  $\lim_{\pi \uparrow \tilde{\pi}_{12}} \frac{d^2W}{d\pi^2} < 0$ . Using

$$\begin{aligned} \frac{d^2W}{d\pi^2} &= -(1-\mu)[s'(u_{P,l}; \theta_l) - s'(u_{A,l}; \theta_l)] \\ &\quad + (\theta_h - \theta_l) \frac{(s''(u_{P,l}; \theta_l))^2 - s'(u_{P,l}; \theta_l)s'''(u_{P,l}; \theta_l)}{(s''(u_{P,l}; \theta_l))^2} \frac{\mu(1-\mu)}{1-\pi} \\ &\quad + (\theta_h - \theta_l) \frac{s'(u_{P,l}; \theta_l)}{s''(u_{P,l}; \theta_l)} \frac{\mu(1-\mu)}{(1-\pi)^2} + (\theta_h - \theta_l) \frac{s'(u_{A,l}; \theta_l)}{s''(u_{A,l}; \theta_l)} \frac{\mu}{\pi^2} \\ &\quad - (\theta_h - \theta_l) \frac{(s''(u_{A,l}; \theta_l))^2 - s'(u_{A,l}; \theta_l)s'''(u_{A,l}; \theta_l)}{(s''(u_{A,l}; \theta_l))^2} \frac{\mu}{\pi}, \end{aligned}$$

we have that

$$\begin{aligned} \lim_{\pi \uparrow \tilde{\pi}_{12}} \frac{d^2W}{d\pi^2} &= (\theta_h - \theta_l) = \frac{(s''(u_l^S; \theta_l))^2 - s'(u_l^S; \theta_l)s'''(u_l^S; \theta_l)}{(s''(u_l^S; \theta_l))^2} \lim_{\pi \uparrow \tilde{\pi}_{12}} \left( \frac{\mu(1-\mu)}{1-\pi} - \frac{\mu}{\pi} \right) \\ &\quad + (\theta_h - \theta_l) \frac{s'(u_l^S; \theta_l)}{s''(u_l^S; \theta_l)} \lim_{\pi \uparrow \tilde{\pi}_{12}} \left( \frac{\mu(1-\mu)}{(1-\pi)^2} + \frac{\mu}{\pi} \right), \end{aligned}$$

which from  $\lim_{\pi \uparrow \tilde{\pi}_{12}} \left( \frac{\mu(1-\mu)}{1-\pi} - \frac{\mu}{\pi} \right) = 0$ ,  $s' > 0$ , and  $s'' = -C''' < 0$  transforms to

$$(\theta_h - \theta_l) \frac{s'(u_l^S; \theta_l)}{s''(u_l^S; \theta_l)} \lim_{\pi \uparrow \tilde{\pi}_{12}} \left( \frac{\mu(1-\mu)}{(1-\pi)^2} + \frac{\mu}{\pi} \right) < 0.$$

Since this inequality is always satisfied  $W$  is locally increasing in  $\pi$  at  $\tilde{\pi}_{12} - \varepsilon$  with  $\varepsilon > 0$  sufficiently small.

# References

- Anderson, S. P. and Renault, R. (2000), Consumer Information and Firm Pricing: Negative Externalities from Improved Information, *International Economic Review* 31, 721-741.
- Armstrong, M. and Rochet, J.-C. (1999), Multi-dimensional Screening: A User's Guide, *European Economic Review* 43, 959-979.
- Armstrong, M. and Vickers, J. (2001), Competitive Price Discrimination, *Rand Journal of Economics* 32, 579-605.
- Bar-Isaac, H., Caruana, G., and Cunat, V. (2010), Information Gathering and Marketing, *Journal of Economics and Management Strategy* 19, 375-401.
- Baron, D.P. and Besanko, D. (1984), Regulation and Information in a Continuing Relationship, *Information Economics and Policy* 1, 267-302.
- Bergemann, D. and Välimäki, J. (2006), Information in Mechanism Design, in: Blundell, R., Newey, W., and T. Persson (eds.), *Proceedings of the 9th World Congress of the Econometric Society*, Cambridge: Cambridge University Press.
- Clay, K.B., Sibley, D.S., and P. Srinagesh (1992), Ex Post vs. Ex Ante Pricing: Optional Calling Plans and Tapered Tariffs, *Journal of Regulatory Economics* 4, 115-138.
- Courty, P. and Li, H. (2000), Sequential Screening, *Review of Economic Studies* 67, 697-717.
- Crémer, J. and Khalil, F. (1992), Gathering Information before Signing a Contract, *American Economic Review* 82, 566-578.
- Dai, C., Lewis, T., and G. Lopomo (2006), Delegating Management to Experts, *Rand Journal of Economics* 37, 503-520.
- Eliasz, K. and Spiegler, R. (2008), Consumer Optimism and Price Discrimination, *Theoretical Economics* 3, 459-497
- Gabaix, X. and Laibson, D. (2006), Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets, *Quarterly Journal of Economics* 121, 505-540.
- Grubb, M. (2009), Selling to Overconfident Consumers, *American Economic Review* 99, 1770-1807.
- Inderst, R. (2004), Contractual Distortions in a (Labor) Market with Frictions, *Journal of Economic Theory* 116, 155-176.
- Janssen, M. and Moraga-González, J.-L. (2004), Strategic Pricing, Consumer Search and the Number of Firms, *Review of Economic Studies* 71, 1089-1118.

- Johnson, J. and Myatt, D. (2006), On the Simple Economics of Advertising, Marketing, and Product Design, *American Economic Review* 93, 756-784.
- Lambrecht, A., Seim, K., and Skiera, B. (2007), Does Uncertainty Matter? Consumer Behavior under Three-Part Tariffs, *Marketing Science* 26, 698-710.
- Lewis, T. and Sappington, D. (1994), Supplying Information to Facilitate Price Discrimination, *International Economic Review* 95, 309-327.
- Lewis, T. and Sappington, D. (1997), Information Management in Incentive Problems, *Journal of Political Economy* 105, 796-821.
- Matthews, S. and Persico, N. (2007), Information Acquisition and Refunds for Return, PIER Working Paper 07-021.
- Miravete, E.J. (1996), Screening Consumers Through Alternative Pricing Mechanisms, *Journal of Regulatory Economics* 9, 111-132.
- Miravete, E.J. (2005), The Welfare Performance of Sequential Pricing Mechanisms, *International Economic Review* 46, 1321-1360.
- Narayanan, S., Chintagunta, P.K., and Miravete, E. (2007), The Role of Self-Selection and Usage Uncertainty in the Demand for Local Telephone Service, *Quantitative Marketing and Economics* 5, 1-34.
- Riordan, M.H. and Sappington, D.E. (1987), Awarding Monopoly Franchises, *American Economic Review* 80, 375-387.
- Rochet, J.-C. and Stole, L.A. (2002), Nonlinear Pricing with Random Participation, *Review of Economic Studies* 69, 277-311.
- Rochet, J.-C. and Stole, L.A. (2003), The Economics of Multidimensional Screening, in: Dewatripont, M., Hansen, L.P., and Turnovsky, S.J., *Advances in Economics and Econometrics: Theory and Applications - Eight World Congress*. New York: Cambridge University Press.
- Thaler, R.H. and C.R. Sunstein (2008). *Nudge: Improving Decisions about Health, Wealth and Happiness*. Yale University Press: New Haven.
- Uthemann, A. (2005), Competitive Screening of Customers with Non-Common Priors, mimeo.
- Varian, H. (1980), A Model of Sales, *American Economic Review* 70, 651-659.