Selling Service Plans to Differentially Informed Customers*

Roman Inderst†
University of Frankfurt and LSE

Martin Peitz‡
University of Mannheim

May 2008

Abstract

We characterize a monopolist's optimal offer of service plans when only some subscribers know their demand (type) already at the contracting stage, while others may learn their demand (type) only after incurring some costs, if at all. While informed customers purchase simpler tariffs, those who are still uninformed purchase tariffs that subsequently allow them to more flexibly adjust their consumed quantity of the service. The presence of uninformed costumers makes it more costly for the firm, in terms of rent left to consumers, to offer the most basic package, which is purchased by informed low-demand customers. Consequently, the firm makes this package relatively unattractive, resulting in a very low quantity of the consumed service. We find that uninformed customers benefit from the presence of informed customers even though information only helps to predict a customer's own demand (type). However, welfare may be lower if there are more informed customers or if acquiring information at the contracting stage becomes less costly for uninformed customers.

Keywords: Nonlinear Pricing, Price discrimination; Multidimensional screening; Heterogeneous information; Information acquisition.

*Martin Peitz gratefully acknowledges financial support for the Deutsche Forschungsgemeinschaft (SFB TR 15).

†Johann Wolfgang Goethe-University, IMFS, Mertonstrasse 17, 60054 Frankfurt am Main, Germany. E-mail: inderst@finance.uni-frankfurt.de.

‡Department of Economics, University of Mannheim, D-68131 Mannheim, Germany. Email: martin.peitz@googlemail.com. Also affiliated to: CEPR, CESifo, ENCORE, and ZEW.

1 Introduction

In this paper we consider subscribers' choice between different service plans when at least some of them do not yet know their future demand. Customers' choice could be between different fixed or mobile telephone call plans or contracts for the supply of electricity or other utilities, where tariff choice and consumption are temporarily separated. Some customers may have only little recollection of their past usage of the service or may subscribe for the first time. Likewise, for some customers future demand may generally be less predictable.[1] Once signed up for the service, however, also previously uninformed customers will learn their respective level of demand over the duration of the contract.[2]

We analyze the pricing problem of a monopolistic firm. This gives rise to a multi-dimensional screening problem as customers may have high or low demand and may also be informed or uninformed about their respective demand "type". As is well known, if all customers were *ex-ante* uninformed about their future demand, then the optimal menu would specify first-best consumption levels and would allow the firm to extract all consumer rent. As is also well known, if all customers already knew their demand type at the outset, then the consumption level of customers with low demand would be downwards distorted, provided they are served at all.

With both informed and uninformed customers present, informed customers purchase simpler tariffs, while those who are still uninformed subscribe to tariffs that subsequently allow to more flexibly adjust the consumed quantity of the service.[3] If the firm wants to ensure that all customers ultimately purchase a strictly positive level of services, then contracts for *all* low-demand customers are more distorted than in the two benchmark cases: both the contracts for informed low-demand customers, compared to the standard screening benchmark, and those for uninformed low-demand customers, compared to the benchmark where all customers were uninformed. As usual, the firm optimally trades off surplus maximization with rent extraction, albeit now the relevant, binding constraints are those *across* informed and uninformed customers.

[1] For instance, depending on life circumstances as well as housing conditions, a customer's demand for electricity may be more variable than that of other customers.

[2] We should stress that also uninformed consumers are not "naive" or "unsophisticated" in any way other than their current ignorance about future demand.

[3] 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 some customers are originally less certain about their future demand than others.

The presence of uninformed customers makes it optimal to make the "basic" package, which is intended for informed low-demand customers, particularly unattractive, resulting in a very low consumed quantity. This is due to the importance of the incentive constraints across informed and uninformed customers. To see this, note first that uninformed customers have also the option to pick any of the contracts designed for informed customers. Their "safest" choice is to select the low-demand type's contract, which would still give them strictly positive consumer rent if they ultimately have a higher willingness to pay (though then their level of consumption is inefficiently low). By making this alternative less attractive for uninformed costumers, the firm can extract a higher price.

In addition, also the rent that is left to informed high-demand customers depends on how attractive this "basic" package still is. However, we show that the binding constraint is *not* always that between the two offers designed for informed customers. Instead, informed high-type customers must be prevented from pretending to be uninformed, which opens up a more indirect channel through which the offer designed for informed low-demand customers affects the rent obtained by informed customers with high demand.

In an extension of the model, uninformed customers can learn their future demand (type) already at the stage of contracting, albeit only after incurring some costs. These costs could simply involve the time and effort spent on going through past bills or thinking ahead about their future consumption needs. If these costs are sufficiently low, then this alternative option for uninformed customers additionally constrains the firm. Intuitively, as these costs become smaller, contracts designed for informed and uninformed customers become more similar.

As their costs of information acquisition become smaller, uninformed customers benefit, even though in equilibrium they do not make use of this option. An uninformed customer also benefits if more of the other customers are informed, even though a customer's information only relates to her own demand (and not, say, to some "shared" aspects such as the availability of different, competitive offers).

As there are either more informed customers or as the costs of information acquisition decrease for uninformed customers, the impact on informed customers' utility and welfare is, in general, ambiguous. With the deregulation of many utilities, including fixed line telephone, electricity, or gas, public agencies have set up internet services to assist households with their decision making. (For instance, they may provide "calculators" that force households to key in an expected demand profile and, thereby, calculate their expected bill for a given tariff.[4]) These policies may both increase the number of informed customers and lower the information acquisition costs for still uninformed customers. Our analysis provides a mixed picture of the implications of such a policy, at least under the circumstances that our model depicts. In particular, those customers who are already informed may be worse off, while also welfare may be lower.[5]

The extension of our model where uninformed customers can acquire information links our paper to the literature on mechanism design with costly information acquisition.[6] Most related here is the seminal paper by Crémer and Khalil (1992). Applied to a procurement setting, the paper considers optimal contracting with two (cost) types for a single agent and the possibility that the agent can learn his type before signing a contract. In our setting, instead, the key characteristic is the *simultaneous* presence of both informed and uninformed agents (customers).[7]

That customers may learn more about their actual willingness to pay for a good after signing an initial contract has also been recognized in the "sequential screening" literature (Courty and Li, 2000).[8] More recently, in Matthews and Persico (2007) customers can also become, albeit again at a cost, earlier informed about their willingness to pay. Besides the fact that in our model informed and uninformed customers coexist, our contribution differs also in that we focus on multi-unit purchases and thus on the optimal design of non-linear contracts.[9]

[4] In addition, these websites often offer price comparison services as well.

[5] This suggests that such a policy may also undermine the incentives to become informed. Cf. the discussion in the Conclusion.

[6] Bergemann and Välimäki (2006) survey some of the literature on mechanism design with endogenous information acquisition.

[7] There is also a strand of the literature in which the *principal* (i.e., the firm in our model) has information or can at costs 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).

[8] Cf. also Baron and Besanko (1984), Riordan and Sappington (1987), as well as Miravete (1996, 2005). Miravete (1996) is of particular interest as this paper also considers non-linear pricing: Consumers have *ex-ante* knowledge about some demand type, which together with some additional "shock" generates their willingness to pay at the time of consumption.

[9] In Lewis and Sappington (1997) there are both informed and uninformed agents, though there the focus is on how to elicit from the informed agent (more) effort that goes into information acquisition about some state that is of relevance for the principal. Somewhat more closely related, in Dai and Lewis (2003) producers differ initially in the precision with which they can later forecast their costs of production. In our setting, however, the better information that some consumers have *ex-ante* creates also *ex-ante* heterogeneity in a second dimension: low- and high-demand types. (Consequently, in our model offers to both informed and uninformed consumers will be distorted, while in their model only the menu offered to

There is also a small but growing literature that combines demand uncertainty with behavioral "biases" such as overconfidence, procrastination, projection bias, etc. In Grubb (2007) customers underestimate the variability of their future demand. While they may differ in their prior estimate of having lower or higher demand, they do not differ with respect to how knowledgeable they are with respect to future demand. In Uthemann (2005) customers have biased priors about having low or high demand later, similar to Eliaz and Spiegler (2006), where they have in addition time-inconsistent preferences. In all these papers, contract design is driven by firms' attempt to extract profits through catering to customers' distorted beliefs.

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 discussion, while Section 5 extends the analysis by allowing uninformed customers to acquire information, albeit at costs, before choosing from the offered contracts. Section 6 concludes. All proofs are relegated to the Appendix.

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 to have in mind an application to mobile call plans in what follows.

The firm has constant marginal cost $\tilde{c}$. A customer of (real-valued) demand type θ, 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 $\lim_{q\to\infty} \tilde{u}'(q) = 0$, which together imply that the first-best level of service will be both finite and strictly positive for any choice $\theta > 0$ and (finite) $\tilde{c}$. We also suppose that $\widetilde{u}$ is twice continuously differentiable.

Before proceeding with the description of the model, it is useful to rephrase the customer's choice problem. Instead of choosing quantity q, we suppose that the customer

the less knowledgeable producer is inefficient.)

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 realize the gross utility θu of a customer of type θ the firm must incur cost of $C(u)$, where the properties of $\tilde{u}$ imply that C is strictly increasing and strictly convex with $C'(0) = \lim_{u\downarrow 0} C'(u) = 0$. Surplus thus becomes $s(u;\theta) := \theta u - C(u)$, which for θ_i is uniquely maximized by some bounded and strictly positive value u_i^{FB}, $i = h, l$. Note that $0 < u_l^{FB} < u_h^{FB}$.

The *ex-ante* probability with which the customer has high demand for the particular service is given by $\mu \in (0,1)$. (We normalize the mass of all customers to one.) Our key departure from the extant literature is that a customer's *a priori* knowledge about his demand type constitutes a second dimension of customer heterogeneity. Precisely, we suppose that only the fraction π of customers know their type at the state of contracting, while the fraction $1-\pi$ share at this stage only the common prior beliefs. Later, at the stage of consumption, all customers are, however, equally informed about their type. Whether or not a customer already knows his type at the stage of contracting as well as the nature of the respective type are all the customer's private information. (In Section 5 an uninformed customer may also learn his type early, albeit only at costs.)

Without loss of generality we can restrict consideration to the following set of offers by the firm. For *ex-**A**nte* informed customers, the firm designates at most two different consumption profiles $u_{A,i}$ and respective total transfers $t_{A,i}$, where $i = l, h$. For the only *ex-**P**ost* informed customers the firm specifies instead a contract consisting out of at most two options: $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$. Hence, the firm offers a menu of contracts. Each customer decides which, if any, contract to sign. After the contract stage uninformed customers learn their demand type and choose their preferred service level. Contracts $(u_{A,i}, t_{A,i})$ specify a fixed allowance. Instead, the contract $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$ allows the customer still the flexibility to choose. The menu can be decomposed as follows: the customer pays $t_{P,l}$ for an allowance up to $u_{P,l}$; if she then wants to consume more, incremental costs are $\Delta_t = t_{P,h} - t_{P,l}$ for the additional allowance $\Delta_u = u_{P,h} - u_{P,l}$. This contract is thus similar to a three-part tariff.

3 Analysis of the Optimal Contract

3.1 The Firm's Program

With only informed customers the firm would face a standard screening problem in choosing the 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, $IR_{A,h}$, becomes $V_{A,l} \geq 0$. It is well-known that both constraints bind at the optimal offer, that all other constraints can be ignored, and that the high type's consumed level of service is first best with $u_{A,h} = u_h^{FB}$. Furthermore, using that high-demand customers realize a rent equal to $u_l^S(\theta_h - \theta_l)$, the firm optimally distorts the low type's consumption value, $u_l^S < u_l^{FB}$. It holds that

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

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 $\mu \geq \theta_l/\theta_h$.

As a second benchmark, suppose that with $\pi = 0$ there would only be uninformed customers. Consumption profiles of both types are then efficient, $u_{P,i} = u_i^{FB}$. Moreover, in this case the customer's individual rationality constraint must now be only 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. This leaves some degree of freedom to specify the optimal transfers, which together have to satisfy IR_P as well as both incentive compatibility constraints once the respective customer has learnt her type: $IC_{P,i}$ with $V_{P,i} \geq \theta_i u_{P,j} - t_{P,j}$. For instance, one possibility is to adjust transfers according to the incurred incremental costs: $t_{P,h} - t_{P,l} = C(u_h^{FB}) - C(u_l^{FB})$.

With both informed and uninformed customers present, the firm faces an additional set of incentive compatibility constraints *across* the respective offers, which we denoted by subscripts A and P. Regarding the incentives of informed types to mimic an uninformed customer, we will show that only the respective constraint for the high-demand type needs to be considered. Given incentive compatibility of the menu offered to uninformed customers, the informed high-demand customer will thus not prefer to pretend to be uninformed if $V_{A,h} \geq V_{P,h}$. We refer to this incentive compatibility constraint *across* informed and uninformed customers as $ICC_{A,h}$.

Looking the other way, for an uninformed customer the alternative to accepting the menu $\{(u_{P,i}, t_{P,i})\}_{i=l,h}$ is to pick one of the (at most two) different contracts that are offered to informed customers, $(u_{A,i}, t_{A,i})$. Following again the standard procedure, we will first consider the relaxed program where we only consider the alternative to mimic the informed *low* type: This yields the constraint ICC_P with

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

We will then show that under the optimal contract the uninformed customer indeed (strictly) prefers not to mimic the informed *high* type.[10]

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

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

subject to the following set of constraints: (i) The downward incentive compatibility constraints for both informed and uninformed customers $IC_{A,h}$ and $IC_{P,h}$; (ii) the individual rationality constraints for the informed low type $IR_{A,l}$ and the uninformed customer IR_P; (iii) and the two cross incentive compatibility constraints, namely for the uninformed customer ICC_P and the informed high type $ICC_{A,h}$. In addition, note that all u must be non-negative.[11]

3.2 Solution

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 $u_{\cdot,i} > 0$. Here, we encounter two cases, to which we refer to as Cases 1 and 2. Subsequently, we

[10] It is also useful to note that if an acceptable contract to the informed low type is offered, then clearly ICC_P *strictly* implies IR_P.

[11] To save space we have chosen not to first write out explicitly the full program. Note, however, that in the relaxed program the following constraints are ignored: the downward incentive compatibility constraints; the individual rationality constraint for the high type; the cross incentive compatibility 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.

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.[12]

Characterization if all customers are served

In this case, we obtain the following characterization for the optimal contracts.

Proposition 1 *The optimal offer under which all customers purchase a positive level of services has the following properties:*
Case 1) If $\pi \geq \frac{1}{2-\mu}$, the firm offers the same contracts to informed and uninformed customers. These are standard contracts with $u_{\cdot,h} = u_h^{FB}$ and $u_{\cdot,l} = u_l^S$.
Case 2) If instead $\pi < \frac{1}{2-\mu}$ holds, then only high-demand customers receive the same contract regardless of whether they are informed or not, which satisfies $u_{P,h} = u_{A,h} = u_h^{FB}$. Instead, the contract for the informed low type is more distorted than that for the uninformed low type as $u_{A,l} < u_l^S < u_{P,l} < u_l^{FB}$.

Recall for Case 2 that u_l^S denotes the distorted consumption level for low-demand types under a standard screening contract (i.e., for $\pi = 1$).

The key to an understanding of 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 or informed customers. Recall first the benchmark $\pi = 0$, where there are uninformed customers only. In this case, the first-best allocation results. In the presence of informed customers, what distorts $u_{P,l}$ is the incentive compatibility constraint $ICC_{A,h}$, which requires that an informed high-demand customer does not want to mimic an uninformed customer (and subsequently pick from the menu the offer designed for the high type).

Take next the other benchmark: $\pi = 1$, where there are informed customers only. There, the contract for the low type is distorted and given by u_l^S. In the presence of uninformed customers, i.e., with $\pi < 1$, the still *lower* consumption level $u_{A,l} < u_l^S$ follows again from the binding incentive constraints across informed and uninformed customers, $ICC_{A,h}$ and ICC_P.

We provide next more details. The uninformed customer's best alternative option, as captured by the constrained ICC_P, is to sign up instead for the contract of the informed

[12] Here and in what follows, we do not characterize separately the (non-generic) cases where the firm is indifferent between different offers (i.e., here with $\pi = \frac{1}{2-\mu}$).

low type. In this case, the uninformed customer will make *ex-post* a positive surplus if she turns out to have high demand: as $V_{A,l} = 0$ this rent is $(\theta_h - \theta_l)u_{A,l}$. Next, the level of $u_{A,l}$ determines also the rent that the informed high type earns, albeit the level of this rent is *not* determined, as would be standard, by the binding constraint $IC_{A,h}$. Instead, we show that in Case 2 of Proposition 1 the constraint $IC_{A,h}$ remains slack. The utility of the informed high type is instead determined from the binding *cross* constraint $ICC_{A,h}$: $V_{A,h} = V_{P,h}$. It is through this *indirect* channel, together with the binding incentive constraint $IC_{P,h}$ for the uninformed customers' menu and the binding cross constraint ICC_P for the uninformed customer, that the level of $u_{A,l}$ affects also the informed high type's utility $V_{A,h}$.

Taking sum of how $u_{A,l}$ affects the rents of the uninformed as well as the informed high type and trading this off with the surplus $s(u_{A,l}; \theta_l)$ that is realized with informed low-demand customers, we show in the proof of Proposition 1 that $u_{A,l}$ solves

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

Comparing this to (1) confirms that $u_{A,l} < u_l^S$ holds in Case 2, given that there $\pi < \frac{1}{2-\mu}$.

Turning to the uninformed customer's low consumption level $u_{P,l}$, in Case 2 it is again only indirectly, via the cross incentive constraint of the informed high type, $ICC_{A,h}$, that a higher level for $u_{P,l}$ increases customers' rents, more specifically that of informed high-type customers. Taking this into account, $u_{P,l}$ optimally trades off surplus maximization with rent extraction if

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

Comparing this to (1) confirms now that $u_{P,l} > u_l^S$ holds in Case 2, where $\pi < \frac{1}{2-\mu}$.

Cases where not all customers are served

Before commenting more on the so far obtained characterization of contracts, we first have to complete the description of the full solution to the firm's program. While so far we assumed that the firm wants to ensure that all customers purchase a strictly positive quantity, the firm may in fact sometimes exclude some low-demand customers so as to extract more rent from all remaining customers. 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, then at most two further cases may arise:*
Case 3) Only high-type customers purchase a positive level of services, $u_{\cdot,h} = u_h^{FB}$*, and realize zero customer surplus.*
Case 4) Both informed and uninformed high-type customers receive again the same, first-best contract, while now also low-type uninformed customers receive a contract stipulating $u_{P,l} < u_l^{FB}$*.*

Of particular interest is Case 4. Here, in order to extract more rent from uninformed customers, the firm is no longer satisfied with making the contract offered to informed low-demand customers very unattractive (through a low level of $u_{A,l}$), but it chooses instead to no longer offer these customers an acceptable contract. Uninformed customers will then no longer receive a positive rent. The optimal choice of $u_{P,l}$ for uninformed customer thus trades off surplus maximization with rent extraction from informed high-demand customers. This rent depends on $u_{P,l}$ again through the binding *cross* constraint $ICC_{A,h}$ (in addition to $IC_{P,h}$). The resulting trade-off is the same as in Case 2, which is why $u_{P,l}$ is again determined from the first-order condition (3). Recall that this implies, in particular, that $u_{P,l}$ is strictly decreasing in both the fraction of informed and the fraction of high-demand customers.

As is intuitive, an uninformed customer generates always (weakly) more revenues for the firm compared to an informed customer. This holds both in expectation over high- and low-demand types and for the respective low-demand customers. (Recall that high-type customers always obtain the same *ex-post* contract with $u_{\cdot,h} = u_h^{FB}$ and $t_{A,h} = t_{P,h}$.)

Corollary 1 *Suppose Cases 2 or 4 apply. Then the firm realizes always 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 customers' side, it is from Corollary 1 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: $V_{P,l} < 0$.[13] Clearly, from an *ex-ante* perspective, uninformed low-demand customers would thus have been better advised to

[13] Cf. also Miravete (1996).

purchase instead the "basic" tariff $(u_{A,l}, t_{A,l})$. However, given their own initial demand uncertainty, the offer designed for uninformed customers was equally attractive (recall that ICC_P is binding) as it entailed also the option to make use of an additional allowance $u_h^{FB} - u_{P,l} > 0$ at an incremental price $t_{P,h} - t_{P,l}$ smaller than the respective utility increment $\theta_h(u_h^{FB} - u_{P,l})$.

Full characterization of the solution to the firm's problem

Which of the different characterized cases 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 μ, the fraction of informed customers π is low, then informed low-demand customers are excluded (Case 4). Otherwise, all customers are served, with Case 1 applying for intermediate values of π and Case 2 for high values.
ii) Suppose instead that $\mu \geq \theta_l/\theta_h$: Then for given μ all low-type customers are excluded if π is sufficiently high (Case 3). For lower values of π, however, only informed low-demand customers are excluded (Case 4).

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 μ and π that determine which of the four cases apply are given explicitly in the proof of Proposition 3. We next provide more intuition for the case distinction in Proposition 3.

The impact of the fraction μ 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 to extract more rents from the former. That is, moving upwards in Figure 1, we move from Cases 1 and 2 to Case 3 and 4, respectively.

Next, for high μ it is also intuitive that it becomes optimal to no longer exclude uninformed low-type customers, but only informed low types if there are sufficiently few informed customers altogether (i.e., as we move to the left in Figure 1). Interestingly, while for high μ a reduction in π thus leads to less exclusion, we can observe the opposite for the case of low μ. There, as we move to the left while staying in the lower part of

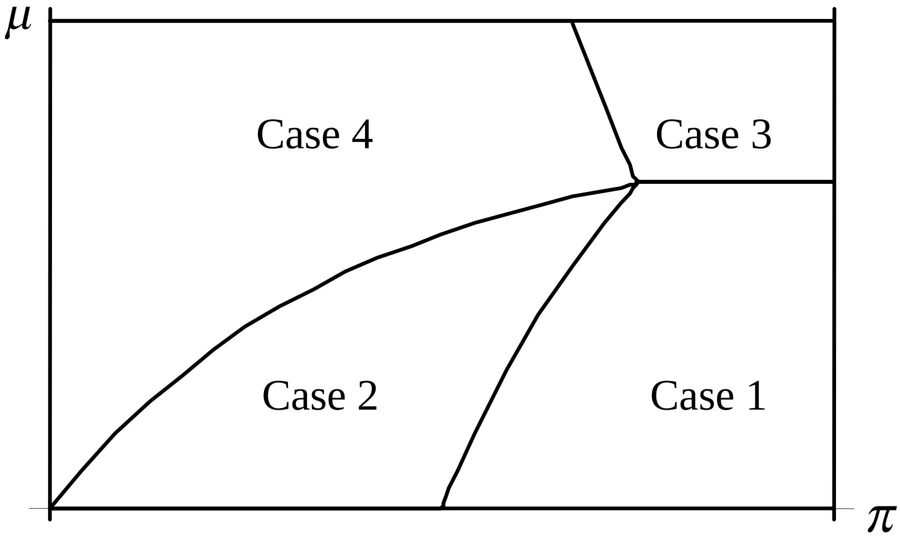

Figure 1: Optimal contracts for $\theta_h = 3/4$ and $\theta_l = 1/2$.

Figure 1 (low μ), the offer made to informed low-demand customers becomes increasingly distorted in an attempt to extract more rents from uninformed customers, who account for an increasingly large fraction. As π becomes too low, informed low-demand customers no longer purchase a positive quantity.

While Proposition 3 states how we move between cases as we decrease or increase π, it does not report on how the respective intervals of values π, for which the different cases apply, change. In a compact way this can be seen from inspecting Figure 1, given that the respective results apply generally and not only for the chosen numerical example. In the rest of this Section, we provide some more formal characterization.

For this we have to introduce some additional notation for the boundaries that separate the different cases. 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 $\widetilde{\pi}_{12} = \frac{1}{2-\mu}$. Applying a similar notation, we have that $\widetilde{\pi}_{24}$ separates Cases 2 and 4, while $\widetilde{\pi}_{34}$ separates Cases 3 and 4. Note that $\widetilde{\pi}_{24}$ is determined from the requirement that $u_{A,l} = 0$ holds in Case 2, where $u_{A,l}$ is strictly decreasing in μ but strictly increasing in π. This implies that $\widetilde{\pi}_{24}$ is indeed upward sloping as a function of μ, as depicted in Figure

1. Finally, the boundary between Cases 3 and 4, $\widetilde{\pi}_{34}$, is obtained from setting $u_{P,l} = 0$ in Case 4. As $u_{P,l}$ is more distorted as there are more informed customers and more high-type customers, $\widetilde{\pi}_{34}$ is indeed strictly decreasing in μ, as again depicted in Figure 1.[14]

3.3 Further Discussion

One insight from our analysis is that serving informed customers with low demand comes at high opportunity costs for the firm, namely in terms of lost profits with both informed high-demand customers and uninformed customers. The firm should thus make the corresponding "basic" contract $(u_{A,l}, t_{A,l})$ relatively unattractive, in particular if the fractions of uninformed customers or high-demand customers are relatively high. This is formally captured by the following Corollary, which follows from implicitly differentiating (2) and (3).

Corollary 2 *If all customers purchase a strictly positive level of services and if* $\pi < \frac{1}{2-\mu}$ *(so that Case 2 of Proposition 1 applies), then as the fraction of informed customers increases (higher* π*), the higher is* $u_{A,l}$ *and the lower is* $u_{P,l}$*. In addition, both* $u_{A,l}$ *and* $u_{P,l}$ *are strictly lower as there are more high-demand customers (higher* μ*).*

The final part of Corollary 2, where the comparative analysis is made with respect to μ, is analogous to results obtained for the standard (one-dimensional) screening model.

Finally, it is useful to summarize the results in Corollary 2 together with those in Propositions 2 and 3 as follows.

Corollary 3 *Suppose Cases 2 or 4 apply. Then as the fraction of uninformed customers or of high-demand customers increases (lower* π *or higher* μ*, 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* $u_{A,l} = 0$*).*

Once the way incentive constraints bind in our specific model has been worked out, Corollary 3 follows intuitively from standard principles of models of screening. From an

[14] A formal statement of the monotonicity of the different boundaries is part of the proof of Proposition 3.

ex-ante perspective, informed low-demand types 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, the distortion "at the bottom" increases as the probabilities of the "higher types" increase (specifically, through an increase in μ or π).[15]

4 Comparative Analysis

To analyze the effect of an increase of the share of informed customers π, is interesting for two reasons. First, in the light of 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 many recently deregulated industries aims to encourage customers to become more knowledgeable, including about their own demand profile.[16] The comparative analysis in π may help to shed more light on the implications of such policies.

Impact on Uninformed Customers

We have seen above that the expected service level of uninformed customers *decreases* in the share of informed customer, π, since $u_{P,l}$ decreases and $u_{P,h} = u_h^{FB}$ is constant in π. However, it turns out that the expected utility of uninformed customers *increases* as more customers become informed.

Corollary 4 *As the fraction of informed customers increases (higher π), uninformed customers' utility increases.*

Hence, as some additional customers become informed, which increases π, those who remain uninformed benefit from their presence. Note here that a customer's information is only with respect to her own demand type. This is different, for instance, in models

[15] Though consumers differ along two dimensions in our model, 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. This is different in standard problems of multi-dimensional screening (cf. Amstrong and Rochet 1999).

[16] 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).

with search and shopping costs, where the presence of customers who are better informed about rivals' offers brings down expected prices, from which all customers benefit (cf. Varian 1980; or more recently Janssen and Moraga-González 2004).[17]

Corollary 4 holds strictly for Case 2 and follows immediately from the proof of Proposition 1. Formally, this holds as we know that $u_{A,l}$ increases in π and that through the binding constraint ICC_P this leads to a higher rent for the uninformed customer. Recall also that Case 2 applies for intermediate values of π, provided that μ is not too high. (Cf. also Figure 1.) For low values of π, where Case 4 applies instead, uninformed customers realize zero rent, while for high π, where Case 1 applies and uninformed customers realize the highest rent, we know that a further increase in π does not further affect contracts and thus utilities.[18]

The intuition for why in Case 2 uninformed customers benefit from the presence of more informed customers can be restated in the following, more straightforward way. As the fraction of informed customers, π, increases, the firm optimally puts less weight on reducing the rent left to uninformed customers and more weight on increasing the surplus realized with informed low-demand customers. (Of course, this uses that the constraint ICC_P binds.)

The insight from Corollary 4 may be interesting in the light of frequent claims that more informed or sophisticated customers are cross-subsidized at the costs of less informed customers. For instance, in Gabaix and Laibson (2006) this holds, albeit under competition, if some customers are knowledgeable about their future demand of an "add-on service", while other customers are naive in that they are not aware of this. In a monopolistic context and with perfectly rational customers, Corollary 4 provides a different benchmark, namely one where the presence of informed customers benefits uninformed customers.[19]

[17] 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.

[18] For high μ, where only Cases 3 and 4 apply as π changes, uninformed customers always realize zero utility.

[19] 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 full market is 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 this case applied to our model must await future research.

Impact on Informed Customers

We turn next to the impact that an increase in π has on *informed* costumers and on welfare. Here, the situation is more complicated as the impact is generally ambiguous. To obtain nevertheless some insights, we first discuss one case where the effect is clear cut, specifically that of Case 4. Subsequently, we explore cases where the impact is ambiguous. Note throughout the analysis that informed low-demand customers always realize zero utility, implying that the comparative analysis focuses on the utility of informed high-demand customers, $V_{A,h}$.

In Case 4 only informed low-demand customers are not served. As π increases, the firm focuses more on rent extraction from informed customers and less on preserving surplus that is realized with uninformed customers (specifically, with uninformed low-demand types through the choice of $u_{P,l}$).

Corollary 5 *Take Case 4, where only informed low-demand customers are not served. As the fraction of informed customers, π, increases, this reduces informed customers' utility.*

We turn now to Case 2. (In Cases 1 and 3, π has no effect on contracts and thus utilities.) In Case 2, the utility of informed high-demand customers, $V_{A,h}$, depends (positively) on both $u_{A,l}$ and $u_{P,l}$, namely through the binding constraint ICC_P and due to the two binding constraints $IC_{P,l}$ and $ICC_{A,h}$.[20] These are, as we know from Corollary 2, differently affected by a change in π. We can show that the set of parameters for which a marginal increase in π has a positive effect on $V_{A,h}$ and the set of parameters for which it has a negative effect are both non-empty (cf. proof of Corollary 6.) In particular, the case where the effect is positive may at first be surprising given that for higher π we would expect the firm to place more weight on extracting rent from informed high-demand customers.

To shed more light on this we use the particular functional specification that C''' is zero.[21] With this specification, we obtain for the contracts in Cases 2 and 4 the explicit

[20]More precisely, from the proof of Proposition 1 we can use that $V_{A,h} = \mu(\theta_h - \theta_l)u_{A,l} + (1-\mu)(\theta_h - \theta_l)u_{P,l}$.

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

solutions

$$\begin{aligned}
u_{A,l} &= \max\left\{0, \frac{1}{2c}\left(\theta_l - \frac{\mu}{1-\mu}\frac{1-\pi+\pi\mu}{\pi}(\theta_h-\theta_l)\right)\right\}, \\
u_{P,l} &= \frac{1}{2c}\left(\theta_l - \mu\frac{\pi}{1-\pi}(\theta_h-\theta_l)\right).
\end{aligned}$$

From the proof of Proposition 1 we have next that $V_{A,h} = \mu(\theta_h - \theta_l)u_{A,l} + (1-\mu)(\theta_h - \theta_l)u_{P,l}$, which is, as noted above, increasing in both $u_{A,l}$ and $u_{P,l}$. (Recall that by the binding ICC_h this is also the utility realized by an uninformed high-type customer such that $V_{A,h} = V_{P,h} = V_h$.) Differentiating $V_{A,h}$ w.r.t. π while substituting for $u_{A,l}$ and $u_{P,l}$ explicitly, we obtain that $dV_{A,h}/d\pi > 0$ if and only if

$$\pi < \widehat{\pi} := \frac{\sqrt{\mu}}{1+\sqrt{\mu}-\mu}. \tag{4}$$

That is, those customers who are already informed benefit if previously uninformed customers also become informed, in case there are presently not too many informed customers.[22] Whether in Case 2 we have indeed a (lower) range of values π for which $dV_{A,h}/d\pi > 0$ and a (higher) range of values for which the opposite holds with $dV_{A,h}/d\pi < 0$ depends on whether $\widehat{\pi}$, as defined in (4), falls into the interval (π_{24}, π_{12}). (Recall that for given $\mu < \theta_l/\theta_h$ this interval describes the whole range of values π for which Case 2 applies.) We can show that this is the case if and only if μ is not too large.

Corollary 6 *Take Case 2, where all customers purchase a positive quantity. The effect of an increase in the fraction of informed customers π on existing informed customers is, in general, ambiguous. Suppose that $C''' = 0$. Then as π increases, existing informed customers are strictly better off if both π and the fraction of high-type customers, μ, are not too high.*

Corollary 6 confirms that the more immediate intuition, namely that an increase in informed customers induces the firm to extract more rent from them and thus hurts existing informed customers, indeed applies if the fraction of informed customers *and* the fraction of high-demand customers are both already high. Otherwise, we have from Corollaries 4 and 6 that an increase in π benefits *all* customers and thus unambiguously also increases consumer surplus.

[22] The analysis for the general case, where we do not impose $C''' = 0$, supports this conclusion. There, we can still show that $dV_{A,h}/d\pi > 0$ holds for small and that $dV_{A,h}/d\pi < 0$ holds for high π.

As a final note, the ambiguity of informed customers' utility with respect to π can best be understood by noting that a change in π increases both the fraction of high- and that of low-type customers. Here, the effect on the latter is stronger the lower is μ. Recall now the re-interpretation of our setting in terms of a standard, one-dimensional screening model. With low μ, an increase in π would increase the fraction of the "bottom type" (i.e., informed high-demand customers) more than the fraction of the "top type" (i.e., informed high-demand customers), while reducing the fraction of the "intermediate type" (i.e., uninformed customers).

Impact on Welfare

Recall first that for Cases 1 and 3 a marginal change in π does not affect contracts. Next, recall that in Case 4 only informed low-demand customers are not served. As π increases, the firm focuses more on rent extraction from informed customers, leading to a more distorted level of $u_{P,l}$. There is also a second, more direct effect through which welfare is reduced as π increases: as more customers become informed, more low-demand types end up not being served instead of consuming (at least) the quantity $u_{P,l}$.

Corollary 7 *Take Case 4, where only informed low-demand customers are not served. As the fraction of informed customers, π, increases, welfare decreases.*

We now turn to Case 2. Here, the impact on welfare is generally ambiguous. Welfare is affected trough three channels.

$$\begin{aligned}\frac{dW}{d\pi} &= -(1-\mu)[s(u_{P,l};\theta_l) - s(u_{A,l};\theta_l)] \\ &\quad +(1-\mu)(1-\pi)s'(u_{P,l};\theta_l)\frac{du_{P,l}}{d\pi} \\ &\quad +(1-\mu)\pi s'(u_{A,l};\theta_l)\frac{du_{A,l}}{d\pi}. \end{aligned} \tag{5}$$

The first line of (5) captures again the direct effect: now as π increases, less low-demand costumers consume the higher, more efficient level $u_{P,l}$ and more consume the lower, less efficient level $u_{A,l}$. The terms in the second and third line of (5) concern, instead, the effect working through a change in contracts. As noted previously, these effects are conflicting as $u_{P,l}$ increases but $u_{A,l}$ decreases. In Appendix 2 we use the specification that C''' is zero to derive parameter regions for which $dW/d\pi$ is positive and parameter regions for which it is negative.

5 Information Acquisition

5.1 Extending the Model

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 get a better estimate of her future demand. To allow for this possibility, we stipulate in what follows that at the contracting stage also uninformed customers can observe their demand type, albeit only after incurring private disutility $k > 0$.[23]

The game between firm and customers can then be described as follows: At stage 1, the firm proposes a set of contracts, as previously, 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 to modify the constraint of an uninformed customer. The alternatives for an uninformed customer, next 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: Cases 1 and 2 with $u_{\cdot,i} > 0$. Moreover, while the full program is solved in the proof of the subsequent Proposition 4, in the main text we confine ourselves to the most salient issues.

[23] As noted in the Introduction, this brings our paper close to Crémer and Khalil (1992), though with the difference that in their setting there is only a single agent (customer) and thus no role for price discrimination between already informed customers and those who can become informed, albeit only at costs.

5.2 Analysis

If, in equilibrium, the uninformed customer did acquire information, the firm would only face informed customers and thus a standard screening problem. The resulting simple menu would then clearly deprive customers of the incentives to acquire information.[24] Recall next that the standard screening offer (derived for the case with only informed customers) was indeed still optimal in our previous analysis if the fraction of informed customers π was sufficiently large (see Case 1 in Proposition 1). Intuitively, introducing in this case the possibility for uninformed customers to acquire information does not change results. The remaining case is that of Case 2, where π is sufficiently high and where previously the offer to informed low-demand customers was more distorted: $u_{A,l} < u_{P,l}$.

With the additional option to acquire information, we show in the proof of Proposition 4 that the uninformed customers' incentive compatibility constraint becomes now

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

where we have already used that, as we can show, $V_{\cdot,h} = V_h$.[25] 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.

Take now 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

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

holds. Note that this is trivially always 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 only ignore the new constraint

[24] 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.

[25] More formally, this follows as $ICC_{A,h}$ again binds. Next to this, we have also used that $V_{A,l} = 0$ and that $V_{P,l} \leq 0$. (Note that $V_{P,l} \leq 0$ is not already implied by $V_{A,l} = 0$ as we again solve the relaxed program, where we ignore the incentive compatibility constraint of the informed low type.)

that arises from the possibility of information acquisition only if the respective costs k are sufficiently high. Otherwise, the firm has to adjust its offer.

Proposition 4 *Suppose uninformed customers can at the stage of contracting 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, given that the new constraint does not bind;
Case 2a) If instead $\pi < \frac{1}{2-\mu}$, then the contract specified in Case 2 is still optimal provided that k is sufficiently large, i.e., if it satisfies (7);
Case 2b) If $\pi < \frac{1}{2-\mu}$ and k is small such that it violates (7), then the optimal offer has still 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.

5.3 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, we show that the difference between the respective values $u_{P,l} > u_{A,l}$ is now pinned down by the binding condition (7):

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

This implies, in particular, that for $k \to 0$ 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: $u_{\cdot,l} \to u_l^S$. 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 for why the firm previously made the (most basic) offer $u_{A,l}$ so unattractively low. This leads us to the following Corollary, which is proved in Appendix 1.

Corollary 8 *As the costs of information acquisition k decrease, the difference $u_{A,l} - u_{P,l} > 0$ decreases according to (8) as $u_{A,l}$ is weakly increasing and $u_{P,l}$ is weakly decreasing in k. For $k \to 0$ we have that $u_{\cdot,l} \to u_l^S$.*

From Corollary 8 contracts for informed customers become thus more efficient and contracts for uninformed customers less efficient as k decreases.

The only relevant case is Case 2b. Corollary 8 mirrors our previous comparative analysis in terms of π, the fraction of informed customers. There, we also analyzed how a change in π affects the utility of both informed and uninformed customers, as well as welfare and aggregate consumer surplus. We showed there that an increase in π benefits uninformed customers. While intuitively uninformed customers also benefit from a reduction in their own costs of information acquisition, given that this forces the firm to make the mor attractive offers, they may now exert a *negative* externality on other, informed customers. While the impact on informed customers, as well as welfare and total consumer surplus is generally ambiguous, for the case with $C''' = 0$ we can make more progress. In analogy to Corollary 6 we can show that, in the current case, lower costs of acquiring information hurt informed customers if their fraction is already relatively large.

Corollary 9 *In Case 2b, uninformed customers always 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 π is sufficiently small and (ii) welfare always increases.*

From Corollary 9, we have for the special case where C''' is zero that welfare is always strictly higher as k decreases. This contrasts with our previous comparative analysis in π. In both cases we had a trade-off in terms of a higher $u_{A,l}$ and a lower $u_{P,l}$. A key difference is, however, that with a change in π also the size of the two customer groups, informed and uninformed, was changed. The respective negative effect on welfare is absent as we change k.

6 Conclusion

For many subscription services tariff choice and consumption are temporarily separated. When signing a contract, customers may thus still be uncertain about their future level of demand. This paper considers the contract design problem of a monopolist facing both uninformed customers and customers who are already informed about their demand (type) at the contracting stage. In an extension we also allow for the possibility that uninformed customers can acquire information at costs.

Given the presence of both informed and uninformed customers as well as informed costumers with high or low demand, our problem is one of multidimensional price discrimination. The specific structure of our problem allows, however, to obtain a full characterization of optimal contracts, depending on the composition of the firm's market in terms of both informed and high-demand customers. These parameters determine the distortion of contracts offered to both informed and uninformed low-demand customers as well as whether all customers are served in the first place or not.

A first finding is that the presence of uninformed costumers makes it more costly for the firm, in terms of rent left to customers, to offer the "basic" package, which is designed for informed low-demand customers. Consequently, the firm optimally makes this package relatively unattractive, resulting in a very low service level; or it may more often than would otherwise be the case decide to exclude these customers.

We find that the presence of informed customers benefits uninformed customers even though information is only about a customer's own demand. In terms of consumer rent, in our model it is thus not the case that informed customers free-ride (through being "cross-subsidized") on uninformed costumers; rather, it is uninformed customers who free ride on the better information of other customers. The impact of having more informed costumers (or, likewise, of reducing the costs of information acquisition) on already informed customers and welfare is, however, generally ambiguous.

In future work it could be interesting to endogenize the differential information at the contracting stage. If customers have different costs of acquiring information, we would suggest that those with low costs become informed, while those with higher costs stay uninformed. The firm's design of the price discriminating offer would thus determine also the fraction of customers who are informed, rather than that being an exogenous variable.[26]

Appendix 1: Relegated 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, we have that:*

[26] See also Bar-Isaac et al. (2007), who analyze the decision of the firm to facilitate information acquisition for consumers with heterogeneous preferences.

(i) $u_{A,h} = u_h^{FB}$*;*

(ii) $u_{P,h} = u_h^{FB}$*;*

(iii) $V_{A,l} = 0$*;*

(iv) while the constraint ICC_P *is binding.*

This is shown as follows. 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. This establishes the claim in Step 1.

Step 2: *If the optimal contract menu has the property that all customers take contracts with strictly positive quantities, then* $ICC_{A,h}$ *must be binding, i.e.,* $V_{A,h} = V_{P,h}$*.*

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 $\mu V_{P,h} + (1-\mu)V_{P,l} = \mu\left(\theta_h u_{A,l} - t_{A,l}\right) = \mu V_{A,h}$. 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) is then equal to $(1-\pi)\mu(\theta_h - \theta_l)u_{A,l}$. 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.

Hence we have shown that $ICC_{A,h}$ must be binding in the optimal contract menu.

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 the level of 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

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

Therefore, the total expected rent that goes to customers is

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

implying that $u_{P,l}$ maximizes

$$(1-\mu)\left[(1-\pi)s(u_{P,l};\theta_l) - \pi\mu(\theta_h - \theta_l)u_{P,l}\right], \tag{10}$$

while $u_{A,l}$ maximizes

$$\pi(1-\mu)s(u_{A,l};\theta_l) - (1-\pi)\mu(\theta_h - \theta_l)u_{A,l} - \pi\mu\mu(\theta_h - \theta_l)u_{A,l}. \tag{11}$$

Step 4: We now establish a condition that the constraint $IC_{A,h}$ is indeed slack, as claimed in the previous step. If $IC_{A,h}$ is slack, i.e., $V_h > (\theta_h - \theta_l)u_{A,l}$, we have

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

which is equivalent to $u_{P,l} > u_{A,l}$.

To compare $u_{P,l}$ and $u_{A,l}$ from (10) and (11), respectively, note that after setting up the first-order conditions and rearranging terms, we have that $u_{P,l} > u_{A,l}$ 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 that $u_{A,l} = u_{P,l} = u_l^S$.

Step 5: For parameters such that $IC_{A,h}$ is not slack, we then have to analyze the situation in which both, $ICC_{A,h}$ *and* $IC_{A,h}$ are binding. 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 leading to $u_{\cdot l} = u_l^S$.

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 $u_{\cdot,h} = u_h^{FB}$ and $t_{\cdot,h} = \theta_h u_h^{FB}$ (Case 4). This leaves us with only one remaining case: Case 3 where only informed low-type customers are excluded.

Note here first that ICC_P becomes irrelevant, but that now the *ex ante* individual rationality constraint IR_P becomes binding: $\mu V_{P.h} + (1-\mu) V_{P.l} = 0$. Next, 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 $V_{P,l} \geq 0$. But then 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 $\widetilde{\pi}_{12} := \frac{1}{2-\mu}$ to separate Case 1 from Case 2. Recall next also 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

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

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

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

which is strictly decreasing in μ.

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 ($\widetilde{\pi}_{12}$, $\widetilde{\pi}_{24}$, and $\widetilde{\pi}_{34}$) 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 6. From the proof of Proposition 1 we can use that $V_{A,h} = \mu(\theta_h - \theta_l)u_{A,l} + (1-\mu)(\theta_h - \theta_l)u_{P,l}$. (For brevity we only consider changes in π for the original case with $k = \infty$.) As from implicit differentiation of (2) and (3) we obtain that

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

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

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

As we make no further assumptions on the curvature $s''(u;\theta) = -C''(u)$, we can not determine generally when (12) holds. However, we can obtain results for the boundaries where Case 2 applies. More immediately, note that for high values of π close to the "right boundary" $\widetilde{\pi}_{12}$ (cf. Figure 1), i.e., for π is close to $1/(2-\mu)$, (12) does not hold, given that in this case the right-hand side of (12) exceeds one, while both $u_{A,l}$ and $u_{P,l}$ become close to $u_{S,l}$ and the left-hand side of (12) thus close to one. Towards the "left boundary" of Case 2, where π is close to $\widetilde{\pi}_{24}$, we can instead show that (12) holds, provided that, in addition, μ is small.[27] Note that from this discussion we also obtain immediately that for the special case where C''' is zero, condition (12) holds if and only if π is not too large, namely[28]

$$\frac{(1-\mu)^2}{\mu} < \frac{(1-\pi)^2}{\pi^2}.$$

In the specification $C''' = 0$ it remains to be shown that $\widehat{\pi} \in (\pi_{24}, \pi_{12})$ for μ sufficiently small. We first check that $\widehat{\pi} < \widetilde{\pi}_{12}$ for all μ. This inequality holds if

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

which is equivalent to $\mu(1-\sqrt{\mu}) < 1-\sqrt{\mu}$, which is indeed always satisfied.

Next consider the inequality $\widehat{\pi} > \widetilde{\pi}_{24}$. Recall that $\widetilde{\pi}_{24}(\mu) = \frac{(\theta_h-\theta_l)\mu}{(1-\mu)[\theta_l+(\theta_h-\theta_l)\mu]}$. Thus $\widehat{\pi} > \widetilde{\pi}_{24}$ is equivalent to

$$\frac{\sqrt{\mu}}{1+\sqrt{\mu}-\mu} > \frac{(\theta_h-\theta_l)\mu}{(1-\mu)[\theta_l+(\theta_h-\theta_l)\mu]}$$

which can be rewritten as

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

This inequality is satisfied for $\mu = 0$ and violated for $\mu \to \theta_l/\theta_h$. There is a critical value such that this the left-hand side is equal to the right-hand side. For smaller values the inequality is satisfied, whereas for larger values it is violated. This follows from the fact that the numerator is increasing in μ and the denominator is decreasing. **Q.E.D. (of Corollary 6)**

---

[27] More formally, at $\pi = \widetilde{\pi}_{24}(\mu)$ we have that 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 $\mu \to 0$.

[28] To mark this area in Figure 1, we would have to additinally introduce the function $\overline{\pi} = \frac{1-\mu}{1-\mu+\sqrt{\mu}}$, which cuts $\widetilde{\pi}_{24}$ from above and the horizontal axis to the left of $\pi = \widetilde{\pi}_{12}(0) = 1/2$.

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

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

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. Note that we assume throughout the proof that it is optimal for the firm to ensure that all customers purchase a positive level of services.

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

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

implying that ICC'_P holds from (7).

In what follows we focus on the case where the condition (7) does not hold such that ICC'_P must bind.

Step 2: Note now first that from the same arguments as in the proof of Proposition 1 we have that $u_{\cdot,h} = u_h^{FB}$ 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

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

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 now also 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

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

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

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

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

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

contradicting (15). This proves our claim.

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

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

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

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

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 (18) (together also with $V_{\cdot,h} = V_h$), note that ICC_P becomes

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

Suppose, by contradiction, that ICC_P does then 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 then 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

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

If $IC_{P,h}$ was *also* not binding, then the firm could benefit from simply reducing $V_h = V_{\cdot,h}$ (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 $u_{\cdot,i} > 0$, this yields the first-order condition

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

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 instead $u_{P,l}$ as the remaining variable service level, 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 $y := \frac{k}{\mu}\frac{1}{(1-\mu)(\theta_h-\theta_l)}$, 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 we have by assumption that there must be a (positive) solution with also $u_{A,l} > 0$, for the characterization it remains to be shown uniqueness, which in turn holds if the left-hand side of (21) is strictly monotonic. Using $du_{A,l}/du_{P,l} = 1$ this follows immediately from strict concavity of the surplus function.

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 ICC_P' just starts to bind, (7) is satisfied with equality, yielding condition (8). **Q.E.D (of Proposition 4.)**

Proof of Corollary 8. We claim that $u_{A,l}$ is strictly decreasing and $u_{P,l}$ strictly increasing in k, where also $u_{\cdot,l} \to u_l^S$ for $k \to 0$ 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

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

For the convergence (and continuity) results, note that we can substitute $y = 0$ for the case of $k = 0$. **Q.E.D (of Corollary 8.)**

Proof of Corollary 9. 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 8, $u_{A,l}$ increases as response to a decrease in k, the uninformed customer's expected utility thus indeed strictly increases.

In the case where k has an impact, 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 $\mu > \pi$. If this is the case, lower information acquisition costs lead to higher rents for informed customers.

We next analyze the effect of a reduction in information acquisition costs on welfare. 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 inequality is always satisfied if C''' is zero since $u_{P,l} > u_{A,l}$. **Q.E.D. (of Corollary 9)**

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 $\pi = \widetilde{\pi}_{24} + \varepsilon$ for $\varepsilon > 0$ sufficiently small and $\mu > 0$ sufficiently small. Note that as $\pi \downarrow \widetilde{\pi}_{24}$, we have $u_{A,l} \to 0$ and thus $s(u_{A,l};\theta_l) \to 0$. 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 $\pi \downarrow \widetilde{\pi}_{24}$ we can substitute π by $\frac{(\theta_h-\theta_l)\mu}{(1-\mu)[\theta_l+(\theta_h-\theta_l)\mu]}$, where

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

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 $\pi \to \widetilde{\pi}_{24}$. Then

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

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

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

which may be positive or negative.

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

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

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

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

we have that

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

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

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

Since this inequality is always satisfied W is locally increasing in π at $\widetilde{\pi}_{12} - \varepsilon$ 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. (2007), Information Gathering and Marketing, mimeo.

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.

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.

Eliaz, K. and Spiegler, R. (2006), Contracting with Diversely Naive Agents, *Review of Economic Studies* 73, 689-714.

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. (2007), Selling to Overconfident Consumers, *American Economic Review*, forthcoming.

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, 1089-118.

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, *Quant. Marketing 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.

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.