

Selling customer information to competing firms

Francesco Clavorà Braulin* and Tommaso Valletti†

September 2016

Abstract

We consider a data broker that holds precise information about customer preferences. The data broker can sell this data set either exclusively to one of two differentiated competing firms, or to both of them. If a downstream firm obtains the data set, it can practice personalized pricing, else it has to offer a uniform price to customers. The first-best allocation can be achieved when data are sold non exclusively, but this never arises in equilibrium. The data broker instead sells the data set exclusively either to the high quality firm or to the low quality firm rival, according to their quality-adjusted cost differential. This leads to inefficient allocations.

Keywords: personalized pricing, vertical product differentiation, exclusivity.

JEL: D43, L13, L42, L82.

*University of Rome Tor Vergata, francesco.clavora@gmail.com

†Imperial College London, University of Rome Tor Vergata and CEPR,
t.valletti@imperial.ac.uk

Selling customer information to competing firms

Francesco Clavorà Braulin* and Tommaso Valletti[†]

September 2016

Abstract

We consider a data broker that holds precise information about customer preferences. The data broker can sell this data set either exclusively to one of two differentiated competing firms, or to both of them. If a downstream firm obtains the data set, it can practice personalized pricing, else it has to offer a uniform price to customers. The first-best allocation can be achieved when data are sold non exclusively, but this never arises in equilibrium. The data broker instead sells the data set exclusively either to the high quality firm or to the low quality firm rival, according to their quality-adjusted cost differential. This leads to inefficient allocations.

Keywords: personalized pricing, vertical product differentiation, exclusivity.

JEL: D43, L13, L42, L82.

*University of Rome Tor Vergata, francesco.clavora@gmail.com

[†]Imperial College London, University of Rome Tor Vergata and CEPR,
t.valletti@imperial.ac.uk

1 Introduction

Recent developments in digital markets have led to the emergence of firms that achieve large turnovers based on business models which involve the collection of personal data. The Google search engine and the Facebook social network are the most known examples. Data are also said to facilitate personalized pricing. Indeed, by using data about their clients, a company receives better information about their purchasing habits and is able to assess their willingness to pay for a given good or service. Provided that it has market power, the company that holds customer data would be able to set different prices for the different customer groups it has identified thanks to the data collected.

An extensive literature in economics has analyzed the impact of price discrimination in competitive settings (see, e.g., Thisse and Vives (1988), Liu and Serfes (2006), Chen and Zhang (2009), and Shy and Stenbacka (2016)). Personal information about consumers can intensify competition among retailers since they compete market-by-market instead of at higher levels of aggregation. In related works, studies on behavior-based price discrimination consider the case where past behavior about customers is used to make targeted offers in future periods (see, e.g., Fudenberg and Tirole (2000), Esteves (2010), and Fudenberg and Villas-Boas (2012)). The literature has mostly tackled cases where detailed customer information is either available or not available symmetrically among competing firms (see Taylor and Wagman (2014) for a survey).

In this paper we take a different point of departure, assuming that the ability of retail firms to directly collect and process customer information is limited. Instead, they have to make recourse to the information collected by an intermediary, a data broker, that then sells it to the downstream firms. Indeed, recent years have seen a surge in the emergence of data intermediaries, such as Acxiom or Teradata, which collect and analyze data for third parties. Some data brokers have entered into partnerships with Facebook or Google, in order to help them improve the targeting of their advertising offers.

This brings the question of the extent to which a data broker that has collected a unique data set will want to sell the data set to all competing downstream firms, or exclusively to some of them. In the former case, the fee is earned more than once, but in the second case it is likely that the data set is more valuable to the firm that gets it. This idea is studied by Montes et al. (2016) who, among other things, show an exclusivity result under various settings. While they model retail competition in a Hotelling fashion, the purpose of this paper is to see if the exclusivity result holds also with another

workhorse in Industrial Organization, vertical product differentiation. This model has the added advantage, in case exclusivity is offered, to identify precisely which firm should be granted it: either the low-quality firm or the high-quality firm.

In order to solve the model, we must consider asymmetric cases where one retail firm does not have information (and thus offers uniform prices) while the rival firm can target individual customers by practicing first-degree price discrimination. These asymmetric analyses are instrumental to derive all possible payoffs and determine the fee that competing firms would be willing to pay when the data set is put on offer. To the best of our knowledge, this interim analysis of mixed pricing between informed and uninformed vertically-differentiated firms is also new in the literature.

2 The model

There are three types of agents: consumers, two competing retailers supplying goods of different quality, and a data broker that holds information about consumer preferences.

We describe consumer preferences first. Consumers buy one unit of a product at most, and their preferences follow a standard model of vertical product differentiation. A type θ buying a product of quality q_i at a price p_i enjoys a net utility of $U(q_i, p_i; \theta) = \theta q_i - p_i$, where θ is uniformly distributed between $\underline{\theta}$ and $\bar{\theta}$ with unit density.

Two differentiated retailers, denoted as $i = L, H$, compete in the downstream market, with $q_L < q_H$. The firms also have different marginal costs of production, with $c_L < c_H$. It will be useful at times to use the notation $q_H - q_L = \Delta_q$ and $c_H - c_L = \Delta_c$.

In order to focus on a more interesting welfare analysis, we posit that

$$\underline{\theta} \leq \theta^W \equiv \frac{\Delta_c}{\Delta_q} \leq \bar{\theta}. \quad (1)$$

This assumption on the quality-adjusted cost differential implies that it is efficient to serve the top end of the market ($\theta^W \leq \theta \leq \bar{\theta}$) with the high-quality product and the bottom end of the market ($\underline{\theta} \leq \theta < \theta^W$) with the low-quality product.

Finally, the data broker holds a data set with precise information about consumers. If a downstream firm is given the data set, then it knows the precise type of each customer in the data set and can practice first-degree price discrimination. Else it just knows the distribution of types, and can offer a uniform price.

The game proceeds in sequential stages. First the data broker allocates the data to one or both downstream firms (more on this in Section 3). Then retail firms, given the data in their possession, compete for consumers in the final market: firms set basic prices if they do not have information; otherwise, in a subsequent stage, they offer a tailored price. This sequence has relevance in the asymmetric case and is further discussed below. Finally, consumers buy the product and consume it. We look for the subgame perfect Nash equilibrium of this game.

2.1 No firm has information

Imagine no retailer has precise customer information. Then competition is in uniform prices. This is standard, with the exception that we consider firms with different marginal costs, whereas the canonical model involves firms with identical costs, often normalized to zero (Tirole, 1988). Following standard procedures, the indifferent consumer is defined as

$$\theta^* q_H - p_H = \theta^* q_L - p_L \implies \theta^* = \frac{p_H - p_L}{q_H - q_L}.$$

Firms' profits are $\pi_L = (\theta^* - \underline{\theta})(p_L - c_L)$ and $\pi_H = (\bar{\theta} - \theta^*)(p_H - c_H)$. Taking the first-order conditions, the equilibrium in prices is given by

$$p_H = c_H + \frac{1}{3} (\Delta_q(2\bar{\theta} - \underline{\theta}) - \Delta_c),$$

$$p_L = c_L + \frac{1}{3} (\Delta_q(\bar{\theta} - 2\underline{\theta}) + \Delta_c).$$

The indifferent consumer then simplifies to

$$\theta^* = \frac{\bar{\theta} + \underline{\theta}}{3} + \frac{\Delta_c}{3\Delta_q}. \quad (2)$$

It is straightforward to confirm that (1) guarantees that there is an interior solution with $\underline{\theta} \leq \theta^* \leq \bar{\theta}$.

Finally, we report the expressions for the profits

$$\pi_H^{N,N} = \frac{(\Delta_q(2\bar{\theta} - \underline{\theta}) - \Delta_c)^2}{9\Delta_q}, \quad \pi_L^{N,N} = \frac{(\Delta_q(\bar{\theta} - 2\underline{\theta}) + \Delta_c)^2}{9\Delta_q},$$

where the superscripts (N, N) refer to the case where no firm has information.

Consumer surplus is equal to

$$\begin{aligned} CS_H &= \int_{\theta^*}^{\bar{\theta}} (\theta q_H - p_H) d\theta = \\ &= \left(\frac{\Delta_q(2\bar{\theta} - \underline{\theta}) - \Delta_c}{18\Delta_q} \right) \left(\left(3\underline{\theta} + \frac{\Delta_c}{\Delta_q} \right) q_H + 2((2\bar{\theta} - \underline{\theta})q_L - 2c_H - c_L) \right), \end{aligned}$$

$$\begin{aligned} CS_L &= \int_{\underline{\theta}}^{\theta^*} (\theta q_L - p_L) d\theta = \\ &= \left(\frac{\Delta_q(\bar{\theta} - 2\underline{\theta}) + \Delta_c}{18\Delta_q} \right) \left(\left(3\bar{\theta} + \frac{\Delta_c}{\Delta_q} \right) q_L - 2((\bar{\theta} - 2\underline{\theta})q_H + 2c_L + c_H) \right). \end{aligned}$$

It follows that total surplus is given by

$$CS = \frac{q_L(\bar{\theta}^2 - \underline{\theta}^2)}{2} - c_L(\bar{\theta} - \underline{\theta}) + \frac{1}{18} \left(\frac{\Delta_c^2}{\Delta_q} - \Delta_c(10\bar{\theta} - 8\underline{\theta}) - \Delta_q(2\bar{\theta}^2 - 14\bar{\theta}\underline{\theta} + 11\underline{\theta}^2) \right).$$

2.2 Both firms have information

In this scenario firms are able to compete for every consumer on an individual basis, making targeted offers $p_H(\theta)$ and $p_L(\theta)$ to type θ . We concentrate on equilibria where firms cannot make offers below their costs. Then each firm serves consumers efficiently, that is, firm L serves types $\underline{\theta} \leq \theta < \theta^W$ offering a price that just matches the best offer that firm H can make ($p_H(\theta) = c_H$), hence

$$p_L(\theta) = c_H - \theta\Delta_q.$$

Similarly, firm H serves types $\theta^W \leq \theta \leq \bar{\theta}$ offering a price that just matches the best offer that firm L can make ($p_L(\theta) = c_L$), that is

$$p_H(\theta) = c_L + \theta\Delta_q.$$

Firms' profits are given by

$$\begin{aligned} \pi_H^{I,I} &= \int_{\theta^W}^{\bar{\theta}} (c_L + \theta\Delta_q - c_H) d\theta = \frac{(\bar{\theta}\Delta_q - \Delta_c)^2}{2\Delta_q}, \\ \pi_L^{I,I} &= \int_{\underline{\theta}}^{\theta^W} (c_H - \theta\Delta_q - c_L) d\theta = \frac{(\underline{\theta}\Delta_q - \Delta_c)^2}{2\Delta_q}, \end{aligned}$$

where the superscripts (I, I) refer to the case where both firms have information.

Consumer surplus is

$$CS_H = \int_{\theta^w}^{\bar{\theta}} (\theta q_H - p_H(\theta)) d\theta = \left(\bar{\theta} - \frac{\Delta_c}{\Delta_q} \right) \left(\left(\bar{\theta} + \frac{\Delta_c}{\Delta_q} \right) \frac{q_L}{2} - c_L \right),$$

$$CS_L = \int_{\underline{\theta}}^{\theta^w} (\theta q_L - p_L(\theta)) d\theta = \left(\frac{\Delta_c}{\Delta_q} - \underline{\theta} \right) \left(\left(\underline{\theta} + \frac{\Delta_c}{\Delta_q} \right) \frac{q_H}{2} - c_H \right).$$

Total consumer surplus therefore satisfies

$$CS = \frac{1}{2}(\bar{\theta}^2 q_L - \underline{\theta}^2 q_H) + (\underline{\theta} c_H - \bar{\theta} c_L) - \frac{\Delta_c^2}{2\Delta_q}.$$

2.3 Asymmetric cases

Consider now a situation where only one of the two firms ($i = L, H$) acquired the information and consequently it can identify precisely every type of consumers. Therefore the informed duopolist is able to charge a tailored price $p_i(\theta)$, while the rival firm must offer a single uniform price p_j to all consumers. The timing of the pricing game is as follows: the uninformed firm moves first, setting a basic price, and the informed firm then responds to the rival. Big data, especially in online markets, give a firm the ability to engage in first-degree price discrimination. This ability is in turn linked with the capacity to set prices in real time, when the transaction is being executed. It is the firm with data that can eventually match the price of its rival, not the other way around: the uninformed firm does not have this option. We additionally remark that a sequential setting of prices solves a problem of equilibrium non-existence that would otherwise arise with simultaneous pricing. When prices are set simultaneously and information is held asymmetrically, there is no equilibrium in pure strategies. Provided that the firm with information sets its prices so as to make customers indifferent between buying from itself and from the rival, the uninformed firm would then want to deviate and attract those consumers. For any candidate pure strategy equilibrium, the uninformed firm has always a profitable deviation. A sequential setting solves this non-existence issue.

We proceed to study two distinct scenarios, depending on which firm has bought the data and thus is informed.¹

¹It would still be interesting for future research to look at mixed strategies equilibria with simultaneous pricing.

2.3.1 Low-quality informed firm

In this case the low quality firm sets a personalized price $p_L(\theta)$. Prices are again restricted to be above or at most equal to the marginal cost, implying that the lowest possible price for the low-quality firm is $p_L(\theta) = c_L$. The indifferent consumer is found from $U(q_H, p_H; \theta) = U(q_L, c_L; \theta)$, resulting in

$$\theta_L^* = \frac{p_H - c_L}{q_H - q_L}.$$

All consumers with $\theta \geq \theta_L^*$ are served by firm H at a uniform price p_H , while all consumers with $\theta < \theta_L^*$ are supplied by firm L at a tailored price that just undercuts firm H , that is

$$p_L(\theta) = p_H - \theta \Delta_q.$$

Anticipating this, the profit maximization problem of firm H is given by $\max_{p_H} (\bar{\theta} - \theta_L^*)(p_H - c_H)$. Taking the first order condition with respect to p_H , the uniform price charged by firm H is equal to

$$p_H = \frac{c_H + c_L}{2} + \frac{\bar{\theta} \Delta_q}{2},$$

while the tailored price set by the low-quality firm simplifies to

$$p_L(\theta) = \frac{c_H + c_L}{2} + \left(\frac{\bar{\theta}}{2} - \theta \right) \Delta_q,$$

which is decreasing in θ . The expression for the indifferent consumer is

$$\theta_L^* = \frac{\bar{\theta}}{2} + \frac{\Delta_c}{2\Delta_q}. \quad (3)$$

Again, (1) guarantees that there is an interior solution with $\underline{\theta} \leq \theta_L^* \leq \bar{\theta}$. Profits of the duopolists in this first asymmetric scenario are

$$\begin{aligned} \pi_H^{N,I} &= \int_{\theta_L^*}^{\bar{\theta}} (p_H - c_H) d\theta = \frac{(\bar{\theta} \Delta_q - \Delta_c)^2}{4\Delta_q}, \\ \pi_L^{I,N} &= \int_{\underline{\theta}}^{\theta_L^*} (p_L(\theta) - c_L) d\theta = \frac{((\bar{\theta} - 2\underline{\theta})\Delta_q + \Delta_c)^2}{8\Delta_q}, \end{aligned}$$

where the first subscript refers to the information available to a given firm, and the second subscript to the information available to its rival.

Consumer surplus is

$$\begin{aligned} CS &= \int_{\underline{\theta}}^{\theta_L^*} (\theta q_L - p_L(\theta)) d\theta + \int_{\theta_L^*}^{\bar{\theta}} (\theta q_H - p_H) d\theta \\ &= \frac{(\bar{\theta} - \underline{\theta})}{2} (\bar{\theta} q_L + \underline{\theta} q_H - c_H - c_L). \end{aligned}$$

2.3.2 High-quality informed firm

The analysis mirrors the previous subsection, where now it is the high quality firm to set a tailored price $p_H(\theta)$, which cannot exceed c_H . The indifferent consumer is

$$\theta_H^* = \frac{c_H - p_L}{q_H - q_L}.$$

All consumers with $\theta < \theta_H^*$ are supplied by firm L at a uniform price p_L , while all consumers with $\theta \geq \theta_H^*$ face a personalized price set by firm H just to undercut firm L ,

$$p_H(\theta) = p_L + \theta \Delta_q.$$

The profit maximization problem of firm L is $\max_{p_L} (\theta_H^* - \underline{\theta})(p_L - c_L)$, from which we obtain

$$p_L = \frac{c_H + c_L}{2} - \frac{\underline{\theta} \Delta_q}{2},$$

while the tailored price is

$$p_H(\theta) = \frac{c_H + c_L}{2} + \left(\theta - \frac{\underline{\theta}}{2} \right) \Delta_q,$$

which is increasing in θ . The expression for the indifferent consumer is

$$\theta_H^* = \frac{\underline{\theta}}{2} + \frac{\Delta_c}{2\Delta_q}, \quad (4)$$

which thanks to (1) always lies in the interior interval $\underline{\theta} \leq \theta_H^* \leq \bar{\theta}$. Profits of the duopolists are

$$\begin{aligned} \pi_H^{I,N} &= \int_{\theta_H^*}^{\bar{\theta}} (p_H(\theta) - c_H) d\theta = \frac{((2\bar{\theta} - \underline{\theta})\Delta_q - \Delta_c)^2}{8\Delta_q}, \\ \pi_L^{N,I} &= \int_{\underline{\theta}}^{\theta_H^*} (p_L - c_L) d\theta = \frac{(\underline{\theta}\Delta_q - \Delta_c)^2}{4\Delta_q}, \end{aligned}$$

while consumer surplus is

$$\begin{aligned} CS &= \int_{\underline{\theta}}^{\theta_H^*} (\theta q_L - p_L) d\theta + \int_{\theta_H^*}^{\bar{\theta}} (\theta q_H - p_H(\theta)) d\theta \\ &= \frac{(\bar{\theta} - \underline{\theta})}{2} (\bar{\theta} q_L + \underline{\theta} q_H - c_H - c_L). \end{aligned}$$

Note how the expression for consumer surplus in aggregate is the same in the two asymmetric cases.

3 Selling strategy of the data holder

We assume that the data holder has all the bargaining power and is able to extract all surplus from the transaction. Three scenarios can arise: either consumer data are sold to both firms, or they are sold exclusively, in which case they can further be sold either to H or to L . In order to determine the price that the data holder will be able to extract, one can think of an auction with externalities, as described by Jehiel and Moldovanu (2000).

If the data holder is willing to sell to both firms, the total price that the seller is able to obtain is given by $P^{Both} = (\pi_H^{I,I} - \pi_H^{N,I}) + (\pi_L^{I,I} - \pi_L^{N,I})$, that is, the sum of the differentials between each firm's own profit when the information is symmetrically acquired and the own profit when only the competitor has acquired information. This is equal to

$$P^{Both} = \frac{((\bar{\theta}\Delta_q - \Delta_c)^2 + (\underline{\theta}\Delta_q - \Delta_c)^2)}{4\Delta_q}.$$

Under exclusivity, the price is given by the differential between the profit with exclusive information and the profit without information when instead the competitor has it, that is, $P_i = (\pi_i^{I,N} - \pi_i^{N,I})$. When the information is at exclusive disposal of the low quality firm, the price will be

$$P_L = \pi_L^{I,N} - \pi_L^{N,I} = \frac{(\bar{\theta}\Delta_q + \Delta_c)^2 - 2\underline{\theta}(2\bar{\theta} - \underline{\theta})\Delta_q^2 - 2\Delta_c^2}{8\Delta_q},$$

and when it is sold only to the high quality firm, the price is

$$P_H = \pi_H^{I,N} - \pi_H^{N,I} = \frac{(\underline{\theta}\Delta_q + \Delta_c)^2 + 2\bar{\theta}(\bar{\theta} - 2\underline{\theta})\Delta_q^2 - 2\Delta_c^2}{8\Delta_q}.$$

We are now in a position to state our main result, which shows that the data seller always prefers to sell information exclusively to one downstream firm.

Proposition 1 *The data holder always has an incentive to sell consumer information exclusively to one downstream competitor. If $\underline{\theta} \leq \frac{\Delta_c}{\Delta_q} \leq \frac{(\bar{\theta} + \underline{\theta})}{2}$, information is sold to the high quality producer; if $\frac{(\bar{\theta} + \underline{\theta})}{2} < \frac{\Delta_c}{\Delta_q} \leq \bar{\theta}$ information is sold to the low quality producer.*

Proof. The proof consist of two steps. First, we compare exclusive prices, where the sign of $P_H - P_L$ depends only on the expression $(\bar{\theta} + \underline{\theta})\Delta_q - 2\Delta_c$. Therefore, whenever $\underline{\theta} \leq \Delta_c/\Delta_q < (\bar{\theta} + \underline{\theta})/2$ it turns out that $P_H > P_L$;

the reverse inequality holds instead for the range $(\bar{\theta} + \underline{\theta})/2 < \Delta_c/\Delta_q \leq \bar{\theta}$. Second, in each relevant range, we compare the price of exclusivity with the total price when selling to both firms. Performing this comparison, it turns out that $P_H > P^{Both}$ for values of $\Delta_c/\Delta_q < (4\bar{\theta} + \underline{\theta})/5$, which is always true in the relevant range $\underline{\theta} \leq \Delta_c/\Delta_q < (\bar{\theta} + \underline{\theta})/2$. Similarly, $P_L > P^{Both}$ when $\Delta_c/\Delta_q > (\bar{\theta} + 4\underline{\theta})/5$, which is always true in the relevant range $(\bar{\theta} + \underline{\theta})/2 < \Delta_c/\Delta_q \leq \bar{\theta}$. We can conclude that information is always sold on an exclusive basis. ■

4 Welfare analysis and concluding remarks

We conclude this paper by making a welfare assessment of the equilibria that we have found. Under the maintained assumption (1), everybody should be supplied a product. Indeed, we have shown that the market is always covered either in the symmetric equilibria (with or without information), or in the asymmetric ones where only one firm has information. Hence what remains to be done is to determine the product allocation in the various cases. This is straightforward to do as the analysis boils down to the comparisons among the indifferent consumers, as described in various cases by (2), (3) and (4).

Proposition 2 *Selling data to both firms is first best but it never arises in equilibrium. Selling data exclusively always generates inefficiencies. In particular, when data are sold exclusively to firm H (respectively, firm L), then too many customers are supplied the high quality product (respectively, the low quality product).*

Proof. Denote as $\tilde{\theta}$ the indifferent consumer under any allocation of data. Social welfare is given by

$$\begin{aligned} W &= \int_{\underline{\theta}}^{\tilde{\theta}} (\theta q_L - c_L) d\theta + \int_{\tilde{\theta}}^{\bar{\theta}} (\theta q_H - c_H) d\theta = \\ &= \tilde{\theta} \Delta_c - \frac{\tilde{\theta}^2}{2} \Delta_q + \bar{\theta} \left(\frac{\bar{\theta}}{2} q_H - c_H \right) - \underline{\theta} \left(\frac{\underline{\theta}}{2} q_L - c_L \right). \end{aligned}$$

It is then $\partial W / \partial \tilde{\theta} = \Delta_c - \Delta_q \tilde{\theta}$ and $\partial^2 W / \partial \tilde{\theta}^2 = -\Delta_q < 0$. Welfare is a concave function of $\tilde{\theta}$ which is obviously maximized when $\tilde{\theta} = \theta^W \equiv \frac{\Delta_c}{\Delta_q}$. This is also the same allocation achieved when both firms have information (see Section 2.2). When information is sold exclusively to firm H, which we know from Proposition 1 happens when $\underline{\theta} \leq \frac{\Delta_c}{\Delta_q} \leq \frac{(\bar{\theta} + \underline{\theta})}{2}$, then from (4) we have that

$$\theta_H^* < \theta^W \iff \frac{\underline{\theta}}{2} + \frac{\Delta_c}{2\Delta_q} < \frac{\Delta_c}{\Delta_q} \iff \underline{\theta} < \frac{\Delta_c}{\Delta_q},$$

which is always strictly true in the range $\underline{\theta} < \frac{\Delta_c}{\Delta_q} \leq \frac{(\bar{\theta} + \underline{\theta})}{2}$. Hence, “too many” customers buy the high-quality product compared to the efficient allocation. The case when data are sold exclusively to firm L is perfectly symmetric and the proof is omitted. ■

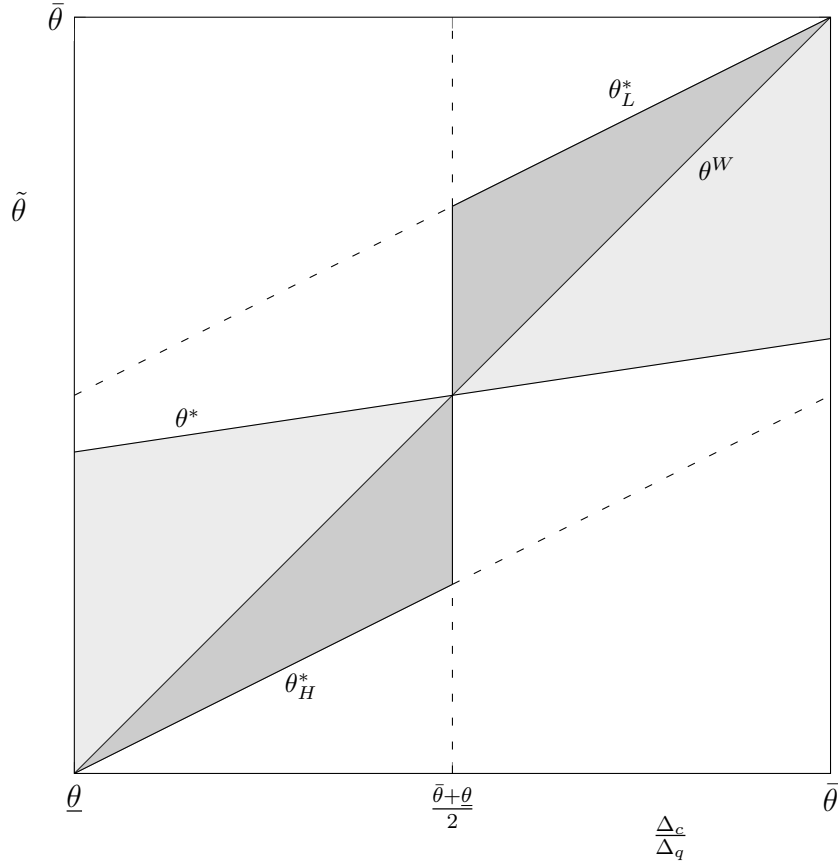


Figure 1: Customer allocation

Figure 1 describes the allocation of customers in the various cases, as a function of the quality-adjusted cost differential, Δ_c/Δ_q . The first best lies on the 45° line, with customers above (below) the line consuming the high quality (low quality) product. The Figure also depicts θ_H^* and θ_L^* , in the relevant range of parameters when exclusivity to one particular firm arises in equilibrium, as described by Proposition 1. Note that, due to the fact that the welfare function is concave in the indifferent consumer, with a maximum at θ^W , inefficiencies are quite marked when Δ_c/Δ_q takes intermediate values, as this is when the distance between the indifferent type θ_i^* and θ^W reaches

its maximal value. The dark shaded area between the curves is therefore a measure of the welfare inefficiencies, for a given value of Δ_c/Δ_q .

The Figure also plots θ^* from (2), when no firm has information. The allocation would be inefficient here too, as given by the light shaded area, but with the opposite pattern, since differences from θ^W become small for intermediate values of Δ_c/Δ_q (and there is no inefficiency for the special value $\frac{\Delta_c}{\Delta_q} = \frac{(\bar{\theta} + \underline{\theta})}{2}$).²

We can therefore obtain some policy conclusions. The first best in our model would be achieved by banning exclusivity contracts, leaving the data seller only with the option of selling to both rivals. This conclusion depends on the data seller having access to the data in any case. If instead the data seller had to invest resources for acquiring or putting together the data set in the first place, it is conceivable that information would not be collected at all if exclusivity was banned, as ex post profits would then be lower. In this situation the more relevant comparison would be the one between allowing exclusivity (which would lead to data collection) and having no information at all. While there is generally no clear-cut comparison between these regimes, the concavity of the welfare function implies that no information at all is still preferred to data exclusivity when the quality-adjusted costs differential (Δ_c/Δ_q) takes intermediate values. On the contrary, exclusivity generates more total surplus when Δ_c/Δ_q is either small or large - ignoring the costs of data acquisition. It would be interesting in future research to explore the dynamic incentives of the upstream firm to acquire data information.

References

Acquisti, A., C. Taylor and L. Wagman (2015). The Economics of Privacy, *Journal of Economic Literature*, forthcoming.

Chen, Y. and T. Zhang (2011). Equilibrium Price Dispersion with Heterogeneous Searchers, *International Journal of Industrial Organization*, Vol. 29(6), 645-654.

²We concentrate our attention on total welfare, given by the sum of consumer surplus and firms' profits. As the debate on price discrimination is often related to consumer privacy, one could also conduct a policy discussion around consumer surplus alone (see Acquisti et al., 2015). This is readily doable in our framework, as we have obtained all the corresponding expressions for CS in the various cases. The detailed analysis is omitted in order to save on space, but it is straightforward to show that, mirroring welfare in Proposition 2, CS is higher when both retail firms have information compared to a case with exclusivity. We can additionally show that the worst-case scenario for consumers is when there is no information at all, as this is when competition (in uniform prices) is weakest.

- Fudenberg, D. and J. Tirole (2000). Customer Poaching and Brand Switching, *RAND Journal of Economics*, Vol. 31(4), 634-657.
- Fudenberg, D. and J. Villas-Boas (2012). Price Discrimination in the Digital Economy, in M. Peitz and J. Waldfogel (Eds.), *The Oxford Handbook of the Digital Economy*, Oxford University Press.
- Jehiel, P. and B. Moldovanu (2000). Auctions with Downstream Interaction among Buyers, *RAND Journal of Economics*, Vol. 31(4), 768-791.
- Liu, Q. and K. Serfes (2009). Customer Information Sharing among Rival Firms, *European Economic Review*, Vol. 50(6), 1571-1600.
- Montes, R., W. Sand-Zantman and T. Valletti (2016). The Value of Personal Information in Markets with Endogenous Privacy, mimeo.
- Shy, O. and R. Stenbacka (2016). Customer Privacy and Competition, *Journal of Economics & Management Strategy*, forthcoming.
- Taylor, C. and L. Wagman (2014). Customer Privacy in Oligopolistic Markets: Winners, Losers, and Welfare, *International Journal of Industrial Organization*, Vol. 34, 80-84.
- Thisse, J. and X. Vives (1988). On The Strategic Choice of Spatial Price Policy, *American Economic Review*, Vol. 78(1), 122-137.
- Tirole, J. (1988). *The Theory of Industrial Organization*, MIT Press.

