CESifo WORKING PAPERS

7078
2018

May 2018

Selling Strategic Information in Digital Competitive Markets

David Bounie, Antoine Dubus, Patrick Waelbroeck

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7078*
Category 14: Economics of Digitization

Selling Strategic Information in Digital Competitive Markets

Abstract

This paper investigates the strategies of a data broker in selling information to one or to two competing firms that can price-discriminate consumers. The data broker can strategically choose any segment of the consumer demand (information structure) to sell to firms that implement third-degree price-discrimination. We show that the equilibrium profits of the data broker are maximized when (1) information identifies the consumers with the highest willingness to pay; (2) consumers with a low willingness to pay remain unidentified; (3) the data broker sells two symmetrical information structures. The data broker therefore strategically sells partial information on consumers in order to soften competition between firms. Extending the baseline model, we prove that these results hold under first-degree price-discrimination.

JEL-Codes: D400, D800, L500, D430.

Keywords: data broker, information structure, price-discrimination.

David Bounie
Telecom Paris Tech / Economics and Social Sciences
46 rue Barrault
France – 75634 Paris Cedex 13
david.bounie@telecom-paristech.fr

Antoine Dubus
Telecom Paris Tech / Economics and Social Sciences
46 rue Barrault
France – 75634 Paris Cedex 13
antoine.dubus@telecom-paristech.fr

Patrick Waelbroeck
Telecom Paris Tech / Economics and Social Sciences
46 rue Barrault
France – 75634 Paris Cedex 13
patrick.waelbroeck@telecom-paristech.fr

May 17, 2018
We would like to thank Yann Balgobin, Marc Bourreau, Romain De Nijs, Alexis Larousse, Qihong Liu, Johannes Paha, Martin Quinn, Konstantinos Serfes, as well as participants at the MaCCI Annual Conference, the 7th Annual Meeting of the French Economic Association, the Workshop of Industrial Organization in the Digital Economy, the CREST Workshop on Platforms and E-commerce, and the Paris Seminar on Digital Economics for useful remarks and comments. Antoine Dubus acknowledges financial support from the Chair Values and Policies of Personal Information of Institut Mines Télécom, Paris. Patrick Waelbroeck thanks for insightful discussions the members of the Chair Values and Policies of Personal Information of Institut Mines Télécom, Paris.

1 Introduction

Digital technologies transform modern economies, with far reaching effects on productivity, employment, innovation, and growth (McAfee and Brynjolfsonn, 2012). In 2016, the digital economy accounted for USD 1.2 trillion in the United States, or 6.5% of GDP, according to the Commerce Department's Bureau of Economic Analysis. The sector that includes network infrastructure, e-commerce and digital media grew at an average annual rate of 5.6% per year from 2006 to 2016, compared to 1.5% growth in the overall economy. A similar trend is also at work in China and the European Union.[1]

The e-commerce and digital media sectors are dominated by U.S. companies, such as Facebook, Apple, Amazon, and Google. These companies base their business models on personal data collection and store traces left by Internet users who visit their online websites. What is less known, however, is that these large companies also share or acquire information from data brokers that also gather information about millions of people from offline or other online sources.[2] The recent Facebook scandal involving Cambridge Analytica has precisely revealed to the public the troubled relations between Facebook and data brokers.[3]

Data brokers indeed collect all sorts of information on consumers from publicly available online and offline sources (such as names, addresses, revenues, loan default information, and registers). Data brokers are major actors in the data economy. There are more than 4000 data brokers operating in a market valued around USD 156 billion per year (Pasquale, 2015). In a study of nine data brokers in 2014,[4] the Federal Trade Commission found that data brokers have information "on almost every U.S. household and commercial transaction. [One] data broker's database has information on 1.4 billion consumer transactions and over 700 billion aggregated data elements; another data broker's database covers one trillion dollars in consumer transactions; and yet another data broker adds three billion new records each month to its databases."[5]

Data brokers therefore have a lot of information that they can sell to help firms learn more about their customers, to target ads, or to price-discriminate consumers. They can sell different

[1] European Commission, Digital Agenda for Europe
[2] Business Insider, Facebook is quietly buying information from data brokers about its users' offline lives, Dec. 30, 2016.
[3] Washington Post, Facebook, longtime friend of data brokers, becomes their stiffest competition, 29 March 2018.
[4] Acxiom, CoreLogic, Datalogix, eBureau, ID Analytics, Intelius, PeekYou, Rapleaf, and Recorded Future.
[5] Federal Trade Commission, 2014, Data brokers: A Call for Transparency and Accountability.

information to a subset of firms in a market and can therefore shape competition between firms. Our paper is concerned with three main questions. First, does a data broker choose to sell information to all market participants or only to a subset of them? Second, does a data broker sell information on all consumers or only on some of them, such as those who are the most profitable for firms? Third, how do the strategies of a data broker change when the precision of information improves?

Previous research has extensively investigated the role of information on markets and its effects on competition. A first strand of the literature focuses on the acquisition of information by firms but does not specifically investigate the strategic role of an intermediary selling this information. Firms may acquire or not information, but the process of information acquisition is supposed to be exogenous (see Radner et al. (1961); Ponssard (1979); Vives (1984); Van Zandt (1996)). More recently, Liu and Serfes (2004) more explicitly study firm's incentives acquire customer-specific information of a given quality level. In their approach, information is a partition of a mass of consumers distributed in different segments on a Hotelling unit line. Firms have the choice either to acquire all the segments of information or to acquire none of them. Two effects of information acquisition on the competitiveness of product markets are emphasized (following Thisse and Vives (1988), (Corts, 1998), Ulph and Vulkan (2000), Liu and Serfes (2004), Taylor (2004), Cooper et al. (2005), Encaoua and Hollander (2007), Jentzsch et al. (2013), Taylor and Wagman (2014), or Matsumura and Matsushima (2015)). First, firms that acquire information on how consumers value their products can extract more surplus (*rent extraction effect*), which in turn increases their profits. Second, firms that have information fight more intensively for each consumer, whereas without information, they would have set a uniform price and compete less fiercely. Competition between firms is therefore intensified and profits are lower (*competition effect*). Liu and Serfes (2004) show that firms acquire all the segments of information, and that the competition effect dominates the rent extraction effect.

A second strand of the literature analyzes the strategies of a data broker in providing competing firms with information on consumers.[6] Braulin and Valletti (2016) study vertically differentiated products, for which consumers have hidden valuations. The data broker can sell to firms information on these valuations. Montes et al. (2018) consider information allowing firms com-

[6] For a more general literature on intermediary gatekeepers, see also Baye and Morgan (2001), Wathieu (2002), Pancras and Sudhir (2007), Weeds (2016), and Chiou and Tucker (2017).

peting à la Hotelling to first-degree price-discriminate consumers. The data broker sells either information on all consumers, or no information at all. Information is sold through a second-price auction mechanism (as in Jehiel and Moldovanu (2000)). Firms are initially uninformed. When the data broker sells information by auction exclusively to one firm, the valuation of each firm is driven by its outside option, which is, for each firm, being uninformed while its competitor is informed. Thus, firms are ready to bid aggressively since being the only uninformed competitor would lower their profits. Montes et al. (2018) find that the data broker exclusively sells information to one of the competitors.

To summarize, the aforementioned papers assume that firms have only the choice to acquire information on all consumers or on none of them. They implicitly restrict the set of possible information structures to an all-or-nothing choice. Based on this assumption, they find that firms acquire all the segments of information and that the data broker sells information exclusively to one of the competitors.

We first show that the 'all-or-nothing' assumption is never optimal for the data broker. Selling all information strengthens the competitive effect of information and biases welfare and profits in equilibrium. Second, in our setting, the data broker sells information to both firms; selling information to only one firm is never optimal for the data broker. We show in particular that the result of Montes et al. (2018) on the optimality of selling information to only one firm is driven by the information-selling mechanism that they use. We finally show that our results hold under third-degree, and under first-degree price-discrimination where consumers are perfectly identified.

We merge the two strands of the literature by introducing a data broker in the model of Liu and Serfes (2004). In our setting, the data broker can sell information that partitions the Hotelling unit line into segments of arbitrary sizes to one or two competing firms. In other words, the data broker has the choice to sell all the segments of information, no information at all, or only some segments of information. The possibility to purchase only a subset of all available demand segments is in line with real-world marketing strategies where firms purchase information on groups of consumers who share specific characteristics (market segmentation). By doing so, firms can identify demand segments that are likely to be the most profitable. Firms that acquire segments of information can set specific prices on each segment of the unit line. Contrary to

Braulin and Valletti (2016) and Montes et al. (2018), the outside option of a firm that wants to buy information from the data broker is a market configuration where both firms are uninformed. In other words, when information is sold to only one firm, there is no threat that, in case that the firm does not acquire information, its competitor does.

Using this setting, we show that it is optimal for the data broker to sell segments that are located closest to firms, but not segments that are located in the middle of the Hotelling line. This partition allows firms to better extract surplus from consumers with the highest willingness to pay and to keep consumers with a low willingness to pay unidentified in order to soften competition between firms. Selling more information than in this optimal setting generates two welfare effects. On the one hand, having better informed firms increases the surplus of unidentified consumers, and of identified consumers with the highest willingness to pay. On the other hand, identified consumers with the lowest willingness to pay lose surplus, as they were unidentified without information acquisition by the firm. Total consumer surplus is increased when firms are informed, even though some consumers suffer from higher prices due to price-discrimination. Firms are in a prisoner's dilemma: more information reduces their profits as they compete more fiercely. They still buy information as being the only uninformed firm would induce even lower profits.

The remainder of the paper is organized as follows. In Section 2 we describe the model, and in Section 3 we characterize the optimal structure of information. In Section 4, we provide the equilibrium of the game, and we discuss the effects of information acquisition on welfare. We conclude in Section 5.

2 Model set-up

We consider a game involving a data broker, two firms (noted $\theta = 1, 2$), and a mass of consumers uniformly distributed on a unit line $[0, 1]$. The data broker collects information about consumers who buy products from the competing firms at a cost that we normalize to zero. Firms can purchase information from the data broker in order to price-discriminate consumers.[7] In Section 4, we first analyze third-degree price-discrimination, then we extend the analysis to first-degree price-discrimination.

[7] The marginal production costs are also normalized to zero.

The two firms are located at 0 and 1 of the unit line and sell competing products to consumers. A consumer located at x derives a gross utility V from consuming the product, and faces a linear transportation cost with value $t > 0$. A consumer buys at most one unit of the product, and we assume that the market is fully covered, that is, all consumers buy the product.[8] Let p_1 and p_2 denote the prices set by Firm 1 and Firm 2, respectively. A consumer located at x receives the following utility:

$$\begin{cases} U(x) = V - tx - p_1, \text{ if he buys from Firm 1}, \\ U(x) = V - t(1-x) - p_2, \text{ if he buys from Firm 2}, \\ U(x) = 0, \text{ if he does not consume.} \end{cases}$$

In the following sections, we define the information structure, the profits of the data broker and of the firms, and the timing of the game.

2.1 Information structure

Firms know that consumers are uniformly distributed on the unit line, but without further information, they are not able to identify their location. Therefore, firms do not know the degree to which consumers value their products and cannot price-discriminate them.[9] Firms can acquire an information structure from a monopolist data broker at a cost w. The information structure consists of a partition of the unit line of n segments or arbitrary sizes. These segments are constructed by unions of elementary segments of size $\frac{1}{k}$, where k is an exogenous integer that can be interpreted as the quality of information. Even though, the data broker can sell any such partition, it is useful to define a reference partition $\mathcal{P}_{ref}$ that includes k segments of size $\frac{1}{k}$. Figure 1 illustrates the reference partition that includes all segments of size $\frac{1}{k}$. Liu and Serfes (2004) consider that the data broker can only sell (or not) the reference partition $\mathcal{P}_{ref}$ to competing firms. A major contribution of this paper is to show that the optimal partition sold by the data broker is not the reference partition $\mathcal{P}_{ref}$.

[8] If the market is not covered, the competition effect that we identify is weakened, and new issues related to customer churn and customer acquisition arise.
[9] This assumption is also made by Braulin and Valletti (2016) and Montes et al. (2018).

Figure 1: Reference partition $\mathcal{P}_{ref}$

We introduce further notations. We note $\mathcal{S}$ the set composed of the $k-1$ endpoints of the segments of size $\frac{1}{k}$: $\mathcal{S} = \{\frac{1}{k}, .., \frac{i}{k}, .., \frac{k-1}{k}\}$. Consider the mapping, i.e., a bijection, that associates to any subset $\{\frac{s_1}{k}, .., \frac{s_i}{k}, .., \frac{s_{n-1}}{k}\} \in \mathcal{S}$ a partition $\{[0, \frac{s_1}{k}], [\frac{s_1}{k}, \frac{s_2}{k}], .., [\frac{s_{n-1}}{k}, 1]\}$, where $s_1 < .. < s_i < .. < s_{n-1}$ are integers lower than k. We write $\mathbb{P}$ as the target set of the mapping: $M : \mathcal{S} \to \mathbb{P}$, this set is composed of all possible partitions of the unit line generated by segments of size $\frac{1}{k}$. Thus, $\mathbb{P}$ is the sigma-field generated by the elementary segments of size $\frac{1}{k}$. In particular, $\mathcal{P}_{ref}$ and $[0, 1]$ are included in $\mathbb{P}$.

The data broker can sell any partition $\mathcal{P}$ of the set of partitions $\mathbb{P}$, as for instance, a partition starting with one segment of size $\frac{1}{k}$, and another segment of size $\frac{2}{k}$, as illustrated in Figure 2.

Figure 2: Example of a partition of the unit line

A firm having information of the form $\{[0, \frac{s_1}{k}], [\frac{s_1}{k}, \frac{s_2}{k}], .., [\frac{s_{n-1}}{k}, 1]\}$ will be able to identify whether consumers belong to one of the segments of the set and charge them a corresponding price. Namely, the firm will charge consumers on $[0, \frac{s_1}{k}]$ with p_1, consumers on $[\frac{s_i}{k}, \frac{s_{i+1}}{k}]$ with p_{i+1}, and similarly for each segment.

Contrary to the existing literature, we allow the data broker to sell a partition different from $\mathcal{P}_{ref}$. As a matter of fact, it can sell any information structure belonging to $\mathbb{P}$. However, we rule out information structures that generate uncertainty regarding the location of the elementary segment of size $\frac{1}{k}$ to which a consumer belongs. As an illustration, suppose that $k = 8$ so that the finest partition consists of 8 segments of size $\frac{1}{8}$. Suppose also that the data broker sells a partition constructed from 3 segments in the following way. The first element of the partition includes segments 1 and 3 which have a size of $\frac{1}{8}$ and that are located at the extremities of the unit line. The second element of the partition is segment 2 of size $\frac{6}{8}$, located in the middle of the line. The information structure is therefore the partition $\{\{1,3\},2\}$. Segments 1 and 3 are not connected and are therefore excluded from our analysis.

2.2 Strategies and timing

The data broker can sell any partition $\mathcal{P}_\theta$ to Firm θ. As a matter of fact, starting from any pairs of partitions, we will show that when the data broker decides to sell information to both firms, it will sell the same partition. We write the generic form of the profits for a partition $\mathcal{P}$,[10] noting NI (resp. I) when a firm is not informed (resp. informed). Also, we write whether a firm and its competitor are informed or not with the couple (A, B) where $A, B \in \{I, NI\}$. For instance, (I, NI) refers to a situation in which Firm θ is informed and Firm $-\theta$ is uninformed. For any information structure, we need to compute the profits for three possible configurations since $\pi^{NI,I}_{\mathcal{P},\theta} = \pi^{I,NI}_{\mathcal{P},\theta}$: $\{\pi^{NI,NI}_{\mathcal{P},\theta}, \pi^{I,NI}_{\mathcal{P},\theta}, \pi^{I,I}_{\mathcal{P},\theta}\}$.

Firms simultaneously set their prices on the unit line when they have no information or on each segment of the partition when they are informed. Each firm knows whether the competitor is informed or not, and the structure of the partition $\mathcal{P}_{-\theta}$.[11] Firms acquire information at a price depending on how much information increases their profits. This value of information varies according to whether or not the competitor purchases information. We consider the profits in equilibrium for any partition $\mathcal{P}_\theta$ of the unit line.

The data broker extracts all surplus from competing firms and maximizes the difference between the profits of an informed firm and those of an uninformed firm. The data broker profit function can be written as

$$\Pi = \begin{cases} \Pi_1 = w_1 = \max_{\mathcal{P}\in\mathbb{P}}\{\pi^{I,NI}_{\mathcal{P},\theta} - \pi^{NI,NI}_{\theta}\}, \\ \text{if the data broker sells information to only one firm,} \\ \\ \Pi_2 = 2w_2 = 2\max_{\mathcal{P}\in\mathbb{P}}\{\pi^{I,I}_{\mathcal{P},\theta} - \pi^{NI,I}_{\mathcal{P},\theta}\}, \\ \text{if the data broker sells information to both competitors.} \end{cases} \tag{1}$$

[10] We drop the subscript θ when there is no confusion.
[11] This assumption is also standard in Braulin and Valletti (2016) and Montes et al. (2018).

The partition proposed by the data broker depends on whether information is sold to one firm or to both firms. We define Π_1 as the maximum of the first part of Eq. (1), and Π_2 as the maximum of the second part of Eq. (1).

For any partition $\mathcal{P}$ composed of n segments, Firm θ maximizes its profits with respect to the prices on each segment, denoted by the vector $\mathbf{p}_\theta = (p_{\theta 1}, .., p_{\theta n}) \in \mathbb{R}^n_+$. The profit function of the firms can be written as follows:

$$\pi_{\mathcal{P},\theta} = \sum_{i=1}^{n} d_{\theta i}(\mathbf{p}_\theta, \mathbf{p}_{-\theta}) p_{\theta i}. \tag{2}$$

The timing of the game is the following:

– Stage 1: the data broker chooses the optimal partition, and whether to sell information to one firm or to two firms.

– Stage 2: firms compete and price-discriminate consumers if they acquire information.

3 Optimal information structure

It is clear from the timing of the game described above that the equilibrium prices charged to consumers and the profits of the firms in stage 2 depend, first, on the optimal partition sold by the data broker in stage 1, and second, on the strategy of the data broker to serve either one or two firms on the market. As a consequence, the data broker has to calculate the prices of any possible information structure that can be sold to firms.

In this section, we prove in Theorems 1 and 2 that we can restrict the analysis to particular information structures that are optimal for the data broker. We first analyze the case where the data broker chooses to sell information to only one firm, i.e., the case of an exclusive selling. Second, we characterize the optimal information structure when the data broker sells information to both firms. We find that the data broker sells a partition $\mathcal{P}(\mathbf{p}_1, \mathbf{p}_2)$ that identifies consumers close to the firm up to a cutoff point $\frac{j}{k}$, and that leaves consumers unidentified in the remaining segment. In Section 4, we calculate the number j^* of segments where consumers are identified in the optimal information structure; j^* will depend on the strategy of the data broker (whether it sells information to one or to two firms). We finally discuss at the end of this section how information acquisition affects competition between firms.

3.1 Information is sold to only one firm

When information is sold exclusively to Firm 1 (without loss of generality), the profit-maximizing information structure for the data broker has the following features. Theorem 1 shows that the data broker sells information on all segments up to a point $\frac{j}{k}$, and leaves a large segment of unidentified consumers after that point. In the rest of the paper, we refer to the consumers located on the j segments of size $\frac{1}{k}$ as the *identified consumers*; the remaining consumers located beyond the j segments of size $\frac{1}{k}$ are referred to as the *unidentified consumers*. Figure 3 illustrates Theorem 1.

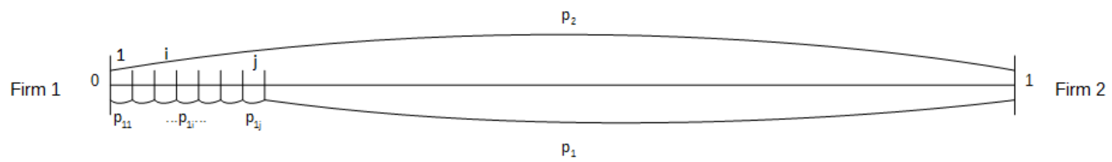

Figure 3: Selling information to one firm: Firm 1 informed

Firm 2 has no information and sets a unique price p_2 over the unit line. Firm 1 can identify consumers on each segment on the left (indexed by $i = 1, .., j$), of size $\frac{1}{k}$. Firm 1 can price-discriminate consumers and sets different prices on each segment, with p_{1i} the price on the ith segment from the origin. Firm 1 sets a price p_1 on the last segment.

Theorem 1. *Let* $\boldsymbol{p}_1 \in \mathbb{R}_+^{j+1}$ *and* $p_2 \in \mathbb{R}_+$. *The profit-maximizing information structure* $\mathcal{P}^*(\boldsymbol{p}_1, p_2)$ *divides the unit line into two segments:*

- *The first segment (closest to the firm buying information) is partitioned in* j *segments of size* $\frac{1}{k}$.
- *Consumers in the second segment of size* $1 - \frac{j}{k}$ *are unidentified.*

Proof: See Appendix A.1.

The proof proceeds in the following way. Consider any information structure. First, we show that the data broker finds it profitable to re-order segments and reduce their size to $\frac{1}{k}$ so that the firm has more information on consumers closest to his product. Second, the data broker can soften competition between firms by leaving a segment of unidentified consumers in the middle.

Theorem 1 makes an important contribution to the existing literature that assumes that the data broker either always sells all information segments to firms, or sells no information at all (Braulin and Valletti, 2016; Montes et al., 2018). We show that this assumption is questionable since selling all segments, i.e., the reference partition of the unit line, is not optimal.

3.2 The data broker sells information to both firms

When information is sold to both firms, the profit maximizing information structure for the data broker has the same features as the optimal partition described in Theorem 1. Theorem 2 shows first that the data broker sells to each firm information on all segments up to a point $\frac{j}{k}$ to Firm 1 and $\frac{j'}{k}$ to Firm 2. Then, it is established that in equilibrium, the data broker sells the same information structure to both firms, that is, $\frac{j}{k} = \frac{j'}{k}$. The remaining consumers are unidentified.

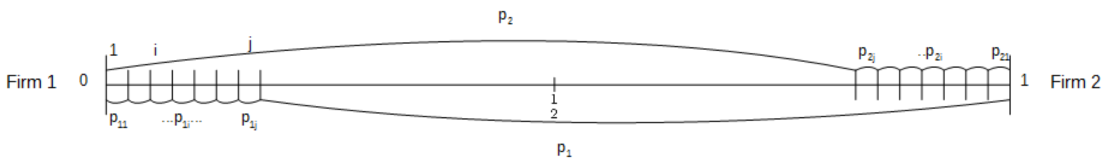

Figure 4: Selling information to both firms

Figure 4 illustrates Theorem 2. Firm 1 (resp. Firm 2) is informed and sets prices p_{1i} (resp. p_{2i}) on each segment of size $\frac{1}{k}$ closest to his location until $\frac{j}{k}$ (resp. $1 - \frac{j}{k}$). After that point, Firm 1 (resp. Firm 2) sets a unique price p_1 (resp. p_2) for the rest of the unit line.

When information is sold to both firms, we rule out situations where firms compete and share demand segments at the extremities of the unit lines. We assume that the data broker does not sell segments allowing firms to poach consumers. We analyze the condition under which both firms have positive demands on a given segment $[\frac{s_i}{k}, \frac{s_{i+1}}{k}]$:

$$C_1 : \ \frac{s_i}{k} \leq \frac{p_2 + t}{2t} \quad \text{and} \quad \frac{p_2 + t}{2t} \leq \frac{2s_{i+1} - s_i}{k}$$

The first part of condition C_1 guarantees that there is a positive demand for Firm 1, while the second part guarantees a positive demand for Firm 2. Inequalities in condition C_1 are expressed as a function of p_2 without loss of generality.

Except for the segment in the middle of the line, we exclude segments located before $\frac{1}{2}$, where Firm 2 has positive demand (and similarly for Firm 1). Thus, we assume that $\frac{p_2+t}{2t} \geq \frac{2s_{i+1}-s_i}{k}$, which is achieved by setting $p_2 = 0$ in the previous inequality (the lowest possible value for p_2): $\frac{1}{2} \geq \frac{2s_{i+1}-s_i}{k}$. Figure 5 illustrates a situation that is ruled out by Assumption A.1.

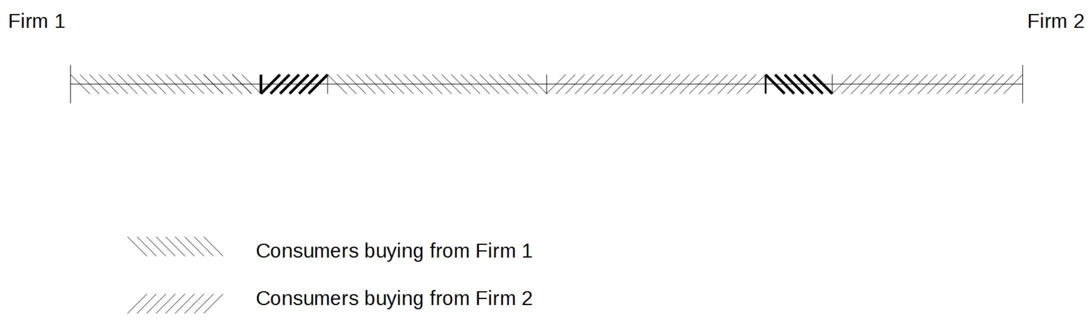

Figure 5: Illustration of a case ruled out by Assumption A.1

Assumption A. 1. ***(No consumer poaching condition)***

When the data broker sells a partition $\mathcal{P} = \{[0, \frac{s_1}{k}], ..., [\frac{s_i}{k}, \frac{s_{i+1}}{k}], .., [\frac{s_n}{k}, 1]\}$ *to Firm 1 and* $\mathcal{P}' = \{[0, \frac{s'_n}{k}], ..., [\frac{s'_{i+1}}{k}, \frac{s'_i}{k}], .., [\frac{s'_1}{k}, 1]\}$ *to Firm 2, the segments verify:* $2\frac{s_{i+1}}{k} - \frac{s_i}{k} \leq \frac{1}{2}$ *and* $2\frac{s'_{i+1}}{k} - \frac{s'_i}{k} \leq \frac{1}{2}$ *for* $i = 0, .., n-2$.[12]

Under Assumption A.1, the optimal partition is similar to the one found in the case of exclusive selling, i.e. when one firm acquires information. The optimal information structure has the following features.

Theorem 2. *Under Assumption A.1, the data broker sells to Firm 1 (resp. Firm 2) a partition with two different types of segments:*

a) There are j *(resp.* j'*) segments of size* $\frac{1}{k}$ *on* $[0, \frac{j}{k}]$ *(on* $[1-\frac{j'}{k}, 1]$ *for Firm 2) where consumers are identified.*

b) Consumers in the second segment of size $1 - \frac{j}{k}$ *are unidentified.*

c) $j = j'$.

[12] We note by convention $s'_0 = s_0 = 0$.

Proof: See Appendix A.1.

The proof proceeds in a similar way as the proof of Theorem 1. We consider any partition satisfying Assumption A.1. We show that the data broker always finds it more profitable to sell segments of size $\frac{1}{k}$. Using the profit function in equilibrium, we then show that selling the same information structure to both firms is optimal, that is $\frac{j}{k} = \frac{j'}{k}$.

Thus, the data broker sells the same information structure to both competitors. This result differs from Belleflamme et al. (2017), where two firms compete on a market for a homogeneous product. Firms can acquire information on their customers to price-discriminate them. They show that firms do not acquire information with the same precision, and a data broker selling information will thus strategically lower the precision of information for one firm.

3.3 Competitive effects of information acquisition

We now interpret how information acquisition affects competition between firms. To do that, we analyze the impact of the acquisition of an additional segment to the optimal partition on their respective profits and prices. More specifically, we compare the changes of prices and profits when Firm 1 acquires an optimal partition $\mathcal{P}$ with the last segment located at $\frac{j}{k}$, and when Firm 1 acquires $\mathcal{P}'$ with the last segment located at $\frac{j+1}{k}$. In the following discussion, Firm 2 remains uninformed.

Purchasing an additional segment will have several impacts on the profits of both firms:

a) Firm 1 price-discriminates consumers on $[\frac{j}{k}, \frac{j+1}{k}]$, which increases its profits.

b) Firm 1 lowers its price on $[\frac{j+1}{k}, 1]$, which increases the competitive pressure on Firm 2. In reaction to this increased competition, Firm 2 lowers its price on the whole unit line ($p_2' < p_2$). The competitive pressure on Firm 1 is increased on the whole unit line as the price of Firm 2 decreases, which has a negative impact on the profits of Firm 1.

The optimal size of the segments where consumers are identified therefore depends on the two opposite effects of information acquisition on firm profits. Following Theorems 1 and 2, it is clear that selling all segments to competing firms is not optimal.

In the following section, we detail the resolution of the game by taking into account the optimal information structure established in Theorems 1 and 2. An informed firm can distinguish

$j + 1$ segments.

4 Model resolution

In this section, we solve the game by backward induction. We compute the equilibrium prices and profits of Firm 1 and 2 using the optimal partition described in Theorems 1 and 2. Then, we analyze whether the data broker sells information to one firm or to both competitors.

4.1 Stage 2: price-setting firms

We note $d_{\theta i}$ the demand of Firm θ on the ith segment. An informed Firm θ maximizes the following profit function with respect to $p_{\theta 1}, .., p_{\theta j}$, and p_θ:

$$\pi_\theta = \sum_{i=1}^{j} d_{\theta i} p_{\theta i} + p_\theta d_\theta. \quad (3)$$

When $j = 0$,[13] the firm does not distinguish any consumer on the unit line, and sets a uniform price as in the standard Hotelling model. An uninformed Firm θ maximizes $\pi_\theta = p_\theta d_\theta$ with respect to p_θ.

The data broker only sells segments of size $\frac{1}{k}$ that are located closest to Firm θ. This partition allows firms to better extract surplus from consumers with the highest willingness to pay. By keeping a segment of unidentified consumers, the data broker softens competition between firms.

Theorems 1 and 2 show that the optimal partition sold by a data broker is therefore not composed of equal-size segments on the whole unit line $[0, 1]$. In this respect, our model can be seen as a generalization of Liu and Serfes (2004) who only consider segments of equal size.

Using Theorems 1 and 2, we characterize the sub-game perfect equilibria for the optimal structure of information by backward induction. There are three cases to consider. In the first case, firms have no information. In the second case, the data broker sells information to one firm. In the third case, the data broker sells information to both firms.

[13] By convention $\sum_{i=1}^{0} d_{\theta i} p_{\theta i} = 0$.

4.1.1 The data broker does not sell information

In this case, firms have no information on consumers and compete in the standard Hotelling framework. Firm θ sets $p_\theta = t$ in equilibrium, and the equilibrium demand is $d_\theta = \frac{p_{-\theta}-p_\theta+t}{2t}$. The profits of Firm θ is $\pi_\theta = \frac{t}{2}$.

4.1.2 The data broker sells information to one firm

Without loss of generality, we assume that only Firm 1 is informed. Firm 1 can distinguish $j+1$ segments of the consumer demand, with j an integer lower than k. Firm 1 price-discriminates by setting a price for each segment p_{1i}. Firm 2 has no information, and sets a uniform price p_2.

Firm 1 maximizes $\pi_1 = \sum_{i=1}^{j} d_{1i}p_{1i} + p_1 d_1$ with respect to p_1, and p_{1i} for $i = 1, .., n$. Firm 2 maximizes $\pi_2 = p_2 d_2$ with respect to p_2.

Profits maximization leads to the prices given in Lemma 1 that we will use to compute the profits of the data broker in Lemma 3.

Lemma 1. *The market equilibrium when the data broker chooses a partition of j segments of size $\frac{1}{k}$ on $[0, \frac{j}{k}]$ and one segment of unidentified consumers on $[\frac{j}{k}, 1]$ are:*

– *Firm 1 obtains the whole demand on each segment $i = 1, .., j$, and*

$$p_{1i} = 2t[1 - \frac{i}{k} - \frac{1}{3}\frac{j}{k}].$$

– *Firms compete on the segment of unidentified consumers, and the prices are*

$$p_1 = t[1 - \frac{4}{3}\frac{j}{k}], \quad \textit{and} \quad p_2 = t[1 - \frac{2}{3}\frac{j}{k}].$$

Proof: See Appendix A.2.

The uniform prices p_1 and p_2 set by both firms both decrease with j. This is the price effect of the intensified competition due to more information on the market. It has a negative effect on firms profits, which is the only effect for Firm 2 that cannot price-discriminate consumers. However, Firm 1 benefits from more information as one more segment of information allows her to charge consumers on this segment with price p_{1i}. Prices for identified consumers p_{1i} decrease with j as a result of the increased competition.

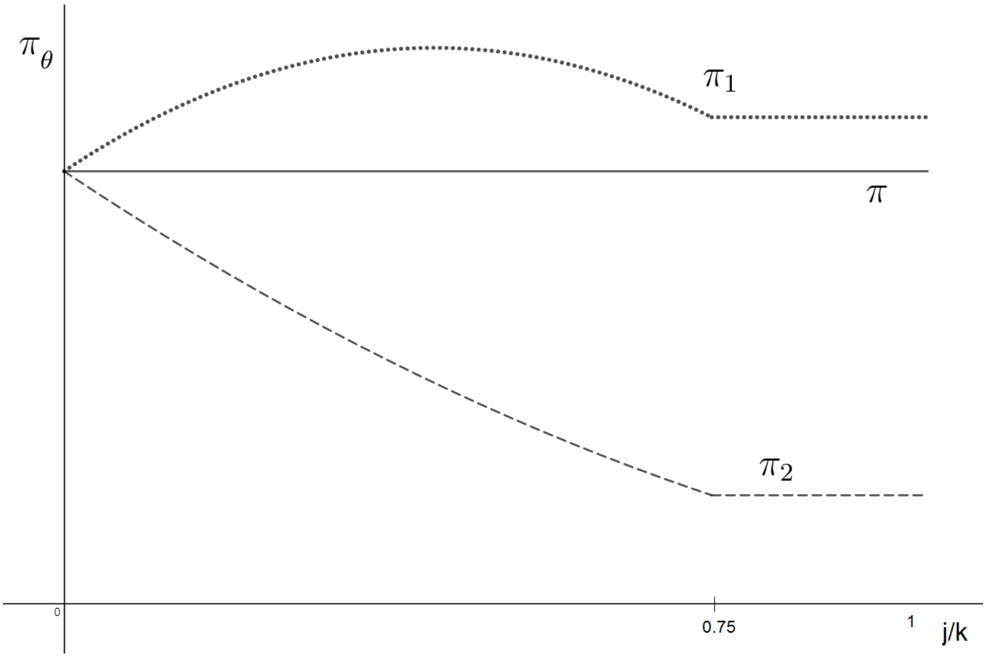

Figure 6: Selling information to one firm: Firms' profits with $t = 1$ and $k = 200$

Figure 6 displays the profits of the firms when only one of them is informed (formulas are given in Appendix A.2). π is the profit of firms in the standard Hotelling framework. On the horizontal axis, the limit between identified and unidentified consumers is given by $\frac{j}{k}$. Firm 1 is informed and makes profits π_1 that depend on $\frac{j}{k}$. Firm 2 is uninformed and makes profits π_2.

Figure 6 illustrates the impacts of information acquisition on firm profits when one firm acquires j segments of size $\frac{1}{k}$ on its closest consumers. We observe that profits of the informed firm follow an inverse U-shaped curve on $[0, \frac{3}{4}]$: more information increases profits of the informed firm when the surplus extraction effect dominates the resulting intensified competition. The profits reach a maximum and then decrease in a second phase. At this point, more information leads to more competition, which dominates the surplus extraction effect, and thus, reduces the profits of the informed firm. The uninformed firm is always harmed when its competitor acquires information and its profits always decrease with j. This is due to the increased competition stemming from a more informed competitor. Comparing firms profits with information to the ones obtained in the standard Hotelling case, we see that the profits of the informed firm (resp. uninformed firm) are always higher (resp. lower) than the profits without information. On $[\frac{3}{4}, 1]$, more information does not change profits and acquiring information on these consumers does not increase profits.

4.1.3 The data broker sells information to both firms

We have shown in Theorem 2 that when the data broker sells information to both firms, it sells the same information structure consisting of j segments of size $\frac{1}{k}$ of identified consumers, and one segment of unidentified consumers. More specifically, Firm 1 can identify j segments, $\{[\frac{i-1}{k}, \frac{i}{k}]\}$ with $i = 1, .., j$ and $j \in \mathbb{N}^*$, and Firm 2 identifies the segments $\{[1 - \frac{i}{k}, 1 - \frac{i-1}{k}]\}$. This leaves a segment of unidentified consumers in the middle of the line $[0, 1]$ where both firms compete. At the extremities of the unit line, both firms price-discriminate identified consumers, as described in Figure 4 of Section 3.2.

Lemma 2 gives the equilibrium prices and profits that we will use to compute the profits of the data broker in Lemma 3.

Lemma 2. *The equilibrium when both firms are informed is characterized by:*

- *For each segment* $i = 1, .., j$, $p_{\theta i} = 2t[1 - \frac{j}{k} - \frac{i}{k}]$.
- *For the segment between* j *and* k*, where firms compete,* $p_\theta = t[1 - 2\frac{j}{k}]$.

Proof: See Appendix A.3.

Similarly to Lemma 1, the prices p_1 and p_2 set by both firms on the share of consumers they cannot identify decrease with j. Prices for identified consumers $p_{\theta i}$ also decrease with j. More information increases competition between firms, which reduces the prices set by firms. However, as information allows firms to identify more consumers, they can charge them with $p_{\theta i}$ instead of p_θ, which has a positive effect on their profits.

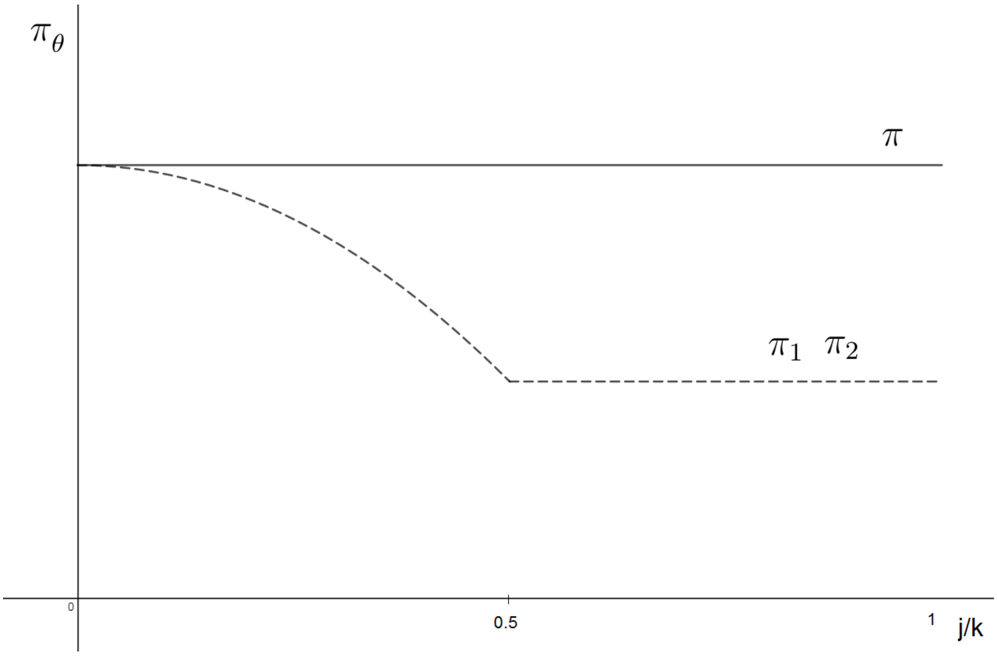

Figure 7: Selling information to both firms: Firms' profits with $t = 1$ and $k = 200$

Figure 7 plots the profits of firms when both of them are informed π_θ as a function of j. On the horizontal axis, $\frac{j}{k}$ is the limit between identified and unidentified consumers $\frac{j}{k}$ (formulas are given in Appendix A.3). When both firms are uninformed, their profits π are given by the standard Hotelling formulas of Section 4.1.1. When both firms acquire information, their profits always decrease with j, and reach a minimum when the data broker sells information on all segments of size $\frac{1}{k}$ on $[0, \frac{1}{2}[$.[14] Beyond $\frac{1}{2}$, more information does not affect the market and profits do not change.

Figure 7 confirms that firms acquiring information face a prisoner's dilemma as in Ulph and Vulkan (2000) and (Stole, 2007). Profits are lower when both firms are informed than when both firms are uninformed (standard Hotelling framework). Both competing firms acquire information even though it leads to a more competitive market because a firm that remains uninformed would have even lower profits if the competitor is informed (see Figure 6).

[14] Consumers at $\frac{1}{2}$ are naturally excluded by Assumption 1.

4.2 Stage 1: profits of the data broker

The data broker can choose among the set of allowable partitions that we have proved to be optimal. The data broker compares the three different outcomes analyzed in Stage 2: selling no information, selling information to only one firm or selling information to both competitors. When no information is sold, the data broker makes no profits, and we refer to this case as the outside option.

Using Lemma 1 and Lemma 2, we maximize the profits of the data broker with respect to j, first when only one firm is informed, and second when both firms are informed. Using Theorems 1 and 2, profits are straightforward to compute, following the mechanism explained in Section 2.2, and are given in Lemma 3.

Lemma 3. *The profits of the data broker are:*

– *When the data broker sells information to only one firm:*

$\Pi_1(j) = w_1(j) = \pi_\theta^{I,NI}(j) - \pi_\theta^{NI,NI} = \frac{2jt}{3k} - \frac{7t}{9}\frac{j^2}{k^2} - \frac{tj}{k^2}$.

– *When the data broker sells information to both competitors:*

$\Pi_2(j) = 2w_2(j) = 2[\pi_\theta^{I,I}(j) - \pi_\theta^{NI,I}(j)] = 2[\frac{2jt}{3k} - \frac{11j^2t}{9k^2} - \frac{jt}{k^2}]$.

4.3 Characterization of the equilibrium

We characterize in this section the number of segments of information sold to firms when only one firm is informed and when both firms are informed. We then compare the profits of the data broker in both cases, and we show that the data broker always sell information to both competitors in equilibrium.

4.3.1 The optimal number of segments

Using Lemma 3, we first find the optimal values of j when one or both firms are informed, then we compare the profits in both situations.[15]

Lemma 4.

[15] For the proof of Lemma 4, we assume that j is defined over $\mathbb{R}$, and the resulting j chosen by the data broker is the integer part of j^*.

– *When one firm buys information, the data broker sets*

$$j_1^* = \frac{6k-9}{14}.$$

– *When both firms buy information, the data broker sets*

$$j_2^* = \frac{6k-9}{22}.$$

4.3.2 Optimal choice of the data broker

From Lemma 4, we can finally calculate the optimal choice of the data broker by comparing his profits when it sells information to one firm or to both firms:

Proposition 1. *Suppose A.1, the data broker optimally sells information to both firms:*

$$\Pi_2^* \geq \Pi_1^*.$$

Proof of Proposition 1: The proof is straightforward. We compare the profits of the data broker when it sells information to one firm or to both firms. The difference of both profits gives $\Pi_2(j_2^*) - \Pi_1(j_1^*) = \frac{(12k^2-36k+27)t}{308k^2}$, *which is positive for any* $k \geq 2$. □

Proposition 1 states that the profits of the data broker are higher when it sells information to both firms rather than to one firm. Proposition 1 contrasts with the result established by Montes et al. (2018) who find that it is always optimal to sell information to only one firm.

Two main reasons explain this difference. First, following Jehiel and Moldovanu (2000), Montes et al. (2018) assume that firms acquire information through a second-price auction with negative externalities. The negative externalities result from the fact that for a firm, losing the auction means that its competitor has the option to buy information. Thus, the outside option of the bidder, which has an impact on the prices that firms are willing to pay, is less favorable when the competitor has the option to purchase information.

To assess the impact of the selling mechanism, we analyze how our main results change with a second-price auction mechanism. First, profit with the optimal partition (j maximal for each price) is higher in Montes et al. (2018) than in our model:

$$\underbrace{\tilde{w}_1 = (\pi_\theta^{I,NI} - \pi_\theta^{NI,I})_{j=\tilde{j}^*} = \frac{4t}{9} - \frac{2t}{3k} + \frac{t}{4k^2}}_{\text{Price with the mechanism used in Montes et al. (2018)}} > \underbrace{w_1 = (\pi_\theta^{I,NI} - \pi_\theta^{NI,NI})_{j=j^*} = \frac{t}{7} - \frac{3t}{7k} + \frac{9t}{28k^2}}_{\text{Price with our mechanism}}$$

The data broker can therefore extract more surplus in Montes et al. (2018) than in our framework.

Second, Montes et al. (2018) implicitly assume that it is optimal for the data broker to sell all segments to firms, which increases in turn competition. In our framework, the data broker finds it optimal to leave low-valuation consumers unidentified, i.e., those who have a low willingness to pay, to soften competition between competing firms.

To conclude, it is straightforward to show that using the auction selling mechanism of Montes et al. (2018) with the optimal information structure characterized in Theorems 1 and 2 leads to an equilibrium with an exclusive sale (one firm is informed, while the other is uninformed). Moreover, assuming that the data broker sells all information segments instead of the optimal partition found in Theorems 1 and 2 leads to an equilibrium with a non-exclusive sale (where both firms are informed). As a result, using a non-optimal information structure does not change the nature of the equilibrium (non-exclusive sale), but using a second-price auction modifies the optimal choice of the data broker (it sells information to only one firm, regardless of whether it is using the optimal information structure or not). Therefore, the assumption related to the selling mechanism is crucial to understand information acquisition, and may have regulatory implications: more information is acquired when the data broker is forced to write an exclusive contract that guarantees that the information sold to a firm will not be sold to another firm if the offer is declined.

4.3.3 Welfare analysis

We analyze in this section two effects of information acquisition on total welfare. First, firms are in a prisoner's dilemma and both suffer from information acquisition compared to a situation without information acquisition. Second, consumer surplus increases. Overall, the total welfare remains constant.[16]

These results are detailed in Corollary 1.

[16] This result is due to our market coverage assumption; as there is no surplus created by a market expansion, we are in a classical zero-sum game model.

Corollary 1.

– *Firm profits in equilibrium are lower than the profits in the standard Hotelling model:*

$$\Delta\pi_\theta(k) = \pi_\theta^{I,I}(k) - \pi_\theta^{NI,NI}(k) < 0.$$

– *Consumer surplus is higher than in the standard Hotelling model:*

$$\Delta CS(k) > 0.$$

– *Depending on their willingness to pay, consumers gain or lose surplus when both firms are informed. Compared to the uniform price $p_\theta^{NI,NI} = t$ set in the standard Hotelling model without information:*[17]

  – *On $[0, \frac{5k-9}{22k}]$ consumers pay a higher price: $p_{\theta i}^{I,I} \geq p_\theta^{NI,NI}$, and are identified.*
  – *On $[\frac{5k-9}{22k}, \frac{1}{2}]$ consumers pay a lower price: $p_{\theta i}^{I,I}, p_\theta^{I,I} \leq p_\theta^{NI,NI}$.*
    * *Consumers on $[\frac{5k-9}{22k}, \frac{6k-9}{22k}]$ are identified.*
    * *Consumers on $[\frac{6k-9}{22k}, \frac{1}{2}]$, are unidentified.*

– *Total surplus remains constant in the market: $\Delta CS(k) + \Delta\pi_\theta(k) = 0$.*

Firms are therefore in a prisoner's dilemma that can be explained as follows. Information acquisition has two opposite effects on firm profits. First, more information allows firms to better extract the surplus of consumers who have a high willingness to pay. Second, firms also compete more intensively for each consumer, which decreases their profits. Overall, firm profits are lower when they both acquire information: the consumer surplus that firms can extract with information does not offset the loss of profits due to a tougher competition.

The existing literature overestimates the effects of data brokers on prices paid by consumers. Indeed, when firms have information on each consumer, they compete more intensively, resulting in lower prices. For instance in Baye and Morgan (2001), firms end up competing à la Bertrand, making zero profits in equilibrium. Our model shows on the contrary that a data broker has

[17] We consider the prices on $[0, \frac{1}{2}]$, the prices on the rest of the line can be found directly by symmetry.

always incentives to soften competition, which increases prices. Consumers are thus relatively worse off when the data broker behaves strategically.

Information acquisition by competing firms has however a positive effect on consumer surplus. Due to increased competition, unidentified consumers located in the middle of the Hotelling line benefit from lower prices even though firms extract more surplus from identified consumers. Overall, information acquisition still benefits consumers even though firms price-discriminate high valuation consumers.

Finally, turning to consumer identification, the share of unidentified consumers, $\frac{6k-9}{11k}$, increases with information quality k. As information on consumers obtained by a data broker becomes more precise, the share of identified consumers increases. Similarly, the share of consumers with a lower level of utility compared with the standard Hotelling model increases with k. As information becomes more precise, the share of consumers losing utility increases. Overall, the gain in consumers' surplus ΔCS decreases with k.[18]

4.3.4 First-degree price-discrimination

We finally study in this last part how first-degree price-discrimination impacts the data broker's strategies. Three reasons motivate this analysis. First we generalize the model with third-degree price-discrimination to test the robustness of our results. Second, as Montes et al. (2018) focus on first-degree price-discrimination, considering it allows us to compare our results with theirs. Third, as digital technologies allow for better information collection and better classification and targeting of consumers, which increase in turn the quality of information, equilibrium under perfect consumer recognition is important to consider.

We show that our model with third-degree price-discrimination converges to a model with first-degree price-discrimination, which is a special case of our baseline model developed in the previous section when $k \rightarrow +\infty$. We show that under first-degree price-discrimination, the data broker sells to each firm an information structure that is similar to the one found in Theorem 2: one segment of consumers is fully identified, and consumers on the other segment are unidentified.

[18] For $k \geq 10$.

Corollary 2. *When $k \to +\infty$, firms first-degree price-discriminate and the data broker sells to both firms an information structure characterized for Firm 1 (and symmetrically for Firm 2) as:*

– *on $[0, \frac{3}{11}]$, consumers are identified.*

– *on $[\frac{3}{11}, 1]$, consumers are unidentified.*

Proof: See Appendix A.5.

From Lemma 2, it is straightforward to show that the profits and consumer surplus under third-degree price-discrimination converge to their corresponding values under first-degree price-discrimination: $\pi_\theta^{I,I} \xrightarrow[k\to\infty]{} \frac{103}{242}t$ and $\Delta CS \xrightarrow[k\to\infty]{} \frac{18}{121}t$. Additionally, consumers on $[\frac{5}{22}, \frac{17}{22}]$ benefit from lower prices when firms are informed.

5 Conclusion

Data brokers are major players in the Internet economy. They collect and treat a huge amount of data on consumers that they can choose to sell to firms for various objectives, including price-discrimination. Data brokers can affect competition on a market by deciding to which firms they want to sell information, and by choosing the amount and the quality of information that they sell. Their role has therefore been scrutinized by regulators and legislators (Crain, 2018).

However, despite the intense debates in the last years, few economic studies have been carried out on data brokers. This article contributes to this developing literature. In particular, we challenge a simple assumption traditionally made in previous research regarding which information structure is chosen by data brokers (and acquired by firms): firms are supposed to purchase either all available information on consumers or no information at all. This 'all-or-nothing' assumption, although simple in appearance, is not in line with current marketing practices, and has strong implications on market equilibrium. First, firms are interested in acquiring only information on specific socio-demographic groups, and not on all potential consumers. Second, this assumption leads to a market equilibrium in which the competitive pressure is stronger than in our model.

We developed a model in which the data broker can choose among a large set of possible information structures to sell to firms. Extending the setting of Liu and Serfes (2004), we proved that the 'all or nothing' strategy is not optimal and never occurs in equilibrium. The optimal information structure segments consumers into two groups: on the one hand, consumers with the highest willingness to pay are identified, and, on the other hand, low-valuation consumers remain unidentified. This strategy allows the data broker to soften competition between firms, and to increase its profits. This information structure is optimal regardless of the market configuration (whether information is sold to only one firm or to both firms).

The optimal selling strategy for the data broker is clearly affected by the mechanism used to sell information. Using a second-price auction with externalities increases competition between firms for acquiring information because losing the auctions has a higher opportunity cost since the competitor might win the auction (Braulin and Valletti, 2016; Montes et al., 2018). In this case, selling information to only one firm is optimal for the data broker. Real-time bidding auctions are second price auctions and for some advertising spaces for which competing firms bid, losing the auction might imply that the competitor wins. In this setting, the assumption used by Montes et al. (2018) might be justified. However, for the vast majority of online auctions, bidders are not competitors, and even when competitors are bidding for the same auction space, they might not know it. In this case, our assumption that the price of information is driven by an outside option where both firms are uninformed is better suited, and leads to an equilibrium where both firms acquire information. Moreover, programmatic buying technologies, such as real-time bidding do not guarantee prices, contextual placement, or impression volume to advertisers, who might prefer direct sales.[19] Direct sales involve human bilateral negotiations for which the identities of competitors has no impact, and negative externalities due to losing the auction play a minor role. Using a selling mechanism where the outside option for a firm is a situation in which both firms are uninformed, is therefore also important to consider.

Overall, selling information in a second-price auction with externalities or using our selling mechanism has strong regulatory implications. Indeed, consumer welfare is directly impacted by the price that they pay and by the amount of personnal data that are sold by data brokers. Our results suggest that it is more important to overview how data brokers sell consumer data to firms rather than setting up a committee to monitor exclusive deals between data brokers and firms, such as recommended by Montes et al. (2018).

[19] In addition, directly sold display ads see higher viewability rates than programmatic display ads (Bounie et al., 2017).

Our model can be used to address two recent privacy issues. First, our model suggests that new privacy policies in the European Union could increase consumer surplus. Stronger privacy protection in Europe means that firms now distinguish coarser consumers segments, which lowers the precision of information structures modeled by k. When k decreases, the share of unidentified consumers increases. Overall, consumer surplus increases with privacy protection regulation. Second, the share of identified consumers is higher when both firms are informed than when only one firm is informed. Therefore, consumer privacy is impacted by the amount of personnal data collected and sold by data brokers: selling information to all firms reduces the share of unidentified consumers.[20]

Finally, our model could be improved in several directions. First, we assumed that the data broker can extract all rents from firms acquiring information. This assumption could be relaxed with profit-sharing rules (such as Nash bargaining), without qualitatively changing our results. Second, we assumed that the market is covered. Further research should relax this assumption to consider market expansion effects.

A Appendix

A.1 Proof of Theorems 1 and 2

In Appendix A.1, we show that the data broker optimally sells a partition that divides the unit line into two segments. The first segment identifies the closest consumers to a firm and is partitioned in j segments of size $\frac{1}{k}$. The second segment is of size $1 - \frac{j}{k}$ and leaves unidentified the other consumers. We first establish this claim when the data broker sells information to only one firm (Proof 1.a), and second when it sells to both firms (Proof 1.b).[21]

Proof of Theorem 1: the data broker sells information to only one firm

The data broker can choose any partition in $\mathbb{P}$ to sell to Firm 1 (without loss of generality). There are three types of segments to consider:

– Segments A, where Firm 1 is in constrained monopoly;

[20] See Acquisti et al. (2016) for a review of the literature on privacy.

[21] All along the proofs, we refer to Liu and Serfes (2004) who prove the continuity and concavity of the profit functions on the segments.

– Segments B, where Firms 1 and 2 compete;[22]

– Segments C, where Firm 1 gets no demand.

We proceed in three steps. In step 1 we analyze the segments of type A and show that on any such segment, it is optimal to sell a partition that divides the segment into segments of size $\frac{1}{k}$. In step 2, we show that all segments of type A are located closest to Firm 1. In step 3 we analyze segments of type B and we show that it is always more profitable to sell a coarser partition on segments where firms compete. Therefore, segment B has only one segment of size $1 - \frac{j}{k}$ where $\frac{j}{k}$ separates segments A and B. Finally, we can discard segments of type C because information on consumers on these segments does not increase profits.

Step 1: We analyze segments of type A where Firm 1 is in constrained monopoly, and show that reducing the size of segments to $\frac{1}{k}$ is optimal.

Consider any segment $[\frac{i}{k}, \frac{i+l}{k}]$ with l, i integers verifying $i + l \leq k$, such that Firm 1 is in constrained monopoly on this segment. We show that selling a finer partition of this segment increases the profits of Firm 1. To prove this claim, we establish that Firm 1 profits is higher with a finer partition $\mathcal{P}'$ with two segments : $[\frac{i}{k}, \frac{i+1}{k}]$ and $[\frac{i+1}{k}, \frac{i+l}{k}]$ than with a coarser partition $\mathcal{P}$ with one segment $[\frac{i}{k}, \frac{i+l}{k}]$.

Figure 8: Step 1: segments of type A

Figure 8 shows on the left panel the profits for a coarse segment of type A, and on the right, finer segments of type A. We compare profits in both cases to show that the finer segmentation is more profitable for firms. We write $\pi_1^A(\mathcal{P})$ and $\pi_1^{AA}(\mathcal{P}')$ the profits of Firm 1 on $[\frac{i}{k}, \frac{i+l}{k}]$ for respectively partitions $\mathcal{P}$ and $\mathcal{P}'$.

[22] A segment $[\frac{i}{k}, \frac{i+l}{k}]$ is of type B if and only if it verifies condition C_1 : $\frac{i}{k} \leq \frac{p_2+t}{2t}$ and $\frac{p_2+t}{2t} - \frac{l}{k} \leq \frac{i+l}{k}$ for any integers $0 \leq i \leq k-1$ and $1 \leq l \leq k-i$. These restrictions are naturally imposed by the structure of the model.

First, profits with the coarser partition is: $\pi_1^A(\mathcal{P}) = p_{1i}d_1 = p_{1i}\frac{l}{k}$. The demand is $\frac{l}{k}$ since Firm 1 gets all consumers by assumption; p_{1i} is such that the indifferent consumer x is located at $\frac{i+l}{k}$:

$$V - tx - p_{1i} = V - t(1-x) - p_2 \implies x = \frac{p_2 - p_{1i} + t}{2t} = \frac{i+l}{k} \implies p_{1i} = p_2 + t - 2t\frac{i+l}{k},$$

with p_2 be the price charged by (uninformed) Firm 2. This price is only affected by strategic interactions on the segments where firms compete, and therefore are not influenced by the pricing strategy of Firm 1 on type A segments.

We write the profit function for any p_2, replacing p_{1i} and d_1:
$\pi_1^A(\mathcal{P}) = \frac{l}{k}(t + p_2 - \frac{2(l+i)t}{k})$
Secondly, using a similar argument, we show that the profits on $[i, i+l]$ with partition $\mathcal{P}'$ is:
$\pi_1^{AA}(\mathcal{P}') = \frac{1}{k}(t + p_2 - \frac{2(1+i)t}{k}) + \frac{l-1}{k}(t + p_2 - \frac{2(l+i)t}{k})$
Comparing $\mathcal{P}$ and $\mathcal{P}'$ shows that profits using the finer partition increases by $\frac{2t}{k^2}(l-1)$, which establishes the claim.

By repeating the previous argument, it is easy to show that the data broker will sell a partition of size $\frac{l}{k}$ with l segments of equal size $\frac{1}{k}$.

Step 2: We show that all segments of type A are closest to Firm 1 (located at* 0 *on the unit line by convention).

There are two cases to analyze: first, a segment of type B is closest to Firm 1 and is adjacent to a segment of type A, and second, a segment of type A is closest to Firm 1 and is adjacent to a segment of type B.

The two cases are shown in Figure 9 and correspond respectively to the partitions $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{P}}'$. The curved line represents the demand of Firm 1, which does not cover type B segments. In partition $\tilde{\mathcal{P}}$, a segment of type B of size $\frac{l}{k}$ is followed by a segment of type A of size $\frac{1}{k}$. We know that segments of type A have at least one segment of size $\frac{1}{k}$ and therefore, this segment can be followed by a segment of type A or B. We show that segments of type A are always located closest to Firm 1 by proving that it is always optimal to change partition starting with segments of type B with a partition starting with segments of type A like in partition $\tilde{\mathcal{P}}'$. We show this claim by analyzing the change between two segments, one of type A, and one of type B. The rest of the segment remains unchanged. Without loss of generality, segments of type A are of size $\frac{1}{k}$ and are located at $\frac{u_i-1}{k}$, and segments of type B, are located at $\frac{s_i}{k}$ and are of size $\frac{l_i}{k}$.[23]

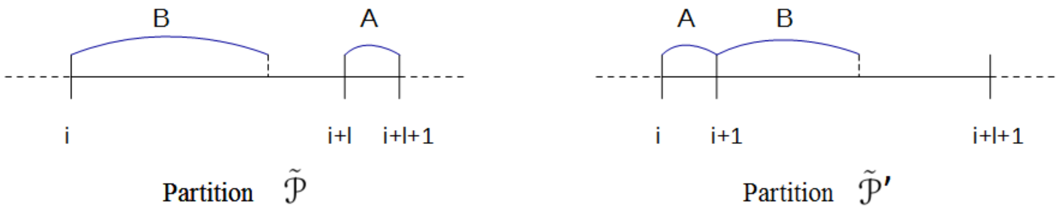

Figure 9: Step 2: relative position of type A and type B segments

A segment of type B located at $[\frac{i}{k}, \frac{i+l}{k}]$ is non null (has a size greater than $\frac{1}{k}$), if the following restrictions imposed by the structure of the model are met: respectively positive demand and the existence of competition.

Condition 1. *For any integers* $0 \leq i \leq k-1$ *and* $1 \leq l \leq k-i-1$.

$$C_1 : \quad \frac{i}{k} \leq \frac{p_2+t}{2t} \quad \textit{and} \quad \frac{p_2+t}{2t} - \frac{l}{k} \leq \frac{i+l}{k}$$

We can rewrite profits of Firm 1 as the sum of two terms, the first term presents the profits on segments of type A where prices are denoted by p'_{1i}, and the second term the profits on segments of type B, where prices are denoted by p_{1i}.

There are h segments of type A, of size $\frac{1}{k}$. On each of these segments, the demand is $\frac{1}{k}$.

There are n segments of type B. We find the demand for Firm 1 on these segments using the location of the indifferent consumer:

$$d_{1i} = x - \frac{s_i}{k} = \frac{p_2 - p_{1i} + t}{2t} - \frac{s_i}{k}$$

We can rewrite the profits as:

$$\pi_1(\mathcal{P}) = \sum_{i=1}^{h} p'_{1i}\frac{1}{k} + \sum_{i=1}^{n} p_{1i}[\frac{p_2 - p_{1i} + t}{2t} - \frac{s_i}{k}] \tag{4}$$

[23] With u_i and s_i integers below k.

Profits of Firm 2 are generated on segments of type B, where the demand for Firm 2 is

$$d_{2i} = \frac{s_i + l_i}{k} - x = \frac{p_{1i} - p_2 - t}{2t} + \frac{s_i + l_i}{k}$$

Profits of Firm 2 can be written therefore as

$$\pi_2(\mathcal{P}) = \sum_{i=1}^{n} p_2[\frac{p_{1i} - p_2 - t}{2t} + \frac{s_i + l_i}{k}] \tag{5}$$

Firm 1 maximizes profits $\pi_1(\mathcal{P})$ with respect to p_{1i} and p'_{1i}, and Firm 2 maximizes $\pi_2(\mathcal{P})$ with respect to p_2, both profits are strictly concave.

The prices in equilibrium are:

$p'_{1i} = t + p_2 - 2\frac{u_i t}{k}$,

$p_{1i} = \frac{p_2+t}{2} - \frac{s_i t}{k} = \frac{t}{3} + \frac{2t}{3n}[\sum_{i=1}^{n}[\frac{s_i}{2k} + \frac{l_i}{k}]] - \frac{s_i t}{k}$,

and

$p_2 = -\frac{t}{3} + \frac{4t}{3n}\sum_{i=1}^{n}[\frac{s_i}{2k} + \frac{l_i}{k}]$.

We can now compare profits with $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{P}}'$. When we move segments of type B from the left of segments of type A to the right of segment of type A, it is important to check that Firm 1 is still competing with Firm 2 on each segment of type B, and that Firm 1 is still in constrained monopoly on segments of type A. The second condition is met by the fact that price p_2 is higher in $\tilde{\mathcal{P}}'$ than in $\tilde{\mathcal{P}}$. The first condition is guaranteed by $C_1 : \frac{p_2+t}{2t} - \frac{l_i}{k} \leq \frac{s_i+l_i}{k}$ for some segments s_i of size l_i. By abuse of notation, let s_i denote the segment located at $[\frac{s_i}{k}, \frac{s_i+l_i}{k}]$, which corresponds to segments of type B that satisfy these condition. Let $\tilde{s}_i$ denote the m segments ($m \in [0, n-1]$) of type B located at $[\frac{\tilde{s}_i}{k}, \frac{\tilde{s}_i+\tilde{l}_i}{k}]$ that do not meet these condition, and are therefore turned into segments of type A.

Noting $\hat{p}_2$ and $\hat{p}_{1i}$ the prices with $\tilde{\mathcal{P}}'$, we have:

$$\begin{aligned} \hat{p}_2 &= \frac{4t}{3(n-m)}[-\frac{n}{4} + \sum_{i=1}^{n}[\frac{s_i}{2k} + \frac{l_i}{k}] + \frac{m}{4} + \frac{1}{2k} - \sum_{i=1}^{m}\frac{\tilde{s}_i}{2k}] \\ &= p_2 + \frac{4t}{3(n-m)}[\frac{3mp_2}{4t} + \frac{1}{2k} + \frac{m}{4} - \sum_{i=1}^{m}\frac{\tilde{s}_i}{2k}] \end{aligned} \tag{6}$$

$$\hat{p}_{1i} = p_{1i} + \frac{1}{2}\frac{4t}{3(n-m)}[\frac{3mp_2}{4t} + \frac{1}{2k} + \frac{m}{4} - \sum_{i=1}^{m}\frac{\tilde{s}_i}{2k}]$$

for segments of type B where condition C_1 holds;

$$\hat{p}_{1i} = p_{1i} + \frac{1}{2}\frac{4t}{3(n-m)}[\frac{3mp_2}{4t} + \frac{1}{2k} + \frac{m}{4} - \sum_{i=1}^{m}\frac{\tilde{s}_i}{2k}] - \frac{t}{k}$$

for segments of type A, where condition C_1 is violated.

Let us compare the profits between $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{P}}'$. Let π_1^{BA} denote the profits of Firm 1 with $\tilde{\mathcal{P}}$, and π_1^{AB} the profits of Firm 1 with $\tilde{\mathcal{P}}'$ on $[\frac{i}{k}, \frac{i+l+1}{k}]$. To compare profits that result by moving segment located at $\frac{i+l}{k}$ to $\frac{i}{k}$ (A to B), we proceed in two steps. First we show that profits increase on $[\frac{i}{k}, \frac{i+l+1}{k}]$ increase, and that p_2 increases as well; and secondly we show that profits on the other type B segments of the line also increase.

First we show that profits increase on $[\frac{i}{k}, \frac{i+l+1}{k}]$, that is, we show that $\Delta\pi_1 = \pi_1^{AB} - \pi_1^{BA} \geq 0$

$$\begin{aligned}\Delta\pi_1 =& \pi_1^{AB} - \pi_1^{BA} \\ =& \frac{1}{k}[\hat{p}_2 - 2\frac{it}{k} - p_2 + 2\frac{i+l}{k}t] + \hat{p}_{1i}[\frac{\hat{p}_2 - \hat{p}_{1i} + t}{2t} - \frac{i+1}{k}] - p_{1i}[\frac{p_2 - p_{1i} + t}{2t} - \frac{i}{k}]\end{aligned} \quad (7)$$

By definition, $\tilde{s}_i$ verifies C_1 thus $\frac{\tilde{s}_i}{k} \leq \frac{p_2+t}{2t}$, which allows us to establish that $\frac{4t}{3(n-m)}[\frac{3mp_2}{4t} + \frac{1}{2k} + \frac{m}{4} - \sum_{i=1}^{m}\frac{\tilde{s}_i}{2k}] \geq \frac{2t}{3nk}$. It is then immediate to show that

$$\Delta\pi_1 \geq \frac{t}{k}[1 - \frac{1}{3n}][\frac{2}{k}\frac{3nl+1}{3n-1} - \frac{p_2}{2t} - \frac{1}{2} - \frac{1}{6nk} + \frac{i}{k} + \frac{1}{2k}] \quad (8)$$

Also, by assumption, firms compete at $\frac{i}{k}$ with $\tilde{\mathcal{P}}$, which implies that C_1 is verified, and in particular, $\frac{p_2+t}{4t} - \frac{i}{2k} \leq \frac{l}{k}$.

Thus:

$$\Delta\pi_1 \geq \frac{t}{k}[1 - \frac{1}{3n}][\frac{2}{k}\frac{3nl+1}{3n-1} - \frac{2l}{k} - \frac{1}{6nk} + \frac{1}{2k}] \geq 0 \quad (9)$$

The profits on segment $[\frac{i}{k}, \frac{i+l+1}{k}]$ are higher with $\tilde{\mathcal{P}}'$ than with $\tilde{\mathcal{P}}$.

Second we consider the profits of Firm 1 on the rest of the unit line. We write the reaction functions for the profits on each type of segments, knowing that $\hat{p}_2 \geq p_2$.

For segments of type A: $\frac{\partial}{\partial p_2}\pi_{1i}^A = \frac{\partial}{\partial p_2}(\frac{1}{k}[t + p_2 - 2\frac{u_i t}{k}]) = \frac{1}{k}$: a higher p_2 increases the profits.

For segments of type B: $\frac{\partial}{\partial p_2}\pi_{1i}^B = \frac{\partial}{\partial p_2}(p_{1i}[\frac{p_2 - p_{1i} + t}{2t} - \frac{s_i}{k}]) = \frac{\partial}{\partial p_2}(\frac{1}{2t}[\frac{p_2+t}{2} - \frac{s_i t}{k}]^2) = \frac{1}{2t}[\frac{p_2+t}{2} - \frac{s_i t}{k}]$, which is greater than 0 since $\frac{p_2+t}{2} - \frac{s_i t}{k}$ is the expression of the demand on this segment, which is positive under C_1.

Thus for any segment, the profits of Firm 1 increase with $\tilde{\mathcal{P}}'$ compared to $\tilde{\mathcal{P}}$.

Result 1. *By iteration, we conclude that type A segments are always at the left of type B segments.*

Step 3: We now analyze segments of type B where firms compete. Starting from any partition but the trivial partition, of size $\frac{j}{k}$, we show that it is always more profitable to sell a coarser partition.

Since there are only two possible types of segments (A and B) and that we have shown that segments of type A are the closest to the firms, segment B is therefore further away from the firm as shown in Figure 3.

We prove this result by showing that if Firm 1 has a partition of two segments where she competes with Firm 2, a coarser partition produces a higher profits. We compute the profits of the firm on all the segments where firms compete, and compare the two situations described below with partition $\hat{\mathcal{P}}$ and partition $\hat{\mathcal{P}}'$.

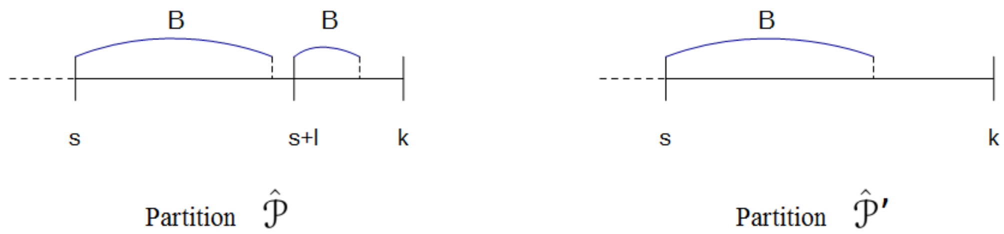

Figure 10: Step 3: segments of type B

$\hat{\mathcal{P}}$ partitions the segment $[\frac{i}{k}, 1]$ in two segments $[\frac{i}{k}, \frac{i+l}{k}]$ and $[\frac{i+l}{k}, 1]$, while $\hat{\mathcal{P}}'$ only gives segment $[\frac{i}{k}, 1]$. We compare the profits of the firm on the segments where firms compete and we show that $\hat{\mathcal{P}}'$ induces higher profits.

There are $n + 1$ segments of type B where firms compete initially with partition $\hat{\mathcal{P}}$. Since partition $\hat{\mathcal{P}}'$ is coarser than partition $\hat{\mathcal{P}}$, the segment located at $\frac{i+l}{k}$ is removed. n of the initial segments of type B remain, and are not necessarily of type B any more.

We proved in step 2 that prices can be written as:

$$p_2 = -\frac{t}{3} + \frac{4t}{3(n+1)} \sum_{i=1}^{n+1} [\frac{s_i}{2k} + \frac{l_i}{k}], \tag{10}$$

$$\begin{aligned} p_{1i} &= \frac{p_2 + t}{2} - \frac{s_i t}{k} \\ &= \frac{t}{3} + \frac{2t}{3(n+1)} \sum_{i=1}^{n+1} [\frac{s_i}{2k} + \frac{l_i}{k}] - \frac{s_i t}{k}. \end{aligned} \tag{11}$$

Let p_{1s} and p_{1s+l} be the prices on the penultimate and on the last segments when the partition is $\hat{\mathcal{P}}$.

$$p_{1s} = \frac{p_2 + t}{2} - \frac{st}{k}, \ \ p_{1s+l} = \frac{p_2 + t}{2} - \frac{s+l}{k} t,$$

$\hat{p}_2$ and $\hat{p}_{1s}$ are the prices of Firm 2 and of Firm 1 on the last segment of partition $\hat{\mathcal{P}}'$.

Again condition C_1 might be violated. We note $\tilde{s}_i$ the m segments where it is the case. We then have:

$$\begin{aligned} \hat{p}_2 &= \frac{4t}{3(n-m)} [-\frac{n-m}{4} + \sum_{i=1}^{n} [\frac{s_i}{2k} + \frac{l_i}{k}] - \sum_{i=1}^{m} \frac{\tilde{s}_i}{2k}] \\ &= \frac{4t}{3(n-m)} [-\frac{n+1}{4} + \sum_{i=1}^{n+1} [\frac{s_i}{2k} + \frac{l_i}{k}] + \frac{m+1}{4} - \sum_{i=1}^{m} \frac{\tilde{s}_i}{2k} - \frac{s+l}{2k}] \\ &= p_2 + \frac{4t}{3(n-m)} [\frac{3(m+1)p_2}{4t} + \frac{m+1}{4} - \sum_{i=1}^{m} \frac{\tilde{s}_i}{2k} - \frac{s+l}{2k}] \\ &\geq p_2 + \frac{4t}{3(n-m)} [\frac{3}{4t} p_2 + \frac{m p_2}{2t} + \frac{1}{4} - \frac{s+l}{2k}] \end{aligned}$$

$$\hat{p}_{1s} = \frac{\hat{p}_2 + t}{2} - \frac{st}{k}$$

We can write the profits of Firm 1 on type B segments of partition $\hat{\mathcal{P}}$ only. We can prevent ourselves from considering segments where Firm 1 is a monopolist since, as we will show, p_2 increases when changing from partition $\hat{\mathcal{P}}$ to $\hat{\mathcal{P}}'$, and thus Firm 1's profits on these segment will increase.

$$\begin{aligned}\pi_1(\hat{\mathcal{P}}) &= \sum_{i=1,s_i\neq\tilde{s}_i}^{n} p_{1i}[\frac{p_2+t}{4t}-\frac{s_i}{2k}] + \sum_{i=1}^{m} p_{1i}[\frac{p_2+t}{4t}-\frac{\tilde{s}_i}{2k}] + p_{1s+l}[\frac{p_2+t}{4t}-\frac{s+l}{2k}] \\ \pi_1(\hat{\mathcal{P}}') &= \sum_{i=1,s_i\neq\tilde{s}_i}^{n} \hat{p}_{1i}[\frac{\hat{p}_2+t}{4t}-\frac{s_i}{2k}] + \sum_{i=1}^{m} \frac{\tilde{l}_i}{k}[\hat{p}_2+t-2t\frac{\tilde{s}_i+\tilde{l}_i}{k}]\end{aligned} \tag{12}$$

We compare the profits of Firm 1 in both cases in order to show that $\hat{\mathcal{P}}'$ induces higher profits:

$$\begin{aligned}\Delta\pi_1 &= \sum_{i=1,s_i\neq\tilde{s}_i}^{n} \hat{p}_{1i}[\frac{\hat{p}_2+t}{4t}-\frac{s_i}{2k}] - \sum_{i=1,s_i\neq\tilde{s}_i}^{n} p_{1i}[\frac{p_2+t}{4t}-\frac{s_i}{2k}] \\ &+ \sum_{i=1}^{m} \frac{\tilde{l}_i}{k}[\hat{p}_2+t-2t\frac{\tilde{s}_i+\tilde{l}_i}{k}] - \sum_{i=1}^{m} p_{1i}[\frac{p_2+t}{4t}-\frac{\tilde{s}_i}{2k}] - p_{1s+l}[\frac{p_2+t}{4t}-\frac{s+l}{2k}] \\ &= \frac{t}{2}\sum_{i=1,s_i\neq\tilde{s}_i}^{n}[\frac{\hat{p}_2+t}{2t}-\frac{s_i}{k}]^2 - \frac{t}{2}\sum_{i=1,s_i\neq\tilde{s}_i}^{n}[\frac{p_2+t}{2t}-\frac{s_i}{k}]^2 \\ &+ \frac{t}{2}\sum_{i=1}^{m}\frac{\tilde{l}_i}{k}[2\frac{\hat{p}_2+t}{t}-4\frac{\tilde{s}_i+\tilde{l}_i}{k}] - \frac{t}{2}\sum_{i=1}^{m}[\frac{p_2+t}{2t}-\frac{\tilde{s}_i}{2k}]^2 - \frac{t}{2}[\frac{p_2+t}{2t}-\frac{s+l}{k}]^2\end{aligned} \tag{13}$$

We consider the terms separately. First,

$$\begin{aligned}&\frac{t}{2}\sum_{i=1,s_i\neq\tilde{s}_i}^{n}[\frac{\hat{p}_2+t}{2t}-\frac{s_i}{k}]^2 - \frac{t}{2}\sum_{i=1,s_i\neq\tilde{s}_i}^{n}[\frac{p_2+t}{2t}-\frac{s_i}{k}]^2 \\ =&\frac{t}{2}\sum_{i=1,s_i\neq\tilde{s}_i}^{n}[[\frac{2}{3(n-m)}[\frac{3}{4t}p_2+\frac{mp_2}{2t}+\frac{1}{4}-\frac{s+l}{2k}]]^2 + [\frac{p_2+t}{2t}-\frac{s_i}{k}][\frac{4}{3(n-m)}[\frac{3}{4t}p_2+\frac{mp_2}{2t}+\frac{1}{4}-\frac{s+l}{2k}]]] \\ \geq&\frac{t}{2}[\frac{p_2+t}{2t}-\frac{s+l}{k}]\frac{4}{3}[\frac{3}{4t}p_2+\frac{mp_2}{2t}+\frac{1}{4}-\frac{s+l}{2k}]\end{aligned} \tag{14}$$

Second, on segments changing from type B to type A when partition changes from $\hat{\mathcal{P}}$ to $\hat{\mathcal{P}}'$:

$$\frac{t}{2}\sum_{i=1}^{m}\frac{\tilde{l}_i}{k}[2\frac{\hat{p}_2+t}{t}-4\frac{\tilde{s}_i+\tilde{l}_i}{k}] - \frac{t}{2}\sum_{i=1}^{m}[\frac{p_2+t}{2t}-\frac{\tilde{s}_i}{2k}]^2 \tag{15}$$

On these m segments, C_1 is violated for the price $\hat{p}_2$ but not for p_2:

$$\frac{\tilde{s}_i+\tilde{l}_i}{k} \geq \frac{p_2+t}{2k}-\frac{\tilde{l}_i}{k} \quad \text{and} \quad \frac{\hat{p}_2+t}{2k}-\frac{\tilde{l}_i}{k} \geq \frac{\tilde{s}_i+\tilde{l}_i}{k}$$

thus:

$$2\frac{\tilde{l}_i}{k} \geq \frac{p_2+t}{2k} - \frac{\tilde{s}_i}{k} \quad \text{and} \quad \frac{\hat{p}_2+t}{2k} - 2\frac{\tilde{l}_i}{k} \geq \frac{\tilde{s}_i}{k}$$

By replacing $\tilde{s}_i$ by its upper bound value and then $\tilde{l}_i$ by its lower bound value we obtain:

$$\frac{t}{2}\sum_{i=1}^{m}\frac{\tilde{l}_i}{k}[2\frac{\hat{p}_2+t}{t} - 4\frac{\tilde{s}_i+\tilde{l}_i}{k}] - \frac{t}{2}\sum_{i=1}^{m}[\frac{p_2+t}{2t} - \frac{\tilde{s}_i}{2k}]^2 \geq 0 \tag{16}$$

Getting back to the profits difference, we obtain:

$$\begin{aligned}\Delta\pi_1 &\geq \frac{t}{2}[\frac{p_2+t}{2t} - \frac{s+l}{k}]\frac{4}{3}[\frac{3}{4t}p_2 + \frac{mp_2}{2t} + \frac{1}{4} - \frac{s+l}{2k}] - \frac{t}{2}[\frac{p_2+t}{2t} - \frac{s+l}{k}]^2 \\ &\geq \frac{t}{2}[\frac{p_2+t}{2t} - \frac{s+l}{k}][\frac{p_2}{2t} + \frac{s+l}{3k} - \frac{1}{6}]\end{aligned} \tag{17}$$

$p_2 = -\frac{t}{3} + \frac{4t}{3(n+1)}\sum_{i=1}^{n+1}[\frac{s_i}{2k} + \frac{l_i}{k}]$ reaches a minimum when the data broker sells the reference partition $\mathcal{P}_{ref}$ to Firm 1, which consists of segments of size $\frac{1}{k}$. Indeed, it is immediate to see that, in a first step, changing from a partition $\hat{\mathcal{P}}'$ with segments of size strictly greater than $\frac{1}{k}$ to a finer partition $\hat{\mathcal{P}}''$ where only s_h and l_h (defined such that $\frac{s_h}{2} + l_h \geq \frac{s_i}{2} + l_i$ for $i = 1, .., n+1$) change: $s''_h = s_h + 1$ and, consequently, $l''_h = l_h - 1$ lowers price p_2. Iterating this procedure until $l_h = 1$, we reiterate it for s_{h-1} and l_{h-1} such that $\frac{s_h}{2} + l_h \geq \frac{s_{h-1}}{2} + l_{h-1} \geq \frac{s_i}{2} + l_i$. After having iterated the first step, we consider a second step, where we add new segments $s_l = min_i\{s_i\} - 1$, and $l_l = 1$. Again it is immediate to see that this second step always lowers p_2. By iteration, the structure converge to the reference partition. We can conclude that p_2 is minimal with the reference partition and $p_2 \geq \frac{t}{2}$. This result extends Liu and Serfes (2004).

This proves that $\Delta\pi_1 \geq 0$, because the first bracket of Equation 17 is positive given C_1, and the second bracket of Equation 17 is positive since $p_2 \geq \frac{t}{2}$.

We have just established that it is always more profitable for the data broker to sell a partition with one segment of type B instead of two smaller segments of type B at the right of the unit line.

Conclusion

These three steps prove that any partition of the line is dominated by an optimal partition composed of two segments, as illustrated in Figure 3. The first segment is composed of j segments of size $\frac{1}{k}$ located at $[0, \frac{j}{k}]$, and the second segment is composed of unidentified consumers, and is located at $[\frac{j}{k}, 1]$.

□

Proof of Theorem 2: the data broker sells symmetrical information to both firms

Part a: optimal structure of information when the data broker sells information to both firms

We prove that the partition described in Theorem 2 is optimal when information is sold to both firms. For each firm, the partition divides the unit line into two segments. The first segment identifies the closest consumers to a firm and is partitioned in j segments of size $\frac{1}{k}$. The second segment is of size $1 - \frac{j}{k}$ and leaves unidentified the other consumers.

Three types of segments are defined as before:

- Segments A, where Firm 1 is in constrained monopoly;
- Segments B, where Firms 1 and 2 compete;
- Segments C, where Firm 1 gets no demand.

In order to prove the optimality of the partition, we use Assumption A.1, to show that the unit line is composed of one segment where firms compete, located at the middle of the line, and segments where firms are monopolists, located close to them. As we will show, the optimal partition under this assumption is similar to the optimal partition when the data broker sells information to one firm.

The profits of the data broker when it sells information to both firms is the difference between firms' profits when they are informed and their outside option: $\Pi_2 = 2(\pi^{I,I}_{\mathcal{P},\theta} - \pi^{NI,I}_{\mathcal{P},\theta})$.

Firm θ buys a partition composed of segments of type A and one segment of type B. To show that a partition in which type A segments are of size $\frac{1}{k}$ is optimal, we prove that $1)$, such a partition maximizes $\pi^{I,I}_{\mathcal{P},\theta}$, and $2)$ such a partition does not change $\pi^{NI,I}_{\mathcal{P},\theta}$.

$1)$: a partition which maximizes $\pi^{I,I}_{\mathcal{P},\theta}$ is necessarily composed of type A segments of size $\frac{1}{k}$.

The proof of this claim is similar to step 1 of Proof 1.a: the price of the competing firm $-\theta$ does not change when Firm θ gets more precise information on type A segments, and since Firm θ can target more precisely consumers with this information, its profits increase.

2): changing from a partition with type A segments of arbitrary size to a partition where type A segments are of size $\frac{1}{k}$ does not change $\pi^{NI,I}_{\mathcal{P},\theta}$.

Assumption A.1 implies that, even when only one firm is informed, the unit line is divided in type A and type B segments. It is immediate to show that the profits of the uninformed firm do not depend on the fineness of type A segments. As a result, $\Pi_2 = 2(\pi^{I,I}_{\mathcal{P},\theta} - \pi^{NI,I}_{\mathcal{P},\theta})$ is maximized when segments of type A have a size of $\frac{1}{k}$.

We can deduce that the optimal partition is composed of two segments, sold to each firm. For Firm 1, the first segment is partitioned in j segments of size $\frac{1}{k}$, and is located at $[0, \frac{j}{k}]$. The second segment is of size $1-\frac{j}{k}$, located at $[\frac{j}{k}, 1]$ and is composed of unidentified consumers. For Firm 2, the first segment is partitioned in j' segments of size $\frac{1}{k}$, and is located at $[1-\frac{j'}{k}, 1]$. The second segment is of size $1-\frac{j'}{k}$, located at $[0, 1-\frac{j'}{k}]$ and is composed of unidentified consumers.

□

Part b: the data broker sells symmetrical information to both firms

We show now that selling symmetrical information is optimal for the data broker, that is, that in equilibrium, $j = j'$.

We compute the prices and profits in equilibrium when both firms are symmetrically informed, with the optimal partition found above.

Firm 1 is a monopolist on the j segments of size $\frac{1}{k}$ in $[0, \frac{j}{k}]$ and Firm 2 has information on $[1-\frac{j'}{k}, 1]$. On $[\frac{j}{k}, 1]$ Firm 1 sets a unique price p_1 and gets demand d_1, similarly on $[0, 1-\frac{j'}{k}]$ Firm 2 sets a unique price p_2 and gets demand d_2.

We write in step 1 the prices and demands, in step 2 we give the profits, and solve for the prices and profits in equilibrium in step 3.

Step 1: expression of the prices and demands.

Firm $\theta = 1, 2$ sets a price $p_{\theta i}$ for each segment of size $\frac{1}{k}$, and a unique price p_θ on the rest of the unit line. The demand for Firm θ on type A segments is $d_{\theta i} = \frac{1}{k}$. The corresponding prices are computed using the indifferent consumer located on the right extremity of the segment, $\frac{i}{k}$. For Firm 1:

$V - t\frac{i}{k} - p_{1i} = V - t(1-\frac{i}{k}) - p_2 \implies \frac{i}{k} = \frac{p_2 - p_{1i} + t}{2t} \implies p_{1i} = p_2 + t - 2t\frac{i}{k}$.

p_2 is the price set by Firm 2 on the left side segments.

Prices set by Firm 2 on the right side segments are found in a similar way and are given by $p_{2i} = p_1 + t - 2t\frac{i}{k}$.

Let denote d_1 the demand for Firm 1 (resp. d_2 the demand for Firm 2) where firms compete. It is found in a similar way as when information is sold to one firm, which gives us $d_1 = \frac{p_2 - p_1 + t}{2t} - \frac{j}{k}$ (resp. $d_2 = 1 - \frac{j'}{k} - \frac{p_2 - p_1 + t}{2t}$).

Step 2: expression of the profits.

The profits of the firms are:
$\pi_1 = \sum_{i=1}^{j} d_{1i}p_{1i} + d_1 p_1 = \sum_{i=1}^{j} \frac{1}{k}(p_2 + t - 2t\frac{i}{k}) + (\frac{p_2 - p_1 + t}{2t} - \frac{j}{k})p_1$
and
$\pi_2 = \sum_{i=1}^{j'} d_{2i}p_{2i} + d_2 p_2 = \sum_{i=1}^{j} \frac{1}{k}(p_1 + t - 2t\frac{i}{k}) + (\frac{p_1 - p_2 + t}{2t} - \frac{j'}{k})p_2$.

Step 3: prices and profits in equilibrium.

We now compute the optimal prices and demands, using first order conditions on π_θ with respect to p_θ. Prices in equilibrium are
$p_1 = t[1 - \frac{2}{3}\frac{j'}{k} - \frac{4}{3}\frac{j}{k}]$
and
$p_2 = t[1 - \frac{2}{3}\frac{j}{k} - \frac{4}{3}\frac{j'}{k}]$
Replacing these values in the above demands and prices gives
$p_{1i} = 2t - \frac{4}{3}\frac{j't}{k} - \frac{2}{3}\frac{jt}{k} - 2\frac{it}{k}$,
$p_{2i} = 2t - \frac{4}{3}\frac{jt}{k} - \frac{2}{3}\frac{j't}{k} - 2\frac{it}{k}$,
$d_1 = \frac{1}{2} - \frac{2}{3}\frac{j}{k} - \frac{1}{3}\frac{j'}{k}$
and
$d_2 = \frac{4}{3}\frac{j'}{k} - \frac{1}{2} - \frac{1}{3}\frac{j}{k}$.
Profits are:

$$\begin{aligned}
\pi_1^* &= \sum_{i=1}^{j} \frac{2t}{k}[1 - \frac{i}{k} - \frac{1}{3}\frac{j}{k} - \frac{2}{3}\frac{j'}{k}] + (\frac{1}{2} - \frac{2}{3}\frac{j}{k} - \frac{1}{3}\frac{j'}{k})t[1 - \frac{2}{3}\frac{j'}{k} - \frac{4}{3}\frac{j}{k}] \\
&= \frac{t}{2} - \frac{7}{9}\frac{j^2t}{k^2} + \frac{2}{9}\frac{j'^2t}{k^2} - \frac{4}{9}\frac{jj't}{k^2} + \frac{2}{3}\frac{jt}{k} - \frac{2}{3}\frac{j't}{k} - \frac{jt}{k^2}
\end{aligned}$$

$$\pi_2^* = \sum_{i=1}^{j'} \frac{2t}{k}[1 - \frac{i}{k} - \frac{1}{3}\frac{j'}{k} - \frac{2}{3}\frac{j}{k}] + (\frac{1}{2} - \frac{2}{3}\frac{j'}{k} - \frac{1}{3}\frac{j}{k})t[1 - \frac{2}{3}\frac{j}{k} - \frac{4}{3}\frac{j'}{k}]$$
$$= \frac{t}{2} - \frac{7}{9}\frac{j'^2 t}{k^2} + \frac{2}{9}\frac{j^2 t}{k^2} - \frac{4}{9}\frac{jj't}{k^2} + \frac{2}{3}\frac{j't}{k} - \frac{2}{3}\frac{jt}{k} - \frac{j't}{k^2}$$

The data broker maximizes the difference

$$\Delta\pi_\theta(j, j') = \pi_{\mathcal{P},\theta}^{I,I} - \pi_{\mathcal{P},\theta}^{NI,I}$$
$$= -\frac{7}{9}\frac{j'^2 t}{k^2} - \frac{4}{9}\frac{jj't}{k^2} + \frac{2}{3}\frac{j't}{k} - \frac{j't}{k^2}$$

At this stage, straightforward FOCs on j and j' confirm that, in equilibrium, $j = j'$. The fact that the solution is a maximum is directly found using the determinant of the Hessian matrix.

□

A.2 Proof of Lemma 1

We compute the prices and profits in equilibrium when information is sold to one firm. By convention, Firm 1, following the optimal structure found in Proof 1.a.

Firm 1 owns a partition of $[0, \frac{j}{k}]$ that includes j segments of size $\frac{1}{k}$, and has no information on consumers on $[\frac{j}{k}, 1]$. Again, firms face three types of segments:

We write in step 1 the prices and demands, in step 2 we give the profits, and solve for the prices and profits in equilibrium in step 3.

Step 1: expression of the prices and demands.

Type A segments are of size $\frac{1}{k}$, and the last one is located at $\frac{j-1}{k}$. Firm 1 sets a price p_{1i} for each segment $i = 1, .., j$ and is in constrained monopoly: $d_{1i} = \frac{1}{k}$. Prices on each segment are determined by the indifferent consumer of each segment located at its right extremity, $\frac{i}{k}$:[24]

$V - t\frac{i}{k} - p_{1i} = V - t(1 - \frac{i}{k}) - p_2 \implies \frac{i}{k} = \frac{p_2 - p_{1i} + t}{2t} \implies p_{1i} = p_2 + t - 2t\frac{i}{k}$.

[24] Assume it is not the case. Then, either p_{1i} is higher and the indifferent consumer is at the left of $\frac{i}{k}$, which is in contradiction with the fact that we deal with type A segments, or p_{1i} is lower and since the demand remain constant, the profits are not maximized.

The rest of the unit line is a type B segment. Firm 1 sets a price p_1 and competes with Firm 2. Firm 2 sets a unique price p_2 for all the consumers on the segment $[0, 1]$. We note d_1 the demand for Firm 1 on this segment. d_1 is found considering the indifferent consumer:

$V - tx - p_1 = V - t(1 - x) - p_2 \implies x = \frac{p_2 - p_1 + t}{2t}$ and $d_1 = x - \frac{j}{k} = \frac{p_2 - p_1 + t}{2t} - \frac{j}{k}$.

Firm 2 sets p_2 and the demand, d_2, is found similarly to d_1, and $d_2 = 1 - \frac{p_2 - p_1 + t}{2t} = \frac{p_1 - p_2 + t}{2t}$.

Step 2: computation of profits.

The profits of both firms can be written as follows:

$\pi_1 = \sum_{i=1}^{j} d_{1i} p_{1i} + d_1 p_1 = \sum_{i=1}^{j} \frac{1}{k}(p_2 + t - 2t\frac{i}{k}) + (\frac{p_2 - p_1 + t}{2t} - \frac{j}{k})p_1$.

$\pi_2 = d_2 p_2 = \frac{p_1 - p_2 + t}{2t} p_2$.

Step 3: prices and profits in equilibrium.

We solve the prices and profits in equilibrium. First order conditions on π_θ with respect to p_θ give us $p_1 = t[1 - \frac{4}{3}\frac{j}{k}]$ and $p_2 = t[1 - \frac{2}{3}\frac{j}{k}]$. By replacing these values in profits and demands we can deduce that: $p_{1i} = 2t[1 - \frac{i}{k} - \frac{1}{3}\frac{j}{k}]$, $d_1 = \frac{1}{2} - \frac{2}{3}\frac{j}{k}$ and $d_2 = \frac{1}{2} - \frac{1}{3}\frac{j}{k}$.

The profits are[25]

$$
\begin{aligned}
\pi_1^* &= \sum_{i=1}^{j} \frac{2t}{k}[1 - \frac{i}{k} - \frac{1}{3}\frac{j}{k}] + \frac{t}{2}(1 - \frac{4}{3}\frac{j}{k})^2 \\
&= \frac{t}{2} + \frac{2jt}{3k} - \frac{7t}{9}\frac{j^2}{k^2} - \frac{tj}{k^2} \\
\pi_2^* &= \frac{t}{2} + \frac{2t}{9}\frac{j^2}{k^2} - \frac{2}{3}\frac{jt}{k}.
\end{aligned}
\tag{18}
$$

□

A.3 Proof of Lemma 2

We compute the prices and profits in equilibrium when both firms are symmetrically informed, with the optimal partition found in Proof 1.b.

Firm 1 is a monopolist on the j segments of size $\frac{1}{k}$ in $[0, \frac{j}{k}]$ and Firm 2 has symmetric information on $[1 - \frac{j}{k}, 1]$. On $[\frac{j}{k}, 1]$ Firm 1 sets a unique price p_1 and gets demand d_1, similarly on $[0, 1 - \frac{j}{k}]$ Firm 2 sets a unique price p_2 and gets demand d_2.

[25] For $p_{1i} \geq 0 \implies \frac{j}{k} \leq \frac{3}{4}$. Profits are equal whatever $\frac{j}{k} \geq \frac{3}{4}$.

We do not go through the computation of prices and demand which are already described in Proof 1.c, and we directly give the prices and profits in equilibrium.

Prices in equilibrium are $p_1 = p_2 = t[1 - 2\frac{j}{k}]$, $p_{\theta i} = 2t[1 - \frac{j}{k} - \frac{i}{k}]$ and $d_\theta = \frac{1}{2} - \frac{j}{k}$.

Profits are:[26]

$$\begin{aligned}\pi_\theta^* &= \sum_{i=1}^{j} \frac{2t}{k}[1 - \frac{i}{k} - \frac{j}{k}] + \frac{1}{2}(1 - 2\frac{j}{k})^2 t \\ &= \frac{t}{2} - \frac{j^2}{k^2}t - \frac{jt}{k^2}.\end{aligned}$$

□

A.4 Proof of Lemma 4

Using the profits from Lemma 3, we determine the optimal size j_1^* of the segments of type A when the data broker only sells information to Firm 1, by maximizing profits with respect to j. When the data broker sells information to both firms, we determine the optimal size j_2^* of type A segments in a similar way.

1) Optimal partition, selling to one firm.

The profits of the data broker when it sells to one firm are:

$$\Pi_1(j) = \frac{2jt}{3k} - \frac{7t}{9}\frac{j^2}{k^2} - \frac{tj}{k^2}.$$

FOC on j leads to the following maximizing value: $j_1^* = \frac{6k-9}{14}$ and:

$$\Pi_1^* = \frac{t}{7} - \frac{3t}{7k} + \frac{9t}{28k^2}.$$

2) Optimal partition, selling to both firms.

We maximize the profit function with respect to j sold to Firm 1 as we assume by symmetry that j' sold to Firm 2 verifies $j' = 1 - j$. The profits of the data broker when both firms are informed are:

[26] For $\frac{j}{k} < \frac{1}{2}$. Profits are equal as soon as $\frac{j}{k} > \frac{1}{2}$.

$$\Pi_2(j) = 2w^2 = 2[\frac{2jt}{3k} - \frac{11j^2t}{9k^2} - \frac{jt}{k^2}].$$

FOC on j leads to $j_2^* = \frac{6k-9}{22}$ and:

$$\Pi_2^* = \frac{2t}{11} - \frac{6t}{11k} + \frac{9t}{22k^2}.$$

□

A.5 Proof of Corollary 2

We generalize the results to first-degree price-discrimination, and show that profits and the optimal structure are equivalent to third-degree price-discrimination.

We prove that the optimal structure when firms first-degree price-discriminate is similar to the structure when firms third-degree price-discriminate. We first characterize the information structure under first-degree price-discrimination, then we determine the optimal partition.

When a firm first-degree price-discriminates, for instance on a segment $[\frac{l_1}{k}, \frac{l_2}{k}]$ with $l_1 \leq l_2$ integers lower than k, two types of segments are defined. On type A' segments, the firm sets a personalized price for each consumer, here $[\frac{l_1}{k}, \frac{l_2}{k}]$. On type B' segments, the firm sets a homogeneous price on each segment, here a price p_1 on $[0, \frac{l_1}{k}]$ and a price p_2 on $[\frac{l_2}{k}, 1]$. If there are n segments of type B', then the firm sets n prices $p_1, .., p_n$, one on each of these segments.

The optimal partition is composed of two segments: on $[0, l]$ ($l \in [0, 1]$) consumers are perfectly identified, and on $[l, 1]$, consumers are unidentified. The proof of this result is not detailed here, as it is similar to the proof of Theorem 1 in Appendix A.1.

It remains to show that on the first segment $[0, l]$, profits under third-degree price-discrimination converge to the profits under first-degree price-discrimination when $k \to \infty$, and to find the optimal size of these segments.

Step 1: Profits under third-degree price-discrimination converge to profits under first-degree price-discrimination

We prove that when one firm is informed (e.g.. Firm 1), profits when $k \to \infty$ correspond to profits under first-degree price-discrimination.

First we write the profits of Firm 1 under first-degree price-discrimination, then we show that when $k \to \infty$ third-degree price-discrimination tends to first-degree price-discrimination. We consider information that divides the unit line into two segments.

Firm 1's profits under first-degree price-discrimination.

We write $\lim_{k\to\infty} \frac{j}{k} = l$. Under first-degree price-discrimination, Firm 1 sets personalized prices on $[0, l]$, and a single price on $[l, 1]$. Firm 2 sets a single price on the unit line: $p_2 = t - \frac{2}{3}l$ (similarly to Lemma 1).

$\pi_1^{FP} = \int_0^l p_1(x) + \frac{t}{2}(1 - \frac{4}{3}l)^2$.

$p_1(x)$ verifies $V - tx - p_1(x) = V - t(1-x) - p_2 \implies p_1(x) = 2t[1 - x - \frac{1}{3}l]$.

We thus have $\pi_1^{FP} = \int_0^l 2t[1 - x - \frac{1}{3}l]\mathrm{d}x + \frac{t}{2}(1 - \frac{4}{3}l)^2$.

We show that third-degree price-discrimination profits converge to first-degree price-discrimination profits.

Consider Equation 18. We want to prove that the sum $\sum_{i=1}^{lk} \frac{2t}{k}[1 - \frac{i}{k} - \frac{1}{3}\frac{j}{k}]$ converges to the profits of first-degree price-discrimination when $k \to \infty$, that is:

$$\lim_{k\to\infty} \sum_{i=1}^{lk} \frac{2t}{k}[1 - \frac{i}{k} - \frac{1}{3}\frac{j}{k}] = \int_0^l 2t[1 - x - \frac{1}{3}l]\mathrm{d}x.$$

Let $f(i) = \frac{2t}{k}[1 - \frac{i}{k} - \frac{1}{3}\frac{j}{k}]$. It is immediate to see that f is decreasing and continuous on $[0, \infty[$, we can thus write: $\int_{i-1}^{i} f(z)\mathrm{d}z \geq f(i) \geq \int_i^{i+1} f(z)\mathrm{d}z$.

Summing each term from 1 to lk we get: $\int_0^{lk} f(z)\mathrm{d}z \geq \sum_{i=1}^{lk} f(i) \geq \int_1^{lk+1} f(z)\mathrm{d}z$.

We have $\int_1^{lk+1} f(z)\mathrm{d}z = \int_0^{lk} f(z)\mathrm{d}z + \int_{lk}^{lk+1} f(z)\mathrm{d}z - \int_0^1 f(z)\mathrm{d}z$.

$$\begin{cases} \lim_{k\to\infty} \int_{lk}^{lk+1} f(z)\mathrm{d}z = \lim_{k\to\infty} \int_{lk}^{lk+1} \frac{2t}{k}[1 - \frac{z}{k} - \frac{1}{3}\frac{j}{k}]\mathrm{d}z = 0. \\ \lim_{k\to\infty} \int_0^1 f(z)\mathrm{d}z = \lim_{k\to\infty} \int_0^1 \frac{2t}{k}[1 - \frac{z}{k} - \frac{1}{3}\frac{j}{k}]\mathrm{d}z = 0. \end{cases} \tag{19}$$

Thus we have: $\lim_{k\to\infty} \int_0^{lk} f(z)\mathrm{d}z \geq \lim_{k\to\infty} \sum_{i=1}^{lk} f(i) \geq \lim_{k\to\infty} \int_0^{j} f(z)\mathrm{d}z$.

By the sandwich theorem we have : $\lim_{k\to\infty} \sum_{i=1}^{lk} f(i) = \lim_{k\to\infty} \int_0^{lk} f(z)\mathrm{d}z = \int_0^l 2t[1 - x - \frac{1}{3}l]\mathrm{d}x$ the last equality is immediate by substitution. The profits under third-degree price-discrimination converge to the profits under first-degree price-discrimination when $k \to \infty$ (thus when quality $\frac{1}{k} \to 0$).

It is straightforward to establish the same result when the data broker sells information to both firms.

Step 2: Optimal size of the segment of identified consumers.

We compute the profits of Firm 1 when the data broker sells to both firms information that allows them to first-degree price-discriminate. We find the following profits: $\pi_1^{FP;I,I} = \int_0^l 2t[1 - x - l]\mathrm{d}x + \frac{t}{2}(1-2l)^2 = \frac{t}{2} - l^2 t$.

The profits of Firm 1 when only Firm 2 is informed are, similarly to the third-degree price-discrimination case: $\pi_1^{FP;NI,I} = \frac{t}{2} + \frac{2t}{9}l^2 - \frac{2t}{3}l$.

The profits of the data broker are then: $\Pi_2 = \frac{2}{3}lt - \frac{11}{9}l^2 t$, maximized with $l^* = \frac{3}{11}$.

□

Bibliography

Alessandro Acquisti, Curtis Taylor, and Liad Wagman. The economics of privacy. *Journal of Economic Literature*, 54(2):442–92, 2016.

Michael R Baye and John Morgan. Information gatekeepers on the internet and the competitiveness of homogeneous product markets. *American Economic Review*, pages 454–474, 2001.

Paul Belleflamme, Wing Man Wynne Lam, and Wouter Vergote. Price discrimination and dispersion under asymmetric profiling of consumers. 2017.

David Bounie, Valérie Morisson, and Martin Quinn. Do you see what I see? ad viewability and the economics of online advertising. *mimeo*, 2017.

Francesco Clavorà Braulin and Tommaso Valletti. Selling customer information to competing firms. *Economics Letters*, 149:10–14, 2016.

Ambarish Chandra and Mara Lederman. Revisiting the relationship between competition and price discrimination. *American Economic Journal: Microeconomics*, 2016.

Lesley Chiou and Catherine Tucker. Search engines and data retention: Implications for privacy and antitrust. 2017.

James C Cooper, Luke Froeb, Daniel P O'Brien, and Steven Tschantz. Does price discrimination intensify competition? implications for antitrust. *Antitrust Law Journal*, 72(2):327–373, 2005.

Kenneth S Corts. Third-degree price discrimination in oligopoly: all-out competition and strategic commitment. *The RAND Journal of Economics*, pages 306–323, 1998.

Matthew Crain. The limits of transparency: Data brokers and commodification. *New Media & Society*, 20(1):88–104, 2018.

David Encaoua and Abraham Hollander. First-degree discrimination by a duopoly: pricing and quality choice. *The BE Journal of Theoretical Economics*, 7(1), 2007.

FTC. Data brokers: A call for transparency and accountability. 2014.

Harold Hotelling. Stability in competition. *The economic journal*, 39(153):41–57, 1929.

Philippe Jehiel and Benny Moldovanu. Auctions with downstream interaction among buyers. *Rand journal of economics*, pages 768–791, 2000.

Nicola Jentzsch, Geza Sapi, and Irina Suleymanova. Targeted pricing and customer data sharing among rivals. *International Journal of Industrial Organization*, 31(2):131–144, 2013.

Qihong Liu and Konstantinos Serfes. Quality of information and oligopolistic price discrimination. *Journal of Economics & Management Strategy*, 13(4):671–702, 2004.

Toshihiro Matsumura and Noriaki Matsushima. Should firms employ personalized pricing? *Journal of Economics & Management Strategy*, 24(4):887–903, 2015.

Andrew McAfee and Erik Brynjolfsonn. Race against the machine. *MIT Sloan Management School of Management*, 2012.

Rodrigo Montes, Wilfried Sand-Zantman, and Tommaso Valletti. The value of personal information in online markets with endogenous privacy. *Management Science*, 2018.

Joseph Pancras and K Sudhir. Optimal marketing strategies for a customer data intermediary. *Journal of Marketing research*, 44(4):560–578, 2007.

Frank Pasquale. The black box society: The secret algorithms that control money and information. *Harvard University Press*, 2015.

Jean-Pierre Ponssard. The strategic role of information on the demand function in an oligopolistic market. *Management Science*, 25(3):243–250, 1979.

Roy Radner et al. *The evaluation of information in organizations*. Management Science Research Group, University of California, 1961.

Lars A Stole. Price discrimination and competition. *Handbook of industrial organization*, 3: 2221–2299, 2007.

Curtis Taylor and Liad Wagman. Consumer privacy in oligopolistic markets: Winners, losers, and welfare. *International Journal of Industrial Organization*, 34:80–84, 2014.

Curtis R Taylor. Consumer privacy and the market for customer information. *RAND Journal of Economics*, pages 631–650, 2004.

Jacques-Francois Thisse and Xavier Vives. On the strategic choice of spatial price policy. *The American Economic Review*, pages 122–137, 1988.

David Ulph and Nir Vulkan. *Electronic commerce and competitive first-degree price discrimination*. University of Bristol, Department of Economics, 2000.

Timothy Van Zandt. Hidden information acquisition and static choice. *Theory and Decision*, 40 (3):235–247, 1996.

Xavier Vives. Duopoly information equilibrium: Cournot and bertrand. *Journal of economic theory*, 34(1):71–94, 1984.

Luc Wathieu. Privacy, exposure and price discrimination. *Division of Research, Harvard Business School*, 2002.

Helen Weeds. Tv wars: Exclusive content and platform competition in pay tv. *The Economic Journal*, 126(594):1600–1633, 2016.