Information Economics and Policy 68 (2024) 101102

Contents lists available at ScienceDirect

Information Economics and Policy

journal homepage: www.elsevier.com/locate/iep

Strategic data sales with partial segment profiling ☆

Flavio Delbono [a], Carlo Reggiani [b,c], Luca Sandrini [d,e,*]

[a] Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy
[b] European Commission JRC, Sevilla, Spain
[c] Department of Economics, University of Manchester, United Kingdom
[d] ZEW - Leibniz Centre of European Economic Research, Mannheim, Germany
[e] QSMS Research Centre, Budapest University of Technology and Economics, Hungary

ARTICLE INFO

JEL classification:
D43
K21
L11
L13
L41
L86
M21
M31

Keywords:
Data markets
Personalised pricing
Price discrimination
Oligopoly
Selling mechanisms

ABSTRACT

We analyse the incentives of a data broker to sell consumer-level data that enable personalised pricing to compete with firms when only a fraction of consumers — centred around one firm that we label "central" — are profiled. The central firm can potentially benefit from the data more than the rival ones ("peripheral"). We show that the data broker may decide not to sell the dataset to the central firm and instead trade with its peripheral competitors. In particular, we identify a strategic reaction of competitors that want to prevent that data increase competition.

1. Introduction

Data gathering, sharing and usage are widespread in today's digital economy. The use of mobile phones and other connected devices has resulted in the continuous generation of massive amounts of data. Many businesses are demanding access to harvest the data and exploit their potential. Data intermediaries such as data brokers and marketing agencies have contributed to the production, collection and sharing of data and are experiencing sustained success. For instance, estimates suggest that the data brokerage sector is expected to grow at an annual rate of 6.8 per cent until 2031 (Transparency Market Research, 2022). These intermediaries are less known to the general public due to the business-to-business nature of their activity. Indeed, the leading companies in the sector, which include Experian, CoreLogic, Epsilon, and Acxiom, collect and exchange data from a wide variety of sources, and whose contracting and business practices are not always transparent (FTC, 2014).

The contribution of data to the economy goes beyond the size of the sector: for example, the total impact of the data market on the EU's economy in 2017 was estimated to be 335.6 billion euros, corresponding to 2.4 per cent of total GDP (Frontier Technology Quarterly, 2019). More generally, data are mostly non-rival and, as such, their use and re-use can generate positive externalities and boost growth (Jones and Tonetti, 2020). Notwithstanding these positive aspects, data transfers can also pose risks for consumers. Besides the well-known individual privacy concerns (Acquisti et al., 2016), data exchanges can also affect market competition. Whereas in the past personalised pricing in com-

☆ We are grateful to the Editor and two anonymous reviewers for the useful comments and constructive criticism. We thank Bruno Carballa-Smichowski, Néstor Duch-Brown, Clara Graziano, Leonardo Madio, Bertin Martens, and Andrew Rhodes for helpful discussion and comments. This article has been presented at the 11th bi-annual Postal Economics Conference on E-commerce, Digital Economy, and Delivery Services, the ASSET 2022 annual conference, the EARIE 2023 annual conference, and the Jornadas de Economia Industrial 2023. We are grateful to all participants. The usual disclaimer applies. The views and opinions expressed in this article are the authors' and do not necessarily reflect those of the European Commission.

* Corresponding author.
E-mail addresses: flavio.delbono@unibo.it (F. Delbono), carlo.reggiani@manchester.ac.uk (C. Reggiani), luca.sandrini@zew.de (L. Sandrini).

https://doi.org/10.1016/j.infoecopol.2024.101102
Received 8 April 2024; Received in revised form 9 August 2024; Accepted 15 August 2024
Available online 22 August 2024

petitive markets has been mostly thought as favourable for consumers by both the academic (Thisse and Vives, 1988) and policy (OECD, 2018) literature, recently circumstances have been identified where this is not necessarily the case and competition may be softened even if firms can engage in personalised pricing (Rhodes and Zhou, 2024).

We contribute to the discussion on the competitive effects of data and personalised pricing by analysing the strategic incentives of a data broker and firms to trade a database that includes consumer information but that only covers a partial segment of the market. In other words, we model a situation in which only consumers with a specific preference profile are in the database, and firms have different valuations for it.[1]

"Partial segment profiling", as we define this situation, is indeed ubiquitous. Data brokers can only collect information about people who leave tracks of their behaviour.

For example, when a new product from a company is launched, experts and influencers review it and post the description online. A fraction of consumers access these reviews, thus revealing their interest in the good. Some other consumers go beyond by liking and sharing posts about the product on social media, and so on. Instead, consumers who do not read or watch the reviews, and do not engage with any other content related to the product, do remain anonymous to the data broker.

Situations of this kind are frequent in a number of markets. Videogames are systematically reviewed before the launch on magazines and streaming platforms such as Twitch and YouTube (Brunt et al., 2020). Books are reviewed on magazines, newspapers, and on online bookshops (Waldfogel and Reimers, 2015; Reimers and Waldfogel, 2021). Also, authors present their work at events that reach a fraction of the market, arguably the most interested part. High-tech products such as smartphones and PCs are tested and reviewed online by digital experts and influencers (Fainmesser et al., 2021; Han et al., 2023). Consumers who are very close in preference to the good object of the reviews reveal their tastes by accessing those reviews.[2]

Partial segmentation can also derive from inferring preferences through potential correlates.[3] Consider the following example in the market for illustrated novels. Suppose there is a book store that sells Japanese manga, a second one that sells Marvel comics, and a third that sells graphic novels, and so on. There is a data broker who previously collected data of consumers of Japanese animated movies (such as demographics, preferred genres, and potentially beyond) and created an audience. The high correlation between this audience and the potential clientele of the first bookstore, specialising in Japanese manga, gives it high incentives to purchase the dataset. However, other book stores could still make use of the database, though less efficiently.

In a context of partial segmentation, we then address the following main research questions. What are the *incentives of a data broker* to sell profiling information about a segment of consumers to one or more of the market competitors? In particular, how is the data broker selling this information, and to whom? Are the data sold exclusively or to more than one of the market competitors? Which of the downstream firms ends up buying the data, and what are the implications for the market outcomes? Do firms have incentives to use the whole data or would they want to further limit its partial reach and, if so, how does the data broker react?

We address the above research questions using a simple model with firms that engage in price competition and one data broker. The consumers are located in a circular city (Vickrey, 1964; Salop, 1979) with an exogenous number of (initially) three firms. The data broker has information on the location of consumers in the arc of the city around one of the firms. As location captures the preferences of consumers, the data can be used to personalise prices and, hence, price discriminate (Thisse and Vives, 1988). Clarity of exposition motivates the choice of presenting the three firms setting with a fixed arc of profiled consumers as the main illustration of our general findings. However, in section 5, we consider arcs of different lengths. Moreover, as we show in section 6, all the main results appear to be robust in the presence of more than three firms in the market.

Our analysis yields three main innovative insights. First, we show that under partial segment profiling, the identity of which firm accesses consumer data plays an important role.

This is because owning the list makes a company more aggressive in pricing, and this is particularly true for the firm whose entire segment is profiled (in what follows, we will refer to this firm as *central*). As a result, social welfare and consumer surplus are relatively high when the firms whose entire segments are not profiled are those that access the list (we will refer to these firms as *peripheral*). On the contrary, aggregate profits are enhanced if one of those firms can access consumer information exclusively. Welfare is instead minimised if it is the central firm that has such exclusive access. Overall, the distribution of surplus is affected by the pricing regime resulting from the access to consumer information.

Second, the selling mechanism influences the outcome of the game, as it alters the ability of the data broker to extract the surplus from downstream companies. In particular, we analyse the strategic incentives of the data broker if she sells information either via an auction with and without a reserve price or via a take-it-or-leave-it offer (TIOLI). Consistently with the literature, we show that auctioning the information maximises the data broker's ability to extract surplus from the bidders.

The intuition for this finding lies in the *strategic reaction* of competing firms to the use of data. The possession of data for consumers close to the firm and the ability to personalise the price offers make rival firms particularly aggressive in pricing. This limits the benefit of obtaining the data for the firm. The strategic price reaction of competing firms is less pronounced when the data are handed to the two competing firms neighbouring the one whose market arc has been profiled by the data broker. This implies that the willingness to bid of the two rivals is higher than the one of the firm for which the data seems tailored.

Third, when considering arcs of different lengths, we show that our main findings are robust for shares of profiled consumers that are both higher and lower than in the baseline model. There are also a number of other mechanisms that come into play when the length of the list changes. For example, we also notice that if the length of the profiled arc is chosen by the data broker, in the presence of constant data gathering costs, it would be optimal to profile a more limited share of the consumers in the arc between a firm and its two neighbours. Indeed, such a reduced list would increase the profits of the buying firms and of the data broker.

Finally, in Section 7 we show that allowing the data broker to sell less data could increase its profit compared to the baseline model. A reduced segment of profiled consumers leads to a softer downstream price competition and increase the firms' willingness to pay. The end result is a further segmented demand, in which "islands of profiled consumers" receiving personalised offers are surrounded by consumers paying uniform posted prices. This configuration is reminiscent of Abrardi et al. (2024a). We relate our setting and results to that article in depth in the literature review below, yet it is important to note the following. In Abrardi et al. (2024a), the profit maximising data broker may create islands of non-profiled consumers between firms. In their setting, islands arise if product differentiation is high, whereas the data sold

[1] This is consistent with the fact that the collection of consumer data is always incomplete. Data brokers analyse interactions on social media, browsing history and behaviour across different websites. This information may not be available, for example, for consumers who deny the use of cookies, only engage with particular contents due to limited attention, do not use social media or specific sites at all, and so on. All the data gathered are then analysed and profiles are grouped into audiences, which include individuals sharing similar characteristics, both personal details or preferences. For a detailed description, see Tucker and Neumann (2020).

[2] We thank an anonymous reviewer for suggesting us these examples.

[3] Miklós-Thal et al. (2023) provide a detailed analysis of the role of correlation between data and the ability of firms to infer consumers' preferences.

would cover all the market otherwise. In our case, instead, the islands result from the lack of willingness of firms to use parts of the list, even if it was given to them for free.

To sum up, the main contributions of this article are, first, to highlight the ability of the data broker to commit to sell information to a number of firms and, second, to show that the data broker may want to sell to firms that do not benefit the most from it (the peripheral ones), but that can preempt another firm (the central one) from gaining a strong competitive advantage.

Moreover, in generating the above results, our analysis also confirms a number of findings from the literature on personalised pricing and price discrimination in oligopolistic markets. In particular, information asymmetry in the market for the final good generates inefficiencies from the welfare perspective. Indeed, the efficient outcome is reached when either all firms access the database or none of them does. Yet, in these two polar cases, either firms (when no firm has the database) or consumers (when all firms have it) receive the majority of this surplus.

Partial segment profiling is a stylised and somewhat extreme modelling feature. Clearly, there are alternative ways to specify the limited profiling data that are as easily justifiable as ours. For example, Conitzer et al. (2012) and Montes et al. (2019), *inter alios*, consider active consumers that can costly avoid profiling, whereas Casadesus-Masanell and Hervas-Drane (2015), Hidir and Vellodi (2021), Xu and Dukes (2022) and Ali et al. (2023) allow consumers to control the amount of information to be revealed. The partial nature of the data coverage that we consider can be thought of as being the result of a marketing study on a particular segment of the market or, alternatively, as data gathered on the previous or potential clientele of one of the firms competing in the downstream market. Our results, then, apply when the partially profiled segment has preferences for a specific firm.

Related literature. Recent years have witnessed a growing interest in the economic impact of data in markets and, in particular, data sharing and trading.[4] There is a wide literature on privacy and its market implications (Acquisti and Varian, 2005; Liu and Serfes, 2006; Choe et al., 2017; Choi et al., 2019; Ichihashi, 2020; Clavorà Braulin, 2023; Anderson et al., 2023; Laub et al., 2023, *inter alios*). As discussed above, a number of these articles have considered incomplete profiling as a direct result of consumers' actions.[5] In our model, instead, the data broker can only achieve partial segment profiling due to the (exogenous) costs or limitations they directly face as, for example, data access only through engagement with a product or its reviews or previously collected information for other purposes, or even a stronger local or market-specific privacy rules and regulations.

Other articles have considered the impact of personalised pricing when firms have asymmetric access to consumer information. Gu et al. (2019) study the effect of exclusive information that enables personalised offers on the incentives to act as price leader in the market. Belleflamme et al. (2020) focus on asymmetric precision on the profitability of price discrimination. They find that as long as the two firms are not identically able to profile consumers, they can both charge prices above the marginal cost. We also model personalised pricing, but the asymmetric access to the information is endogenous. Further, the information is only about consumers that have an innate preference for a specific firm.

Choe et al. (2024) analyse the incentives of a *data rich* company to strategically share a portion of the database to soften competition and increase surplus extraction. They use a Hotelling model and show that the data-rich firm is willing to share information about a segment of the market sufficiently close, but not too much, to the rival. By doing so, the data-rich company isolates the group of consumers that are closest to the rival and induces it to raise the uniform price in order to extract more surplus from them. Consequently, the data rich firm enjoys less competitive pressure in the other segments of the market where it holds a competitive advantage given by data. We differentiate from this article mainly because we investigate the incentives of an external data broker to trade data access with competing firms, hence shutting down the strategic motives of the data broker in the market for the final good.

This article also contributes to the literature on data brokers' incentives. Bergemann et al. (2022), Ichihashi (2021), Gu et al. (2022), and Abrardi et al. (2024b) analyse upstream competition (or lack of) between data brokers that combine data to sell to downstream competing companies. Montes et al. (2019) model privacy concerned consumers and find that a data broker always has the incentive to sell data exclusively to a competing duopoly firm. Bounie et al. (2021) study a spatial duopoly and characterise the optimal partition of a consumer database. Through partitioning, the data broker offers non-overlapping information to both firms, leaving a uniform price segment in the middle. The former segment allows firms to enhance their profits, whereas the latter is characterised by fiercer competition. Given this trade-off, the data broker eventually sells only one partition to one firm exclusively. In our case, the presence of the third firm implies that the uniform price segment is not necessarily extremely competitive. This feature makes it profitable to sell partial information to more than one firm. We also focus on the sale of an exogenously given partial segment of consumers and, unlike them, we do not study the optimal partition.

A closely related article is Abrardi et al. (2024a), who consider endogenous entry in a circular model and show that a data broker can adjust the size of the database sold to each firm to soften competition and to regulate entry level. Our framework differs from Abrardi et al. (2024a) in two main directions. First, we consider a database whose size is exogenously determined, and it is centred around one of the firms. Abrardi et al. (2024a) let the data broker adjust at will the audience generated by data analytics around all the firms. Second, we do not consider entry, and we focus on the short-term effects of firms acquiring and using the database.

In this light, our approach complements the analysis in Abrardi et al. (2024a) as it shows an unidentified incentive of firms that have to bid for accessing the database, that is, the strategic reaction of the firms that want to prevent a better-positioned rival to fully exploit the competitive advantage granted by exclusive access to data. This intuition is partially in line with more general settings featuring auctions with negative externalities as, for example, Jehiel et al. (1996) and Jehiel and Moldovanu (2000). The complementary perspective is also reflected in the welfare implications of our respective findings. In their setting, endogenising entry leads the data sale to increase welfare and decrease consumer surplus as entry is reduced. In our model, the equilibrium non-exclusive sale of data implies a welfare loss driven by increased mismatch costs; at the same time, it implies a consumer gain linked to more competitive pricing compared to a situation with no data at all.

Finally, Kim et al. (2019) and Martens et al. (2021) also study data sharing in a Salop model with three firms: the former in the context of data-driven mergers, whereas the latter focuses on platforms. Martens et al. (2021) assume that only the platform knows the locations of the firms and, as a result, may bias consumer recommendations. Instead, in Kim et al. (2019), like in our article, the relevant information is the location of consumers. In their article, all consumers in the market are profiled, and in a pre-merger equilibrium, the data are sold exclusively. Instead, we focus on situations in which the information held by the data broker only covers a particular segment of the market. The main implication is that exclusive selling of non-divisible information is never the optimal strategy.

[4] For a detailed survey of the literature, see Goldfarb and Tucker (2019), Bergemann and Bonatti (2019), and Pino (2022).

[5] For instance, recently Ali et al. (2023) analyse the effects of personalised pricing powered by access to consumers' data on welfare and consumer surplus. They investigate the topic assuming that partial profiling derives from the consumer's incentives to use privacy control strategically — i.e., to reveal only a subset of information to a chosen firm.

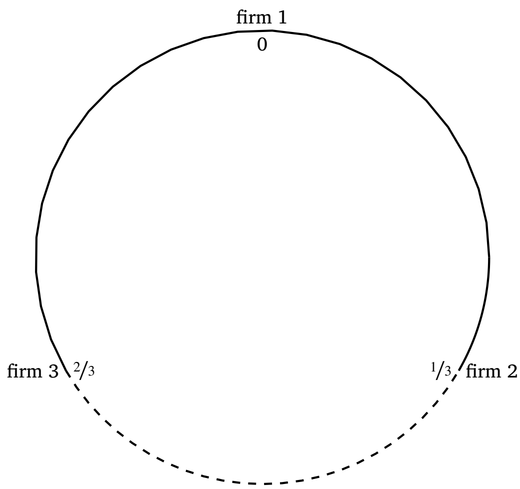


Fig. 1. The Salop model with three firms. The dashed line represents the anonymous segment and the full line the profiled one.

Structure. The rest of the article is structured as follows. Section 2 illustrates the model and its main assumptions before providing some preliminary results on price competition under different allocations of the partial information to downstream firms. Section 3 studies the data broker's sale of the dataset. The next sections extend our baseline results in several directions. Section 4 discusses other possible ways of selling the data and the optimal selling mechanism. Section 5 allows for different sizes of the profiled arc of consumers. Section 6 evaluates the impact of a higher number of firms. Section 7 shows how the data broker's ability to sell less data can increase its profit compared to the baseline model. Finally, Section 8 discusses the results and their implications. Unless otherwise stated, the proofs are in Appendix A.

2. The model and preliminary results

The market. Consider a market with one data broker (she) and three competing retailers $i = 1, 2, 3$. Consumers are uniformly distributed on the unit circle (Vickrey, 1964; Salop, 1979) and their location is denoted as x. Firms are located equidistantly at $y_1 = 0$, $y_2 = 1/3$, and $y_3 = 2/3$. Consumers can demand at most one unit of the good. The utility of a consumer x for the good of firm i is:

$$U(x, y_i) = v - t|x - y_i| - p_i, \tag{1}$$

where v is the good's valuation, t is the unit transport cost, and p_i is the price. We assume that v is sufficiently high for the market to be fully covered and $v > t$ is sufficient (but not necessary) for that to be the case. For simplicity, there are no variable or fixed costs.

Consumer information and data selling. The data broker possesses information only on some consumers located in the segment between firm $i - 1$ and firm $i + 1$. Without loss of generality, we assume that the data broker has information about consumers located between firm 3 and firm 2. In particular, to begin with, the data broker has information on consumers *on the arc around firm 1*, i.e., $x \in [2/3, 1]$ and $x \in [0, 1/3]$, respectively. For the sake of clarity, we will refer to this arc as the *profiled segment* of the market. Instead, we will refer to the non-profiled arc between firm 2 and firm 3 - i.e., $x \in [1/3, 2/3]$ - as to the *anonymous segment* (Fig. 1). The valuable information in this model is the location of consumers x. The data broker sells the data by auction. In Section 3, we show that an auction is a more profitable strategy for the data broker than making a *take-it-or-leave-it-offer* to a subset of the firms in the market (Montes et al., 2019; Bounie et al., 2021).

Timing. At Stage 0, the data broker costlessly gathers information about the segment between firm 2 and firm 3, i.e., the segments $x \in [2/3, 1]$ and $x \in [0, 1/3]$. At Stage 1, consumer information in the data broker's possession is sold. At Stage 2, firms engage in price competition.[6] As we look for the Subgame Perfect Nash Equilibrium, the game is solved by backward induction.

2.1. Price competition

We relegate the solution of all subgames to the appendix A.1. There are several possible subgames to be considered at Stage 2. We start from two benchmark cases (Section A.1.1): (i) no firm has access to consumer information, (ii) all firms have access to consumer information. As it is standard in spatial competition models, the key determinant of the demand segments is the consumers' indifference between pairs of neighbouring firms. These two subgames are characterised by symmetry: in all cases, the equilibrium indifferent consumers are located exactly halfway between each couple of firms, i.e., $x_{12} = 1/6$, $x_{23} = 1/2$, and $x_{31} = 5/6$, where the subscripts denote the two firms the consumers at that location are indifferent between. If there is no information available, the consumers are indifferent between the posted prices of firms, whereas when all firms have access to the information x_{12} and x_{31} are indifferent between these firms' personalised offers.

We then consider the case in which one firm has exclusive access to the list (Section A.1.2): this firm can be firm 1, whose market segment has been profiled, or one between firm 2 and firm 3. The equilibrium demand configurations for these cases are represented in Fig. A.1.

If firm 1 has access to consumers' information, then it can use it to make personalised offers that expand its demand, i.e., $\tilde{x}_{12} = 5/18 > 1/6$. At the same time, the rivals decrease prices to retain at least part of the consumers in the segment where they compete with firm 1. Indeed, as a result of firm 1 having exclusive access to consumer information, firm 2 and firm 3 become more aggressive in pricing. The equilibrium prices reflect the trade-off between the standard uniform price competition on the anonymous segment and the need to match firm 1's personalised prices on the profiled segment. Firm 1 makes more profit than the competitors thanks to the exclusive information.

In case it is either firm 2 or firm 3 that has access to the list, then the demand configuration is highly asymmetric, and the indifferent consumer is much closer to firm 1 as a result of the personalised offers of, say, firm 2. In that case, $\tilde{x}_{12} = 7/156 < 1/6$. Also in this case firm 1 responds by pricing more aggressively than in the no information case, which leads to lower prices by firms 3 and 2 as a result.

More in details, firm 1 suffers from the competition of firm 2's personalised prices on its own arc and, as a result, decreases its price, which results to be the lowest in equilibrium. This affects firm 3, which posts a higher price, but lower than firm 2 in response. The pricing rankings reflect those of profits: firm 2 benefits the most from exclusive information about firm 1's arc of consumers. Firm 1, in turn, is the most damaged by firm 2 having information about its own market segment. In summary, we note that the competitive effect of a rival holding information about the profiled segment affects firm 1. The shock then propagates to firm 3 and, finally, bites back firm 2 through its own uniform price.

Finally, we consider the cases in which two firms get the consumer information. In one case, the two firms include firm 1, whereas in the other, firm 2 and firm 3 gain access to the data (Section A.1.3). The equilibrium demand configurations for these cases are represented in Fig. A.2.

In case the central firm and a rival have the information (say, firm 1 and firm 2), the third firm with no information (say, firm 3) is the most damaged. The uninformed, in fact, faces fierce competition from

[6] In cases when one firm holds information about consumers on a specific arc, a well-known problem is the existence of a pure strategy Bertrand-Nash equilibrium (see Rhodes and Zhou, 2024, p.25). In order to ensure equilibrium existence, we assume that personalised price schedules are set only after uniform prices are set.

Table 1
Summary of the prices and profits in each subgame of the pricing stage.

	No info (NI)	All info (AI)	Excl 1 (1)	Excl 2 (2)	Both 1 & 2 (12)	Both 2 & 3 (23)
p_1	0.333 t	-	-	0.244 t	-	0.167 t
p_2	0.333 t	0.333 t	0.222 t	0.321 t	0.286 t	0.333 t
p_3	0.333 t	0.333 t	0.222 t	0.308 t	0.238 t	0.333 t
π_1	0.111 t	0.056 t	0.154 t	0.059 t	0.109 t	0.028 t
π_2	0.111 t	0.083 t	0.049 t	0.135 t	0.069 t	0.118 t
π_3	0.111 t	0.083 t	0.049 t	0.095 t	0.057 t	0.118 t
Π	0.333 t	0.222 t	0.253 t	0.289 t	0.235 t	0.264 t
CS	v - 0.417 t	v - 0.306 t	v - 0.361 t	v - 0.388 t	v - 0.333 t	v - 0.361 t
TS	v - 0.084 t	v - 0.084 t	v - 0.108 t	v - 0.099 t	v - 0.098 t	v - 0.097 t

the personalised offers of firm 1 and, as a result, its price is lower than the one of firm 2. Firm 3 also gets the lowest profit, whereas firm 1 benefits from personalised pricing and has the highest profit. As in the case in which only firm 2 has access to the data, the competitive effect of personalised prices hits firm 3 more directly and then propagates to firm 2. However, in this case, the portion of consumers served through personalised offers is actually is larger, which decreases the profitability of firms 3 and firm 2 even further.

If firm 2 and firm 3 have access to the list, the equilibrium prices for non profiled consumers are the same as in the benchmark. The competition between firm 2 and firm 3 for the anonymous segment is not affected by the information. The profiled segment, in fact, is served by both firms through personalised offers. Firm 1 suffers the consequences of this information allocation, as it has to decrease its price to compete with personalised pricing on its own market arc. The lower price of firm 1 is also reflected in much lower profit than the two informed competitors.

2.2. Prices, profits, and welfare

We start with a recap of the results of the pricing stage. Table 1 reports the equilibrium posted prices and firms' and industry profits in all the pricing subgames. Each subgame's label is used as a superscript in the ensuing comparisons and analysis. The table highlights one interesting feature of the presence of personalised pricing on posted prices: no matter what subgame is reached, posted prices are never higher than in the no information benchmark ($t/3$). This underlines the pro-competitive effect of personalised prices, which induces rivals to be more competitive and best respond with lower posted prices.

Proposition 1 provides a comparison of the firm's profits in each of the possible pricing subgames. It is important to recall that if one of the peripheral firms gets the information exclusively, this is firm 2 and not firm 3. Table 1 leads to the following Proposition:

Proposition 1. *The equilibrium profits of each firm in the pricing subgames compare as follows:*

$$\pi_1^1 > \pi_1^{NI} > \pi_1^{12} = \pi_1^{13} > \pi_1^2 = \pi_1^3 > \pi_1^{AI} > \pi_1^{23},$$
$$\pi_2^2 > \pi_2^{23} > \pi_2^{NI} > \pi_2^3 > \pi_2^{AI} > \pi_2^{12} > \pi_2^{13} > \pi_2^1,$$
$$\pi_3^3 > \pi_3^{23} > \pi_3^{NI} > \pi_3^2 > \pi_3^{AI} > \pi_3^{13} > \pi_3^{12} > \pi_3^1.$$

The proposition makes clear that firm 1, whose segment of nearby consumers is profiled, benefits from exclusive use of the list despite the consequent increase in competition intensity. Interestingly, its second best would be that no information is shared or sold. This outcome would be better than sharing the data with firm 2, as it drives all firms to set the highest possible price, whereas sharing the list would entail stronger competitive pressure that is detrimental to profits. In detail, by sharing data with firm 2 rather than having them alone, firm 1 would not be able to fully exploit the potential of the list when competing against firm 2. Ultimately, this negative effect more than compensates for the relatively softer competitive pressure exerted by firm 3.

Similarly, firm 2 greatly benefits from having exclusive access to consumers' information. Intuitively, exclusive access to data means that firm 2 can price discriminate one segment of the market. Firm 1's best reply is to lower her price and be more aggressive against both firm 2's price schedule and firm 3's price. However, price competition does not propagate as if firm 1 had the data since firm 3 faces competition on just one sub-segment of her market.

Finally, it is interesting to notice that the profit of firm 2 when all firms buy the data is higher than its profit when it buys it jointly with firm 1, i.e., $\pi_2^{AI} > \pi_2^{12}$. At the same time, the profit of firm 3 is higher when firms 1 and 2 have both access to the list than when firm 1 has it exclusively, i.e., $\pi_3^{12} > \pi_3^1$.

Welfare analysis. The previous analysis has important implications. From the industry perspective, no information maximises the joint profits, whereas the most competitive subgame is when all firms have access to the list of profiled consumers. When the central firm has access to the information, either exclusively or jointly, the industry profits decrease compared to the case when the rivals do. Exclusive information (for example, to firm 2 or firm 1) leads to higher industry profits than if the same firms share the information with one of the rivals.

As expected, the consumer surplus displays an almost perfectly inverse order. The best scenario is when all firms have access to the list, whereas no information is the less desirable subgame. This result is in line with Parker et al. (2020), who call for a regulatory intervention that facilitates data sharing mechanisms to benefit consumers. In our setting, this can be explained as a consequence of the intense price competition when all firms have access to the information. Interestingly, from a consumer's perspective we note that the exclusive availability of the information to firm 1 is equivalent to the case in which both firm 2 and 3 access it. Indeed, the different allocation of the information does not affect the intensity of competition in each sub-segment of the market.

Finally, the total surplus is maximised in the two benchmark cases of no information and when all firms have access to it. The only difference is that in the former case, the allocation is biased towards the firms, whereas in the latter, it is towards consumers. Moreover, the subgame in which the information is held by the central firm (firm 1) is the least desirable from a welfare perspective. As there are no demand expansion effects and prices constitute transfers between consumers and firms, all the total surplus results are driven by the overall transport costs and the symmetry of the location of the indifferent consumers. To summarise:

Proposition 2. *The industry profits in the pricing subgames compare as follows:*

$$\Pi^{NI} > \Pi^2 = \Pi^3 > \Pi^{23} > \Pi^1 > \Pi^{12} = \Pi^{13} > \Pi^{AI}.$$

As for consumer surplus:

$$CS^{AI} > CS^{12} = CS^{13} > CS^{23} = CS^1 > CS^2 = CS^3 > CS^{NI},$$

and total surplus:

$$TS^{NI} = TS^{AI} > TS^{23} > TS^{12} = TS^{13} > TS^{2} = TS^{3} > TS^{1}.$$

3. The data broker's incentives

We finally focus on the data broker decision. There are several mechanisms that the data broker can employ to sell the data. In particular, building upon Jehiel and Moldovanu (2000), we assume that the data broker sells the list to one or more firms via a system of auctions with a reserve price. In other words, the data broker sets the minimum bid the firms must match in order to win the auction.

We design the auction as follows. First, the data broker chooses how many contracts to sell. If the data broker commits to sell $k \in \{1,2,3\}$ contracts, then data will be purchased by a coalition of exactly k members - or no one if the reserve price is too high. Given the number of contracts available, the data broker sets the reserve price for the auction - i.e., the minimum bid that the coalition of bidders must pay in order to win the auction. Finally, firms place their bids. We define each firm's willingness to pay as the difference in the firms' payoffs if they obtain the list and the counterfactual case in which a rival company purchases the data in their place. The only constraint from the data broker perspective is that it cannot violate the commitment on the number of contracts.[7]

Let us consider the easiest case where the data broker commits to sell $k = 1$ contract. Each firm's willingness to pay is given by the difference between the payoff from being the *informed* company and the payoff from being one of the non-informed. Because the payoffs vary depending on which firm possesses the data, we consider the worst-case scenario of each firm as the exit option. In our example, firm 1 is willing to pay $\pi_1^1 - \pi_1^j$ with $j = 2,3$, whereas both firms 2 and 3 are willing to pay $\pi_j^j - \pi_j^1$.

We now turn to the case in which the data broker commits to sell $k = 2$ contracts. There exist three feasible coalitions: $\{1,2\}$, $\{1,3\}$, and $\{2,3\}$. Each coalition states its willingness to pay, which is defined as the sum of the willingness to pay of each member. Consistently, we measure members' willingness to pay as the difference between the payoff from staying in the winning coalition and the payoff from being the firm outside of the coalition. In our model, for example, the willingness to pay of coalition $\{1,2\}$ can be written as $(\pi_1^{12} - \pi_1^{23}) + (\pi_2^{12} - \pi_2^{13})$. Similarly, the willingness to pay of coalition $\{1,3\}$ can be written as $(\pi_1^{13} - \pi_1^{23}) + (\pi_3^{13} - \pi_3^{12})$. Instead, coalition $\{2,3\}$ is willing to pay $(\pi_2^{23} - \pi_2^{13}) + (\pi_3^{23} - \pi_3^{12})$.

Finally, the case with $k = 3$ contracts. In this scenario, all firms buy the list, or none does, as the data broker cannot lower the number of contracts declared ex-ante. Hence, the coalition is willing to pay $(\pi_1^{AI} - \pi_1^{NI}) + (\pi_2^{AI} - \pi_2^{NI}) + (\pi_3^{AI} - \pi_3^{NI})$.

Using the payoffs in Table 1, the auctions lead to the following data broker profits[8]:

if $k = 1$ then $\quad \pi_{DB}^1 = 0.095t,\ \pi_{DB}^2 = \pi_{DB}^3 = 0.085t,$

if $k = 2$ then $\quad \pi_{DB}^{12} = \pi_{DB}^{13} = 0.094t,\ \pi_{DB}^{23} = 0.123t,$

if $k = 3$ then $\quad \pi_{DB}^{AI} = 0$

Straightforward comparisons yield the following:

Proposition 3. *Assume the selling mechanism is an auction with a reserve price. Then, the profits of the data broker in the pricing subgames compare as follows:*

$$\pi_{DB}^{23} > \pi_{DB}^{1} > \pi_{DB}^{12} = \pi_{DB}^{13} > \pi_{DB}^{2} = \pi_{DB}^{3} > \pi_{DB}^{AI}.$$

Proposition 3 shows how the strategic reaction of competing firms to the use of data tailored to firm 1's clientele is sufficiently strong to induce the data broker to trade data with them. The possession of information about consumers close to the firm, as well as the ability to personalise the price offers, induce rival firms to be particularly aggressive in pricing. This limits the firms' benefit of obtaining the data. The strategic price reaction of competing firms is less pronounced when the data are handed to the two competing firms neighbouring the one whose market arc has been profiled by the data broker. This implies that the willingness to bid of the two rivals is jointly higher than the one of the firm for which the data seems tailored.

Discussion of the result. There are three main economic forces driving the main result in Proposition 3. First, intuitively, the firms that can use the data less efficiently (firms 2 and 3 in our illustrative model) can still use the list to price discriminate consumers. Their cumulative incentives to use the data oppose firm 1's willingness to exclusively exploit the list. Second, both firms 2 and 3 react to firm 1's usage of the list by lowering their price and reducing both their extensive and intensive margins in the profiled segments of the market. Hence, to avoid this loss, they have positive incentives to purchase the data, leaving the main rival uninformed. Again, this opposes firm 1's incentives to buy the list to limit price competition in the profiled segment. Finally, and importantly, the effects of enhanced price competition propagate in the anonymous segment of the market, limiting the intensive margin of firms 2 and 3. Thus, they are both willing to pay for the list to prevent this additional negative effect on their payoffs. This third economic driver is unique to firms 2 and 3. In fact, firm 1 has no interest in what happens in the anonymous segment.

The sum of these three economic forces drives the results in Proposition 3. Notice that the number of firms in the winning coalition plays a role only in the first of the three drivers described above. In particular, because the data broker wants to maximise the value of the list, it wants to sell it to a coalition that can use it entirely. This has more to do with which firms are in the coalition rather than how many. In fact, the data broker has higher incentives to sell the list to firm 1 in exclusivity than to the coalition composed of firms 1 and 2 (or 3).

4. Alternative selling mechanisms

An alternative way to model the auction with a reserve price, in the spirit of Bounie et al. (2020) and Abrardi et al. (2024c), is that the data broker commits to a maximum number of contracts (not the exact number of contracts). By doing so, the data broker is able to extract even more surplus from the downstream firms. To understand it, consider firm 1's problem. The best thing she can do is to buy the list exclusively. The data broker maximises surplus extraction by committing to sell two contracts at most and threatening firm 1 to offer the list to the coalition composed of firms 2 and 3. Similarly, the data broker can offer the two contracts to the coalition of firms 2 and 3, threatening them to offer the list only to firm 1. The main result does not change as $\pi_{DB}^{23} = (\pi_2^{23} - \pi_2^1) + (\pi_3^{23} - \pi_3^1) = 0.137t > 0.127t = \pi_1^1 - \pi_1^{23} = \pi_{DB}^1$.

Assuming that the data broker can set a reserve price for the bids ensures that the auction extracts the highest surplus from the downstream firms. However, the assumption is not crucial for our results to hold. In fact, if we consider a more standard second-price auction, the incentives of the data broker remain unaltered. The only difference is that the coalition composed by firms 2 and 3 will pay a lower price for the information, namely the willingness to pay of the coalition with the

[7] We relax this assumption later and show that results are qualitatively similar.

[8] Notice that, in this section, we implicitly assume that the data broker can extract all the surplus from the firms — i.e., she can set a reserve price equal to the highest willingness to pay. However, for this to be an equilibrium, it must be that the targeted coalition would not reject the offer in equilibrium. In practice, this means that the data broker can only extract up to the second-highest willingness to pay, as a larger reserve price would not be consistent with the individual rationality constraint of the seller. Results remain qualitatively unchanged. We thank an anonymous referee for pointing out this issue.

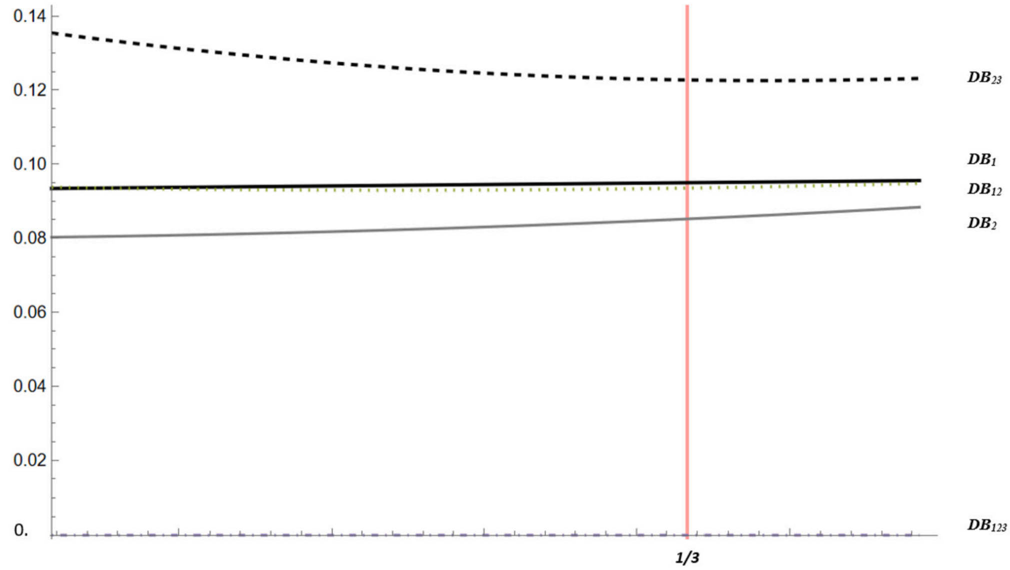


Fig. 2. Willingness to pay and length of the list.

second-highest reservation value -i.e., the coalition composed by firm 1 only.

Finally, a different mechanism that the data broker could use to sell data is through a TIOLI offer. In this case, we define the willingness to pay of the firms as the difference in the profits if they buy the list and the counterfactual case in which they do not. This leads to the following data broker profits:

$$\pi^1_{DB} = \pi^1_1 - \pi^{NI}_1 = 0.043t,$$
$$\pi^2_{DB} = \pi^2_2 - \pi^{NI}_2 = 0.023t,$$
$$\pi^{12}_{DB} = (\pi^{12}_1 - \pi^2_1) + (\pi^{12}_2 - \pi^1_2) = 0.069t,$$
$$\pi^{23}_{DB} = 2(\pi^{23}_2 - \pi^2_3) = 0.047t$$
$$\pi^{AI}_{DB} = (\pi^{AI}_1 - \pi^{23}_1) + 2(\pi^{AI}_3 - \pi^{12}_3) = 0.081t$$

To conclude, assume the selling mechanism is a TIOLI offer; then, the profits of the data broker in the pricing subgames compare as follows:

$$\pi^{AI}_{DB} > \pi^{12}_{DB} > \pi^{23}_{DB} > \pi^1_{DB} > \pi^2_{DB}.$$

Proposition 3 and, more generally, our findings in this section on data selling provide interesting insights. A data broker that has profiled one arc of consumers around a firm never sells the information about consumers *exclusively* to the central firm.

If the data broker adopts an auction or a sequential bargaining selling method, the optimal choice is to sell consumer information not to the central firm (firm 1) but to the two peripheral rivals together, firms 2 and 3. The only scenario in which firm 1 obtains the list is when the data broker chooses a TIOLI offer and sells the data to all firms in the market. This is also the only scenario in which private incentives are aligned with the social optimum (see Section 2.2).

Differently from the case in which the data broker organises auctions, with a TIOLI offer, she is not able to extract all the willingness to pay from the firms. In fact, firms do not internalise the danger that a rival gets the information in their place, which in turn does not trigger any strategic reaction.

The optimal selling mechanism. From the previous analysis, one may naturally question which selling mechanism the data broker prefers in order to maximise profits. Unsurprisingly, the data broker earns larger profits from auctioning the data rather than selling the list using a TIOLI offer. This follows naturally from the fact that auctioning the list allows the data broker to efficiently extract the downstream competing firms' willingness to pay.

Furthermore, focusing on the auction mechanisms presented above, the larger profits are obtained by not committing to a certain number of contracts. In fact, by doing so, the data broker can exert additional pressure on the competing firms by threatening them to sell the list to the firm or coalition of firms that maximises competition intensity in the downstream market.

5. The coverage of the list

A simplifying assumption of the illustrative model is that all consumers located in the arc between firm 3 and firm 2 and centred around firm 1 are profiled. Here, we relax this assumption and extend the model to consider a symmetric arc, of length A on each side of firm 1. In light of the many subcases and the lengthy and often repetitive derivations, the details of the setup and the formal analysis can be found in Web Appendix B. Fig. 2 provides a graphical summary of the main findings by plotting how the willingness to pay of the firms or groups of firms changes with the width of the arc of profiled consumers. As it is apparent from the figure, we focus on a relatively limited interval of the length list around $A = 1/3$, i.e., lists characterised by length $A \in [7/24, 23/66]$.

Intuitively, the analysis suggests that, as long as the arc of profiled consumers is sufficiently large to allow firms 2 and 3 to directly benefit from it, the data broker never sells the list exclusively to one firm. In line with this intuition, Fig. 2 shows that selling to firm 2 and firm 3 is more profitable for the data broker than selling exclusively to firm 1.

Some notes are in order here. We choose this subset of parameters in order to assess the robustness of our main findings and, at the same time, limit the number of cases that have to be considered. Indeed, below the $A = 7/24$ and above the $A = 23/66$ thresholds, some of the assumptions on the location of the indifferent consumers stop to hold, and the analysis needs to consider a number of corner solutions. Besides avoiding the technical complications, our choice seems reasonable in the light of the qualitatively similar nature of the insights. Indeed, the results are qualitatively robust, for example, for lengths of the list lower than $7/24$, but still not too short. In fact, it is only if $A < 1/6$, not represented in the figure, that firm 1 has an advantage in attracting the profiled and close-by consumers. As a result, it has a higher willingness to pay for the list.

The main finding is that as long as the length of the list is sufficient, selling the information to the central firm is never part of the data broker profit maximising strategy, as in our baseline model. Indeed, the list is sold to both firm 2 and firm 3, and this finding confirms that the results in Proposition 3 hold for both shorter and longer lengths of the list.[9]

6. A higher number of firms

A market with three firms is a special case, and we would like to gain insights into cases in which more firms, potentially $n (> 3)$, are active in the market. Regrettably, a full generalization of our illustrative model to n firms proves to be complex. This is due to asymmetric shock on the prices of different firms, which are then asymmetrically transmitted to all other firms.[10] In Web Appendix C, we set up the problem of all the firms and identify the first-order conditions in all subgames of the firm pricing stage. As in our baseline, the possession of the list allows us to identify all consumers between the firms neighbouring firm 1, in this case, firm 2 and firm n.

The model, however, can be fully solved numerically for any number of firms.[11] We can then provide results and confirm the validity of our previous insights for a given number of competing firms: $n = 4$, $n = 5$, and $n = 10$. The monotonicity of these numerical results suggests that similar findings would be obtained for other market sizes.

Some notes about the pricing subgames are in order. First, it is important to establish that *no firm other than 1, 2 and n* has an incentive to buy the list. As all the other firms are located far from the consumers that the data broker has profiled, they cannot extract profits through personalised pricing and, as such, their willingness to pay for the list is zero in all pricing subgames.

Second, all segments apart from those with profiled consumers are characterised by first-order conditions typical of competition *à la* Salop. In particular, in subgames where the list affects *symmetrically* the firms on both sides of firm 1 but not their posted prices (i.e., when firms 2 and n have the list or all three firms have it), all competitors except firm 1 pick Salop prices in equilibrium.

Further, in case firm 1 holds the list exclusively, the price impact of it propagates symmetrically through both firm 2 and firm n and then through to the other competitors. The more challenging cases, instead, are the ones where the list affects *asymmetrically* firms that do not own it: this is the case for subgames where firm 2 has the list exclusively or when both firms 1 and firm 2 acquire it.

What are then the implications for the data broker? As in our illustrative model with three firms, we find that the DB has no incentives to sell the data exclusively. Moreover, our numerical results confirm that, in equilibrium, information is sold symmetrically to the two firms located at the extremes of the list, jointly. In other words, Propositions 3 and the other results in Section 4 hold regardless of the number of competitors in the market. Indeed, all variables change smoothly and monotonically with the number of firms: Table 2 provides a summary of the firms' prices and profits and the data brokers' profits in case of sale through an auction.[12]

Finally, on the basis of Table 2, we can note that the market structure induced by the data broker is such that firms 2 and n have the list. This implies that all $n - 3$ firms located away from the arc of profiled consumers are *unaffected* by the price competition in the profiled segment of the market. As noticed above, these firms play the usual Salop strategy, i.e. their price is t/n. Indeed, the firm suffering from the enhanced price competition induced by the list is firm 1, which ends up charging its customers $t/2n$.

7. Data broker's ability to sell less data

A further question that emerges at this point is whether the downstream firms have the incentives to use all the *potentially valuable* data that they receive from the data broker.[13][14] Furthermore, the data broker may recognise these incentives of the downstream firms, and whether it wants to sell the entire database in its possession (including all consumers in the arc $[0, 1/3]$ and $[2/3, 1]$) is not obvious.

More in detail, suppose that one of the firms has entered in possession of the entire list ($[0, 1/3]$ and $[2/3, 1]$). We show in Web Appendix D that there are cases in which such an informed firm is better off by not price discriminating the consumers close to its location. This is so because such a strategy would make the firm more aggressive in the unprofiled market segment and, hence, decrease its profit. In the Web Appendix, we show that including the consumers with the lowest transportation costs into the set of those who are charged the uniform price would lead firm 2 to increase its profit in case the informed firms are both firm 1 and firm 2. Similarly, firm 2 and firm 3 would benefit from such a strategy if they both had access to the list and firm 1 is uninformed.

The problem with this potential lack of use of the data by the downstream firms is that they cannot credibly commit to do so. Indeed, in our context, any firm has a unilateral *ex-post* incentive to make personalised offers. As a result, in this section we extend the set of actions of the data broker, allowing it to further reduce the scope of the list. In other words, by recognising the incentives of the downstream firms, the data broker can sell a subset of the profiles in the arc $[0, 1/3]$ and $[2/3, 1]$, hence, making it impossible for the firms to use them, which acts as a credible commitment and leads to an enhanced profitability.

The analysis in Web Appendix D shows that profits are enhanced when the data broker has the ability to sell less data in three of the eight possible pricing subgames, i.e., when firm 1 and firm 2 (equivalently, firm 1 and firm 3) and when firm 2 and firm 3 are informed. Fig. D.1 illustrates the further reduced lists sold in the equilibrium of these two subgames.

If firm 1 and firm 2 are informed, a partition like the one depicted in panel (a), covering the arc $[0, x_{12}]$ and $[2/3, 1]$, turns out to be more profitable. In particular, all firms increase their profits, so that the industry total is $0.238t$ (Table D.1), compared to $0.235t$ in the baseline (Table 1). If firm 2 and firm 3 are informed, a partition like the one in panel (b), covering the arcs $[0, x_{12}]$, $[\tilde{x}_{32}, 1/3]$, $[2/3, \tilde{x}_{23}]$ and $[x_{31}, 1]$ is sold by the data broker.[15] In this case, firm 2 and firm 3 increase their profits, and the industry total grows to $0.286t$ (Table D.1), compared to only $0.264t$ in the baseline (Table 1).

Finally, the data broker can then increase its profit through auctioning such a reduced list. In particular, if selling to firm 1 and firm 2, it could obtain $0.093t$, whereas by auctioning to firm 2 and firm 3 $0.140t$. Then, selling to the two peripheral firms remains the data broker's optimal strategy, but it leads to a higher revenue ($0.140t$ compared to $0.123t$).

[9] Notice that the results would hold for lists of longer lengths. However, because some conditions on the location of indifferent consumers in some off-equilibrium scenarios cease to hold for $A > 23/66$, we decided to restrict the focus of this extension to $A \in [7/24, 23/66]$.

[10] We note that the problem has similarities with the case of asymmetric cost shocks in the Salop model, addressed by Syverson (2004) and solved by Alderighi and Piga (2012) under fixed locations and complete information. Yet, there are further types of asymmetry that make our setting even more complicated.

[11] In the case of asymmetric subgames, the procedure can be quite tedious as n grows large.

[12] The results for a TIOLI selling mechanism are equally robust.

[13] We are grateful to an anonymous reviewer for proposing this line of inquiry.

[14] By potentially valuable, we mean data that the informed firms use in the equilibria of our baseline model and, instead, we disregard data that the informed firms could not potentially use — i.e., data of consumers sufficiently far away from the informed firms' locations, such that they cannot possibly attract them.

[15] Note that the case where firm 2 and firm 3 are sold a reduced list is characterised by the appearance of "islands", i.e., segments of demand served through personalised pricing, in between consumers buying through posted prices. This is reminiscent of Abrardi et al. (2024a), whose firms also can extend offers to consumers located on profiled "islands" in between them.

Table 2
The number of firms, prices and profits.

No info	p_1	p_2	p_n	π_1	π_2	π_n	π_{DB}^{Auct}
$n = 3$	0.333t	0.333t	0.333t	0.111t	0.111t	0.111t	0
$n = 4$	0.250t	0.250t	0.250t	0.063t	0.063t	0.063t	0
$n = 5$	0.200t	0.200t	0.200t	0.040t	0.040t	0.040t	0
$n = 10$	0.100t	0.100t	0.100t	0.010t	0.010t	0.010t	0
1, 2 and n	p_1	p_2	p_n	π_1	π_2	π_n	π_{DB}^{Auct}
$n = 3$	-	0.333t	0.333t	0.056t	0.083t	0.083t	0
$n = 4$	-	0.250t	0.250t	0.031t	0.469t	0.469t	0
$n = 5$	-	0.200t	0.200t	0.020t	0.030t	0.030t	0
$n = 10$	-	0.100t	0.100t	0.005t	0.008t	0.008t	0
Excl 1	p_1	p_2	p_n	π_1	π_2	π_n	π_{DB}^{Auct}
$n = 3$	-	0.222t	0.222t	0.154t	0.049t	0.049t	0.095t
$n = 4$	-	0.179t	0.179t	0.092t	0.032t	0.032t	0.058t
$n = 5$	-	0.145t	0.145t	0.060t	0.021t	0.021t	0.038t
$n = 10$	-	0.073t	0.073t	0.015t	0.005t	0.005t	0.0096t
Excl 2 (n)	p_1	p_2	p_n	π_1	π_2	π_n	π_{DB}^{Auct}
$n = 3$	0.244t	0.321t	0.308t	0.059t	0.135t	0.095t	0.086t
$n = 4$	0.183t	0.247t	0.232t	0.033t	0.077t	0.054t	0.046t
$n = 5$	0.146t	0.199t	0.186t	0.021t	0.050t	0.034t	0.029t
$n = 10$	0.073t	0.100t	0.093t	0.005t	0.013t	0.009t	0.0071t
Both 1 & 2 (n)	p_1	p_2	p_n	π_1	π_2	π_n	π_{DB}^{Auct}
$n = 3$	-	0.286t	0.238t	0.109t	0.069t	0.057t	0.094t
$n = 4$	-	0.240t	0.183t	0.062t	0.045t	0.034t	0.058t
$n = 5$	-	0.197t	0.146t	0.040t	0.030t	0.021t	0.038t
$n = 10$	-	0.099t	0.073t	0.010t	0.007t	0.005t	0.0096t
Both 2 & n	p_1	p_2	p_n	π_1	π_2	π_n	π_{DB}^{Auct}
$n = 3$	0.167t	0.333t	0.333t	0.028t	0.118t	0.118t	0.123t
$n = 4$	0.125t	0.250t	0.250t	0.016t	0.066t	0.066t	0.066t
$n = 5$	0.100t	0.200t	0.200t	0.010t	0.043t	0.043t	0.042t
$n = 10$	0.050t	0.100t	0.100t	0.003t	0.011t	0.011t	0.0105t

To conclude, a word of caution. We note, in fact, that the opportunity for the data broker to enact this strategy leads the informed firms and, as a result, the data broker, to achieve higher profits. However, it is not a fully-fledged analysis of the optimal partition of a consumers' list as, for example, in Bounie et al. (2021). Furthermore, the results in this section bear some resemblance with the literature on sleeping patents, which hints at the possibility that a firm might deem excluding a rival from getting innovation is more valuable than using the innovation itself (Gilbert and Newbery, 1982; Vickers, 1985). In our context, the data broker further reduces the list to give the possibility to one or more downstream firms not to use the information and be more profitable. The upstream intervention is, hence, crucial in our setting.

8. Discussion

Not always detailed information is available about all consumers in the market, and often the potential clientele of one firm is better profiled than others. This article has studied the strategic incentives of a data broker to sell this type of partial segment profiling information to competing firms. Indeed, we considered a database of consumers that includes only the potential customers of one of the competing firms. The data can be used to implement personalised pricing.

We analyse these incentives in an oligopoly model of spatial competition where three firms compete in prices à la Salop. In this setting, we find that each of the three firms in the market would benefit from the exclusive use of the data. Interestingly, however, the second-best outcome for the central firm would be that no information is shared or sold. In particular, we find that if access to the information creates an asymmetry between the competitors, the defensive response of firms without the list can be particularly aggressive, with a negative impact on profitability.

The model also allows us to analyse the welfare effects of the data broker's choices. At the aggregate level, firms would be better off when no information is shared, as this is the scenario in which price competition is as soft as possible. Conversely, when all three firms have access to consumer data, price competition in the profiled segment is very fierce. As all firms can set a competitive personalised price for the consumers in the profiled segment, surplus extraction is minimal.

Consumer surplus orderings reflect the above results. When all firms have access to data, the intensity of competition in the profiled segment makes consumers better off on average. Although non-profiled consumers face the usual Salop price, the gains for those on the profiled sub-arc are so high that they outweigh any other scenarios.

Interestingly, from a welfare perspective, there is no difference between all firms having access to data or no firm, as they both represent the first best. However, the two cases are not equivalent in terms of the distribution of social surplus. In particular, the former favours consumers, whereas the latter would be preferred by firms. In other words, a policy-maker considering to mandate data-sharing or not faces a choice between which side of the market to back. This is a typical feature of spatial models à la Salop (1979) - Thisse and Vives (1988), as the one we employ.

The most important findings, however, regard the sale of data. We show that in equilibrium the data broker has the incentives to sell the database to firms other than the one which is most suitable to use it. The equilibrium is also characterised by islands of demand served through personalised offers, alternating with segments where consumers pay uniform prices. Furthermore, the data broker's strategy crucially depends on his ability to extract the surplus from the downstream firms, hence from the selling mechanism adopted. In case data are auctioned, it sells the consumer list to both the firms located at the extremes of the profiled segment. Instead, in the case of a TIOLI offer, the list is sold to all firms in the market. Contrary to previous findings in the literature (Montes et al., 2019; Kim et al., 2019, inter alios), an implication of our findings is that the data broker *never* has the incentives to sell the data exclusively.

Moreover, the data broker's incentives are not aligned with social welfare. In fact, as an auction provides the data broker with higher profits than a TIOLI offer, we shall identify this scenario as the expected equilibrium of the game. Thus, in such equilibrium, the second best is realised, and the central firm needs to price aggressively in order to best respond to the personalised price of the two neighbouring rivals. On the other hand, consumers on the anonymous sub-arc are not affected by the list, and, as a result, they do not suffer from possible brand mismatches.

Our framework is clearly stylised, and it aims to address situations where there is a sharp disparity in the incomplete profiling of consumers: one segment of potential consumers of a firm is accurately profiled, whereas other segments are scantly or not profiled at all. The contribution, then, is to show that in the presence of this type of information structure and more than two firms, the usual incentives to sell data exclusively do not apply. This is due to the softened competition between the peripheral firms and the less aggressive consequent price response of the central firm.

Notwithstanding the recalled limitations, our findings can provide relevant managerial and policy implications. Consider, for example, a data broker that is in possession of data on the potential clientele of a firm. A somewhat counter-intuitive implication of our findings is that unless the portion of consumers profiled is rather limited, the data broker should not approach such a firm first. Instead, the maximum willingness to pay could be extracted by auctioning off the data. In that case, as long as the data cover a sufficient part of the arc around the firm, the closest rivals would acquire the list.

The central firm faces a profitability threat and needs to adopt defensive strategies. For example, if the data are not already available, one option could be to make it harder for a data broker to profile the consumers. One of these strategies could entail making privacy salient on their website to enhance their consumer's attention in releasing data. Another more costly option could be to vertically integrate with the data broker in order to internalise the impacts of data selling. Finally, a regulation mandating data sharing would be in the interest of such a firm.

A relevant policy issue is whether the access to data from upstream firms can advantage some competing firms that have access to it. Martens and Mueller-Langer (2020) point out how sharing real-time digital car data between manufacturers and a network of official dealers can lead to price discrimination and potential foreclosure of independent downstream competitors. In this example, a vehicle transmits data to the manufacturer which, in turn, may have exclusive deals with authorised repair garages. These consumers represent part of the potential clientele of the repair garages, and they are likely to have a relative preference for those linked to the manufacturers through exclusivity or other agreements.

Our work, then, contributes to the understanding of how accessing such information can impact competition between repair garages. In particular, we note that the data broker is not spontaneously willing to sell the information to all the firms in the market, nor exclusively to the company that is best suited to use the data. Thus, the welfare-maximising policy-maker should consider either a ban on data collection and sale (if the goal is to favour aggregate profits over consumer surplus) or a mandatory data sharing regulation, which would not only achieve the maximum exploitation of data but also induce pro-competitive market outcomes.

We shall note, however, that the second policy option changes firms' incentives substantially, and the stakeholders should be aware of it. Whereas selling data to all three firms can be an optimal strategy under certain circumstances, the data broker's *bargaining power* collapses to the minimum if she *must* sell consumer information to all the firms.

Such a regulatory intervention would be welcomed downstream but is likely to face hostile reactions from data holders. Thus, a policy-maker that aims to support this policy might design a tax on data usage to competing firms to redistribute the revenues with the data broker, particularly if the latter has to recover from the costs of collecting data and needs to be incentivised to do so.

We already noted some of the limitations of our circular city spatial framework à la Salop (1979) - Thisse and Vives (1988). These types of models are characterised by localised competition: despite all firms potentially competing with each other, in equilibrium, consumers tend to focus on the firms located most closely to them. Moreover, the demand is inelastic, and all consumers buy if they have sufficiently high valuations for the good. An interesting direction for further research would be to assess how our results would change under non-localised competition and elastic demand (Perloff and Salop, 1985; Chen and Riordan, 2007).

Funding and competing interests

The authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. One of the authors acknowledges the financial support by the National Research, Development and Innovation Office (NKFIH) under grant number OTKA PD-138543.

CRediT authorship contribution statement

Flavio Delbono: Writing – review & editing, Methodology, Conceptualization. **Carlo Reggiani:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Luca Sandrini:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data availability

No data was used for the research described in the article.

Appendix A

A.1. Results on price competition

A.1.1. Benchmark cases

No firm has access to consumer information. Consider the model à la Salop introduced in Section 2. If no firm has access to consumer information, each firm simultaneously sets its prices to maximise profits. In other words, there is price competition *à la* Salop with three firms. For given prices, each firm's demand depends on the location of the consumers who are indifferent between buying from the firm or one of its two neighbours, i.e.:

$$U(x, y_i) = U(x, y_{i-1}) \text{ and } U(x, y_i) = U(x, y_{i+1}),$$

where the utility functions are defined as in equation (1). As a result, the profit function of, for example, firm 1 is:

$\pi_1 = p_1 \left[\left(\frac{t}{3} + \frac{p_2 - p_1}{2t} \right) + \left(\frac{t}{3} + \frac{p_3 - p_1}{2t} \right) \right].$

Standard profit maximisation leads to the following result (proof omitted):

Lemma 1. ***(Salop, 1979)*** *The unique equilibrium in a pricing subgame in which no firm has access to consumers information is characterised by the following prices and profits:*

$$p_i = \frac{t}{3}, \ \ \pi_i = \frac{t}{9}, \ \ i = 1, 2, 3.$$

All firms have access to consumer information. If all firms have access to the information on consumers in the profiled segment of the market, firms will use the information to condition price offers to the consumer's location and price discriminate. In other words, firms can send personalised offers to consumers at each location x on the arc.

This implies that firms are competing fiercely at each location x: as the distance of each firm is the only source of differentiation, Bertrand competition with heterogeneous costs (due to the distance) takes place at each location. Firms charge a non-negative price, as otherwise, they would make a loss and decrease their profit. Hence, the closest firm can attract the consumers by charging a non-negative price that exactly matches the offer of the second closest firm (Thisse and Vives, 1988; Taylor and Wagman, 2014). For example, considering the sub-arc between firm 1 and firm 2, firm 1 can attract all consumers located between $x = 0$ and $x = 1/6$. On that sub-arc, firm 2 cannot offer any price lower than $p_2(x) = 0$. The price schedule for firm 1 can be found by solving for $p_1(x)$ the following:

$$U(x, y_1) = v - tx - p_1(x) = v - t(1/3 - x) = U(x, y_2),$$

leading to: $p_1(x) = t/3 - 2tx$. On the sub-arc between $x = 1/6$ and $x = 1/3$, a similar argument establishes that $p_1(x) = 0$ as the non-negativity constraint binds for firm 1.

Following a similar reasoning, the firms' price schedules on the arc $x \in [2/3, 1]$ and $x \in [0, 1/3]$ are as follows:

$$p_1(x) = \begin{cases} t\,(1/3 - 2x), & \text{if } 0 \le x < 1/6 \\ t\,(2x - 5/3), & \text{if } 5/6 \le x < 1 \\ 0, & \text{otherwise} \end{cases}; \tag{A.1}$$

$$p_2(x) = \begin{cases} t\,(2x - 1/3), & \text{if } \frac{1}{6} \le x < \frac{1}{3} \\ 0, & \text{otherwise} \end{cases}; \tag{A.2}$$

$$p_3(x) = \begin{cases} t\,(5/3 - 2x), & \text{if } 2/3 \le x < 5/6 \\ 0, & \text{otherwise} \end{cases}. \tag{A.3}$$

Despite the access to the data is symmetric, the price schedules (A.1)-(A.2)-(A.3) are clearly different and firms face an asymmetric situation. In particular, firm 1 price discriminates consumers both on its left and its right, whereas firms 2 and 3 can apply personalised schedules only on one side. This feature will play a notable role in the following analysis. The remaining consumers on the anonymous segment, i.e., between firm 2 and firm 3, are offered a uniform price. The indifferent consumer is identified by solving $U(x, y_2) = U(x, y_3)$. Solving the profit-maximisation problem leads to:

Lemma 2. *If all firms have access to consumer information, the equilibrium consists of the price schedules (A.1)-(A.2)-(A.3) and the prices:*

$$p_2 = p_3 = \frac{t}{3}.$$

The firms' profits are, respectively,

$$\pi_1 = \frac{t}{18}, \ \ \pi_2 = \pi_3 = \frac{t}{12}.$$

Proof. As a result of the pricing derived in (A.1)-(A.2)-(A.3), the profits on the profiled segment of the market are:

$$\pi_1^d = \int_0^{1/6} p_1(x)dx + \int_{5/6}^{1} p_1(x)dx = \frac{t}{18}$$

$$\pi_2^d = \int_{1/6}^{1/3} p_2(x)dx = \frac{t}{36} .$$

$$\pi_3^d = \int_{2/3}^{5/6} p_3(x)dx = \frac{t}{36}$$

The remaining consumers on the anonymous segment, i.e., between firm 2 and firm 3, are offered a uniform price. The indifferent consumer is identified by solving $U(x, y_2) = U(x, y_3)$. The firms' profit functions are:

$$\pi_2 = p_2 \left(\frac{1}{6} + \frac{p_3 - p_2}{2t} \right) + \frac{t}{36}, \qquad \pi_3 = p_3 \left(\frac{1}{6} + \frac{p_2 - p_3}{2t} \right) + \frac{t}{36}.$$

Standard calculations lead to the profit-maximising anonymous prices

$$p_2 = p_3 = \frac{t}{3}.$$

We note that there is no positive price that allows firm 1 to attract unprofiled consumers away from firms 2 and 3. Using the price schedules (A.1)-(A.2)-(A.3), and the prices p_2 and p_3, the profits of the firms can be written as

$$\pi_1 = \frac{t}{18}, \ \ \pi_2 = \pi_3 = \frac{t}{12}. \quad \square$$

Proposition 2 illustrates the asymmetric profit impact of the possession of consumer information. Indeed, all firms compete more fiercely for the profiled segment and, as a result, they make less profit than in the no information benchmark (Proposition 1). However, firm 1 is more damaged than firms 2 and 3, as its potential customers are profiled on both sides. The rivals' customers are only profiled on one of their two market segments. The uniform prices paid by the non-profiled consumers in the anonymous segment are relatively high; in fact, they are the same as in the no-information benchmark. The average price paid by profiled consumers is lower than the benchmark, and firm 1 suffers twice from this intensified competition.

A.1.2. Exclusive access to consumer information

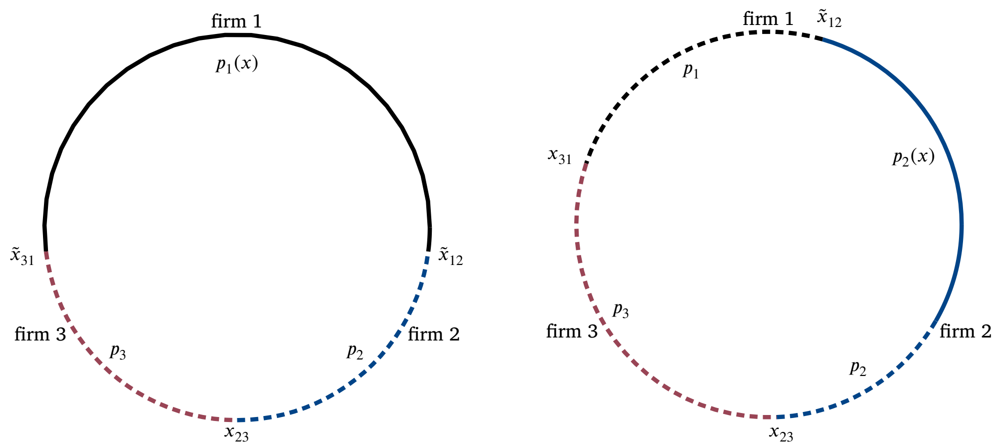
We now consider the subgames in which one firm has access to the information on firm 1's arc of consumers. The firm holding exclusive information can be firm 1 or one between firm 2 and firm 3. We analyse these two possibilities in what follows.

Firm 1 has access to consumer information. If firm 1 has exclusive access to the list, it will use it to personalise offers to the consumers on the profiled segment. Firm 2 and firm 3, instead, can only set uniform prices, p_2 and p_3, for all consumers.

Fig. A.1(a) anticipates the equilibrium demand configuration. Given those prices, firm 1 price schedule is:

$$p_1(x) = \begin{cases} \max\left\{p_2 + t\,(1/3 - 2x), 0\right\}, & \text{if } 0 \le x < 1/3 \\ \max\left\{p_3 + t\,(2x - 5/3), 0\right\}, & \text{if } 2/3 \le x < 1 \end{cases}. \tag{A.4}$$

Denote the consumers for which the price schedule of firm 1 is zero, i.e., $p_1(\tilde{x}_{12}) = p_1(\tilde{x}_{13}) = 0$, as $\tilde{x}_{12} = 1/6 + p_2/2t$ and $\tilde{x}_{13} = 5/6 - p_3/2t$. Assume these consumers lie on the profiled segment, which we verify is the case in equilibrium. Then, the following proposition summarises our main findings in the pricing subgame if firm 1 has exclusive access to information about consumers on its own arc.



(a) Firm 1 has access to consumer information (b) Firm 2 has access to consumer information.

Fig. A.1. Equilibrium demand segments if consumer information is held by two firms. The solid lines represent segments served through list offers, the dashed lines by uniform prices. Firm 1 segments in black, firm 2 in blue, firm 3 in purple. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

Lemma 3. *If firm 1 has exclusive access to consumer information, the equilibrium consists of the price schedules (A.4), with prices and marginal consumers given by:*

$$p_2 = p_3 = \frac{2}{9}t,\ \tilde{x}_{12} = \frac{5}{18},\ \tilde{x}_{13} = \frac{13}{18}.$$

The firms' profits are, respectively,

$$\pi_1 = \frac{25}{162}t,\ \ \pi_2 = \pi_3 = \frac{4}{81}t.$$

Proof. From the price schedule (A.4), the profit function of the firms are:

$$\pi_1 = \int_0^{\tilde{x}_{12}} \left[p_2 + t\,(1/3 - 2x)\right] dx + \int_{\tilde{x}_{13}}^{1} \left[p_3 + t\,(2x - 5/6)\right] dx.$$

$$\pi_2 = p_2\left[\left(\frac{1}{6} + \frac{p_3 - p_2}{2t}\right) + \left(\frac{1}{3} - \tilde{x}_{12}\right)\right].$$

$$\pi_3 = p_3\left[\left(\frac{1}{6} + \frac{p_2 - p_3}{2t}\right) + \left(\tilde{x}_{13} - \frac{2}{3}\right)\right].$$

Standard calculations lead to the profit-maximising prices for the anonymous segment

$$p_2 = p_3 = \frac{2}{9}t.$$

Using these prices and the price schedule (A.4), it is possible to derive the profits of the firms:

$$\pi_1 = \frac{25}{162}t,\ \ \pi_2 = \pi_3 = \frac{4}{81}t.$$

Firm 2 (or firm 3) have no unilateral incentive to increase discretely the price to $p_2^{dev} = t/3$, as this would give profits $\pi_2^{dev}(t/3, 2t/9) = t/27 < 4t/81$. □

Firm 2 or firm 3 have access to consumer information. Consider the case of either firm 2 or firm 3 having exclusive access to information about consumers on the arc around the rival (firm 1), i.e., the profiled segment. Assume that firm 2 has access to the consumers' information without loss of generality. In this case, firm 2 sets a price schedule for the profiled consumers ($x \in [0, 1/3]$) and a price p_2 for non-profiled consumers on the anonymous segment.[16] Firm 1 and firm 3 set uniform prices p_1 and p_3. Given these prices, firm 2 personalised price schedule is:

$$p_2(x) = \max\left\{p_1 + t\,(2x - 1/3), 0\right\}. \tag{A.5}$$

Denote the consumers for which the price schedule of firm 2 is zero, i.e., $p_2(\tilde{x}_{21}) = 0$, as $\tilde{x}_{21} = 1/6 - p_1/2t$, and assume that these consumers lie on the profiled segment, i.e., $\tilde{x}_{21} \in [0, 1/3]$. The equilibrium in the pricing subgame, if firm 2 has exclusive information on consumers on firm 1's arc, can be characterised as follows.

Lemma 4. *If firm 2 has exclusive access to consumer information, the equilibrium consists of the price schedule (A.5) and the prices*

$$p_1 = \frac{19}{78}t,\ \ p_2 = \frac{25}{78}t,\ \ p_3 = \frac{4}{13}t.$$

The marginal consumer is located at $\tilde{x}_{21} = 7/156$. The firms' profits are, respectively,

$$\pi_1 = \frac{361}{6084}t,\ \ \pi_2 = \frac{3275}{24336}t,\ \ \pi_3 = \frac{16}{169}t.$$

Proof. From the price schedule (A.5), the profit function of the firms can be written as:

$$\pi_1 = p_1\left[\tilde{x}_{21} + \left(\frac{1}{6} + \frac{p_3 - p_1}{2t}\right)\right].$$

$$\pi_2 = p_2\left(\frac{1}{6} + \frac{p_3 - p_2}{2t}\right) + \int_{\tilde{x}_{21}}^{1/3} \left[p_1 + t\,(2x - 1/3)\right] dx.$$

$$\pi_3 = p_3\left[\left(\frac{1}{6} + \frac{p_2 - p_3}{2t}\right) + \left(\frac{1}{6} + \frac{p_1 - p_3}{2t}\right)\right].$$

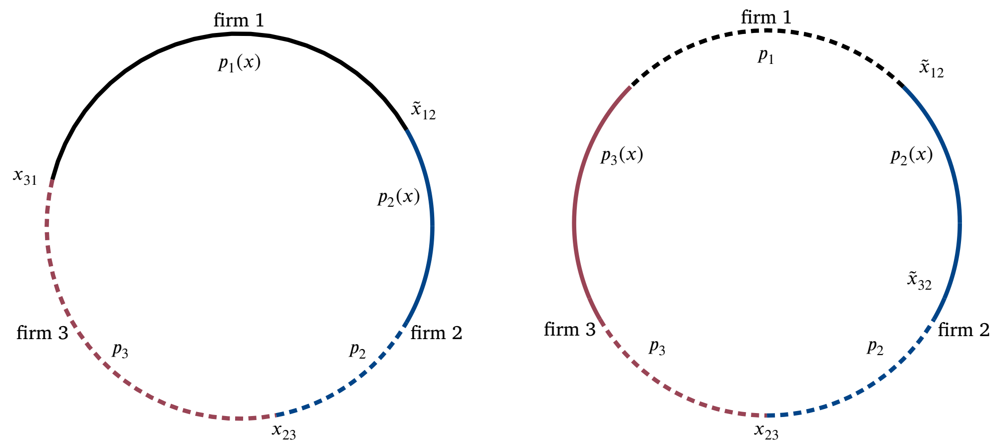
Standard calculations lead to the profit-maximising prices for the anonymous segment

$$p_1 = \frac{19}{78}t,\ \ p_2 = \frac{25}{78}t,\ \ p_3 = \frac{4}{13}t.$$

Using these prices and the price schedule (A.4), it is possible to derive the profits of the firms:

$$\pi_1 = \frac{361}{6084}t,\ \ \pi_2 = \frac{3275}{24336}t,\ \ \pi_3 = \frac{16}{169}t.$$

[16] Note that in principle firm 2 could also reach consumers in the arc $x \in [2/3, 1]$, but this never happens in equilibrium.



(a) Firm 1 and firm 2 have consumer information (b) Firm 2 and firm 3 have consumer information.

Fig. A.2. Equilibrium demand segments if consumer information is held by two firms. The full lines represent segments served through list offers, the dashed lines by uniform prices. Firm 1 segments in black, firm 2 in blue, firm 3 in purple.

Firm 1 has no unilateral incentive to increase discretely the price to $p_1^{dev} = t/3$, as this would give profits $\pi_1^{dev}(t/3, 4t/13) = 2t/39 < 361t/6084$. □

A.1.3. Two firms access consumer information

The final subgames to consider are when a subset of more than one firm but not all have access to the information on firm 1's arc of consumers. The subset can include firm 1 or not, and we will analyse these two cases in turn in what follows.

Firm 1 and 2 have access to consumer information. If firm 1 and firm 2 have access to the information, they can offer personalised prices to consumers on the profiled segment — i.e., firm 1 can send personalised prices to consumers in $x \in [2/3, 1]$ and $x \in [0, 1/3]$, whereas firm 2 can personalise its price schedule to attract consumers in $x \in [0, 1/3]$. Consequently, there will be intense competition between firm 1 and firm 2 for the profiled consumers lying on the sub-arc between them. Clearly, on that segment, neither firm can offer a price lower than its cost (zero in this case) or it would make losses, i.e., $p_i(x) \geq 0$, $\forall x[0, 1/3], i = 1, 2$. This allows to identify the price schedule and the indifferent consumer on that arc.

Fig. A.2(a) anticipates the equilibrium demand configuration. Given the price of firm 3 and the previous observations, the price schedules of firm 1 and firm 2 are, respectively:

$$p_1(x) = \begin{cases} \max\{t(1/3 - 2x), 0\} & \text{if } x \in [0, 1/3] \\ \max\{p_3 + t(2x - 5/3), 0\} & \text{if } x \in [2/3, 1] \end{cases}, \tag{A.6}$$

$$p_2(x) = \max\{t(2x - 1/3), 0\}. \tag{A.7}$$

The consumers for which the price schedule of firm 1 and firm 2 are zero are located $\tilde{x}_{12} = 1/6$. Denote also the consumers for which the price schedule of firm 1 is zero, i.e., $p_1(\tilde{x}_{31}) = 0$, as $\tilde{x}_{31} = 5/6 - p_3/2t$. As $\tilde{x}_{31} \in [2/3, 1]$ holds, the equilibrium pricing subgame if firm 1 and firm 2 have information on the consumers on firm 1's arc can then be characterised as follows.

Lemma 5. *If firm 1 and firm 2 have access to consumer information, the candidate equilibrium consists of the price schedules (A.6)-(A.7) and the prices*

$$p_2 = \frac{2}{7}t, \;\; p_3 = \frac{5}{21}t.$$

The marginal consumers are located at $\tilde{x}_{12} = 1/6$ and $\tilde{x}_{31} = 5/7$. The firms' profits are, respectively,

$$\pi_1 = \frac{193}{1764}t, \;\; \pi_2 = \frac{121}{1764}t, \;\; \pi_3 = \frac{25}{441}t.$$

Proof. We show the above result as follows: from the price schedules (A.6)-(A.7), the profit function of the firms are:

$$\pi_1 = \int_0^{1/6} \left[t(1/3 - 2x)\right] dx + \int_{\tilde{x}_{31}}^{1} \left[p_3 + t(5/3 - 2x)\right] dx.$$

$$\pi_2 = p_2\left(\frac{1}{6} + \frac{p_3 - p_2}{2t}\right) + \int_{1/6}^{1/3} \left[t(2x - 1/3)\right] dx.$$

$$\pi_3 = p_3\left[\left(\frac{1}{6} + \frac{p_2 - p_3}{2t}\right) + \left(\tilde{x}_{31} - 2/3\right)\right].$$

Standard calculations lead to the profit-maximising prices for the anonymous segment

$$p_2 = \frac{2}{7}t, \;\; p_3 = \frac{5}{21}t.$$

Using these prices and the price schedules (A.6)-(A.7), it is possible to derive the profits of the firms:

$$\pi_1 = \frac{193}{1764}t, \;\; \pi_2 = \frac{121}{1764}t, \;\; \pi_3 = \frac{25}{441}t.$$

Firm 2 and firm 3 have no unilateral incentive to increase discretely the price to $p^{dev} = t/3$, as this would give profits $\pi_2^{dev}(t/3, 5t/21) = 5t/126 < 121t/1764$ and $\pi_3^{dev}(t/3, 2t/7) = t/21 < 25t/441$, respectively. □

Firm 2 and firm 3 have access to consumer information. If firm 2 and firm 3 have access to the information, they can offer personalised prices to consumers on the profiled segment ($x \in [2/3, 1]$ and $x \in [0, 1/3]$). All three firms will also offer posted prices p_i.

Fig. A.2(b) illustrates the anticipated equilibrium demand configuration. Given these prices, the schedules for firm 2 and firm 3 are, respectively:

$$p_2(x) = \max\left\{p_1 + t(2x - 1/3), 0\right\}, \tag{A.8}$$

$$p_3(x) = \max\left\{p_1 + t(5/3 - 2x), 0\right\}. \tag{A.9}$$

Denote the consumers for which the price schedule of firm 2 and firm 3 are zero, i.e., $p_2(\tilde{x}_{21}) = p_2(\tilde{x}_{31}) = 0$, as $\tilde{x}_{12} = 1/6 - p_1/2t$ and $\tilde{x}_{31} = 5/6 + p_1/2t$, respectively. Assume that these consumers lie on the profiled arc, which we verify is the case in equilibrium.

Then, the equilibrium in the pricing subgame, if firm 2 and firm 3 have information on the consumers on firm 1's arc, can then be characterised as follows.

Lemma 6. *If firm 2 and firm 3 have access to consumer information and use all the list, the equilibrium consists of the price schedules (A.8) - (A.9) and the prices*

$$p_1 = \frac{t}{6}, \;\; p_2 = \frac{t}{3}, \;\; p_3 = \frac{t}{3}.$$

The marginal consumers are $\tilde{x}_{12} = {}^1\!/_{12}$ *and* $\tilde{x}_{31} = {}^{11}\!/_{12}$. *The firms' profits are, respectively,*

$$\pi_1 = \frac{t}{36}, \;\; \pi_2 = \frac{17}{144}t, \;\; \pi_3 = \frac{17}{144}t.$$

Proof. We prove the above results as follows: from the price schedules (A.8)-(A.9), the profit function of the firms are:

$$\pi_1 = p_1 \left[\tilde{x}_{12} + \left(1 - \tilde{x}_{31}\right)\right].$$

$$\pi_2 = p_2 \left(\frac{1}{6} + \frac{p_3 - p_2}{2t}\right) + \int_{\tilde{x}_{12}}^{1/3} \left[p_1 + t\left(2x - {}^1\!/_3\right)\right] dx.$$

$$\pi_3 = p_3 \left(\frac{1}{6} + \frac{p_2 - p_3}{2t}\right) + \int_{\tilde{x}_{31}}^{1/3} \left[p_1 + t\left({}^5\!/_3 - 2x\right)\right] dx.$$

Standard calculations lead to the profit-maximising prices for the anonymous segment

$$p_1 = \frac{t}{6}, \; p_2 = p_3 = \frac{t}{3}.$$

Using these prices and the price schedules (A.6)-(A.7), it is possible to derive the profits of the firms:

$$\pi_1 = \frac{t}{36}, \; \pi_2 = \pi_3 = \frac{17}{144}t. \quad \square$$

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.infoecopol.2024.101102.

References

Abrardi, L., Cambini, C., Congiu, R., Pino, F., 2024a. User data and endogenous entry in online markets. J. Ind. Econ. https://doi.org/10.1111/joie.12383. In press.

Abrardi, L., Cambini, C., Pino, F., 2024b. Data Brokers Competition, Synergic Datasets, and Endogenous Information Value. SSRN Working Paper 4901441.

Abrardi, L., Cambini, C., Pino, F., 2024c. Regulating data sales: the role of data selling mechanisms. Telecommun. Policy, 102813.

Acquisti, A., Varian, H.R., 2005. Conditioning prices on purchase history. Mark. Sci. 24, 367–381.

Acquisti, A., Taylor, C.R., Wagman, L., 2016. The economics of privacy. J. Econ. Lit. 54, 442–492.

Alderighi, M., Piga, C.A., 2012. Localized competition, heterogeneous firms and vertical relations. J. Ind. Econ. 60, 46–74.

Ali, S.N., Lewis, G., Vasserman, S., 2023. Voluntary disclosure and personalized pricing. Rev. Econ. Stud. 90, 538–571.

Anderson, S., Baik, A., Larson, N., 2023. Price discrimination in the information age: prices, poaching, and privacy with personalized targeted discounts. Rev. Econ. Stud. 90, 2085–2115.

Belleflamme, P., Lam, W.M.W., Vergote, W., 2020. Competitive imperfect price discrimination and market power. Mark. Sci. 39, 996–1015.

Bergemann, D., Bonatti, A., 2019. Markets for information: an introduction. Annu. Rev. Econ. 11, 1–23.

Bergemann, D., Bonatti, A., Gan, T., 2022. The economics of social data. Rand J. Econ. 53, 263–296.

Bounie, D., Dubus, A., Waelbroeck, P., 2020. Market for Information and Selling Mechanisms. SSRN Working Paper 3454193.

Bounie, D., Dubus, A., Waelbroek, P., 2021. Selling strategic information in competitive markets. Rand J. Econ. 52, 283–313.

Brunt, C.S., King, A.S., King, J.T., 2020. The influence of user-generated content on video game demand. J. Cult. Econ. 44, 35–56.

Casadesus-Masanell, R., Hervas-Drane, A., 2015. Competing with privacy. Manag. Sci. 61, 229–246.

Chen, Y., Riordan, M.H., 2007. Price and variety in the spokes model. Econ. J. 117, 897–921.

Choe, C., King, S., Matsushima, N., 2017. Pricing with cookies: behavior-based price discrimination and spatial competition. Manag. Sci. 64, 5669–5687.

Choe, C., Cong, J., Wang, C., 2024. Softening competition through unilateral sharing of customer data. Manag. Sci. 70, 526–543. https://doi.org/10.1287/mnsc.2023.4689.

Choi, J.P., Jeon, D.-S., Kim, B.-C., 2019. Privacy and personal data collection with information externalities. J. Public Econ. 173, 113–124.

Clavorà Braulin, F., 2023. The effects of personal information on competition: consumer privacy and partial price discrimination. Int. J. Ind. Organ. 87.

Conitzer, V., Taylor, C.R., Wagman, L., 2012. Hide and seek: costly consumer privacy in a market with repeat purchases. Mark. Sci. 31, 277–292.

Fainmesser, I.P., Olié Lauga, D., Ofek, E., 2021. Ratings, reviews, and the marketing of new products. Manag. Sci. 67, 6629–7289.

Frontier Technology Quarterly, 2019. Data economy: radical transformation or dystopia? https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/FTQ_1_Jan_2019.pdf. (Accessed 19 August 2024).

FTC, 2014. Data Brokers: A Call for Transparency and Accountability.

Gilbert, R.J., Newbery, D.M., 1982. Preemptive patenting and the persistence of monopoly. Am. Econ. Rev. 72, 514–526.

Goldfarb, A., Tucker, C., 2019. Digital economics. J. Econ. Lit. 57, 3–43.

Gu, Y., Madio, L., Reggiani, C., 2019. Exclusive Data, Price Manipulation and Market Leadership. CESifo Working Paper No. 7853.

Gu, Y., Madio, L., Reggiani, C., 2022. Data brokers co-opetition. Oxf. Econ. Pap. 74, 820–839.

Han, X., Wang, L., Fan, W., 2023. Cost-effective social media influencer marketing. INFORMS J. Comput. 35, 138–157.

Hidir, S., Vellodi, N., 2021. Privacy, personalization, and price discrimination. J. Eur. Econ. Assoc. 19, 1342–1363.

Ichihashi, S., 2020. Online privacy and information disclosure by consumers. Am. Econ. Rev. 110, 569–595.

Ichihashi, S., 2021. Competing data intermediaries. Rand J. Econ. 52, 515–537.

Jehiel, P., Moldovanu, B., 2000. Auctions with downstream interaction among buyers. Rand J. Econ., 768–791.

Jehiel, P., Moldovanu, B., Stacchetti, E., 1996. How (not) to sell nuclear weapons. Am. Econ. Rev., 814–829.

Jones, C.I., Tonetti, C., 2020. Nonrivalry and the economics of data. Am. Econ. Rev. 110, 2819–2858.

Kim, J.-H., Wagman, L., Wickelgren, A.L., 2019. The impact of access to consumer data on the competitive effects of horizontal mergers and exclusive dealing. J. Econ. Manag. Strategy 28, 373–391.

Laub, R., Miller, K.M., Skiera, B., 2023. The economic value of user tracking for publishers. ArXiv preprint arXiv:2303.10906.

Liu, Q., Serfes, K., 2006. Customer information sharing among rival firms. Eur. Econ. Rev. 50, 1571–1600.

Martens, B., Mueller-Langer, F., 2020. Access to digital car data and competition in aftermarket maintenance services. J. Compet. Law Econ. 16, 116–141.

Martens, B., Parker, G., Petropoulos, G., Van Alstyne, M.W., 2021. Towards Efficient Information Sharing in Network Markets. TILEC Discussion Paper No. DP2021-014.

Miklós-Thal, J., Goldfarb, A., Haviv, A.M., Tucker, C., 2023. Digital Hermits. NBER Working Paper 30920.

Montes, R., Sand-Zantman, W., Valletti, T., 2019. The value of personal information in online markets with endogenous privacy. Manag. Sci. 65, 1342–1362. https://doi.org/10.1287/mnsc.2017.2989.

OECD, 2018. Personalised pricing in the digital era. Discussion Paper. Organisation for Economic Co-operation and Development.

Parker, G., Petropoulos, G., Van Alstyne, M.W., 2020. Digital platforms and antitrust. SSRN Working Paper 3608397.

Perloff, J.M., Salop, S.C., 1985. Equilibrium with product differentiation. Rev. Econ. Stud. 52, 107–120.

Pino, F., 2022. The microeconomics of data–a survey. J. Ind. Busin. Econ. 49, 1–31.

Reimers, I., Waldfogel, J., 2021. Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings. Am. Econ. Rev. 111, 1944–1971.

Rhodes, A., Zhou, J., 2024. Personalized pricing and competition. Am. Econ. Rev. 114, 2141–2170.

Salop, S.C., 1979. Monopolistic competition with outside goods. Bell J. Econ. 10, 141–156.

Syverson, C., 2004. Market structure and productivity: a concrete example. J. Polit. Econ. 112, 1181–1222.

Taylor, C., Wagman, L., 2014. Consumer privacy in oligopolistic markets: winners, losers, and welfare. Int. J. Ind. Organ. 34, 80–84. https://doi.org/10.1016/j.ijindorg.2014.02.010.

Thisse, J.-F., Vives, X., 1988. On the strategic choice of spatial price policy. Am. Econ. Rev. 78, 122–137.

Transparency Market Research, 2022. Data broker market report. https://www.transparencymarketresearch.com/data-brokers-market.html. (Accessed 19 August 2024).

Tucker, C., Neumann, N., 2020. Buying consumer data? Tread carefully. Harvard Business Review. https://hbr.org/2020/05/buying-consumer-data-tread-carefully. (Accessed 19 August 2024).

Vickers, J., 1985. Pre-emptive patenting, joint ventures, and the persistence of oligopoly. Int. J. Ind. Organ. 3, 261–273.

Vickrey, W., 1964. Microstatistics. Brace and World.

Waldfogel, J., Reimers, I., 2015. Storming the gatekeepers: digital disintermediation in the market for books. Inf. Econ. Policy 31, 47–58.

Xu, Z., Dukes, A., 2022. Personalization from customer data aggregation using list price. Manag. Sci. 68, 960–980.