CESifo
AREA CONFERENCES 2019

Economics of Digitization

Munich, 22–23 November 2019

Exclusive Data, Price Manipulation and Market Leadership

Yiquan Gu, Leonardo Madio, and Carlo Reggiani

CESifo WORKING PAPERS

7853
2019
September 2019

Exclusive Data, Price Manipulation and Market Leadership

Yiquan Gu, Leonardo Madio, Carlo Reggiani

CESifo

Impressum:

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

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

CESifo *Working Paper No. 7853*
Category 14: Economics of Digitization

Exclusive Data, Price Manipulation and Market Leadership

Abstract

The unprecedented access of firms to consumer level data not only facilitates more precisely targeted individual pricing but also alters firms' strategic incentives. We show that exclusive access to a list of consumers can provide incentives for a firm to endogenously assume the price leader's role, and so to strategically manipulate its rival's price. Prices and profits are non-monotonic in the length of the consumer list. For an intermediate size, price leadership entails a semi-collusive outcome, characterized by supra-competitive prices and low consumer surplus. In contrast, for short or long lists of consumers, exclusive data availability intensifies market competition.

JEL-Codes: D430, K210, L110, L130, L410, L860, M210, M310.

Keywords: exclusive data, price leadership, personalized pricing, price discrimination.

Yiquan Gu
University of Liverpool
Management School
Chatham Street
United Kingdom – Liverpool, L69 7ZH
yiquan.gu@liv.ac.uk

Leonardo Madio
Université Catholique de Louvain
Center for Operations Research and Econometrics (CORE)
Voie du Roman Pays, 34 – L1.03.01
Belgium – 1348 Louvain-la-Neuve
leonardo.madio@uclouvain.be

Carlo Reggiani
University of Manchester
School of Social Sciences-Economics
United Kingdom – Manchester, M13 9PL
carlo.reggiani@manchester.ac.uk

September 2019
We are grateful to Paul Belleflamme, Axel Gautier, Aditya Goenka, Timo Klein, Wynne Lam, Luca Lambertini, Paul Madden, Alessandro Pavan, Yossi Spiegel, Tommaso Valletti, Chris Wallace and seminar participants at the 44th EARIE Conference (Maastricht), 16th ZEW Conference on the Economics of Information and Communication Technologies (Mannheim), University of Liverpool, University of Manchester and Southwestern University of Finance and Economics, for helpful comments and discussion. Leonardo acknowledges financial support from the "MOVE-IN Louvain" Incoming Post-doc Fellowship Programme. The usual disclaimer applies.

1 Introduction

Firms' access to "big data" and the advent of artificial intelligence have revolutionised business models and the interaction of firms in the market. Tech giants such as Amazon, Apple or Google have collected, stored and analysed individual data (*e.g.*, clicks, purchase histories, connected devices) for years. Even brick and mortar companies like the British grocery Tesco or Starbucks cafés have been accumulating data through loyalty reward programmes or exploiting technological innovations. Most of these data are *exclusive* and can potentially be employed to personalise offers and more effective price discrimination.[1] Exclusive data access, hence, is an asset that can deliver an advantage relative to rival firms competing in similar segments, that often have more limited access to consumers' characteristics.

Whether access to data can entail pro or anti-competitive effects is largely debated. On the one hand, the legal literature, based on traditional antitrust analysis, suggests that big data constitute a source of entry barriers or, more generally, competitive advantage (Grunes and Stucke, 2015; Stucke and Grunes, 2016; Graef, 2017). On the other hand, others argue that data are non-rival and non-excludable and access to them does not, *per se*, lead to anti-competitive concerns (Tucker and Wellford, 2014; Lambrecht and Tucker, 2017; Varian, 2018; Tucker, 2019).[2] Moreover, recent studies and policy reports highlighted that the impact of data-driven price discrimination on consumer welfare is mostly ambiguous (Taylor and Wagman, 2014; Bourreau et al., 2017; Bourreau and de Streel, 2018), although "exclusive possession of data, combined with a lack of engagement by consumers, can lead to a lack of competitive pressure" (Furman et al., 2019, p 34).

This article contributes to the debate by further exploring the issue of *exclusive access* to customer data and shows how this can affect firms' *competitive strategies*, in the form of market segmentation, price leadership, and tacit coordination.

We present a model of price competition with horizontal product differentiation in which one firm has access to an exclusive *list* of consumer location. Through this list, which only includes a share of all consumers in the market, the accessing firm (informed) can personalise its offers and price discriminate by matching the utility of the rival's offer for the profiled consumers. Whereas such consumers are anonymous only to the non-accessing firm (uninformed), the rest of the consumers remain anonymous to both firms. This setting allows us studying the impact of accessing exclusive consumer information on the strategic incentives of firms to lead in price competition. In particular, the two firms independently and simultaneously decide on whether to announce its price "early" or "late" (Hamilton and Slutsky, 1990). If their choices coincide, then simultaneous competition follows. Otherwise, the firm that selects "early" announces its price first and acts as a price

[1] Whereas all these practices are not new, "big data" are enabling more and more precise (and often subtle) discrimination, particularly online (Mikians et al., 2012, 2013; Hannak et al., 2014).

[2] A discussion on data-driven incumbency advantage is also provided by Biglaiser et al. (2019).

leader. Then, the other firm sets its price after observing its rival's move.

We find that exclusive possession of data, by enabling personalised pricing, has the potential to dramatically affect market competition and firms' strategic incentives. When the share of identifiable consumers is small, both firms would prefer to be a follower rather than a leader. Being a price follower provides the opportunity to undercut the rival to obtain a higher profit than when leading (d'Aspremont and Gerard-Varet, 1980; Hamilton and Slutsky, 1990).

As the size of the list increases, the ability to price discriminate becomes an important strategic asset for a firm. A *novel mechanism* enabled by the exclusive information, in fact, works through the linkage between the prices. The firm owning the list can induce the rival firm to raise its price by willingly posting a high price itself first. As a result, the rival can profitably undercut the price of the list accessing firm in the anonymous segment and serve it entirely. The list accessing firm can in turn serve the list segment consumers by matching the utility the rival firm offers them at its relatively high price. Such a strategic price manipulation leads to a "win-win" scenario for both firms: the informed firm's price in the anonymous segment is high and so is the rival's undercutting price which in turn benefits the informed firm in the list segment.

The mechanism described can generate enough profit from the identified consumers to induce the list accessing firm to completely abandon the anonymous segment by setting a high price there in the first place. The strategy is also profitable to the uninformed follower, as long as the size of the list is not very large. In the latter case, the anonymous segment is very small and, as a result, the firm without access to data would find it profitable to compete also in the list segment by setting a very low price that cannot be not matched by the informed firm. Therefore, the above mechanism no longer holds when the list grows very large.

There are real world markets in which the conditions exist for the forces highlighted by our model to operate. For instance, in the grocery sector, we observe evidence of exclusive access to data, personalised offers, and evidence of market leadership. The British chain Tesco collects customer data through their Clubcard loyalty programme since the mid-nineties, whereas ASDA, its largest competitor, does not engage in similar activities. These data are used to send targeted offers to consumers in the loyalty programme. There is also evidence that Tesco acts as a price leader in the sector: the chain has been reported to be often the first to amend its pricing strategy and, interestingly, this happens more frequently when prices increase (Seaton and Waterson, 2013; Dobson et al., 2016).

Further, Amazon marketplace has revolutionised e-commerce. The platform has access to a huge volume of exclusive data on its users' browsing histories. This information can be used to personalise its product and price offers, which is often implemented through the "exclusive deals" reserved to Amazon Prime customers.[3] Evidence suggests that in the

[3]Moreover, Amazon's recent acquisition of Whole Foods is likely to give them an informational advantage

US, engineers at Wal-Mart access Amazon's website several million times a day: Wal-Mart relies on computer programs that scan competitors' prices, so it can adjust its listings accordingly (Dastin, 2017).[4] This clearly hints at Amazon's price leadership.

Moreover, the global coffee chain Starbucks uses its loyalty rewards app to collect geo-location and purchasing data.[5] These data are then used to send app notifications and emails with personalised beverage and food offers. Established rival café chains are unlikely to have access to the same depth and precision of information. Notably, Dunkin' Donuts in US is currently updating and reviving its loyalty programme, whereas in the UK a close competitor like Costa is only piloting a mobile ordering service (Tyko, 2019; Comunicaffe, 2019).

The previous examples underline the interactions between informed leaders and uninformed followers' that our model effectively captures. Indeed, exclusive access to a sufficiently large list of customers can affect the leaders' pricing incentives, potentially setting the stage for the tacit coordination mechanism that our results highlight. Such anti-competitive effects do not always arise in equilibrium. We find instead that exclusive access to a very small or very large amount of data has a pro-competitive effect: it feeds market competition by reducing prices and firms' profits.

The main contribution of our analysis is, then, to emphasise that exclusive access to consumer level information *can* lead to pro- or anti-competitive conducts, depending on the size of the list. Unlike most part of the received literature on competitive price discrimination (Thisse and Vives, 1988; Armstrong, 2006; Stole, 2007), consumers' information and segmentation might act as a semi-collusive device that softens price competition, with negative implications for consumer welfare.

These conclusions should draw the attention of policy makers. Whereas the UK CMA (2018) stated that "tacit coordination and personalised pricing are very unlikely to occur together", our results suggest caution. Tacit coordination, in fact, might stem from the joint presence of exclusive information to price discriminate and market leadership. In such cases, mandatory data sharing can be an effective tool to restore market competition, but it should be considered on a case by case basis.

The rest of the article unfolds as follows. In Section 2 we briefly review the most closely related literature to further locate our contribution. Section 3 presents the model. Section 4 studies price competition. Section 5 presents the main results and their implications. Section 6 discusses the findings and provides the policy and competitive implications of the analysis.

---

compared to their closest rivals in the grocery segment, given the huge amount of data accessible through the new parent company. On the Amazon-Whole Foods merger and its data implications see, *e.g.*, Petro (2017).

[4] In the context of e-commerce, scraping the market leaders' prices is not unusual: the Portuguese Autoridade da Concorrência (2019) reports that almost half of the inquired firms systematically engaged in it.

[5] Starbucks went even further recently, experimenting licence plate recognition in its drive through affiliates (Hodgson, 2019).

2 Related literature and contribution

There is a vast literature on competitive price discrimination. A common finding is that, if firms are symmetric, price discrimination leads to lower profits relative to uniform pricing. The results are rooted in the so called “best response asymmetry”: the “strong market” (low elasticity segment) of one firm is the “weak market” (high elasticity segment) of its rival, and vice versa (Thisse and Vives, 1988; Armstrong, 2006; Stole, 2007). These forces are so robust that lead to similar results in models of “behaviour based price discrimination”, whereby firms condition their prices to purchase history (Fudenberg and Tirole, 2000; Villas-Boas, 1999). The literature, then, developed in several different directions.[6]

Here, we focus on one stream of this literature which considers the effect of data endowment on firms' strategies.[7] In the marketing jargon, this has been referred to as “addressability”, that is, the possibility to identify consumers and offer personalised prices (Blattberg and Deighton, 1991). In this regard, Chen and Iyer (2002) and Belleflamme et al. (2017) study competition between firms that have access to addressable segments. Both articles find that price discrimination can be profitable as long as firms have heterogeneous degrees of addressability.

In Chen and Iyer (2002), investments in addressability partition the market into four segments, depending on whether consumers are profiled by both, one or none of the firms. This leads to two countervailing effects: on the one hand, a “surplus extraction” effect, whereby a firm cannot match its rival's price in a given segment; on the other hand, a “competitive effect” driven by the fiercer competition in the common addressable segment. Heterogeneity ensures that the first effect dominates for at least one firm; at the same time, the rival gets low profits as a result.

Belleflamme et al. (2017) consider a homogeneous product and identify three segments of consumers: those profiled by both firms, those only known to the informationally advantaged one and non-addressable consumers. In a mixed strategy, the expected price is higher than marginal cost, and the informationally disadvantaged firm posts the highest price with non-zero probability. This relaxes the competition for consumers profiled only by the advantaged firm.

Although our analysis are related to theirs, there are important differences. For instance, our model has differentiated products, exogenous exclusive consumer data, endogenous timing and pure strategy equilibria. Above all, we reach the result of high prices and profitable price discrimination with a fully covered market: firms' prices do not “cut out”

[6]See Fudenberg and Villas-Boas (2006) and Esteves (2009) for in depth reviews. For more recent contributions see, *e.g.*, Pazgal and Soberman (2008), Zhang (2011), Jing (2016) and Choe et al. (2017).

[7]Other streams of this literature have dealt with related issues such as privacy and its market implications (Taylor, 2004; Casadesus-Masanell and Hervas-Drane, 2015; Shy and Stenbacka, 2016), modelling of the data broker industry (Bergemann and Bonatti, 2015, 2019; Bounie et al., 2018; Gu et al., 2019; Ichihashi, 2019), and data ownership (Dosis and Sand-Zantman, 2019).

consumers, and hence do not alter the elasticity of the different segments (Pazgal et al., 2013). In our model, it is price leadership that offers a credible commitment and allows for firms' "coordination" and complete market segmentation.

Our approach to model consumer profiling resembles Montes et al. (2019) who, on the other hand, address a different question: how does a data broker optimally sell a list of consumers. In their study, firms achieve addressability by buying the list, and the key finding is that revenue of an owner of consumer data is maximised through exclusive sale to one of the competing firms.[8] Montes et al. (2019) allow consumers to be removed from the list.[9] If the costs of preserving privacy is not too high, the consumers left on the list have a higher average willingness to pay. This leads the informed firm to post a high price and, anticipating that, the uninformed one behaves likewise and does not compete for the addressable segment. This outcome is reminiscent of ours, but the forces in operation are different. In Montes et al. (2019), it is consumers' behaviour that enables better segmentation and prices higher than those in the Hotelling benchmark; in our setting, instead, price leadership, together with an intermediate share of consumers on the list, enables both firms to charge an extremely high price.

A second key contribution of our results is to the literature on timing in oligopoly games. The incentives to lead were studied in the pioneering articles of d'Aspremont and Gerard-Varet (1980) and Hamilton and Slutsky (1990). Price competition between identical firms is typically characterised by a second-mover advantage: even in the presence of differentiated products, the follower benefits from the possibility of undercutting (Gal-Or, 1985).[10] In this context, this article is closely related to Van Damme and Hurkens (2004) and Amir and Stepanova (2006). Both articles show that cost asymmetries between firms can overturn firms' strategic incentives. Namely, if costs are sufficiently heterogeneous, the most efficient firm has an incentive to lead. They both employ an extended game where firms decide on an early or late move and achieve the result by invoking the risk-dominance refinement (Harsanyi and Selten, 1988). The main difference is that in Van Damme and Hurkens (2004) deciding to move first requires committing to an action, whereas in Amir and Stepanova (2006) it does not. Recently, Madden and Pezzino (2019) show that the ownership of an essential input can lead to endogenous price leadership without invoking refinements.

Our article complements and adds to this literature by showing that technologically identical firms may still have an incentive to lead in price competition in presence of a heterogeneous access to an immaterial asset, *i.e.*, asymmetric informational endowment. Moreover, we also endogenise the timing of the pricing choice in presence of price discrimination.

[8] Clavorà Braulin and Valletti (2016) use a similar approach in the context of vertically differentiated firms. They also find that in equilibrium exclusive data selling arises.

[9] Chen et al. (2019) also consider active consumers but with a different connotation: all consumers receive personalised offers but some search further for the posted prices.

[10] The importance of market leadership and of informational advantage is also highlighted by Calzolari and Pavan (2006). In their model of sequential contracting, an upstream seller can take advantage of its (exogenous) leadership by revealing or not some information to another seller. In our case, there is no information sharing but the informational advantage can induce a firm to act as a leader while influencing the rival one.

3 The framework

We consider a market in which two profit maximising firms sell a horizontally differentiated product to a unit mass of consumers by competing in prices. Let the two firms, $i = A, B$, be located at 0 and 1 on the Hotelling line, respectively. Each consumer demands at most one unit of the product. Consumers are uniformly distributed on the Hotelling line and indexed by their location $x \in [0, 1]$. Hence, for consumer x the indirect utility from buying product A at price p_A is $V_A = v - p_A - tx$, where v is a consumer's reservation value and t measures transport cost. Likewise, for consumer x the indirect utility from buying product B at price p_B is $V_B = v - p_B - t(1 - x)$. For both firms, we assume constant marginal cost of production which is normalised at zero. Throughout the article, we also assume that v is sufficiently large so that the market is covered.[11]

Turning to consumer information, we assume that one of the firms, without loss of generality, Firm A, has exclusive access to a list of consumers. For ease of exposition, we refer to it as the "informed firm", whereas to Firm B as the "uninformed firm". The list is exogenously given and can be thought as having been previously obtained, through either a data broker or active collection. Although generic data are typically non-rival and often largely available to firms (*e.g.*, data harvesting on the Web), as discussed, customer-specific data can be rendered practically excludable or unequally commercially valuable: this happens both as a reflection of market players' unequal data endowment as well as their heterogeneous analytical capabilities (*e.g.*, Amazon vs. Walmart, Tesco vs. ASDA). By the same token, exclusive access to a list of customers is also rooted in the previous literature: Montes et al. (2019) alongside with Clavorà Braulin and Valletti (2016) and, in the context of auctions, Jehiel and Moldovanu (2000), find that data suppliers have an incentive to provide consumer information through exclusive deals.[12]

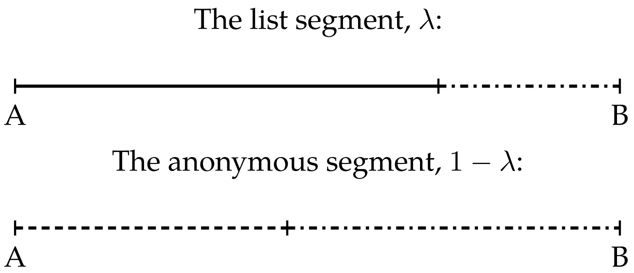
Formally, Firm A's customer data availability is captured by the length of the list $\lambda \in (0, 1)$, which can be interpreted as the percentage of the consumers whose location x is known to A. This information enables price discrimination, as for those consumers who are on the list the firm can make personalised offers conditional on their location, *i.e.*, offer $\widetilde{p}_A(x)$ to consumer x. As a result, Firm A can distinguish two segments of the market depending on whether or not a consumer is on the list. We refer to the former as the "list" segment and the latter the "anonymous" segment. For simplicity, we assume consumers in each segment to be uniformly located. Without any consumer information, Firm B, however, can only set one uniform price for its product. The model is illustrated in Figure 1.

In this setting, we study firms' strategic incentives to lead or to follow in price competition. To this end, consider the following dynamic game. In the first stage, Firms A and B inde-

[11] More specifically, for our analysis, $v > 4t$ is sufficient.

[12] In a looser interpretation, this assumption can also be thought as capturing firms' asymmetric data endowment, where the exclusive part available to one of the firms is crucial for implementing effective personalised pricing.

Figure 1: The market



Note: The figure presents the two segments of the market. λ represents the percentage of consumers who are identified by Firm A and can be offered a personalised price $\widetilde{p}_A(x)$. Those consumers can of course also observe p_B, Firm B's uniform price. For consumers in the anonymous segment, they observe both p_A (Firm A's uniform price for consumers it cannot identify) and p_B but are not offered the personalised price $\widetilde{p}_A(x)$. In this example, Firm A can match the utility Firm B offers with a personalised offer in the list segment up to a certain point (full line). The rest buys from Firm B (dot-dashed line). In the anonymous segment, the usual purchase decisions as in the standard Hotelling model apply.

pendently and simultaneously decide whether to announce their price for the *anonymous segment* early (τ_0) or late (τ_1). Depending on the outcome of this stage, in the second stage, four subgames follow.

In the two outcomes where the timing choices coincide, *i.e.*, (τ_0, τ_0) or (τ_1, τ_1), the firms set their prices in the anonymous segment, p_A and p_B, independently and simultaneously. Then, after observing Firm B's price p_B, Firm A can price discriminate in the list segment by setting a location dependent price schedule $\widetilde{p}_A(x)$ attempting to match the utility that the consumer located at x can obtain by buying from firm B at price p_B.

On the other hand, if the first stage outcome is (τ_0, τ_1), Firm A first sets a price p_A for the anonymous segment and then Firm B chooses p_B after observing p_A. Again in the last stage, Firm A sets a schedule $\widetilde{p}_A(x)$ for the list segment. Finally, when the first stage outcome is (τ_1, τ_0), Firm B first sets a price p_B and then Firm A chooses p_A for the anonymous segment and a schedule $\widetilde{p}_A(x)$ for the list segment. We note that we have assumed Firm A can always set $\widetilde{p}_A(x)$ after observing p_B. This reflects the observation that in this information age, Firm A can update its personalised prices more quickly than Firm B can adjust its uniform price.[13] We solve the game by backward induction starting from simultaneous price competition.

[13]Another justification is that, given Firm A's exclusive access to the consumer list, in equilibrium it is immaterial whether Firm A sets the price schedule $\widetilde{p}_A(x)$ at the same time as p_A (by backward induction) or after actually observing p_B.

4 Price competition subgames

4.1 Simultaneous price competition in the anonymous market

Suppose that either (τ_0, τ_0) or (τ_1, τ_1) has resulted in the first stage. In this case, the two firms set prices simultaneously.

To identify firm profits, we first consider Firm A's price schedule for the list segment in the last stage. Given p_B, Firm A's optimal list segment schedule $\widetilde{p}_A(x)$ should aim to match the next best alternative for consumers in the list segment. That is, Firm A may set $\widetilde{p}_A(x)$ for a consumer located at x in the list segment such that she is indifferent between buying from A or B:

$$\widetilde{p}_A(x) = p_B + t(1-2x). \tag{1}$$

This requires us to consider two cases depending on p_B.

First, if $p_B < t$, Firm A can only attract consumers up to $\bar{x} = {(t+p_B)}/{2t}$. This is because Firm B's price is so low that even if Firm A sets it price at zero for consumers to the right of $\bar{x}$, they would still buy from B due to savings on the transport cost. Thus, in the list segment, the profits of the firms are, respectively,

$$\begin{aligned} \widetilde{\pi}_A &= \int_0^{\bar{x}} \widetilde{p}_A(x)dx = \frac{(t+p_B)^2}{4t}, \\ \widetilde{\pi}_B &= p_B(1-\bar{x}) = \frac{p_B(t-p_B)}{2t}. \end{aligned} \tag{2}$$

For a given pair of uniform prices (p_A, p_B), the firms' respective total profits from both the list and the anonymous segment are,

$$\begin{aligned} \pi_A &= \lambda \frac{(t+p_B)^2}{4t} + (1-\lambda)p_A\left(\frac{1}{2} + \frac{p_B - p_A}{2t}\right), \\ \pi_B &= \lambda \frac{p_B(t-p_B)}{2t} + (1-\lambda)p_B\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right). \end{aligned}$$

Following the standard procedure, the candidate equilibrium is $p_A = \frac{3}{\lambda+3}t$ and $p_B = \frac{3-\lambda}{\lambda+3}t$. Note that $p_B < t$, indeed, holds.

Second, if $p_B \geq t$, then Firm B does not serve the list segment. As a consequence, the firms' best responses are entirely determined in the anonymous segment. Thus, the only candidate equilibrium is $(p_A, p_B) = (t, t)$, as in the standard Hotelling model. However, this is not an equilibrium since a profitable deviation exists for Firm B. Hence, there is no equilibrium if $p_B \geq t$. We show this formally in Appendix A.1.

The following proposition summarises our main findings in the simultaneous pricing subgame.

Proposition 1. *The equilibrium in a simultaneous pricing subgame consists of the unique pair*

$$p_A = \frac{3}{\lambda+3}t, \;\; p_B = \frac{3-\lambda}{\lambda+3}t.$$

The two firms' profits are, respectively,

$$\pi_A = \frac{9(\lambda+1)}{2(\lambda+3)^2}t, \;\; \pi_B = \frac{(3-\lambda)^2}{2(\lambda+3)^2}t.$$

Proof: See Appendix A.1.

Proposition 1 already provides interesting insights. To begin with, it demonstrates that if prices are posted simultaneously, exclusive access to consumer data makes the market more competitive. Compared to the standard Hotelling outcome ($\lambda = 0$), Firm A's access to consumer information makes Firm B price more aggressively, in order to serve some of those consumers who are included in the list. As prices are strategic complements, this also results in a lower posted price charged by Firm A.

For Firm B, the price effect triggered by the rival's access to the list dominates, and it obtains a lower profit than it would obtain had it not been the list. In contrast, Firm A's profit increases with the length of the list, λ, thanks to its ability to price discriminate consumers in that segment. This observation highlights both the absolute and relative benefit of having access to consumer information, compared to the uninformed firm.

The above reasoning is supported by simple comparative statics, which shows that an increase in λ feeds fiercer competition and both posted prices, p_A and p_B, go down. Between p_A and p_B, Firm B's price decreases even faster: the uninformed firm has to compete more aggressively to attract consumers in the wider list segment and in the shrinking anonymous one. Whereas more intense competition is usual in a symmetric scenario, leading to prisoner's dilemma situations in the choice of price discrimination (*e.g.*, Fudenberg and Tirole, 2000, *inter alios*), this also emerges in case of exclusive access to data. Increased price competition, however, still grants a profit advantage to the informed firm.

4.2 Firm B as the price leader

We now consider the subgame following (τ_1, τ_0) where Firm B first sets its posted price p_B, and then Firm A chooses its pricing after observing it. To organise our analysis, we proceed in two steps. First, we address Firm A's optimal schedule for the list segment and its best response in the anonymous segment after observing p_B (Lemma 1). Second, we present Firm B's optimal price at the beginning of this subgame (Proposition 2).

We start by characterising Firm A's best responses in the list segment and in the anonymous segment, respectively. When p_B is a viable option for consumer x, Firm A shall try to set $\widetilde{p}_A(x)$ to make her indifferent between buying A and B. However, as discussed in Section 4.1, if p_B is so low, such that Firm A has no chance of luring the consumer at x, any permissible price would be a best response. On the other hand, when p_B is so high that it is itself not a viable option for consumer x, $\widetilde{p}_A(x)$ shall make her indifferent between buying A or not buying at all. Lemma 1 summarises these results and, in addition, presents Firm A's best response in the anonymous segment, which also depends on the level of p_B.

Lemma 1. *For given $p_B \geq 0$ and $x \in [0, 1]$, Firm A's optimal list segment schedule is*

$$\widetilde{p}_A(x) \begin{cases} = v - tx & \text{if } p_B \geq v - t(1-x) \\ = p_B + t(1-2x) & \text{if } \max\{(2x-1)t, 0\} < p_B < v - t(1-x) \\ \in [0, +\infty) & \text{if } 0 \leq p_B \leq \max\{(2x-1)t, 0\} \end{cases} .$$

For a given $p_B \geq 0$, Firm A's best response price in the anonymous segment is

$$p_A(p_B) = \begin{cases} v - t, & \text{if } p_B \geq v \\ p_B - t, & \text{if } 3t \leq p_B < v \\ \frac{t}{2} + \frac{p_B}{2}, & \text{if } 0 \leq p_B < 3t \end{cases} .$$

Proof: See Appendix A.2.

On the basis of Firm A's best responses (Lemma 1), we can now characterise Firm B's possible profits and its optimal choice when leading in price competition. First note that, when $p_B \geq 3t$, Firm A's best response is to match B's offer in the list segment and undercut entirely in the anonymous segment. Thus, firm B obtains neither demand nor profits, and such prices are clearly (weakly) dominated. So, we can focus on $p_B < 3t$.

Yet, there are two cases to be considered. If $t < p_B < 3t$, the price is relatively high and Firm B receives no demand in the list segment. Given A's best response in Lemma 1 and the profits in the anonymous segment, a candidate equilibrium is identified. If $p_B \leq t$, we proceed similarly. In this case, Firm B serves all those list consumers to whom Firm A is unable to match its offer, and a share of the anonymous segment. As long as the length of the list is not too short, an interior solution exists and a second candidate equilibrium is identified. Thus, the question for Firm B is when to set a price below or above t. By comparing Firm B's profits associated to these two candidate equilibrium prices, we can identify the share of consumers in the list for which Firm B is indifferent between choosing a relatively high and a low price. The details of these results are in Proposition 2 and its proof in Appendix A.3.

Proposition 2. *When Firm B leads in the price competition, the subgame perfect Nash equilibrium in this pricing subgame is as follows.*

(i) If $\lambda \leq 3/5$, then the equilibrium prices are

$$p_A = \frac{5}{4}t, \;\; p_B = \frac{3}{2}t,$$

with respective profits

$$\pi_A = \frac{25 + 23\lambda}{32}t, \;\; \pi_B = \frac{9}{16}(1 - \lambda)t.$$

(ii) If $\lambda > 3/5$, then

$$p_A = \frac{5 + \lambda}{4(1 + \lambda)}t, \;\; p_B = \frac{3 - \lambda}{2(1 + \lambda)}t,$$

with respective profits

$$\pi_A = \frac{(5 + \lambda)^2}{32(1 + \lambda)}t, \;\; \pi_B = \frac{(3 - \lambda)^2}{16(1 + \lambda)}t.$$

Proof: See Appendix A.3.

The results in Proposition 2 have an intuitive interpretation. If the share of consumers on the list is not too large ($\lambda \leq 3/5$), Firm B can charge a relatively high price ($p_B > t$) as a leader. In this case, giving up the list segment is not too costly and it is more than compensated by the higher profits made on the anonymous segment. Firm A, in fact, follows and responds with a lower yet relatively high price ($p_A > t$). On the other hand, as the list segment becomes important ($\lambda > 3/5$), Firm B would rather set a low price ($p_B < t$) to make sure that it attracts demand from the list segment too.

We note further that when the list is relatively small, prices are not affected by how many consumers can be actually profiled. However, an increase in the list size increases the share of market demand received by the informed firm, and hence increases its profit and decreases that of the uninformed firm. In contrast, if the list is sufficiently comprehensive, $\lambda > 3/5$, increasing the list size makes Firm B increasingly concerned of list segment profits. Hence, induces it to place more weight on the low price designed to fend off A's matching offers in the list segment. This in turn feeds the competitive pressure on p_A in the anonymous segment, which then lowers firms' profits.

4.3 Firm A as the price leader

Consider the subgame following (τ_0, τ_1) where Firm A chooses its price in the anonymous segment p_A first, and then Firm B sets its price p_B after observing p_A. Finally, Firm A sets the price schedule in the list segment, after observing p_B. To present the results in this section, let us first define two critical values of list size. Namely, let

$$\lambda_1 = \frac{8v - 9t - 4\sqrt{2}\sqrt{(v-3t)(2v-3t)}}{16v - 23t},$$
$$\lambda_2 = \frac{2\left(2v - t + 2\sqrt{(v-2t)(v-t)}\right)}{8v - 7t},$$

where $0 < \lambda_1 < \lambda_2 < 1$. Deriving the price equilibrium when Firm A leads is intricate, and the technical details are provided in Appendix A.4. The steps involved in demonstrating the results are, however, the following.

First, we identify the expressions of Firm B's profit function for any possible level of the posted prices, p_A and p_B. There are a number of cases to be considered as, for a given leader price p_A, the choice of p_B could result in the follower to face demand: (i) from no segment of consumers; (ii) from part of the anonymous segment; (iii) from part of the anonymous and the list segment; (iv) from all of the anonymous segment; (v) from all of the anonymous segment and part of the list segment.

Second, we derive Firm B's best response function. Intuitively, being the follower Firm B can always undercut Firm A in the anonymous segment. Hence, if Firm A posts a relatively low price, Firm B's best response also involves a low price, in which firms split both segments of the market. If, instead, Firm A posts a relatively high price, Firm B can undercut in the anonymous segment and still post a fairly high price compared to the Hotelling model. In this case, either the two firms split both segments of the market or Firm B may be able to serve the anonymous segment on its own.

Third, we can finally state Firm A's profit function taking into account Firm B's best responses. The optimal choice of the leader depends on the share of consumers on the list, λ. In particular, the thresholds identified at the beginning of this subsection, λ_1 and λ_2, result from comparing the profits of the leader in the identified candidate price equilibria. In fact, at the first stage of this subgame, Firm A picks the price that maximises its profit, anticipating Firm B's best responses.

Following the above steps, we can state these results.

Proposition 3. *Consider the subgame following (τ_0, τ_1) where Firm A leads in the price competition. The subgame perfect Nash equilibrium in this subgame is as follows.*

(i) If $0 < \lambda < \lambda_1$, then the equilibrium prices are

$$p_A = \frac{3-\lambda}{1-\lambda}\frac{t}{2}, \;\; p_B = \frac{5-3\lambda}{1-\lambda}\frac{t}{4},$$

with respective profits

$$\pi_A = \frac{(9-7\lambda)(1+\lambda)}{16(1-\lambda)}t, \;\; \pi_B = \frac{(5-3\lambda)^2}{32(1-\lambda)}t.$$

(ii) If $\lambda_1 \leq \lambda \leq \lambda_2$, then

$$p_A = v, \;\; p_B = v - t,$$

with respective profits

$$\pi_A = \lambda(v-t), \;\; \pi_B = (1-\lambda)(v-t).$$

(iii) If $\lambda_2 < \lambda < 1$, then

$$p_A \in \left[\frac{2+\lambda}{2\lambda}t, v\right], \;\; p_B = \frac{2-\lambda}{2\lambda}t,$$

with respective profits

$$\pi_A = \frac{(2+\lambda)^2}{16\lambda}t, \;\; \pi_B = \frac{(2-\lambda)^2}{8\lambda}t.$$

Proof: See Appendix A.4

The results in Proposition 3 are important for the rest of the article. A key feature is that the equilibrium price and profits are *non-monotonic* in the share of consumers profiled in the list, λ.

To see this, we consider each case in more detail. First, the list is relatively short ($0 < \lambda < \lambda_1$). In this case, Firm B's best response is

$$p_B(p_A) = \begin{cases} \frac{(1-\lambda)p_A+t}{2} & \text{if } 0 < p_A \leq \frac{t}{\sqrt{1-\lambda}} \\ \frac{p_A+t}{2} & \text{if } \frac{t}{\sqrt{1-\lambda}} < p_A \leq 3t \\ p_A - t & \text{if } 3t < p_A \leq v \end{cases} . \tag{3}$$

As explained above, Firm B always has an incentive to undercut in the anonymous segment, but (3) shows that the amount and the implications depend on the leader's price. Indeed, by anticipating this, the informed firm sets a uniform price $p_A = \frac{3-\lambda}{1-\lambda}\frac{t}{2}$. This is a relatively low price, and it enables Firm A to attract consumers in the anonymous segment. At the same time, setting an even lower price, $p_A \leq \frac{t}{\sqrt{1-\lambda}}$, would lead Firm B to best respond according to the first line of (3). However, this would hurt profits, and hence it cannot be an equilibrium. A further alternative for Firm A would be to set a very high price, *i.e.*, $p_A > 3t$, in the third line of (3). However, this would imply letting Firm B serve the whole anonymous segment. As the latter segment is relatively small, such a high price would not be sufficient to maximise profits of Firm B. As a result, in this first case, the equilibrium is characterised by relatively low posted prices, $p_A = \frac{3-\lambda}{1-\lambda}\frac{t}{2}$ and $p_B = \frac{5-3\lambda}{1-\lambda}\frac{t}{4}$, respectively.

Second, consider an intermediate size of the list ($\lambda_1 \leq \lambda \leq \lambda_2$). Unlike the above case, Firm A finds now profitable to set a very high price ($p_A > 3t$) and, indeed, pushes it to its maximum, *i.e.*, the consumers' reservation value v. Whereas this implies giving up the anonymous segment, the relatively large share of consumers in the list makes it worthwhile. In fact, as Firm B's best response is to charge $p_B = v - t$, Firm A is guaranteed to serve all the list segment. As the price posted by Firm B is also high, Firm A's personalised offers are rather effective in extracting surplus from consumers. The outcome is profitable for both firms, with profits being $\pi_A = \lambda(v - t)$ and $\pi_B = (1 - \lambda)(v - t)$, respectively.

Hence, for an intermediate length of the list, exclusive data enable the price leader to achieve a manipulative yet profitable outcome. The informed firm, in fact, chooses such a high price that no consumer ever pays. However, this strategy encourages the follower to undercut just enough and secure the whole anonymous segment. The informed firm can then "cash in" through personalised offers that attract all the consumers on its list. Indeed, the mechanism generates an endogenous market segmentation.[14] In this case, the classical "best response asymmetry" of price competition is pushed to the limit: the breaking point is reached and each firm serves a fully separated segment.

Finally, a large share of consumers are on the list ($\lambda_2 \leq \lambda \leq 1$). As the list size increases beyond λ_2, the manipulative mechanism above cannot be sustained. The reason lies in the fact that, as noted above, Firm B makes profit only on the anonymous segment, which becomes proportionally smaller as a higher percentage of consumers are profiled. Therefore, Firm B's best response is no longer the high price, $p_B = v - t$, but a lower price. More in detail,

$$p_B(p_A) = \begin{cases} \frac{(1-\lambda)p_A + t}{2} & \text{if } 0 < p_A < \frac{2+\lambda}{2\lambda}t \\ \frac{2-\lambda}{2\lambda}t & \text{if } \frac{2+\lambda}{2\lambda}t \leq p_A \leq v \end{cases}.$$

As a result, by setting $p_B = \frac{2-\lambda}{2\lambda}t$ and suppose that $p_B < p_A - t$, Firm B can serve all the anonymous segment but also a fraction of the listmarket. The combined profit is larger than $(1 - \lambda)(v - t)$. If that is the case, Firm A is indifferent between any price above $\frac{2+\lambda}{2\lambda}t$, i.e., all the equilibrium prices are consistent with $p_B < p_A - t$.

Moreover, it can be shown that Firm A does not have an incentive to decrease its price too much. When $p_A \leq t$, Firm B's best response is $p_B(p_A) = \frac{(1-\lambda)p_A+t}{2}$. The resulting candidate equilibrium features very low prices, but the wide list does not justify for Firm A competing too fiercely in posted prices to attract the relatively small share of consumers in the anonymous segment.

[14] This market segmentation is reminiscent of the telecom regulation literature on universal service obligations, see *e.g.*, Valletti et al. (2002) and Gautier and Wauthy (2010).

Figure 2: Normal Form Representation of the Game

(i) Short list ($0 < \lambda \leq \lambda_1$)

		Firm B	
		τ_0	τ_1
Firm A	τ_0	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$	$\frac{(9-7\lambda)(1+\lambda)}{16(1-\lambda)}t$, $\frac{(3\lambda-5)^2}{32(1-\lambda)}t$
	τ_1	$\frac{(23\lambda+25)}{32}t$, $\frac{9(1-\lambda)}{16}t$	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$

(ii) Relatively short list ($\lambda_1 < \lambda \leq 3/5$)

		Firm B	
		τ_0	τ_1
Firm A	τ_0	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$	$\lambda(v-t), (1-\lambda)(v-t)$
	τ_1	$\frac{(23\lambda+25)}{32}t$, $\frac{9(1-\lambda)}{16}t$	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$

(iii) Relatively long list ($3/5 < \lambda \leq \lambda_2$)

		Firm B	
		τ_0	τ_1
Firm A	τ_0	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$	$\lambda(v-t), (1-\lambda)(v-t)$
	τ_1	$\frac{(5+\lambda)^2}{32(1+\lambda)}t$, $\frac{(\lambda-3)^2}{16(1+\lambda)}t$	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$

(iv) Long list ($\lambda_2 < \lambda < 1$)

		Firm B	
		τ_0	τ_1
Firm A	τ_0	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$	$\frac{(\lambda+2)^2}{16\lambda}t$, $\frac{(2-\lambda)^2}{8\lambda}t$
	τ_1	$\frac{(5+\lambda)^2}{32(1+\lambda)}t$, $\frac{(\lambda-3)^2}{16(1+\lambda)}t$	$\frac{9(1+\lambda)}{2(3+\lambda)^2}t$, $\frac{(3-\lambda)^2}{2(3+\lambda)^2}t$

The matrix represents the payoffs of Firm A and B when leading, following, or acting simultaneously for different lengths of the list.

5 Incentive to lead and welfare implications

In this section we study the incentive to lead or to follow in the whole game, and how this crucially relates to the share of consumers on the list. We distinguish four cases using critical values identified in Propositions 2 and 3. For each of them, in Figure 2, we provide the normal form representation of the game, and hence, summarise the strategic situation faced by the firms.

A well understood and recurring result in the literature on price competition with endogenous timing is that a firm normally prefers to be the follower rather than the leader (d'Aspremont and Gerard-Varet, 1980; Hamilton and Slutsky, 1990, *inter alios*). This holds true also in our setting for the uninformed firm: indeed, the following lemma confirms that Firm B's profits are always higher when it follows than when it leads.

Lemma 2. *For all $0 < \lambda < 1$, Firm B earns more profits in (τ_0, τ_1) than in (τ_1, τ_0).*

Proof: See Appendix A.5.

The question at this point, however, is to uncover Firm A's incentives to lead. Focusing on the normal form game in Figure 2 above, further payoff comparisons allow us to state the following results.

Proposition 4. *(i) For any given length of the list, the game has two, and only two, pure strategy Nash equilibria in the first stage: (τ_0, τ_1) and (τ_1, τ_0).*

(ii) If the list is sufficiently long, $\lambda > \hat{\lambda} = \frac{25t}{32v-55t}$,

(a) Firm A's profit is higher as a price leader than as a price follower, and

(b) the equilibrium (τ_0, τ_1) is payoff dominant.

Proof: See Appendix A.6.

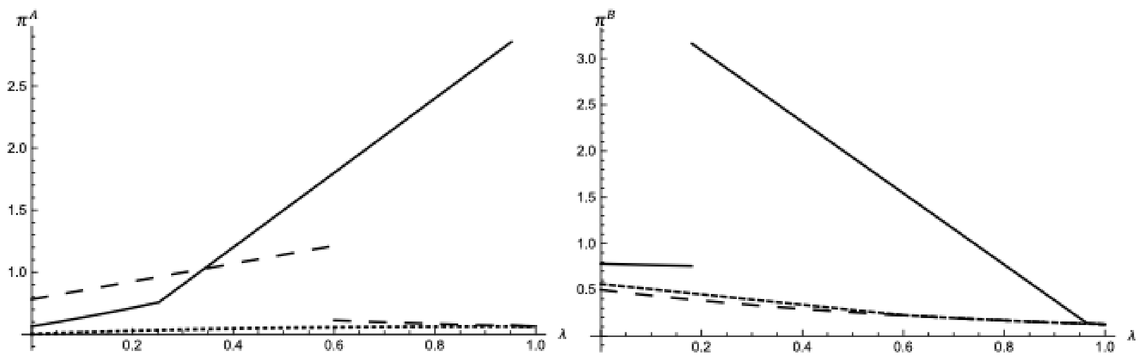
Proposition 4 contains the main findings of the article. First, and consistently with the received literature on price competition with endogenous timing, the normal form game in Figure 2 always has two pure strategy equilibria, (τ_0, τ_1) and (τ_1, τ_0). No matter what is the share of the consumers on the list of Firm A, one of the firm leads and the other follows. This standard outcome extends to the context of price discrimination enabled by consumer level information.

Second, if the share of consumers on the list is relatively small ($0 < \lambda \leq \hat{\lambda}$), the usual logic of price competition applies: both the firms would be better off as followers rather than leaders. Firms clearly face a coordination problem and it is hard to predict which firm ends up leading. Relying on a refinement like payoff/Pareto dominance (Harsanyi and Selten, 1988) does not lead to a conclusion in this case.

As the list becomes more comprehensive, however, incentives and strategies do change. By part (ii) of Proposition 4, a clear payoff ranking between the outcomes can be found as long as the share of consumers profiled exceeds the threshold $\hat{\lambda} = {}^{25t}/_{32v-55t}$. In that case, access to consumer information enables a novel mechanism, operating through the linkage between the prices. The firm owning the list benefits from leading, as it can induce the rival firm to raise its price, which in turn can make price discrimination in the list segment more profitable.

More importantly, if the list consists of an intermediate share of consumers, $\hat{\lambda} < \lambda \leq \lambda_2$, the payoff dominant outcome is characterised by high prices and profits. In this interval, as discussed in section 4.3, the firms "specialise" on different segments of the

Figure 3: Length of the list and firms' profits ($v = 4$, $t = 1$)



The figure presents the profits of Firm A (left) and Firm B (right) in the different subgames, as a function of the length of the list λ. The solid line identifies the case in which Firm A leads the game, whereas the dashed line the case in which it follows. The dotted line identifies a simultaneous price competition.

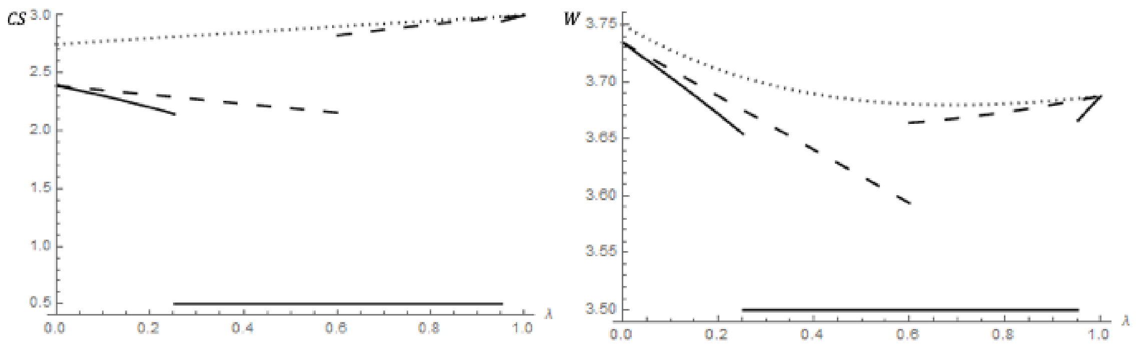
market. This result is innovative and contributes to the literature. The access to exclusive consumer information and the consequent price discrimination does give the informed firm an incentive to lead in price competition, provided that only an intermediate share of consumers is profiled. Unlike Van Damme and Hurkens (2004) and Amir and Stepanova (2006) endogenous price leadership stems from an immaterial asset, rather than productive efficiency.

Interestingly, the highlighted mechanism works as a credible commitment of Firm A towards Firm B: by doing so, Firm A anticipates and allows Firm B's undercutting, but starting from a high initial price. This way, Firm A manipulates B's strategies and reaps high profits from consumers in the list segment. Crucially, this is also in the interest of Firm B and constitutes a semi-collusive device in which both firms charge high prices.

Last but not least, if $\lambda_2 < \lambda < 1$, the list segment becomes too appealing and the uninformed firm competes to attract consumers there. As a result, it reduces its price to such a level that cannot be matched by the rival's offers for some consumers in the list. Despite the drop in profits, Firm A's incentive to lead remains.

Figure 3 illustrates the above findings through an example, where $v = 4$ and $t = 1$. Firm A's profit (left panel) when leading the price competition is depicted as a solid line. The dashed line represents Firm A's profit when it follows. The dotted line is Firm A's profit in simultaneous price competition. As it can be seen in the graph, A's profit in simultaneous price competition is always below that in a sequential price competition. Moreover, A's profit as a leader is higher than that as a follower when $\lambda > \hat{\lambda} = 0.342$. In contrast, by the same logic as in a traditional symmetric sequential price game, Firm B's profit (right panel) when moving as a follower (solid line) is always higher than that in any of the other two scenarios.

Figure 4: Consumers surplus and social welfare ($v = 4$, $t = 1$)



The figure presents consumer surplus (left) and social welfare (right) in the different subgames, as a function of the list size λ. The solid line identifies the case in which Firm A leads the game, whereas the dashed line the case in which it follows. The dotted line identifies a simultaneous price competition.

5.1 Welfare implications

The above findings have important implications not only for firm profits but also for consumer surplus and social welfare. We define the latter as the sum of the industry profit and consumer surplus. The next proposition summarises the main result.

Proposition 5. *Both consumer surplus and social welfare are lower when the informed firm leads.*

Proof: See Appendix A.7

Proposition 5 indicates that an informed price leader would be socially inefficient relative to when firms set prices simultaneously or the uninformed firm leads. Interestingly, the fall in social welfare is driven by the very high transport costs imposed to consumers. In fact, for an intermediate length of the list, each firm exclusively serves one of the segments of the market. As a result, there is substantial mismatch between consumer taste and the product they buy, e.g., consumers with strong preferences for Firm B would buy the product from Firm A if the latter firm profiles them and makes a personalised offer. This results in an inefficient product allocation. Such mismatch is much less pronounced when the uninformed firm leads or when firms compete simultaneously.

Figure 4 graphically summarises the above discussion and provides further insights. The left panel presents the consumer surplus, whereas the right panel presents the social welfare. The solid line indicates the scenario in which Firm A leads the game, the dashed line the case in which it follows, and finally the dotted line the case of a simultaneous price competition.

First, we note that the regime that maximises consumer surplus and social welfare is that of simultaneous competition which, however, never arises as a Nash equilibrium. This reason for this welfare observation is that price discrimination intensifies market competition and simultaneity does not give rise to perfect specialisation of market segments.

Second, the difference between Firm A leading or following is minimised at the extremes, that is, when a very small or a very large percentage of consumers are profiled. When the size of the list is very small, there is no payoff dominant Nash equilibrium. In this case, it would be socially optimal to have the uninformed firm to take the leader's role. On the contrary, for a very comprehensive list, Firm A leads and this is inefficient. However, competition intensifies and the difference between the regimes narrows for both consumer surplus and social welfare. We further note that consumer surplus is maximal when (almost) everyone is profiled as this entails a very large pro-competitive effect.

Last but not least, for an intermediate size of the list, that is, when firms coordinate on high prices, consumer surplus reaches its minimum. Such a result is driven both by high mismatch costs (*i.e.*, transport costs) and the very high prices set by firms in the different segments of the market. Clearly, the former effect is the main driver of the sharp reduction in social welfare, which also reaches its minimum.

6 Concluding remarks

Competition policy authorities and scholars from different fields (*e.g.*, economics, law, computer science) have been increasingly paying attention to the development of data analytics and the large amount of data harvesting. Whether the latter should be received as good or bad news for consumers is controversial. We contribute to the ongoing discussion by studying how access to exclusive data, jointly with the possibility to engage in differential pricing, impacts on market competition and firms' strategies.

Specifically, we look at the incentives of competing firms to lead or follow in price competition. If a sufficient share of consumers is identified, the ability to price discriminate becomes an important strategic asset, and the informed firm has an incentive to lead in the price game. Moreover, the firm with access to data can even manipulate the follower's pricing decision and induce a quite high price in the anonymous segment, which the informed firm can then match in the list segment. To the best of our knowledge, we are the first to show that such a non-cooperative "coordination" on very high prices can take place through the exclusive access to data. This mechanism breaks down when the share of identified consumers is very small or very large, as competitive forces kick back in.

As a matter of fact, this article has implications of interest for policy-makers as well as antitrust scholars and practitioners. We point out that that collecting and using data is not, *per se*, a source of competitive advantage which may entail an upward pressure on prices. Rather, we show that exclusive data availability may entail pro- or anti-competitive conducts depending on the share of consumers profiled. This is because the effect on consumer welfare is U-shaped as prices are non-monotonic in the size of the list. However, there are potentially anti-competitive effects which should ring multiple alarm bells, particularly when the share of profiled consumer is intermediate. In this respect, a possible solution for regulators would be to impose case-by-case mandatory data sharing obligations. For instance, when facing a data holder with a dominant position in a relevant market, a regulator might be willing to enforce more "symmetry" in the access to data, possibly avoiding market coordination and supra-competitive prices. By doing so, the regulator would restore a situation in which price discrimination has a pro-competitive effect.

Our results suggest that data collection may have implications that go beyond privacy as it can alter the firms' strategic incentives. In this context, regulatory interventions that limit firms' data gathering, such as the EU GDPR, the ePrivacy Regulation or the Californian CCPA, may have unintended consequences. For instance, a firm with potential exclusive access to a very large dataset may be constrained by regulation to reduce the number of consumers profiled. This may eventually push the market from a competitive scenario to a semi-collusive one. On the other hand, it might also be the case that a strict regulation drives the market from a semi-collusive outcome to a situation similar to that of price competition with limited or no information. Hence, the competitive consequences of privacy interventions need to be carefully pondered.

Moreover, our analysis shows that large scale data gathering can have nuanced collusive implications. Currently, a topical debate concerns the role of artificial intelligence in feeding collusive outcomes between firms (Calvano et al., 2018, 2019; Klein, 2019; Miklós-Thal and Tucker, 2019) as algorithms may learn to coordinate their pricing strategies. Our conclusion, instead, suggests that tacit coordination may also arise as a result of firms' asymmetric ability to price discriminate. This is particularly worrying as such conduct might not be prosecuted in court. In fact, as pointed out by Harrington (2018), supra-competitive prices, as the ones in our model, are not the result of an overt act of communication and, hence, may slip under the radar.

Further, Heidhues and Kőszegi (2017) show that when firms can exploit consumers' unexpected mistakes, contrary to perfect price discrimination, "naïveté based" discrimination can decrease welfare. Similarly, our article shows that, under certain circumstances, even standard, perfect price discrimination can be highly damaging for consumers and overall welfare. In fact, the usual pro-competitive and welfare enhancing effect of personalised pricing only applies for very large values of the share of profiled consumers.

This article also provides several insights from a managerial perspective. First, the possibility to (exclusively) price discriminate a sufficiently large number of consumers can indeed make a difference and changes the incentive to lead in the market. That is, unlike the received literature, identical firms in terms of fundamentals may want to behave as market leaders. This can happen provided that one firm holds a sufficiently large information advantage. As noted in the Introduction, this is supported by evidence from online and offline markets. Consistent with our findings, firms such as Tesco and Amazon are the informationally strong players in their respective markets and, as such, they can potentially price manipulate their rivals.

Second, from the perspective of the informed firm, there is a clear advantage in profiling the right share of consumers. This share, according to our results, corresponds to the high end of the intermediate list size. Collecting too much data, on the other hand, would intensify market competition with dreadful consequences for profitability. Moreover, from the same perspective, a firm with a very large data endowment may wish to use only its optimal level and throw away the remaining information. For example, a player such as Amazon might want to "cap" the participants to its price discriminating loyalty program (*i.e.*, Prime). However, if too many customers have already been registered and profiled, it might get hard for Amazon to convince its rivals that information use will be capped.

We studied the impact of exclusive access to information in the topical context of competition between retailing firms. The price manipulation mechanism enabled by information and endogenous leadership, however, may apply and provide insights in other settings where agents may have access to exclusive information. For instance, in a directed search labour market (Wright et al., 2017), employers can be highly asymmetric in their information about relevant candidates' characteristics (*e.g.*, specialisation, skills, work location preferences, housing, alternative job offers). Further, a similar mechanism may also apply in financial markets (Pagano and Jappelli, 1993), where some big players feature a competitive advantage over their rivals.

A Appendix

A.1 Proof of Proposition 1

If $p_B \leq t$, given the profit functions, the candidate equilibrium is identified by the following first order conditions:

$$\begin{aligned}
\frac{\partial \pi_A}{\partial p_A} &= (1-\lambda)\left(\frac{1}{2} + \frac{p_B - 2p_A}{2t}\right) = 0, and \\
\frac{\partial \pi_B}{\partial p_B} &= (1-\lambda)\left(\frac{1}{2} + \frac{p_A - 2p_B}{2t}\right) + \lambda\frac{(t-2p_B)}{2t} = 0.
\end{aligned}$$

By solving the system of equations, the equilibrium prices in the anonymous segment are:

$$p_A = \frac{3}{\lambda+3}t, \;\; p_B = \frac{3-\lambda}{\lambda+3}t,$$

whereas the personalised price is simply found by plugging the above p_B into equation (1). The second order conditions also hold. As a result, firms' profits are respectively,

$$\pi_A = \frac{9(\lambda+1)}{2(\lambda+3)^2}t, \;\; \pi_B = \frac{(3-\lambda)^2}{2(\lambda+3)^2}t.$$

If $p_B \geq t$, instead, Firm A by setting (1) attracts all the consumers in the list segment, i.e., $\bar{x} \geq 1$. Hence,

$$\widetilde{\pi}_A = \int_0^1 \widetilde{p}_A(x)dx = p_B \tag{4}$$

and $\widetilde{\pi}_B = 0$. In this case, firms' respective total profits are,

$$\begin{aligned}
\pi_A &= \lambda p_B + (1-\lambda)p_A\left(\frac{1}{2} + \frac{p_B - p_A}{2t}\right), \\
\pi_B &= (1-\lambda)p_B\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right).
\end{aligned}$$

As the firms' best responses are entirely determined in the anonymous segment, i.e., they are equivalent to the standard Hotelling model, the only candidate equilibrium is $(p_A, p_B) = (t, t)$. However, Firm B can profitably deviate by lowering its price from t and attracting additional consumers. This can be seen as there is always a $\delta > 0$ such that:

$$\begin{aligned}
\pi_B(t, t-\delta) &= (1-\lambda)(t-\delta)\left(\frac{1}{2} + \frac{\delta}{2t}\right) + \lambda\frac{\delta(t-\delta)}{2t} \\
&= (1-\lambda)\frac{t-\delta}{2} + \frac{\delta(t-\delta)}{2t} \\
&> (1-\lambda)\frac{t}{2} = \pi_B(t,t).
\end{aligned}$$

Thus, an equilibrium does not exist if $p_B \geq t$. *Q.E.D.*

A.2 Proof of Lemma 1

Let Firm B's price p_B be given. To establish the results, we consider four cases: (i) Firm B's price is very high ($p_B \geq v$), (ii) Firm B's price is relatively high ($v - t \leq p_B < v$), (iii) Firm B's price is relatively low ($t \leq p_B < v - t$) and, finally, (iv) Firm B's price is low ($p_B < t$).

(i) Let $p_B \geq v$. In this case, Firm A can serve the list segment by charging the price schedule: $\widetilde{p}_A(x) = v - tx$. The associated profits obtained from consumers in the list segment are $\widetilde{\pi}_A = v - \frac{t}{2}$. In the anonymous segment, p_B is also so high that Firm A can serve it all by setting $p_A = v - t$.

(ii) Let $v - t \leq p_B < v$. The price of Firm B is still so high that Firm A can match offers on the list segment and serve it all. In particular, there is a threshold location $\widetilde{x}$ such that to the left of it, Firm A is a local monopolist and the outside option is not buying, while to the right of $\widetilde{x}$, Firm B is a viable option and Firm A has to match it. The threshold is implicitly defined by $v - p_B - t(1 - \widetilde{x}) = 0$, implying

$$\widetilde{x} = 1 - \frac{v - p_B}{t}.$$

It then follows that Firm A can attract all the list segment by adopting the following price schedule

$$\widetilde{p}_A(x) = \begin{cases} v - tx & \text{if } 0 < x \leq \widetilde{x} \\ p_B + t(1 - 2x) & \text{if } \widetilde{x} < x \leq 1 \end{cases}.$$

As neither Firm A's price schedule nor its demand in the list segment depends on its own price in the anonymous segment, Firm A's best response in the anonymous segment is unaffected by the presence of the list segment. Moreover, as $v > 4t$, $p_B \geq v - t > 3t$ and hence, Firm A only has to price at $p_B - t$ to fully capture the anonymous segment. To see this, note that the standard best response $t/2 + p_B/2 < p_B - t$ which is unnecessarily low.

(iii) Let $t \leq p_B < v - t$. Such prices are sufficiently low for B to be a potential option for all consumers on the list segment. At the same time, the price is sufficiently high for Firm A to match it and attract all consumers in that segment. In summary, for any p_B in this range, $\widetilde{p}_A(x) = p_B + t(1 - 2x)$ with associated list segment profit $\widetilde{\pi}_A = p_B$ and $\widetilde{\pi}_B = 0$. By the same reasoning as that in the above case, Firm A's best response price in the anonymous segment is

$$p_A(p_B) = \begin{cases} p_B - t & \text{if } 3t \leq p_B < v - t \\ \frac{t}{2} + \frac{p_B}{2} & \text{if } t < p_B < 3t \end{cases}.$$

(iv) Let $p_B < t$. In this case, excluding the possibility of negative prices, Firm A can attract consumers only up to $\bar{x} = \frac{t+p_B}{2t}$. Firm B can serve the remaining consumers on the list segment. In the anonymous segment the standard Hotelling best response applies,

$$p_A(p_B) = \frac{t}{2} + \frac{p_B}{2}.$$

Firm A's best responses in Lemma 1 then result from the above four cases. *Q.E.D.*

A.3 Proof of Proposition 2

It is easy to verify that if $p_B \geq 3t$, Firm B can not obtain demand from any segment and $\pi_B = 0$. These prices are weakly dominated.

If $t < p_B < 3t$, Firm B receives no demand in the list segment, and hence its profit, given A's best response in Lemma 1, is

$$\pi_B = (1 - \lambda)\, p_B \left(\frac{1}{2} + \frac{p_A(p_B) - p_B}{2t} \right).$$

From the first order conditions, an interior solution is identified and the second order conditions hold. The candidate equilibrium price is then $p_B = {}^{3t}\!/_{2}$ with Firm A's best response in the anonymous segment being $p_A = {}^{5t}\!/_{4}$. Their respective profits are

$$\pi_A = \frac{23\lambda + 25}{32} t, \;\; \pi_B = \frac{9\,(1-\lambda)}{16} t. \tag{5}$$

If $p_B \leq t$, Firm B serves those consumers to whom Firm A is unable to match its offer. Given Firm A's best response in both segments, Firm B's profits are

$$\pi_B = \lambda \frac{p_B\,(t - p_B)}{2t} + (1-\lambda) p_B \left(\frac{1}{2} + \frac{p_A(p_B) - p_B}{2t} \right).$$

The first order condition yields $p_B = \frac{3-\lambda}{1+\lambda} \frac{t}{2}$ if $\lambda > {}^{1}\!/_{3}$. Firm A's best response in the anonymous segment then is $p_A = \frac{t}{4} \frac{\lambda+5}{\lambda+1}$ and Firm B's total profits are $\pi_B = \frac{t}{16} \frac{(\lambda-3)^2}{\lambda+1}$. If, on the other hand, $\lambda < {}^{1}\!/_{3}$, the best Firm B can do under the constraint of $p_B \leq t$ is $p_B = t$. However, Firm B's profit by doing so is $\pi_B = (1-\lambda){}^{t}\!/_{2}$ which is lower than that in (5) when $p_B = {}^{3t}\!/_{2}$ instead.

By comparing Firm B's profits in these cases

$$\frac{9\,(1-\lambda)}{16} t \geq \frac{t}{16} \frac{(\lambda-3)^2}{\lambda+1},$$

it follows that the inequality holds provided that $\lambda \leq {}^{3}\!/_{5}$. *Q.E.D.*

A.4 Proof of Proposition 3

To demonstrate the results, we start from Firm B's profit function for different choices of p_B (sub-section A.4.1). We then derive Firm B's best responses for different lengths of the list and finally firm A's optimal pricing strategy for a short, an intermediate and a long list (sub-sections A.4.2, A.4.3 and A.4.4, respectively).

A.4.1 The follower's profit function

As Firm A is leading, p_A is given when Firm B is called to choose its price. There are three possible cases depending on p_A. In each case, the profits of Firm B are as follows.

If $0 < p_A \leq t$, then

$$\pi_B = \begin{cases} 0 & \text{if } p_B > p_A + t \\ (1-\lambda)p_B\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right) & \text{if } t \leq p_B \leq p_A + t \\ \lambda\frac{p_B(t-p_B)}{2t} + (1-\lambda)p_B\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right) & \text{if } 0 < p_B < t \end{cases}. \tag{6}$$

If $p_B > p_A + t$, the price set by Firm B is too large. Hence, it has no demand. Suppose Firm B sets a price $t \leq p_B \leq p_A + t$, then it can obtain a share $\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right)$ of consumers in the anonymous market, and compete with Firm A. Finally, by further lowering the price below t, Firm B can gain extra demand, $\frac{t-p_B}{2t}$, in the list segment.

If $t < p_A \leq 2t$, then

$$\pi_B = \begin{cases} 0 & \text{if } p_B > p_A + t \\ (1-\lambda)p_B\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right) & \text{if } t \leq p_B \leq p_A + t \\ \lambda\frac{p_B(t-p_B)}{2t} + (1-\lambda)p_B\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right) & \text{if } p_A - t < p_B < t \\ \lambda\frac{p_B(t-p_B)}{2t} + (1-\lambda)p_B & \text{if } p_B \leq p_A - t \end{cases}. \tag{7}$$

The expressions of Firm B's profits in the first three lines are identical to those obtained in equation (6). However, as shown in the fourth line, Firm B receives all the demand of the anonymous segment if its price is below $p_A - t$. The difference between the third and fourth line is due to the fact that Firm B prices very aggressively in response to an intermediate price of its rival.

Finally, if $2t < p_A \leq v$, then

$$\pi_B = \begin{cases} 0 & \text{if } p_B > p_A + t \\ (1-\lambda)p_B\left(\frac{1}{2} + \frac{p_A - p_B}{2t}\right) & \text{if } p_A - t < p_B \leq p_A + t \\ (1-\lambda)p_B & \text{if } t \leq p_B \leq p_A - t \\ \lambda\frac{p_B(t-p_B)}{2t} + (1-\lambda)p_B & \text{if } p_B < t \end{cases}. \tag{8}$$

Note that p_A is very large in this case. The first two lines are similar to the previous cases. In the third line, however, Firm B can obtain all the demand in the anonymous segment by undercutting p_A by t, while still setting a price $p_B > t$. Finally, if setting a price below t, Firm B can also attract consumers on the list segment, besides serving the entire anonymous market.

A.4.2 A relatively short list ($0 < \lambda < \lambda_1$)

Firm B's best response function

If $p_A \leq t$, it follows that Firm B's best response is $p_B(p_A) = \frac{(1-\lambda)p_A + t}{2}$. To see this, consider that the candidate best response in the second line of (3) is the standard Hotelling model's best response, $\frac{p_A+t}{2}$, which results in profits

$$\pi_B(p_A) = \frac{(p_A+t)^2(1-\lambda)}{8t}.$$

The best response in the third line is $p_B(p_A) = \frac{(1-\lambda)p_A+t}{2}$, which result in profits

$$\pi_B(p_A) = \frac{(p_A(1-\lambda)+t)^2}{8t}.$$

Hence, π_B is maximised in the third line of (6) as $\frac{(p_A(1-\lambda)+t)^2}{8t} > \frac{(p_A+t)^2(1-\lambda)}{8t}$ for $p_A \leq t$.

Next, consider $t < p_A \leq 2t$. By the discussion above, which still applies in this case, the best response is $\frac{p_A+t}{2}$ in the second line of (7). However, Firm B may now consider to lower its price to a local critical point $p_B(p_A) = \frac{(1-\lambda)p_A+t}{2}$, which is found by maximising the third line of (7). As in this subcase we must have $p_A - t < p_B < t$, a best response equal to $p_B(p_A) = \frac{(1-\lambda)p_A+t}{2}$ is only feasible for $p_A < \frac{t}{1-\lambda}$. If the latter condition is satisfied, Firm B attracts part of the consumers in the list segment and maximises the third line of (7), obtaining profits $\pi_B(p_A) = \frac{(p_A(1-\lambda)+t)^2}{8t}$. Otherwise, for $p_A > \frac{t}{1-\lambda}$, Firm B can increase profit by raising its price up to the upper limit t to obtain $\pi_B(p_A) = \frac{(p_A+t)^2(1-\lambda)}{8t}$.

Moreover, as $\lambda < \lambda_1$, the first order derivative with respect to p_B of the fourth line of (7) is positive. Hence, Firm B can set its maximum price in this subcase which is $p_B = p_A - t$.

Note, however, that the associated profit, $\pi_B(p_A) = \frac{(p_A-t)(2t-\lambda p_A)}{2t}$, is dominated by that obtained by setting a price equal to $\frac{(1-\lambda)p_A+t}{2}$ (third line of (7)) or higher.

All in all, B's best response when $p_A \in (t, 2t]$ can be found by comparing profits in the second and third lines of (7) and it is equal to

$$p_B(p_A) = \begin{cases} \frac{(1-\lambda)p_A+t}{2} & \text{if } t < p_A \leq \frac{t}{\sqrt{1-\lambda}} \\ \frac{p_A+t}{2} & \text{if } \frac{t}{\sqrt{1-\lambda}} < p_A \leq 2t \end{cases},$$

where the critical value, $\frac{t}{\sqrt{1-\lambda}}$, is the price p_A that exactly equalises Firm A's profits.

Finally, in case $p_A > 2t$, the second line of (8) is maximised through a best response equal to $\frac{p_A+t}{2}$, which entails a profit $\frac{(1-\lambda)(t+p_A)^2}{8t}$. According the third line, instead, Firm B covers entirely the anonymous segment and its profit does not depend on p_A. This leads firm B to charge the highest possible price compatible with such demand configuration, that is, $p_B(p_A) = p_A - t$, and resulting profit $(1-\lambda)(p_A - t)$. Also, note that, in the second line, a price $\frac{p_A+t}{2}$ is available only when $p_A \leq 3t$ as $p_B \geq p_A - t$. Otherwise, if $3t < p_A \leq v$, the third line of (8) applies and determines the best response. Moreover, for similar reasons as before, the scenario in the fourth line is dominated.

Overall, Firm B can set a price p_B according to the third line of (8) if $3t < p_A \leq v$, and to the second if $2t < p_A \leq 3t$. Hence, Firm B's best response if $p_A > 2t$ is obtained by comparing profits in the two subcases and is equal to

$$p_B(p_A) = \begin{cases} \frac{p_A+t}{2} & \text{if } 2t < p_A \leq 3t \\ p_A - t & \text{if } 3t < p_A \leq v \end{cases}.$$

To facilitate the analysis, the expression below summarises Firm B's best response for all $0 < \lambda < \lambda_1$:

$$p_B(p_A) = \begin{cases} \frac{(1-\lambda)p_A+t}{2} & \text{if } 0 < p_A \leq \frac{t}{\sqrt{1-\lambda}} \\ \frac{p_A+t}{2} & \text{if } \frac{t}{\sqrt{1-\lambda}} < p_A \leq 3t \\ p_A - t & \text{if } 3t < p_A \leq v \end{cases}.$$

Firm A's optimal pricing

Consider now the first stage of the game where Firm A acts as a leader and chooses p_A. With Firm B's best response given above, Firm A's profit is as follows.

$$\pi_A = \begin{cases} \lambda \frac{\left(t+\frac{t+(1-\lambda)p_A}{2}\right)^2}{4t} + (1-\lambda)p_A\left(\frac{1}{2}+\frac{\frac{t+(1-\lambda)p_A}{2}-p_A}{2t}\right) & \text{if } 0 < p_A \leq \frac{t}{\sqrt{1-\lambda}} \\ \lambda\frac{p_A+t}{2} + (1-\lambda)p_A\left(\frac{1}{2}+\frac{\left(\frac{t+p_A}{2}\right)-p_A}{2t}\right) & \text{if } \frac{t}{\sqrt{1-\lambda}} \leq p_A \leq 3t \\ \lambda(p_A-t)+(1-\lambda)(0) & \text{if } 3t < p_A \leq v \end{cases}. \tag{9}$$

In the first line of (9), p_A is relatively low and will be met with a low response. Hence, Firm A will only be able to serve a fraction of the list segment, with its profit given in (2). However, Firm A also receives a fraction of the anonymous segment. In the second and third line, since the best response of Firm B is above t, Firm A satisfies the demand of all the list segment consumers. Firm A's profit is hence given by (4). Note that in the second line Firm A also serves some consumers in the anonymous market, but not in the third.

The first order derivative of the expression in the first line of (9) is

$$\frac{1-\lambda}{8t}\left[3t(2+\lambda)-(4+3\lambda+\lambda^2)p_A\right].$$

One verifies that it is positive when evaluated at $\frac{t}{\sqrt{1-\lambda}}$. Thus, the best Firm A can achieve in this subcase would be to set $p_A = \frac{t}{\sqrt{1-\lambda}}$. However, by comparing the first and second line of (9), going above such price only increases profits and it cannot be an equilibrium.

Maximising the second line of (9) gives rise to

$$p_A = \frac{3-\lambda}{1-\lambda}\frac{t}{2}, \tag{10}$$

with Firm A's associated profit of

$$\pi_A = \frac{(1+\lambda)(9-7\lambda)}{1-\lambda}\frac{t}{16}. \tag{11}$$

Similarly, maximising the third line of (9) indicates $p_A = v$ as a candidate equilibrium and Firm A's profit is equal to $\lambda(v-t)$. In this case, Firm A gives up serving the anonymous segment and only serves the list one at the highest price.

By comparing the profits associated to the candidate prices $p_A = v$ and $p_A = \frac{3-\lambda}{1-\lambda}\frac{t}{2}$, we find that Firm A sets (10), if, and only if, the profit (11) is larger than $\lambda(v-t)$. This indeed holds if, and only if, $\lambda < \lambda_1$.

As the equilibrium p_A is given in (10), Firm B charges

$$p_B = \frac{t + \frac{t}{2}\frac{3-\lambda}{1-\lambda}}{2} = \frac{5-3\lambda}{1-\lambda}\frac{t}{4},$$

with profit

$$\pi_B = \frac{(3\lambda - 5)^2}{32\,(1-\lambda)}t.$$

A.4.3 An intermediate list ($\lambda_1 \leq \lambda \leq \lambda_2$)

From the result in sub-section A.4.2, it also follows that when λ is above λ_1, Firm A would prefer to set its price at v, earning a profit of $\lambda(v - t)$. In this case, Firm B's best response is $v - t$ with a profit of $(1 - \lambda)(v - t)$.

As the list segment grows, Firm A has no incentive to change pricing as long as Firm B's best response function does not change, which happens when Firm B finds profitable to compete on the list segment as well. Formally, we need to investigate Firm B's profit when firm A sets a high price, namely, (8). As the list grows, Firm B compares the profit of serving the entire anonymous segment, $(1 - \lambda)(v - t)$, with the best it can achieve by decreasing its price and competing for list consumers. This implies maximising the fourth line of (8). The result is found to be $p_B = \frac{2-\lambda}{2\lambda}t$, with associated profit $\frac{(2-\lambda)^2}{8\lambda}t$. The latter profit is larger than $(1 - \lambda)(v - t)$ if, and only if, $\lambda > \lambda_2$. Note also, when $\lambda > \lambda_2$, $p_B = \frac{2-\lambda}{2\lambda}t < t$, consistent with the fourth line of (8).

Hence, when $\lambda_1 \leq \lambda \leq \lambda_2$, Firm A finds it optimal to set a price equal to $p_A = v$, whereas Firm B charges $p_B = v - t$. Equilibrium profits are, respectively, $\pi_A = \lambda(v - t)$ and $\pi_B = (1 - \lambda)(v - t)$. Firm A only serves the list segment and Firm B only serves the anonymous segment.

A.4.4 A long list ($\lambda_2 < \lambda < 1$)

Firm B's best response function

Finally, we consider $\lambda_2 < \lambda < 1$. As shown in sub-section A.4.3, if Firm A's price is high enough such that by setting $p_B = \frac{2-\lambda}{2\lambda}t$, Firm B can ensure the entire anonymous segment demand, then it is indeed Firm B's best response. This is, if $p_A \geq \frac{2+\lambda}{2\lambda}t$. The reason is that, as $\lambda > \lambda_2$, by setting $p_B = \frac{2-\lambda}{2\lambda}t$ and provided that $p_A > p_B + t$, Firm B obtains the whole anonymous segment and a fraction of the list market, with a combined profit larger than $(1 - \lambda)(v - t)$, i.e., the best it can achieve in the anonymous segment alone.

On the other hand, when $p_A \leq t$, Firm B's best response is $p_B(p_A) = \frac{(1-\lambda)p_A+t}{2}$, as before. With λ being relatively large, Firm B's profit when $t < p_A < \frac{2+\lambda}{2\lambda}t$ is as in (7). Note that, as the upper bound $\frac{2-\lambda}{2\lambda}t$ is larger than $p_A - t$, the best Firm B can do in the fourth line is to set $p_A - t$, with a profit of $\frac{(p_A-t)(2t-\lambda p_A)}{2t}$. In the third line, Firm B can set $p_B(p_A) = \frac{(1-\lambda)p_A+t}{2}$ and obtain $\frac{(t+(1-\lambda)p_A)^2}{8t}$. However, since $p_A < \frac{2+\lambda}{2\lambda}t < \frac{t}{1-\lambda}$, $p_B(p_A) = \frac{(1-\lambda)p_A+t}{2}$ is attainable, and its profit is higher than that of $p_A - t$.

To summarise, when $\lambda_2 < \lambda < 1$, Firm B's best response is

$$p_B(p_A) = \begin{cases} \frac{(1-\lambda)p_A+t}{2} & \text{if } 0 < p_A < \frac{2+\lambda}{2\lambda}t \\ \frac{2-\lambda}{2\lambda}t & \text{if } \frac{2+\lambda}{2\lambda}t \leq p_A \leq v \end{cases}. \tag{12}$$

Firm A's optimal pricing

Given Firm B's best response in (12), Firm A's profit is as follows.

$$\pi_A = \begin{cases} \lambda\frac{\left(t+\frac{t+(1-\lambda)p_A}{2}\right)^2}{4t} + (1-\lambda)p_A\left(\frac{1}{2} + \frac{\frac{t+(1-\lambda)p_A}{2}-p_A}{2t}\right) & \text{if } 0 < p_A < \frac{2+\lambda}{2\lambda}t \\ \lambda\frac{\left(t+\frac{2-\lambda}{2\lambda}t\right)^2}{4t} & \text{if } \frac{2+\lambda}{2\lambda}t \leq p_A \leq v \end{cases}, \tag{13}$$

In the first line, Firm A partially serves the anonymous and the list segment of the market. In the second line, by setting a sufficiently large price, Firm A only serves the list segment.

Maximising the first line of (13) gives the following candidate equilibrium price:

$$p_A = \frac{3(2+\lambda)t}{\lambda^2+3\lambda+4} < \frac{2+\lambda}{2\lambda}t,$$

with associated profit of

$$\pi_A = \frac{9}{4}\frac{(1+\lambda)t}{\lambda^2+3\lambda+4}. \tag{14}$$

On the other hand, by setting a price above $\frac{2+\lambda}{2\lambda}t$, Firm A can obtain

$$\pi_A = \frac{(\lambda+2)^2}{16\lambda}t,$$

which is larger than that in (14). Thus, Firm A is indifferent between all prices above $\frac{2+\lambda}{2\lambda}t$ and hence any of them can be a part of subgame perfect equilibrium. *Q.E.D.*

A.5 Proof of Lemma 2

We consider $\Delta_B := \pi_B(\tau_0, \tau_1) - \pi_B(\tau_1, \tau_0)$ in the four different cases identified in Figure 2 depending on the length of the list.

(i) Let $\lambda \in (0, \lambda_1]$. In this case, the difference Δ_B is

$$\left(\frac{(3\lambda-5)^2}{32(1-\lambda)} - \frac{9(1-\lambda)}{16}\right)t = \frac{t}{32}\frac{8-(1-3\lambda)^2}{1-\lambda} > 0.$$

(ii) Let $\lambda \in (\lambda_1, 3/5]$. Δ_B is

$$(1-\lambda)(v-t) - \frac{9(1-\lambda)}{16}t = (1-\lambda)\left(v - \frac{25}{16}t\right) > 0.$$

(iii) Let $\lambda \in (3/5, \lambda_2]$. The difference Δ_B reads $(1-\lambda)(v-t) - \frac{(\lambda-3)^2}{16(1+\lambda)}$ which is strictly positive if $\lambda < \frac{3t+8\sqrt{2}\sqrt{(2v-3t)(v-t)}}{16v-15t}$. However, $\lambda \leq \lambda_2$ and λ_2 is strictly less than this critical value of $\frac{3t+8\sqrt{2}\sqrt{(2v-3t)(v-t)}}{16v-15t}$. Thus, $\pi_B(\tau_0, \tau_1) - \pi_B(\tau_1, \tau_0) > 0$ in this case.

(iv) Let $\lambda \in (\lambda_2, 1)$. The difference Δ_B is

$$\left(\frac{(2-\lambda)^2}{8\lambda} - \frac{(\lambda-3)^2}{16(1+\lambda)}\right)t = \frac{t}{16}\frac{\lambda^3-9\lambda+8}{\lambda^2+\lambda} > 0.$$

Hence, for all $0 < \lambda < 1$, Firm B earns more profits in (τ_0, τ_1) than in (τ_1, τ_0). *Q.E.D.*

A.6 Proof of Proposition 4

We consider the difference between Firm A's profit when following and leading, $\Delta_A := \pi_A(\tau_1, \tau_0) - \pi_A(\tau_0, \tau_1)$ in the four different cases in turn.

(i) Let $\lambda \in (0, \lambda_1]$. $\Delta_A = \frac{(23\lambda+25)}{32}t - \frac{(9-7\lambda)(1+\lambda)}{16(1-\lambda)}t$ which is larger than 0 if and only if $\lambda < \frac{2\sqrt{2}-1}{3}$. However, as $\lambda_1 < \frac{2\sqrt{2}-1}{3}$, $\Delta_A > 0$.

(ii) Let $\lambda \in (\lambda_1, 3/5]$. Then, $\Delta_A = \frac{(23\lambda+25)}{32}t - \lambda(v-t)$ which is larger than 0 if, and only if, $\lambda < \frac{25t}{32v-55t}$. That is,

$$\Delta_A \begin{cases} > 0 & \text{if } \lambda_1 < \lambda < \frac{25t}{32v-55t} \\ = 0 & \text{if } \lambda = \frac{25t}{32v-55t} \\ < 0 & \text{if } \frac{25t}{32v-55t} < \lambda \leq \frac{3}{5} \end{cases}.$$

(iii) Let $\lambda \in (3/5, \lambda_2]$. In this case, $\Delta_A = \frac{(5+\lambda)^2}{32(1+\lambda)}t - \lambda(v-t)$ which is larger than 0 if, and only if, $\lambda < \frac{21t-16v+8\sqrt{2(2v+3t)(v-t)}}{32v-33t}$. However, this critical value is less than $3/5$ and hence $\Delta_A < 0$.

(iv) Let $\lambda \in (\lambda_2, 1)$. Then $\Delta_A = \frac{(5+\lambda)^2}{32(1+\lambda)}t - \frac{(\lambda+2)^2}{16\lambda}t = \frac{(1-\lambda)(\lambda^2+\lambda-8)t}{32\lambda(1+\lambda)} < 0$.

To summarise, Firm A's profit is strictly larger when leading than following, namely, $\Delta_A < 0$, if, and only if, $\lambda > \frac{25t}{32v-55t}$. *Q.E.D.*

A.7 Proof of Proposition 5

The relevant expressions for each subgame are as follows. We denote aggregate profits as Π, consumer surplus as CS, and social welfare as W.

A.7.1 Simultaneous price competition

First, consider the case with simultaneous competition. Industry profits are

$$\Pi = \pi_A + \pi_B = \frac{(18 + \lambda(3 + \lambda))}{2(3 + \lambda)^2} t.$$

Consumer surplus is defined as follows

$$\begin{aligned} CS =& \lambda\Big\{ \int_0^{\bar{x}} [v - tx - \tilde{p_A(x)})]dx + \int_{\bar{x}}^1 [v - t(1 - x) - p_B)]dx \Big\} + \\ & (1 - \lambda)\Big\{ \int_0^{\frac{p_A - p_B}{2t}} [v - tx - p_A)]dx + \int_{\frac{p_A - p_B}{2t}}^1 [v - t(1 - x) - p_B)]dx \Big\}, \end{aligned}$$

with $\bar{x} = (t+p_B)/2t$. Simplifying and rearranging, consumer surplus is equal to

$$CS = v - \frac{(45 + \lambda(21 - 2\lambda))}{4(3 + \lambda)^2} t.$$

As a result, social welfare is equal to

$$W = v - \frac{(9 + \lambda(15 - 4\lambda))}{4(3 + \lambda)^2} t.$$

For ease of exposition, in the next two sequential cases, we report only the final expressions. The derivations are available upon request to the authors.

A.7.2 Firm B leads

Consider when Firm B leads the game and the critical value of $\lambda = 3/5$ (Proposition 2). Depending on the dimension of the list segment, industry profits are

$$\Pi = \begin{cases} \frac{43+5\lambda}{32} t & \text{if } 0 < \lambda < \frac{3}{5} \\ \frac{43+\lambda(3\lambda-2)}{32(1+\lambda)} t & \text{if } \frac{3}{5} < \lambda \leq 1 \end{cases}.$$

Consumer surplus is equal to

$$CS = \begin{cases} v - \frac{103+25\lambda}{64} t & \text{if } 0 < \lambda < \frac{3}{5} \\ v - \frac{103+\lambda(143+\lambda(9+\lambda))}{64(1+\lambda)^2} t & \text{if } \frac{3}{5} < \lambda \leq 1 \end{cases}.$$

As a result, social welfare is

$$W = \begin{cases} v - \frac{17+15\lambda}{64}t & \text{if } 0 < \lambda < \frac{3}{5} \\ v - \frac{17+\lambda(61-\lambda(5\lambda-7))}{64(1+\lambda)^2}t & \text{if } \frac{3}{5} < \lambda \leq 1 \end{cases}.$$

A.7.3 Firm A leads

Consider when list accessing firm, Firm A, leads the game. Note that, in this case, there are two critical values λ_1 and λ_2 as defined in Section 4.3. Industry profits are

$$\Pi = \begin{cases} \frac{43-\lambda(26+5\lambda)}{32(1-\lambda)}t & \text{if } 0 < \lambda < \lambda_1 \\ v - t & \text{if } \lambda_1 < \lambda < \lambda_2 \\ \frac{1}{16}\left(3\lambda - 4 + \frac{12}{\lambda}\right) & \text{if } \lambda_2 < \lambda < 1 \end{cases},$$

whereas consumer surplus is defined as follows

$$CS = \begin{cases} v - \frac{103-25\lambda(2+\lambda)}{64(1-\lambda)}t & \text{if } 0 < \lambda < \lambda_1 \\ \frac{t}{2} & \text{if } \lambda_1 < \lambda < \lambda_2 \\ v - \frac{t}{\lambda} & \text{if } \lambda_2 < \lambda < 1 \end{cases}.$$

As a result, social welfare is

$$W = \begin{cases} v - \frac{17-(15\lambda-2)\lambda}{64(1-\lambda)}t & \text{if } 0 < \lambda < \lambda_1 \\ v - \frac{t}{2} & \text{if } \lambda_1 < \lambda < \lambda_2 \\ v - \frac{(2-\lambda)(2+3\lambda)}{16\lambda} & \text{if } \lambda_2 < \lambda < 1 \end{cases}.$$

Results follow from a direct comparison: both consumer surplus and social welfare are lower when Firm A leads. *Q.E.D.*

References

Amir, R. and Stepanova, A. (2006). Second-mover advantage and price leadership in Bertrand duopoly. *Games and Economic Behavior*, 55(1):1–20.

Armstrong, M. (2006). Recent developments in the economics of price discrimination. In *Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress*. Cambridge University Press.

Autoridade da Concorrência (2019). Digital ecosystems, big data and algorithms. July, 2019.

Belleflamme, P., Lam, W. M. W., and Vergote, W. (2017). Price discrimination and dispersion under asymmetric profiling of consumers. AMSE Working Paper.

Bergemann, D. and Bonatti, A. (2015). Selling cookies. *American Economic Journal: Microeconomics*, 7(3):259–94.

Bergemann, D. and Bonatti, A. (2019). Markets for information: An introduction. *Annual Review of Economics*, 11:1–23.

Biglaiser, G., Calvano, E., and Crémer, J. (2019). Incumbency advantage and its value. *Journal of Economics & Management Strategy*, 28(1):41–48.

Blattberg, R. and Deighton, J. (1991). Interactive marketing: exploiting the age of addressability. *Sloan Management Review*, 33(1):5–15.

Bounie, D., Dubus, A., and Waelbroek, P. (2018). Selling strategic information in competitive markets. *CESifo Working Paper No. 7078*.

Bourreau, M. and de Streel, A. (2018). The regulation of personalised pricing in the digital era. *OECD. DAF/COMP/WD (2018) 150*.

Bourreau, M., De Streel, A., and Graef, I. (2017). Big data and competition policy: Market power, personalised pricing and advertising. *CERRE Project Report.*

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2018). Artificial intelligence, algorithmic pricing and collusion. *CEPR Discussion Paper No. DP13405*.

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2019). Algorithmic pricing what implications for competition policy? *Review of Industrial Organization*, 55(1):155–171.

Calzolari, G. and Pavan, A. (2006). On the optimality of privacy in sequential contracting. *Journal of Economic Theory*, 130(1):168–204.

Casadesus-Masanell, R. and Hervas-Drane, A. (2015). Competing with privacy. *Management Science*, 61(1):229–246.

Chen, Y. and Iyer, G. (2002). Consumer addressability and customized pricing. *Marketing Science*, 21(2):197–208.

Chen, Z., Choe, C., and Matsushima, N. (2019). Competitive personalized pricing. *Management Science*. forthcoming.

Choe, C., King, S., and Matsushima, N. (2017). Pricing with cookies: Behavior-based price discrimination and spatial competition. *Management Science*, 64(12):5669–5687.

Clavorà Braulin, F. and Valletti, T. (2016). Selling customer information to competing firms. *Economics Letters*, 149:10–14.

CMA (2018). Pricing algorithms research, collusion and personalised pricing. October, 2018.

Comunicaffe (2019). Costa coffee rolls out mobile ordering service across 500+ stores nationwide. *International Comunicaffe*, `https://www.comunicaffe.com/costa-coffee-rolls-out-mobile-ordering-service-across-500-stores-nationwide/`. [Last accessed September 11, 2019].

d'Aspremont, C. and Gerard-Varet, L.-A. (1980). Stackelberg-solvable games and pre-play communication. *Journal of Economic Theory*, 23(2):201–217.

Dastin, J. (2017). Amazon trounces rivals in battle of the shopping 'bots'. *Reuters*, `www.reuters.com/article/us-amazon-com-bots-insight/amazon-trounces-rivals-in-battle-of-the-shopping-bots-idUSKBN1860FK`. [Last accessed September 11, 2019].

Dobson, P. W., Kim, S.-H., and Lan, H. (2016). Identifying price-leadership structures in oligopoly. *Mimeo*.

Dosis, A. and Sand-Zantman, W. (2019). The ownership of data. *Available at SSRN 3420680*.

Esteves, R. B. (2009). A survey on the economics of behaviour-based price discrimination. *NIPE Working Paper*, Universidade do Minho.

Fudenberg, D. and Tirole, J. (2000). Customer poaching and brand switching. *RAND Journal of Economics*, pages 634–657.

Fudenberg, D. and Villas-Boas, J. M. (2006). Behavior-based price discrimination and customer recognition. *Handbook on Economics and Information Systems*, 1:377–436.

Furman, J., Coyle, D., Fletcher, A., Marsden, P., and McAuley, D. (2019). Unlocking digital competition. Report of the Digital Competition Expert Panel. March, 2019.

Gal-Or, E. (1985). First mover and second mover advantages. *International Economic Review*, pages 649–653.

Gautier, A. and Wauthy, X. (2010). Price competition under universal service obligations. *International Journal of Economic Theory*, 6(3):311–326.

Graef, I. (2017). Algorithms and fairness: What role for competition law in targeting price discrimination towards ends consumers. *Columbia Journal of European Law*, 24:541.

Grunes, A. P. and Stucke, M. E. (2015). No mistake about it: The important role of antitrust in the era of big data. *Antitrust Source (Apr. 2015), Online*.

Gu, Y., Madio, L., and Reggiani, C. (2019). Data brokers co-opetition. *CESifo Working Papers No. 7523*.

Hamilton, J. H. and Slutsky, S. M. (1990). Endogenous timing in duopoly games: Stackelberg or Cournot equilibria. *Games and Economic Behavior*, 2(1):29–46.

Hannak, A., Soeller, G., Lazer, D., Mislove, A., and Wilson, C. (2014). Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 Conference on Internet Measurement*, pages 305–318. ACM.

Harrington, J. E. (2018). Developing competition law for collusion by autonomous artificial agents. *Journal of Competition Law & Economics*, 14(3):331–363.

Harsanyi, J. C. and Selten, R. (1988). A general theory of equilibrium selection in games. *MIT Press Books*, 1.

Heidhues, P. and Kőszegi, B. (2017). Naïveté-based discrimination. *Quarterly Journal of Economics*, 132(2):1019–1054.

Hodgson, C. (2019). Fast-food chains to trial licence plate recognition in drive-throughs. *Financial Times*, `https://www.ft.com/content/a07906a8-a1c3-11e9-a282-2df48f366f7d`. [Last accessed September 11, 2019].

Ichihashi, S. (2019). Non-competing data intermediaries. *Mimeo*.

Jehiel, P. and Moldovanu, B. (2000). Auctions with downstream interaction among buyers. *RAND Journal of Economics*, pages 768–791.

Jing, B. (2016). Behavior-based pricing, production efficiency, and quality differentiation. *Management Science*, 63(7):2365–2376.

Klein, T. (2019). Autonomous algorithmic collusion: Q-learning under sequantial pricing. *Amsterdam Law School Research Paper*.

Lambrecht, A. and Tucker, C. E. (2017). Can big data protect a firm from competition? *CPI Antitrust Chronicle*, 1(1).

Madden, P. and Pezzino, M. (2019). Endogenous price leadership with an essential input. *Games and Economic Behavior*.

Mikians, J., Gyarmati, L., Erramilli, V., and Laoutaris, N. (2012). Detecting price and search discrimination on the internet. In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, pages 79–84. ACM.

Mikians, J., Gyarmati, L., Erramilli, V., and Laoutaris, N. (2013). Crowd-assisted search for price discrimination in e-commerce: First results. In *Proceedings of the 9th ACM Conference on Emerging Networking Experiments and Technologies*, pages 1–6. ACM.

Miklós-Thal, J. and Tucker, C. (2019). Collusion by algorithm: Does better demand prediction facilitate coordination between sellers? *Management Science*, 65(4):1552–1561.

Montes, R., Sand-Zantman, W., and Valletti, T. (2019). The Value of Personal Information in Online Markets with Endogenous Privacy. *Management Science*, 65(3):1342–1362.

Pagano, M. and Jappelli, T. (1993). Information sharing in credit markets. *Journal of Finance*, 48(5):1693–1718.

Pazgal, A. and Soberman, D. (2008). Behavior-based discrimination: Is it a winning play, and if so, when? *Marketing Science*, 27(6):977–994.

Pazgal, A., Soberman, D., and Thomadsen, R. (2013). Profit-increasing consumer exit. *Marketing Science*, 32(6):998–1008.

Petro, G. (2017). Amazon's acquisition of whole foods is about two things: Data and product. *Forbes*, `www.forbes.com/sites/gregpetro/2017/08/02/amazons-acquisition-of-whole-foods-is-about-two-things-data-and-product`. [Last accessed September 11, 2019].

Seaton, J. S. and Waterson, M. (2013). Identifying and characterising price leadership in British supermarkets. *International Journal of Industrial Organization*, 31(5):392–403.

Shy, O. and Stenbacka, R. (2016). Customer privacy and competition. *Journal of Economics & Management Strategy*, 25(3):539–562.

Stole, L. A. (2007). Price discrimination and competition. *Handbook of Industrial Organization*, 3:2221–2299.

Stucke, M. E. and Grunes, A. P. (2016). *Big Data and Competition Policy*. Oxford University Press (2016).

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

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

Thisse, J.-F. and Vives, X. (1988). On the strategic choice of spatial price policy. *American Economic Review*, pages 122–137.

Tucker, C. (2019). Digital data, platforms and the usual [antitrust] suspects: Network effects, switching costs, essential facility. *Review of Industrial Organization*, 54(4):683–694.

Tucker, D. S. and Wellford, H. (2014). Big mistakes regarding big data. *Antitrust Source, American Bar Association*.

Tyko, K. (2019). Dunkin' testing a big change to dd perks. Is your location participating? *USA Today*, `https://eu.usatoday.com/story/money/2019/04/16/dunkin-looks-change-dd-perks-loyalty-program/3488078002/`. [Last accessed September 11, 2019].

Valletti, T. M., Hoernig, S., and Barros, P. P. (2002). Universal service and entry: The role of uniform pricing and coverage constraints. *Journal of Regulatory Economics*, 21(2):169–190.

Van Damme, E. and Hurkens, S. (2004). Endogenous price leadership. *Games and Economic Behavior*, 47(2):404–420.

Varian, H. R. (2018). Artificial intelligence, economics, and industrial organization. In *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.

Villas-Boas, J. M. (1999). Dynamic competition with customer recognition. *RAND Journal of Economics*, pages 604–631.

Wright, R., Kircher, P., Julien, B., and Guerrieri, V. (2017). Directed search: A guided tour. *National Bureau of Economic Research*.

Zhang, J. (2011). The perils of behavior-based personalization. *Marketing Science*, 30(1):170–186.