// NO.21-007 | 01/2021

DISCUSSION PAPER

// FRANCESCO CLAVORÀ BRAULIN

The Effects of Personal Information on Competition: Consumer Privacy and Partial Price Discrimination

The Effects of Personal Information on Competition: Consumer Privacy and Partial Price Discrimination*

Francesco Clavorà Braulin †

January 2021

Abstract

This article studies the effects of consumer information on the intensity of competition. In a two dimensional duopoly model of horizontal product differentiation, firms use consumer information to price discriminate. I contrast a full privacy and a no privacy benchmark with intermediate regimes in which the firms target consumers only partially. No privacy is traditionally detrimental to industry profits. Instead, I show that with partial privacy firms are always better-off with price discrimination: the relationship between information and profits is hump-shaped. Consumers prefer either no or full privacy in aggregate. However, even though this implies that privacy protection in digital markets should be either very hard or very easy, the effects of information on individual surplus are ambiguous: there are always winners and losers. When an upstream data seller holds partially informative data, an exclusive allocation arises. Instead, when data is fully informative, each competitor acquires consumer data but on a different dimension.

Keywords: price discrimination, data broker, consumer information, privacy

JEL Codes: D43, L11, L13

*I am grateful to Emilio Calvano for his guidance, Vincenzo Denicolò, Alexandre de Cornière, Paul Belleflamme, Wilfried Sand-Zantman, and Carlo Reggiani for insightful comments, and to the audiences at the 3rd Doctoral Workshop on The Economics of Digitization, the 7th PhD Student Workshop on Industrial and Public Economics, Earie 2019, the 11th Lisbon Meetings in Game Theory and Applications, the BiGSEM Doctoral Workshop, the CESifo Area Conference on Economics of Digitization, and the 2020 ASSET Meeting.

†ZEW Mannheim, e-mail: francesco.clavorabraulin@zew.de

1 Introduction

This article studies settings in which firms know something but not everything about consumer preferences and use this partial knowledge to target prices.

Consumer data collection in digital markets is pervasive and it poses several policy-relevant questions. One of the most intriguing is about the effects of consumer data exploitation on competition. Information about consumers is a crucial asset for many digital businesses and it is highly valuable when it allows firms to change their strategies in a profitable way[1]. For instance, it allows accurate consumer profiling which opens up the possibility of making personalized offers based on user characteristics (Stucke, 2018). Firms can therefore provide customized services and personalized recommendations, deliver more targeted advertising or even personalize prices shown to consumers. In particular, motivated by ubiquitous online data collection, personalized pricing is a topical area of research and consumer privacy is a natural concern when firms can more or less accurately target customers. Clearly, firms need information about preferences to implement sophisticated pricing strategies: they can collect consumer data by themselves or can acquire it from data brokers (Montes, Sand-Zantman, and Valletti, 2018).

In this article I focus on the effects of information exploitation on profits and consumer surplus when competing firms price discriminate but privacy can be partially enforced, showing that targeted prices are not necessarily detrimental to industry profits, as instead suggested by large part of the literature on price discrimination in competition; then, I introduce an upstream data seller and I characterize his incentives to sell consumer data to competing firms. The novelty of the article is in addressing price discrimination and data sales among firms in a two dimensional model of horizontal product differentiation in which only one or both dimensions of consumer private information can be collected and put on sale by a monopolistic data broker. Initially I contrast a full privacy benchmark in which the downstream competing firms set uniform prices to another baseline scenario in which there is no privacy and individual price targeting is feasible. Then, in the main part of the article, I study the effects of a partial enforcement of consumer privacy. The intuition is that firms may know some consumers' char-

[1]See `https://www.mckinsey.com/using-big-data-to-make-better-pricing-decisions`

acteristics but not everything about them: for instance, a regulation may prevent firms from exploiting complete consumer profiles by requesting some form of anonymization to preserve privacy or, simply, firms may not be able to infer perfectly each consumer's willingness to pay. Therefore, in contrast to the prevalent one dimensional literature on consumer privacy, I examine the effects of information on profits and the incentives to sell data in a setting in which the information structure is slightly more complex and also more realistic.

Consider as an example two online outlets selling technological products: the first retailer is specialized in MacBooks whereas the second one in personal computers equipped with a Windows operating system. A data broker tracking the technical characteristics of each user's smartphone is likely to successfully infer the brand preferences of the consumers: a user browsing the web through an iPhone could more likely buy another Apple product and this information is valuable to the competing websites. On the other hand, consumers' preferences for a smaller sized but more portable laptop or for a large screen but heavier product may not be so easily observable[2]. Alternatively, as a minimal working example, suppose that each consumer's willingness to pay is entirely captured by two attributes, age and location: individual demographics are easily found in public records or even self-declared whereas information on location can be tracked via the GPS or the IP address of the user's device but it could be concealed.

Despite the fact that price discrimination could be regarded as a theoretical curiosity, there is evidence, although limited, of price discrimination in digital markets (Hindermann, 2018). Consumer targeting revolves around user-based, technical (operating system) and location-based features. Hannak, Soeller, Lazer, Mislove, and Wilson (2014) find limited evidence of price discrimination in the hotel sector, and Hupperich, Tatang, Wilkop, and Holz (2018) find some evidence also in the rental car sector. Dube and Misra (2017) investigate the empirical implications of price discrimination with high-dimensional data on customer features. They rely on experimental data and consider a digital firm which employs machine learning techniques to target prices, showing that profits always increase with price discrimination whereas consumer surplus is almost unchanged. In a competitive setting in which firms have access

[2]Alternatively, it would be interesting to consider a second vertical dimension of information, which would be a natural extension of this model to a setting à la Neven and Thisse (1987).

to consumers' real-time and historical location data, Dube, Fang, Fong, and Luo (2017) find that profitability of price discrimination crucially depends on the competitor's response: a firm enjoys large profit gains when it successfully targets consumers at the rival's location or when there is a price response in the same direction, whereas such gains are mitigated when prices move in different directions. Experimental evidence suggests therefore that profits may increase when price targeting is possible. Strictly related to the feasibility of these sophisticated pricing strategies, also the role of data collectors and data sellers, generally named data brokers, is far-reaching in the markets for information. Access to consumer data increases the ability of firms to segment customers and reach them with personalized offers, and data sellers may find it optimal to discriminate among data buyers on the basis of their willingness to pay for consumer data (Pancras and Sudhir, 2007). For instance, an incentive to grant exclusive access to valuable consumer data to certain partners while foreclosing other firms emerged clearly in the recent Facebook case[3]. Ultimately, strategic behavior of data collectors and sellers and exclusive data access may harm consumers and the society as a whole (Duch-Brown, Martens, and Mueller-Lang, 2017).

Instead, in the literature on consumer privacy, the prevalent idea is that more information induces more intense price competition and therefore a privacy enforcement by means of a ban on price discrimination is counterproductive: when all firms have detailed data on preferences (i.e. no privacy) they are worse-off with price discrimination at the benefit of all consumers. In this model with two dimensional consumer data, no privacy strengthens competition as well. Not surprisingly, the standard argument of Bertrand competition in transportation costs holds also in two dimensions, even though not all consumers get a discount in equilibrium. The main contribution is instead to show that access to partial information on consumer preferences always increases industry profits: firms are better-off with price discrimination in all games in which privacy is partially enforced. Interestingly, when there is full information in the market but one player is exclusively informed, less surplus is extracted from consumers than under partial privacy. This finding is going to be crucial for the characterization of the optimal selling strategy of the data broker, suggesting that exclusive data access is not always the primary competition policy concern.

[3]See `https://www.nytimes.com/2018/12/05/facebook-documents-uk-parliament.html`

The main mechanism that generates a redistribution of surplus from consumers to the firms through price targeting is the inability of competing firms to observe one of the two dimensions of consumer preferences: the standard argument of Bertrand competition in transportation costs that holds with no privacy breaks down under partial privacy. When only one dimension is known, consumers are ranked by the firms into their strong or weak markets only with respect to that piece of information. However, an unobserved dimension implies that the partially identified consumers are still quite heterogeneous: some of them are indeed close to the rival in the product space, but this attribute remains unobserved. Indeed, in equilibrium each firm has an incentive to price high to all customers close in the observed dimension but, whenever discounts are offered at the other outlet, it serves only those with a high willingness to pay in both dimensions, letting those consumers with a poor match in the unobserved dimension to inefficiently switch to the rival firm. In terms of profits, pricing above the Nash equilibrium uniform price more than offsets the loss of some customers that anyway would have a good fit with the product. This uncertainty is crucial: differently from a one dimensional spatial model with observable locations in which a firm pricing too high would immediately lose all consumers at locations targeted by the rival with suitable discounts, in a two dimensional model with partially observable types each firm always keeps at least some customers. The reason is that in a two dimensional model the rival cannot be as aggressive as it would be necessary to reap all the other firm's loyal customers. The candidate price would be too low. These findings are in contrast with standard one dimensional information acquisition games, in which firms end up in a prisoners' dilemma situation: both firms price discriminate and make lower profits than absent price discrimination. In turn, the standard implication is that when the downstream data allocation is induced by an upstream seller, information is awarded exclusively to preserve the supra-competitive profit of the informed firm and, more importantly, to maximize the difference in profits between the informed firm and the uninformed one. In this model, when partial privacy is enforced, exclusivity holds only when the data broker has information on a single dimension. Instead, if the dataset held by the seller is fully informative, then the data broker has an incentive to sell information non exclusively, awarding different dimensions of consumer data to different competitors. Consumers prefer no privacy in aggregate but, if individual targeting is unfeasible, full privacy should be welcomed in the market: intermediate levels of privacy are always detrimental to consumers. The conclusion is that privacy protection should be made either very hard or very easy. However, I also find that the impact of data exploitation on individual consumer surplus is ambiguous across games: some consumers are made worse-off but others are better-off, even under no privacy. The heterogeneity of the privacy consequences for final consumers does not allow to draw a clear cut policy conclusion but it seems to suggest a more nuanced approach to consumer protection.

Related literature. The economic literature on privacy and price discrimination has formalized a trade-off between a rent extraction effect, which reinforces when one downstream firm is exclusively informed, and a competition effect, which is maximized when information on consumer preferences is symmetrically allocated among firms. In the latter case, all personalized prices lie below the level of the Nash equilibrium uniform price. Firms are therefore worse-off when both have the ability to target consumers or, in other words, when there is no privacy (Taylor and Wagman, 2014)[4]. The negative effect on profits of non exclusive information is one of the main findings of the price discrimination literature (Thisse and Vives, 1988; Armstrong, 2006) under best-response asymmetry (Corts, 1998), and it is the crucial driver of data exclusivity: the data seller has an incentive to promote an allocation that maximizes rent extraction[5]. Indeed, the literature has widely studied the data brokers' incentives to sell information when data buyers can exploit acquired information to make targeted price offers. Montes et al. (2018) show that in a horizontally differentiated duopoly an upstream seller has an incentive to induce maximal asymmetry with respect to the downstream access to consumer data, therefore selling data exclusively to one competitor. Price competition is relaxed and the retailers have a strong incentive to become the exclusive winner of consumer data. As a result, the data seller can extract the highest price for information. Kim, Wagman, and Wickelgren (2018) confirm an exclusivity result even when the downstream market is a triopoly. Clavorà Braulin and Valletti (2016) come to the same conclusion in a vertical differentiation duopoly, formalizing the conditions under which it is the high or the low quality firm to exclusively receive consumer data. From a competition policy point of view, the regulators should be

[4]Belleflamme and Vergote (2016) show that consumers may be better-off with no privacy even under monopoly.

[5]For recent literature reviews see Acquisti, Taylor, and Wagman (2016) and Ganuza and Llobet (2018).

therefore concerned about exclusive transfers of consumer information in the data market, because exclusivity is detrimental to the uninformed firms and final consumers. However, a key feature of this strand of the literature on data brokers and consumer privacy is its focus on settings in which price discrimination is perfect: the dataset can be sold either exclusively or not exclusively, but only entirely, implying that firms can perfectly identify consumers tracked by the seller. When the data broker has the possibility to partition the compiled dataset prior to the sale stage the exclusive equilibrium may not arise. Bounie, Dubus, and Waelbroeck (2018) show that the data broker sells information to both competitors: when the dataset can be optimally partitioned, the seller has an incentive to sell symmetric but not overlapping subsets of consumer data to firms in order to soften downstream competition. In a market for a homogeneous product, Belleflamme, Lam, and Vergote (2020) show that a data broker always has an incentive to allocate information to both downstream firms, but only under vertical data differentiation: the quality of the dataset sold to firms must be different, implying that the duopolists can identify consumers with asymmetric but correlated technologies.

This article is methodologically close to Baye, Reiz, and Sapi (2018) and Liu and Shuai (2013). The first article, which builds on Jentzsch, Sapi, and Suleymanova (2013), proposes a two period model where consumers differ both in their geographical position and flexibility in transportation costs. Locations are perfectly observed by firms but individual disutility costs are unknown. However, first period purchases are imperfectly informative about consumer flexibility. They show that firms can be better-off by combining location information with behavioral data for price discrimination purposes when consumers are moderately heterogeneous in flexibility. Firms categorize customers in the same way with respect to the degree of flexibility, and Baye et al. (2018) investigate the effects on profits of combining additional data on flexibility with perfectly observable data on locations. In this article, instead, the type of available information is different: transportation costs are homogeneous across consumers, but outlets rank them differently with respect to both dimensions of private information. In addition, firms are allowed to hold both similar and different data about consumers. Liu and Shuai (2013) propose a static two dimensional model of horizontal product differentiation, in which information allows firms to segment consumers in two groups along each dimension. They find that, when both firms observe only one and the same dimension of private information, partial price discrimination rises industry profits; instead, when firms have partial information but on different dimensions, firms are worse-off. My article builds on their setup, but in my model consumer data is finer so that a continuum of consumers is identified along each dimension; moreover, I consider also the case of perfect price discrimination or, equivalently, no privacy. In contrast to their findings, I find that firms are always better-off under partial price discrimination, regardless of the type of partial information held by the players. Finally, differently from Baye et al. (2018) and Liu and Shuai (2013), I also investigate and characterize the incentives of an upstream data holder to sell data to competing firms. This article is also conceptually close to Shy and Stenbacka (2015), provided that they investigate the impact of different privacy regimes on market outcomes in a model with switching costs in which price discrimination is not possible.

In what follows I firstly set up the theoretical model and illustrate the two benchmark cases. Then I solve for the relevant games with partial privacy and I investigate the relationship between the type of information held by the firms and industry profits. Finally, I characterize the optimal selling strategy of a monopolistic data broker.

2 The model

I consider an augmented version of the linear city model (Hotelling, 1929): goods are horizontally differentiated but the product space is two dimensional. There are three types of players.

Consumers. There is a unitary mass of heterogeneous consumers with unit demand and gross utility $v > 0$ from consumption. Their type is (x, y) where x and y are orthogonal and uniformly distributed over the unit square $[0, 1] \times [0, 1]$. Consumers incur in a linear transportation cost $t > 0$ when buying at price p_i the product of firm i located in (x_i, y_i) so that a type (x, y) receives a net utility

$$u_i = v - p_i - t|x - x_i| - t|y - y_i| \tag{1}$$

where the two dimensions x and y are qualified by the same degree of differentiation.

Firms. Two competing firms $i = 1, 2$ are exogenously located in $(0, 0)$ and $(1, 1)$, respectively. Firms compete in prices and they either set a uniform price p_i absent information about consumers, or they condition their price schedules on the available information. The marginal cost of production is normalized to zero.

The implication is that types with a low (high) realization of both x and y are in firm 1's strong (weak) market, whereas types with a high (low) realization of x and y are in firm 2's strong (weak) market (Corts, 1998). Efficiency requires that consumers (x, y) with $x + y < 1$ buy from firm 1, whereas those with $x + y > 1$ acquire the product from firm 2. These conditions ensure that transportation costs are minimized.

The timing of the game is the following:

1. Firms simultaneously set prices;

2. Consumers observe the offers designed for them and decide which product to buy.

Data seller. Later on, I introduce a data broker in the game to make information acquisition endogenous. The seller observes either only one (equivalently, x or y) or both dimensions of information and gathers it into the datasets X, Y or XY, respectively. The main assumption is that in the latter case the data broker can sell each dimension jointly or separately but, once data is acquired, the buyer receives the dataset entirely[6]. Importantly, prior to the pricing game, the information allocation among the firms becomes common knowledge. Instead, consumers observe only the personalized prices, if any, specifically designed for them.

Two extreme privacy benchmarks are initially discussed: the full privacy ("*fp*") and the no privacy ("*np*") regimes. Then, in the main part of the article, all types of games with partial privacy are investigated. Finally, with the introduction of the data broker, I will solve for the subgame perfect Nash equilibrium of the information acquisition game. Notice that scenarios in which one firm is exclusively awarded the dataset XY will be instrumental for this characterization and therefore these additional subgames will be solved for in Section 4. The market is fully covered in equilibrium whenever $v \geq 2t$, and this assumption is maintained

[6]Optimal partitioning of the dataset prior to the information sale stage is not considered in this article. Bounie et al. (2018) investigate this option and, building on the work of Liu and Serfes (2004), show that both firms receive information in equilibrium, although not overlapping partitions of the original dataset.

throughout the analysis.

Preview of the results. This article is primarily concerned with the effects of information on profits: a hump-shaped relationship between the discrete amount of data available to the firms and industry profits is depicted in Figure 1.

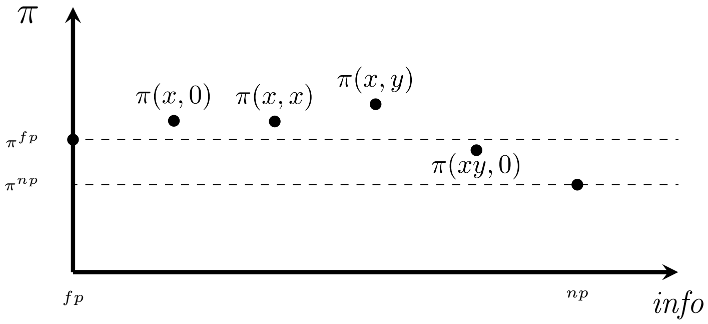

Figure 1: Industry profits.

With partial privacy industry profits lie always above the full privacy level. Instead, whenever there is full exclusivity in the market, profits are slightly eroded. Firms are collectively better-off with intermediate levels of privacy. Additional insights emerge from Figure 2. In any game with partial privacy firms are weakly better-off than under full privacy, regardless of the data they hold. In particular, partial exclusivity does not harm uninformed players, whereas full exclusivity is clearly detrimental to competitors.

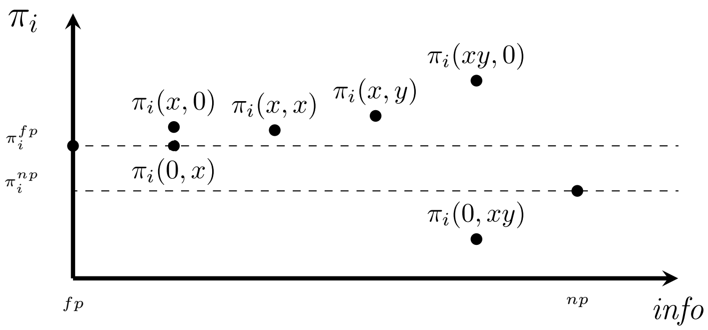

Figure 2: Individual profits.

Full privacy

Suppose that full privacy is enforced and firms compete in uniform prices. At prices p_1 and p_2 there is an indifferent type y for each value of x, which implies that there exists a continuum of marginal consumers defined as

$$v - p_1 - tx - ty = v - p_2 - t(1-x) - t(1-y) \;\; \Rightarrow \;\; \tilde{y}(x) = \frac{p_2 - p_1 + 2t(1-x)}{2t} \qquad (2)$$

Both firms serve a positive fraction of consumers whenever $|p_i - p_j| < 2t$. When the price difference $p_1 - p_2$ is larger (lower) than $2t$ $(-2t)$ then firm 1 faces zero (unitary) demand. Firms' profits are $\pi_1 = p_1 D_1(p_1, p_2)$ and $\pi_2 = p_2 D_2(p_1, p_2)$.

Lemma 1. *When no firm has information, equilibrium prices are* $p_1^{fp} = p_2^{fp} = t$ *and each firm makes a profit equal to* $\pi_i^{fp} = \frac{t}{2}$.

Proof. See Appendix A. ■

Marginal types are located along $y^*(x) = 1 - x$, the bisector of the unit square, and each firm serves half of consumers. All consumers buy from the closest outlet and efficiency is achieved. Industry profits are $\pi^{fp} = t$ whereas consumer surplus writes

$$CS^{fp} = 2\int_0^{1-x}\int_0^1 (v - t - t(x+y))\, f(x) f(y) dx dy = v - \frac{5t}{3} \qquad (3)$$

and total welfare is equal to $v - \frac{2t}{3}$.

No privacy

Consider a setting in which firms observe x and y and set a personalized price $p_i(x, y)$ for each consumer (x, y). This is an extension of Taylor and Wagman (2014) to two dimensional private information. In their case, with consumers uniformly distributed along the unit line, all discriminatory prices are driven downwards because firms fiercely compete for consumers at each location. The closest retailer charges the saving in total transportation costs enjoyed by a customer when buying the product from that firm rather than the farthest retailer (Bhaskar and To, 2004). Price schedules are efficient and the personalized offer to the consumer equidistant from both outlets is equal to the marginal cost. Given that each firm faces a pool of consumers which are relatively closer to the rival's location, retailers have a common incentive to tailor with lower discriminatory prices consumers located in these weak markets. In equilibrium all personalized prices are driven down because location based models are characterized by best response asymmetry in the horizontal dimension (Corts, 1998; Armstrong, 2006).

When there is no privacy, firms compete at the individual level. The indifference condition writes

$$p_1(x,y) + tx + ty = p_2(x,y) + t(1-x) + t(1-y) \tag{4}$$

and a generic consumer (x, y) buys from firm 1 if

$$p_1(x,y) \leq p_2(x,y) + t(1-2x) + t(1-2y). \tag{5}$$

Both firms know exactly the location of each consumer and they can set very aggressive prices to attract distant customers. Firms are willing to price as low as the marginal cost to serve an additional consumer, and they are left only with the possibility to extract the savings in transportation costs over both dimensions granted to close consumers. The standard Bertrand logic therefore applies also in two dimensional models.

Firm 1 has a transportation cost advantage when serving consumers with $x+y < 1$, whereas firm 2 has an advantage over those with $x+y > 1$. In turn, firms set their tailored offers accordingly to $p_1(x,y) = \max\{t(1-2x) + t(1-2y), 0\}$ and $p_2(x,y) = \max\{t(2x-1) + t(2y-1), 0\}$. Equilibrium personalized prices are

$$p_1^{np}(x,y) = 2t(1-(x+y)) \qquad \text{if} \quad x+y < 1 \tag{6a}$$

$$p_2^{np}(x,y) = 2t((x+y)-1) \qquad \text{if} \quad x+y > 1 \tag{6b}$$

and equal to the marginal cost otherwise. The market boundary is the same as in the full privacy benchmark and efficiency is achieved again. Notice that only consumers which are located along the bisector of the unit square effectively pay a zero price. All other consumers pay a positive price that reflects the savings in transportation costs from going to the closest retailer. Provided that the model is two dimensional, some consumers end up being charged the highest feasible price: consumers located precisely at firms' locations get an offer equal to $2t$ and are fully exploited. Nevertheless, even though some prices are larger than in the full privacy benchmark, the competition effect still prevails on the rent extraction effect. Firm 1's profit is

$$\pi_1^{np} = \int_0^{1-x} \int_0^1 p_1^*(x,y)f(x)f(y)dxdy = \frac{t}{3} \tag{7}$$

and symmetrically for the rival[7]. Whenever firms have symmetric access to full information, they would be better-off by committing not to price discriminate. Consumer surplus is larger and equal to $CS^{np} = v - \frac{4t}{3}$ so that total welfare is left unchanged with respect to the full privacy benchmark. However, there are winners are losers among consumers: those with $0 \leq x < \frac{1}{2}$ and $0 \leq y < \frac{1}{2} - x$ are charged a personalized price larger than p_i^* at outlet 1, whereas those with $\frac{1}{2} < x \leq 1$ and $\frac{3}{2} - x < y \leq 1$ pay more at outlet 2. Nevertheless, the mass of consumers paying on average a lower price is larger and therefore profits decrease.

3 Partial privacy

So far privacy was either fully or not enforced at all. Instead, in digital markets, consumers are likely to be partially targeted: firms eventually know something about them but not everything. Data collectors such as Acxiom gather data on several dimensions, from demographics to financial attributes[8], but even though it is realistic to assume that consumers are segmented in various categories, it is also less tenable to argue that an extremely precise profile of each consumer is created. In this context, this is equivalent to assume that an intermediate privacy regime is imposed in the market so that first degree price discrimination is unfeasible.

Symmetric partial information

Suppose that both firms have information on x and are able to tailor prices $p_i(x)$ to this continuum of consumer groups. Consumer (x,y) in group x accepts the personalized price of firm 1 if and only if

$$p_1(x) + tx + ty \leq p_2(x) + t(1-x) + t(1-y). \tag{8}$$

[7]Notice that, under all-out competition in one dimensional models, firms' profit is equal to $\frac{t}{4}$ (Thisse and Vives, 1988). The transition to a two dimensional model induces by construction a larger degree of product differentiation, which is reflected in an increase in profits with respect to a setting à la Thisse and Vives (1988).

[8]See `https://www.acxiom.com/customer-data/infobase/`

The expression for the indifferent consumer writes as in (2), except for the fact that prices are tailored on the realized value of x observed by both firms. Notice that for each x such prices take a specific value p_1 and p_2, and to save on notation this allows us to write the demand of firm 1 at each x as

$$D_1(p_1, p_2; x) = F(y \leq \tilde{y}(x)) = \frac{p_2 - p_1 + 2t(1-x)}{2t} \tag{9}$$

while $D_2(p_1, p_2; x) = 1 - D_1(p_1, p_2; x)$. Similarly to the uniform pricing game, both firms have positive demand whenever $|p_1(x) - p_2(x)| < 2t$, at least for some values of x. Firms maximization problems yield asymmetric best responses $b_1(p_2) = (p_2 + 2t(1-x))/2$ and $b_2(p_1) = (p_1 + 2tx)/2$ that depend on x. Firm 1's best response is strictly monotone decreasing in x and it shifts inwards following an increase in x, whereas firm 2's best response shifts upwards.

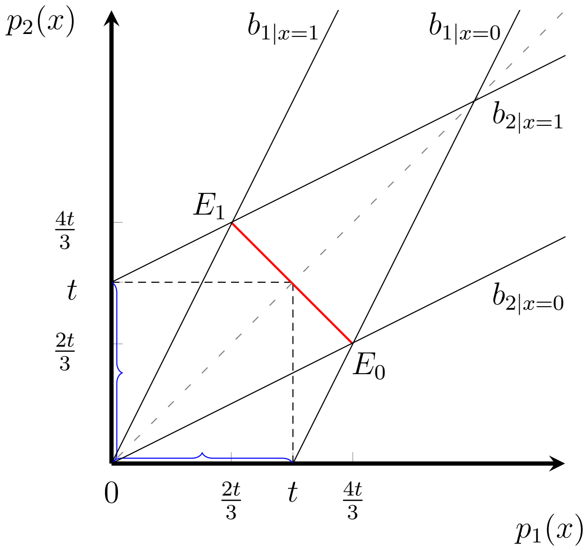

Figure 3: Best responses for $x = 0$ and $x = 1$ and locus of equilibria with discriminatory prices for all values of x (red segment). The blue braces show the price range in a standard one-dimensional model with discrimination.

The equilibrium personalized prices

$$p_1^*(x) = t + \frac{t}{3}(1 - 2x) \tag{10a}$$

$$p_2^*(x) = t + \frac{t}{3}(2x - 1) \tag{10b}$$

lie on the segment joining E_0 to E_1 in Figure 3. This segment is drawn by translating the best responses in the space $p_1(x) \times p_2(x)$, as the observable parameter changes from $x = 0$ to $x = 1$. Each firm charges a maximum price of $\frac{4t}{3} > p^{fp}$ to the closest consumers and a minimum price of $\frac{2t}{3}$ to the farthest ones[9]. Notice that the full privacy price lies in between the highest and the lowest discriminatory prices, which is a well established fact in pricing games characterized by best-response *symmetry*. Stole (2003) argues that without such symmetry this equilibrium feature cannot emerge, which instead this model with best-response *asymmetry* contradicts[10].

Formally, from (9) define the demand elasticity of consumers in group x with respect to firm i's price as $E_{D_i} = -\frac{\partial D_i}{\partial p_i}\frac{p_i}{D_i}$, which can be written as

$$E_{D_1} = \frac{p_1}{p_2 - p_1 + 2t(1-x)} \qquad \text{and} \qquad E_{D_2} = \frac{p_2}{p_1 - p_2 + 2tx} \tag{11}$$

In comparison to the full privacy regime, consumers in the neighbourhood of firm 1 have a more inelastic demand for firm 1's product, while far away consumers have a relatively more elastic demand for it. A symmetric argument applies to firm 2. The market is therefore divided in two regions: (i) for $0 < x < \frac{1}{2}$, firm 1 faces an inelastic demand whereas firm 2 has an elastic demand; (ii) for $\frac{1}{2} < x < 1$, the reverse holds true.

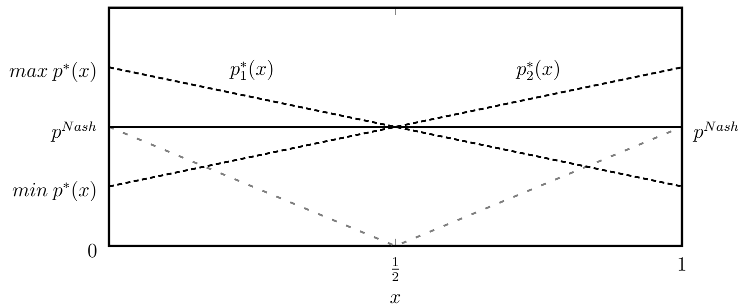

Figure 4: Personalized price schedules in two dimensional (densely dashed) and one dimensional models (loosely dashed).

Accordingly to the elasticity of demand, competing firms rank consumers in an opposite way, and they have contrasting incentives when setting prices conditional on the same information, as intuitively shown in Figure 4. In other words, best response asymmetry holds, but some consumers pay a higher price whereas others benefit from a targeted discount. In

[9]Notice that the average price schedule is equal to the Nash equilibrium uniform price. This is not true in one dimensional spatial models characterized by best-response asymmetry, where the average price decreases.

[10]What we usually expect with best-response asymmetry is an increase or a decrease of all equilibrium prices.

particular, some customers switch between outlets to acquire the product at a lower price.

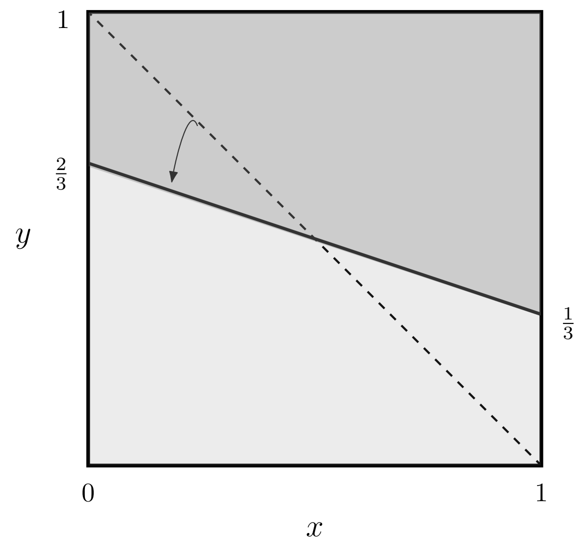

Figure 5: Rotation of the market boundary.

Indeed, in equilibrium the market boundary is $y^*(x) = \frac{2-x}{3}$ and it exhibits an anticlockwise rotation with respect to the full privacy game. Each firm serves always a positive fraction of y types in all groups x. Symmetric partial price discrimination allows firms to serve even some of the most loyal customers of the rival. Two major implications follow: (i) a subset of buyers is charged more than p^{fp} (i.e. groups $x < \frac{1}{2}$ at firm 1 and $x > \frac{1}{2}$ at firm 2), whereas other groups enjoy a discount (i.e. $x > \frac{1}{2}$ at firm 1 and $x < \frac{1}{2}$ at firm 2); (ii) the allocation of consumers among the two outlets is inefficient, given that some consumers switch between firms to benefit from a discounted personalized price, causing a net increase in overall transportation costs. The anticlockwise rotation of the market boundary is driven by types x close to their ideal firm but with an unobservable realized y which is closer to the rival's product characteristics: given that each firm increases its price precisely to close groups x, some members of these categories benefit from switching to the low pricing rival, trading-off a better match in the dimension y with a larger mismatch in the dimension x. Firm's 1 profit[11] is

$$\pi_1(x,x) = \int_0^{\frac{2-x}{3}} \int_0^1 \left(t + \frac{t}{3}(1-2x) \right) f(x)f(y)dxdy = \frac{14t}{27} \tag{12}$$

and symmetrically for firm 2. Each firm makes a larger profit when both competitors obtain

[11] In the following I denote with $\pi_i(a,b)$ the realized firm i's profit when i has information set a and j has b, with a and b taking values in $\{0,x,y\}$. For aggregate outcomes, a (b) refers to firm 1 (2).

access to some consumer data. Symmetric partial information makes firms less aggressive when setting prices. In turn, even tough firms are competing on equal grounds, less privacy is not beneficial to consumers or, at least, not to all of them, differently from what is suggested in the one dimensional literature. Indeed, consumer surplus is

$$CS(x,x) = 2\int_0^{\frac{2-x}{3}} \int_0^1 \left(v - p_1^*(x) - t(x+y)\right) f(x)f(y)dxdy = v - \frac{47t}{27} \tag{13}$$

and it can be easily shown that $CS(x,x) < CS^{fp}$ holds. There is a redistribution of surplus from consumers to the firms even though some buyers acquire the product at a lower price. The discounts offered to low valuation consumers in the x dimension are more than offset by the increase in prices imposed on high valuation consumers who cannot fruitfully switch. Only those located at $x = \frac{1}{2}$ are indifferent with respect to full privacy and consume efficiently. Consumers are overall worse-off, but the impact on individual net utilities is ambiguous. Instead, both firms are strictly better-off and the uncertainty about the other dimension is crucial for the ability of firms to extract more rents from consumers.

Finally, total welfare is equal to $v - \frac{19t}{27}$ and it is lower than in both benchmarks[12]. The distortion in the allocation of consumers among the competing firms causes a surplus loss due to the net increase in transportation costs.

Different partial information

Suppose that both firms have access to partial consumer data but on different dimensions. The data allocation is asymmetric, with the two firms targeting consumers along dimensions x and y, respectively. When the allocation is reversed the analysis is similar. Firm 1 sets a discriminatory price $p_1(x)$ and firm 2 simultaneously sets $p_2(y)$.

Lemma 2. *When firm 1 has information on x and firm 2 has information on y personalized prices are $p_1^*(x) = t(\frac{3}{2} - x)$ and $p_2^*(y) = t(\frac{1}{2} + y)$. Each firm's profit is equal to $\pi_i = \frac{7t}{12}$.*

Proof. See the Appendix A. ■

[12] In one dimensional models there is a redistribution of surplus from firms to consumers, and overall welfare is unchanged. See the characterization of the equilibria with and without privacy in the linear city model of Taylor and Wagman (2014). Moreover, recall that the allocation of consumers in those two cases is efficient and the location of the marginal consumer does not change.

The equilibrium market boundary coincides with the bisector. Access to different partial information restores efficiency. However, firms are able to extract more surplus from consumers than in the symmetric case: industry profits increase up to $\pi(x, y) = \frac{7t}{6}$ whereas consumer surplus is driven down to $CS(x, y) = v - \frac{11t}{6} < CS(x, x)$, the lower bound across all games. Total welfare is maximized but efficiency is achieved at the expenses of consumers.

Recall that under symmetric partial information the market was broadly divided in two regions accordingly to demand elasticities, and one firm's equilibrium price was mirroring the price set by the rival at each x (see Figure 4). Here the elasticities of demand write

$$E_{D_1} = \frac{p_1}{p_2 - p_1 + 2t(1-x)} \qquad \text{and} \qquad E_{D_2} = \frac{p_2}{p_1 - p_2 + 2ty} \tag{14}$$

with E_{D_1} increasing in x and E_{D_2} decreasing in y, with the market divided in four regions. When setting prices, firms now have symmetric incentives in the two regions located along the negatively sloped diagonal, but contrasting incentives in the other two regions.

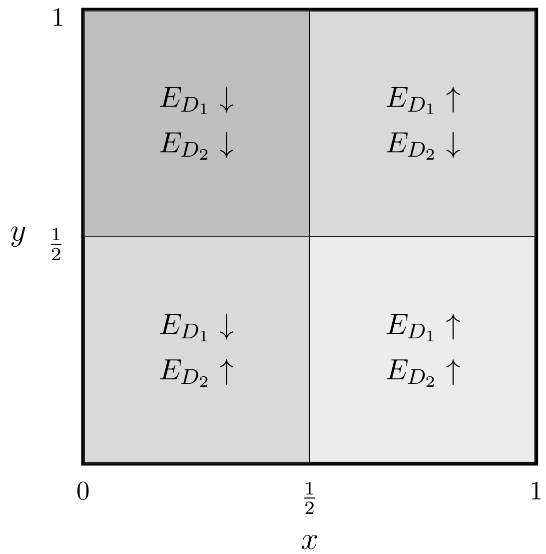

Figure 6: Price elasticity of demand under different partial information.

In other words, along the negatively sloped diagonal, best response asymmetry fails to hold, and firms rank consumers similarly even tough they discriminate on different pieces of private information. When consumer privacy is partially enforced and firms have access to different partial information, there is a mixture of best responses symmetry and asymmetry. As a result, only types with $\frac{1}{2} < x < 1$ and $0 < y < \frac{1}{2}$ get targeted discounts in equilibrium. All the others pay on average a higher price: profits increase even more, at the detriment of consumers.

Partial exclusivity

Suppose that firm 1 exclusively acquires partial consumer data[13] on the x dimension. Consumer (x, y) buys from the informed firm at the personalized price if and only if

$$p_1(x) + tx + ty \leq p_2 + t(1-x) + t(1-y) \tag{15}$$

which yields the following expression for the locus of indifferent consumers:

$$y(x) = \frac{p_2 - p_1(x) + 2t(1-x)}{2t}. \tag{16}$$

The best reply of the informed player is given by $b_1(p_2) = (p_2 + 2t(1-x))/2$, which is defined for all x realizations. Instead, the uninformed player is not able to optimally respond at each x to the price schedule posted by the rival. However, firm 2 can set a uniform price that "on average" is a best reply to the rival's one: in equilibrium the candidate uniform price of firm 2 must be equal to the average candidate price schedule of firm 1.

Lemma 3. *When only firm 1 has partial information the uniform price is $p_2^* = t$ while the personalized price is $p_1^*(x) = t\left(\frac{3}{2} - x\right)$. Profits are equal to $\pi_2 = \frac{t}{2}$ and $\pi_1 = \frac{13t}{24}$.*

Proof. See Appendix A. ■

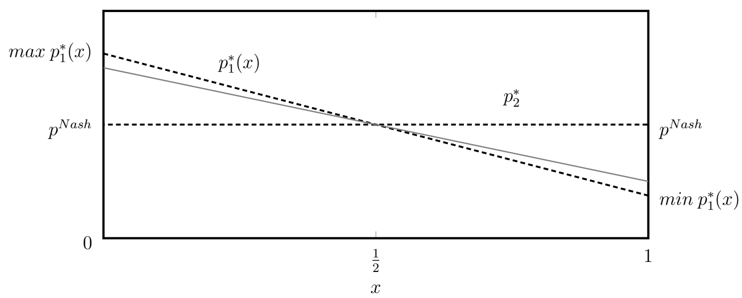

Figure 7: Rotation of the personalized price charged by the exclusively informed firm (densely dashed) with respect to the symmetric partial information game (not dashed).

[13]The analysis is symmetric when it is firm 2 to receive exclusive information and, also, when the y dimension is the acquired item.

The average value of $p_1^*(x)$ is exactly p_2^*. In contrast to the symmetric game with partial price discrimination, the informed firm sets an even higher personalized price in its strong market, but it is also forced to price more aggressively to types x close to the uninformed firm's location, similarly to what happens in the game with different partial information. As a result, the equilibrium market boundary is $y^*(x) = \frac{3-2x}{4}$, and interestingly it rotates clockwise with respect to the symmetric information scenario. In the different partial information regime, the inefficiency is partially mitigated as more consumers buy their preferred product. Provided that $\pi_i(x, 0) > \pi_i(0, x)$, exclusive information naturally gives to the informed player a competitive advantage. However, exclusive access to partial data is not detrimental to the uninformed player: firm 2 realizes the same profit as under full privacy and the standard negative externality caused by exclusive access to data does not arise when data is partially informative. In other words, firm 2's data foreclosure does not result in a "too large" difference between the payoffs of the two players. Consumer surplus is equal to $CS(x, 0) = v - \frac{83t}{48}$ and it is slightly higher than $CS(x, x)$ but it ranks below the benchmarks. Overall, in comparison to the other games with partial privacy, exclusivity does not harm neither the uninformed player nor final consumers but rather it partially restores efficiency.

A welfare analysis of partial privacy

Industry profits. When consumer private types are identified as the pair (x, y) and firms partially observe the consumers' willingness to pay, more information is not necessarily associated with stronger competition in the market.

Usually, when both firms have access to consumer data and compete on equal grounds, competition intensifies and the firms are worse-off with price discrimination. In my model, as long as the competing firms observe only something about consumers, price discrimination is profit enhancing. The first contribution of this article is to show that, with the introduction of a slightly more complex information structure, the ranking of industry profits as a function of the amount of information in the market is non-monotonic:

$$\pi(x, y) > \pi(x, 0) > \pi(x, x) > \pi^{fp} > \pi^{np}. \tag{17}$$

In particular, firms are collectively best-off when they can exploit some but different information: rent extraction is maximized with data differentiation. In addition, whenever it is impossible to create truly comprehensive profiles of consumers, so that partial privacy is effectively imposed in the market, exclusive access to data is shown to be not detrimental to uninformed competitors:

$$\pi_i(x,y) > \pi_i(x,0) > \pi_i(x,x) > \pi_i^{fp} = \pi_i(0,x) > \pi_i^{np}. \tag{18}$$

When two dimensions of product differentiation are considered, it is irrelevant for a firm without access to data whether its rival acquires or not partial information. This ranking will have a crucial implication for the optimal selling strategy of a monopolistic data broker in Section 4.

Consumer surplus. From a consumer protection point of view, the main implication of this analysis is that less privacy in the market is not necessarily beneficial to consumers.

Common wisdom suggests that more information in the market should benefit consumers (i.e. CS^{fp} should be a lower bound on surplus) and, indeed, the literature has widely shown that less privacy is better for all consumers when both competing firms price discriminate. In the limit, no privacy at all is optimal.

Even though my model confirms that the no privacy extreme case benefits consumers, I find that consumer surplus is reduced when intermediate privacy regimes are imposed in the market, showing that in aggregate

$$CS^{np} > CS^{fp} > CS(x,0) > CS(x,x) > CS(x,y). \tag{19}$$

Partial privacy is always detrimental to customers, and data differentiation yields the worst outcome. If a fully informative scenario with individually personalized prices at both outlets is not feasible, then price discrimination should be banned. The negative effects of partially targeted prices call for the following policy advice: privacy protection in digital markets should be made either very hard, promoting a no privacy outcome, or very easy, a second-best alternative. Consumers would collectively prefer no privacy due to the fierce price competition between firms. Otherwise, in case some data about them is collected and exploited, they would instead opt for full privacy rather than intermediate scenarios.

However, despite the clear policy conclusion suggested by this result, completely banning or not regulating at all the use of data may not be beneficial for all consumers.

The effects of different privacy regimes on individual surplus are quite heterogeneous: under partial privacy, it is true that some customers are exploited with higher targeted prices, but others receive a tailored price that is truly a discount with respect to the Nash equilibrium uniform price. Indeed, when firms can partially price discriminate along one dimension only, some personalized prices are above p^{fp} whereas others lie below it. When firms hold the same partially informative data, some consumers strategically but inefficiently switch between outlets to benefit from tailored discounts. When firms hold different partially informative data, efficiency is restored and some consumers receive a discount as well. However, the majority of them is exploited with surcharges. Moreover, notice that even under no privacy, which provides an upper bound on surplus, the consumer-level analysis delivers heterogeneous conclusions: some consumers are in the worst possible scenario, given that the efficient equilibrium prices approach $2t$ for the most captive consumers. There are therefore winners and losers in each scenario, which makes it hard to draw an unambiguous policy conclusion regarding the impact of privacy regulation on specific data subjects.

The most insightful comparison is between the full privacy and the symmetric partial information case, an allocation that induces the largest inefficiency. All other comparisons directly follow from what is shown here. Recall that under full privacy the market boundary is $y^*(x) = 1 - x$ and transportation costs are minimized. When the information regime changes, the market boundary $y^*(x) = \frac{2-x}{3}$ rotates anticlockwise. As shown in Figure 8, in order to benefit from a relatively lower price, some consumers are willing to incur in a larger transportation cost and buy from the farthest outlet. Partial privacy generates a misallocation of consumers among the duopolists: some customers buy the "wrong" product. Therefore, under symmetric partial price discrimination, in contrast to full privacy, we can identify three types of consumers:

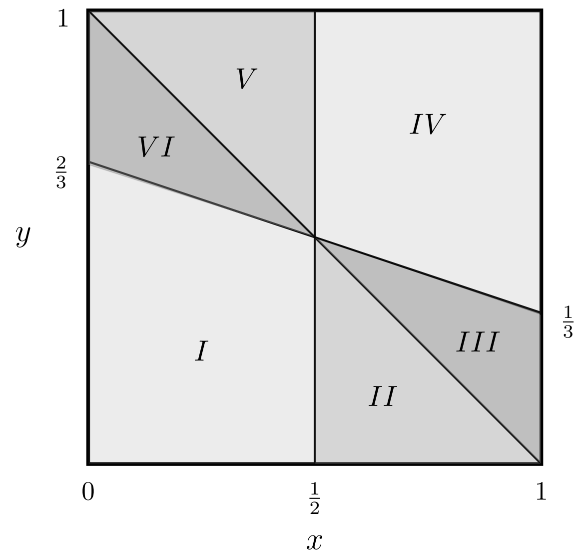

Figure 8: Consumers' allocation among firms: from full privacy to symmetric partial information.

1. Consumers in the sets $I = \left\{(x, y) : 0 \leq x < \frac{1}{2}, 0 \leq y < \frac{2-x}{3}\right\}$ and $IV = \left\{(x, y) : \frac{1}{2} \leq x \leq 1, \frac{2-x}{3} \leq y \leq 1\right\}$ are strictly worse-off but buy from the same firm at a higher price;

2. Consumers in the sets $II = \left\{(x, y) : \frac{1}{2} \leq x \leq 1, 0 \leq y < 1 - x\right\}$ and $V = \left\{(x, y) : 0 \leq x < \frac{1}{2}, 1 - x \leq y \leq 1\right\}$ are strictly better-off and buy from the same firm getting a discount;

3. Consumers in the sets $III = \left\{(x, y) : \frac{1}{2} \leq x \leq 1, 1 - x \leq y < \frac{2-x}{3}\right\}$ and $VI = \left\{(x, y) : 0 \leq x < \frac{1}{2}, \frac{2-x}{3} \leq y < 1 - x\right\}$ switch between firms and buy a mismatched product.

There is an inefficient flow of consumers between outlets. The switchers avoid the high prices of the nearest competing firm, and prefer to get the good from the farthest firm at a sensible discount. However, in aggregate, the positive effect on the switchers' net utility coming from the discount is perfectly offset by the increase in their transportation costs, so that the overall surplus of switchers does not vary. Indeed, as shown in Figure 9, only switchers located along the line $\frac{5-4x}{6}$ are really indifferent between full and partial privacy: switchers located relatively far from the newly chosen retailer incur in an additional transportation cost that outweighs the discount; only the others actually have a net benefit from a lower tailored price. Figure

9 provides a graphical intuition for the net increase in profits: the fraction of losers is clearly larger than the area of winners.

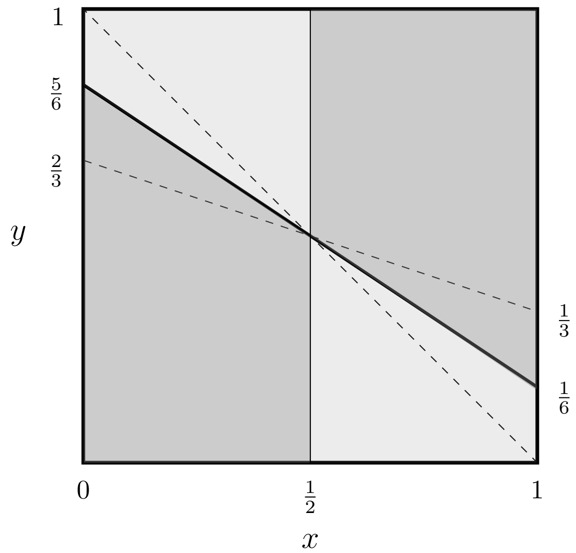

Figure 9: Winners (light grey) and losers (dark grey) when the same partial information is held by both firms.

Firms' interests to acquire and use personal data are partially aligned with consumers' interests. A small fraction of consumers would agree with the disclosure of personal information whereas others would prefer to conceal it. In the one dimensional literature instead these interests are always misaligned, as prices move in one direction only when both firms are informed.

Finally, it is worth to briefly comment on what happens under partial exclusivity and different partial information, in comparison to the symmetric case. Under exclusive access to data, the inefficiency is partially mitigated given that the market boundary rotates clockwise. More consumers buy the right product but, at the informed outlet, some of them are served at an even lower price while other customers are charged more. However, the efficiency gain is sufficient to have a slight increase in consumer surplus. Exactly the opposite happens with partial data differentiation. Efficiency is restored, but total consumer surplus hits a lower bound. Differently from partial exclusivity, now two asymmetrically informed firms target consumers with both high and low personalized prices. Efficiency clearly ensures that total welfare is the same as in the two benchmarks, but the ability of firms to extract surplus from consumers is maximized. From a consumer privacy perspective, having two competing firms endowed with different data yields the worst possible outcome. For completeness, Table 1 reports a summary of all the equilibrium outcomes characterized so far.

	Industry profits	Consumer surplus	Total welfare
Full privacy	t	$v - \frac{5t}{3}$	$v - \frac{2t}{3} \uparrow$
No privacy	$\frac{2t}{3} \downarrow$	$v - \frac{4t}{3} \uparrow$	$v - \frac{2t}{3} \uparrow$
Symmetric partial info	$\frac{28t}{27}$	$v - \frac{47t}{27}$	$v - \frac{19t}{27}$
Different partial info	$\frac{7t}{6} \uparrow$	$v - \frac{11t}{6} \downarrow$	$v - \frac{2t}{3} \uparrow$
Partial exclusivity	$\frac{25t}{24}$	$v - \frac{83t}{48}$	$v - \frac{11t}{16}$

Table 1: Equilibrium outcomes (the arrows identify the maximum and minimum of each column).

4 Monopolistic data broker

Here I study the incentives of a data broker to sell collected data to the downstream firms. The data seller posts a take it or leave it offer for the available information and the seller is naturally assumed to hold all the bargaining power. Payments are made at this stage and the information allocation among downstream firms becomes common knowledge before the simultaneous pricing game. The selling mechanism usually exploited in the literature resembles an auction with downstream externalities (Jehiel and Moldovanu, 2000): when the dataset is exclusively offered to a competing firm which does not buy data, then the rival has the chance to acquire it. However, in a two dimensional model, partial exclusivity does not impose a negative externality on the loser of the auction, differently from standard information acquisition games.

In the following, the data broker is said to hold partially informative data when only x (or equivalently, only y) is collected; data is fully informative when both dimensions are available. The privacy scenarios analyzed so far become now the subgames of the sequential game in which the data broker sells information at a novel first stage.

Partially informative data

Suppose that only one dimension of private information is collectible. Without loss of generality it is assumed to be the x dimension. It could be that y represents sensible data that must be anonymized or that its collection is extremely costly.

	NI	I
NI	$t/2,\ t/2$	$t/2,\ 13t/24$
I	$13t/24,\ t/2$	$14t/27,\ 14t/27$

Table 2: Pricing game with partially informative data.

Proposition 1. *When the dataset is partially informative, the seller sells partial information to only one firm at a price* $P^{EX} = \frac{t}{24}$.

Proof. First of all, notice that the exclusive price[14] is equal to

$$P^{EX} = \pi_i(x, 0) - \pi_i(0, x) = \frac{t}{24}. \tag{20}$$

Suppose that the seller posts a take it or leave it offer $P_i^{NE} > 0$, where P_i^{NE} is the price at which both firms can acquire consumer data. The non exclusive price[15] writes

$$P_i^{NE} = \pi_i(x, x) - \pi_i(0, x) = \frac{t}{54}. \tag{21}$$

[14] Bounie et al. (2018) consider a selling mechanism different from an auction with downstream externalities and write the exclusive price as the difference in profits $\pi_i(x, 0) - \pi_i^{fp}$, where the outside option is the standard Hotelling profit (i.e. our full privacy benchmark). They assume that the seller commits *ex-ante* to an exclusive deal with only one firm, without the possibility to offer the dataset to the rival in case of a rejection. Here these exclusive prices are equivalent.

[15] A more general expression for the price for information would be $P_i = \alpha \times (\pi_i(x, x) - \pi_i(0, x))$ and similarly for the exclusive price. The bargaining power measured by α is set equal to one in the proof.

The seller therefore compares

$$\underbrace{2 \times (\pi_i(x,x) - \pi_i(0,x))}_{P^{NE}=\frac{t}{27}} < \underbrace{(\pi_i(x,0) - \pi_i(0,x))}_{P^{EX}=\frac{t}{24}}. \tag{22}$$

■

In the subgame perfect Nash equilibrium of the partially informative game the data seller awards information exclusively to one downstream competitor.

Fully informative data

When full information is available the question is whether the data seller decides to gather partial or full information about consumers and what is the optimal way of selling this data. It is instrumental for the analysis to characterize the following scenario.

Full exclusivity. Suppose that firm 1 has access to a dataset containing full information on each consumer's willingness to pay while firm 2 is uninformed. When the exclusively informed firm is able to set a different price for each consumer it will optimally set the individual price accordingly to

$$p_1(x,y) = \max\{0, p_2 + t(1-2x) + t(1-2y)\} \tag{23}$$

which directly follows from the indifference condition at prices $p_1(x,y)$ and p_2. The intuition is that the informed firm, having the exclusive advantage of being able to identify individual locations, makes each consumer just indifferent between the two products. Then it is possible to show the following result.

Lemma 4. *When only firm 1 has full information the uniform price is* $p_2^* = \frac{2t}{3}$ *and the personalized price is* $p_1^*(x,y) = 2t\left(\frac{4}{3} - (x+y)\right)$. *Profits are equal to* $\pi_2 = \frac{4t}{27}$ *and* $\pi_1 = \frac{62t}{81}$.

Proof. See Appendix A. ■

Profits are equal to

$$\pi_2(0, xy) = \int_{\frac{4}{3}-x}^{1} \int_{\frac{1}{3}}^{1} p_2^* \, dxdy = \frac{4t}{27} \tag{24a}$$

$$\pi_1(xy, 0) = \int_{0}^{1} \int_{0}^{\frac{1}{3}} p_1^*(x, y) \, dxdy + \int_{0}^{\frac{4}{3}-x} \int_{\frac{1}{3}}^{1} p_1^*(x, y) \, dxdy = \frac{62t}{81}. \tag{24b}$$

The informed firm is not only better-off with price discrimination but a *business stealing* effect emerges: firm 2 is more aggressive but it serves less consumers. This implies that, in contrast to partial exclusivity, full exclusivity is detrimental to the uninformed firm. When only one firm has information on a unique dimension, an exclusive allocation does not impose a negative externality on the uninformed firm. Instead, when only one firm has information on both dimensions, such negative externality is there, as it is standard in the literature on selling data to competing firms. Industry profits lie slightly below the full privacy level and consumer surplus is equal to $CS = v - \frac{5t}{3} = CS^{fp}$. Overall, from a consumer protection point of view, this outcome implies that full exclusivity dominates all games with partial privacy, despite the increased inefficiency in consumer allocation among the outlets.

Given that firm 1 realizes the largest individual payoff across all games, it is not obvious what the seller should do with full information. In order to maximize the price of data, the seller usually induces a negative externality on the eventual loser of the data auction, so to impose the worst outside option on data buyers. If one player does not accept the initial take it or leave it data offer (either for partial or full information), then it is optimal from the point of view of the seller to grant full exclusivity to only one firm, so that the outside option of the information acquisition game is always the lowest possible payoff $\pi_i(0, xy) = \frac{4t}{27}$. But does the seller always have an incentive to induce an exclusive downstream allocation?

Formally, the data broker has access to x and y and can decide to allow individual or group targeting. Competing firms know that they can either receive partial or full information offers. Clearly, a fully informative dataset is always sold exclusively when put on sale as a unique block; otherwise, firms would find themselves in the no privacy scenario. Alternatively, a fully informative dataset can be partitioned along the two dimensions: a downstream partial data allocation could be induced. Notice that partial exclusivity (i.e. $(x, 0)$,$(0, x)$,$(y, 0)$,$(0, y)$) is obviously dominated by full exclusivity: $\pi_i(xy,0) - \pi_i(0,xy) > \pi_i(x,0) - \pi_i(0,x)$. In turn, Proposition 1 implies that also the allocations (x,x) or (y,y) do not arise. Instead, recalling that the industry profit $\pi(x,y)$ is an upper bound on the distribution of profits across all subgames, the seller may find it profitable to induce downstream partial data differentiation.

	NI	I
NI	$t/2,\ t/2$	$4t/27,\ 62t/81$
I	$62t/81,\ 4t/27$	$7t/12,\ 7t/12$

Table 3: Payoff matrix when data is fully informative.

The relevant payoff matrix for the analysis is summarized in Table 3, where in the cell $\{I, I\}$ firms receive different partial information whereas in the cell $\{I, NI\}$ one firm wins the entire dataset. The crucial intuition is that, as long as the data broker can freely decide how much information to sell, the outside option of the data buyers is represented by the payoff of being the uninformed firm when the rival is fully informed. Provided that data is collected and added to the information structure put on sale at no cost, the threat of an exclusive sale that heavily disadvantages the losing firm is credible. Before proving the main result of this section, it is instrumental to characterize another subgame that was not discussed yet.

Lemma 5. *Suppose that both firms observe x but only firm 1 has access to y. Then prices are equal to $p_2^*(x) = tx$ and $p_1^*(x,y) = t(2-x-2y)$ with profits $\pi_2 = \frac{t}{6}$ and $\pi_1 = \frac{7t}{12}$.*

Proof. See Appendix A. ■

As long as both firms hold some data, granting additional information to only one of the players solely has the effect of making the rival more aggressive whereas the fully informed player is just indifferent with respect to the starting allocation. Indeed, this scenario induces an allocation that tends towards the no privacy regime. In turn, consumer surplus is equal to $CS(xy,x) = v - \frac{3t}{4}$ and it lies between the full privacy and the no privacy outcome. Provided that mainly consumers benefit, it is reasonable to expect that the data seller does not have an incentive to promote such downstream data allocation.

Given the comprehensive ranking of individual payoffs characterized so far, it is possible now to describe the optimal selling strategy of the data holder.

Proposition 2. *If full information is collected at no cost, then both firms acquire partial but different information at a price* $P_i^{NE} = \frac{47t}{108}$.

Proof. An offer from the seller includes the type of information put on sale, an individual price for it and, implicitly, whether information is offered exclusively or non exclusively. The sale of full information implies exclusivity, provided that no privacy lowers industry profits. The *ex-ante* exclusive selling price is

$$P^{EX} = \pi_i(xy, 0) - \pi_i(0, xy) = \frac{62t}{81} - \frac{4t}{27} = \frac{50t}{81}. \tag{25}$$

On the other hand, the individual non exclusive selling price for partial information now becomes

$$P_i^{NE} = \pi_i(x, y) - \pi_i(0, xy) = \pi_i(y, x) - \pi_i(0, xy) = \frac{7t}{12} - \frac{4t}{27} = \frac{47t}{108}. \tag{26}$$

The data broker therefore compares

$$2 \times P_i^{NE} = \frac{47t}{54} > P^{EX} \tag{27}$$

which implies that the maximum revenue from information sale is secured by awarding partial information to both downstream competitors. We claim that the optimal selling strategy of the data broker is structured as follows: (1) a non exclusive contract regarding partial information, with x sold to one firm and y to the rival; (2) an exclusive contract regarding full information (x and y) in case contract (1) is not accepted by all buyers.

The proof is articulated in two parts: in the first part, we check that the data broker has no incentives to deviate from the non exclusive contract proposed above, which also proves that Table 3 is indeed the relevant payoff matrix for the analysis; in the second part, we check that the buyers have an unilateral incentive to accept the offer P_i^{NE}.

Part 1. Suppose that a non exclusive contract has been accepted by firms at some positive price P_i possibly different from P_i^{NE}. Payments are made just before the pricing game, once the final data allocation becomes common knowledge. Suppose that the data broker has the option to offer exclusively x to the buyer that has acquired partial information on y. Suppose that firm 1 had initially access to y, and it gets data on x and y whereas firm 2 is left with access to x only. From Lemma 5 the new prices would be $p_2^*(x) = tx$ and $p_1^*(x, y) = t(2-x-2y)$, with respective profits given by $\pi_2 = \frac{t}{6}$ and $\pi_1 = \frac{7t}{12}$. The firm that receives the after-sale exclusive offer gets exactly the same payoff, given that the rival is already partially informed. There is no incentive to acquire additional information about consumer preferences. In turn, the data broker do not have an incentive to deviate from the initial non exclusive contract.

Part 2. We show that a firm cannot gain by unilaterally deviating from the non exclusive contract proposed above, which also implicitly provides the rationale behind the choice of the payoff $\pi_i(0, xy)$ as the outside option in the expression for the relevant price. Suppose that firm 2 does not accept the non exclusive contract about partial information x. The data broker can then simply offer information on x along witg that on y to firm 1. In other words, an exclusive contract is now offered to the not deviating firm. In case of a further rejection of the offer both firms obtain the no information payoff. However, if the exclusive offer is accepted, the deviating firm obtains exactly the outside option that appears in P_i^{NE}. Indeed, firm 1 would have an incentive to acquire full information at a lower *ex-post* exclusive price, which is equal to

$$P^{EX} = \pi(xy, 0) - \pi^{fp} = \frac{62t}{81} - \frac{t}{2} = \frac{43t}{162} \tag{28}$$

and firm 2 gets the payoff of the uninformed firm under full information. Therefore, each firm has an unilateral incentive to accept the non exclusive contract which maximizes the seller's revenue (i.e. $\bar{R} = \frac{47t}{54}$) at the first stage. As usual, given the standard assumption on bargaining power, the price charged for consumer data makes the buyer exactly indifferent between acquiring information and the outside option. When the collection choice of the seller is made endogenous, the outside option of the firms is less favorable, and therefore seller's revenue is larger than the one characterized under exogenously given partial information. ■

The possibility to offer full information allows the data broker to credibly threaten firms with an exclusive downstream data allocation. The threat is credible considering that it is impossible for the duopolists to coordinate and deviate together from the proposed non exclusive contract, as each competitor has an incentive to wait for the deviation of the rival and then get the exclusive offer. Notice also that this result holds independently of whether contracts are offered before or after the actual collection of consumer data. Zero collection costs imply that the seller can always propose a new offer in case of a deviation by one player (i.e. immediately acquire the additional information that is needed to propose the exclusive data package), which makes the outside option of data buyers less favorable.

Costly data collection

In presence of market frictions such as investments in collection capabilities it could become difficult to build a comprehensive dataset to put on sale. Here I interestingly show that the non exclusivity result holds also when data collection costs are positive but not too large. Suppose that the data broker incurs in the fixed costs $k_x > 0$ and $k_y \geq 0$ when collecting information on dimensions x and y respectively. We only require that $k_x + k_y \leq \frac{47t}{54}$, otherwise the data broker would not have enough resources to threaten firms with an exclusive offer, which is the key mechanism used by the seller to extract more rents from the data buyers. Put it differently, the seller has an incentive to actively gather full consumer data as long as the cost of investing in the tracking technology does not exceed the maximum revenue from information sale. The key intuition is that in presence of costly upstream information acquisition, only the cost of collecting an extra dimension really matters for the data broker's incentives once the constraint is satisfied.

Proposition 3. *If $k_x + k_y \leq \frac{47t}{54}$, then the data broker collects full information and each firm acquires partial but different information at a price $P_i^{NE} = \frac{47t}{108}$.*

Proof. See the text. ■

As long as the constraint on costs is satisfied, in the region below the negatively sloped diagonal in Figure 10, the data broker collects full information and obtains the maximum revenue $\bar{R}$ when selling different partial information: the threat of an exclusive deal is credible since the seller has a fully informative dataset, and such threat is leveraged to extract more rents from the firms through the non exclusive contract. What is left to investigate is what happens when the constraint on costs is not satisfied: the data broker cannot collect full information.

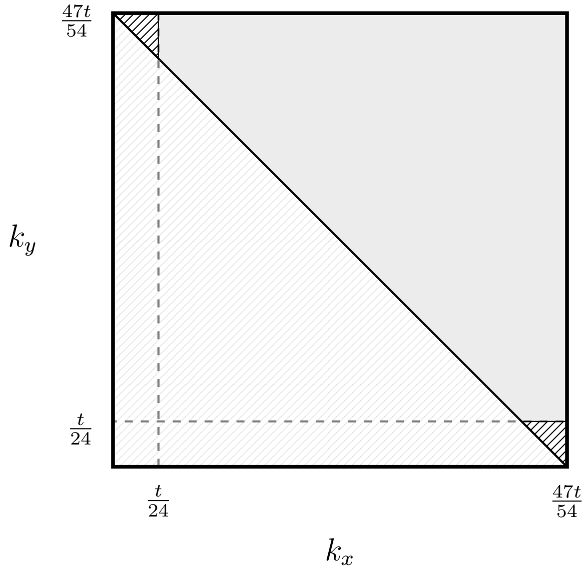

Figure 10: Selling strategy of the data broker when collection costs are positive.

This observation has an immediate impact on the outside option of buyers. Provided that it is common knowledge that a fully exclusive contract cannot be proposed when $k_x + k_y > \frac{47t}{54}$, competing firms are aware of the fact that an unilateral deviation from the contract about partial consumer data now yields $\pi_i(0, x) = \frac{t}{2}$, if we assume that x is collected first in the case of a partially informative dataset. The data broker then prefers to sell partial information exclusively, as long as the cost of collecting a single dimension of consumer data does not exceed the exclusive price characterized in Proposition 1. It is therefore straightforward to show the following result.

Proposition 4. *Suppose that $k_x + k_y > \frac{47t}{54}$. If k_x $(k_y) \leq \frac{t}{24}$, then the data broker collects partial information on dimension x (y) and sells it exclusively to one firm at a price $P^{EX} = \frac{t}{24}$; otherwise, no information is collected.*

Proof. See the text. ■

The punchline is that positive costs of data collection determine which type of information structure is in the hands of the seller; in turn, availability of full or partial information determines whether a negative externality is imposed or not on the loser of the data auction. Accordingly to the prevalent outside option, which is $\pi_i(0, xy) = \frac{4}{27}$ under full information and $\pi_i(0, x) = \pi_i^{fp}$ under partial information, the seller is able to extract more or less rents.

5 Main highlights

Motivated by the huge collection and exploitation of consumer data in digital markets, I study the effects of information on competition when price discrimination is a feasible targeting strategy, and I provide a complete characterization of the impact of different privacy regimes on profits and consumer surplus in a model in which the information structure is slightly more complex and realistic. I also investigate the incentives of a monopolistic data broker to sell data to competing firms. The major implications of this article are twofold: (i) the relationship between information and industry profit is hump-shaped, implying that firms are always better-off with partial price discrimination, and intermediate levels of privacy are in aggregate detrimental to consumers; (ii) accordingly to the type of data held by the seller, both exclusive and non exclusive sales of information can arise in equilibrium.

When the structure of consumer private information is two dimensional (x, y) and firms observe only one dimension, competition is relaxed. Partial price discrimination is profit enhancing and the standard prisoner's dilemma logic does not apply here. No privacy instead, as expected, strengthens price competition. Firms are therefore worse-off in the extreme cases of full or no privacy, and an inverse U-shaped relationship between the quantity of data available to the players and industry profits arises. The inability to observe a portion of relevant information is crucial for the increase in industry profits. When data is partially informative the standard Bertrand competition argument in transportation costs breaks down, and firms find it optimal to increase prices to close consumers given that not all customers but only a minor portion of them would eventually switch to the rival to get a discount. This mechanism works only if the model is two dimensional. Indeed, if the unobserved dimension became irrelevant, we would get again an "all-out" competition outcome. But given the profitability of a particular type of competitive price discrimination, why is there a limited evidence of targeted prices in markets? Omitting here the questionable explanation of the difficulties to detect discrimination in markets, the game theoretical answer is a simple one. Absent an upstream data seller, firms would have to invest autonomously in tracking and data collection capabilities. In turn, even though firms would collectively benefit from gathering partially informative data, full exclusivity still yields the largest firm-level payoff. If it is feasible to gather full information, firms would therefore fail to coordinate to collect and use only partial information.

In terms of consumer protection, this article speaks to the policy debate surrounding consumer privacy: should we promote more or less privacy in the market? In terms of aggregate consumer surplus, my model shows that no privacy is optimal whereas an intermediate regulation is detrimental to consumers. If instead firms are not able to compete on individual basis but partial information is prevalent in the market, then a full privacy regulation should be welcomed. Thus, the policy recommendation is to promote either no privacy or, eventually, full privacy. However, the impact of different regimes on individual consumer surplus is ambiguous. While some consumers would prefer to conceal their personal data from price discriminating firms, others prefer to receive tailored offers, regardless the collection of partially or fully informative data about their preferences: neither it would be beneficial for all consumers to allow individual targeting nor it would be advisable to ban completely the use of consumer data. Therefore, even though we may conclude that privacy protection in digital markets should be made either very hard or very easy, the ambiguity of the effects of price discrimination on individual consumers seems to call for a more nuanced approach to privacy protection. Different types would have diverging preferences when deciding whether to opt for privacy, either full or partial, or for full information disclosure. A regulation that allows each consumer to make an informed choice about disclosure (or concealment) of personal information in digital markets seems to point in the right direction. Indeed, this is the standard adopted in the General Data Protection Regulation (EU2016/679) which assigns to data subjects the right to consent with personal data collection and exploitation. The GDPR is centered around this empowerment of data subjects. To some extent, privacy is granted by default and individuals hold the right to directly enforce their personal privacy if needed. Data brokers have to obtain a clear and affirmative consent from users prior to collection of their personal data. However, many digital services require such consent as a condition *sine qua non* for accessing the service itself, which implies that not giving consent is not really a viable option for users when close substitutes are not at hand. The service terms of many platforms are shown on a take-it-or-leave-it basis, so that the user does not have a real choice.

Finally, the article tries to inform the policy debate on exclusive access to data and whether this is a first order competition policy concern. In the last part, I discuss the incentives of a data broker to induce an exclusive or a non exclusive downstream allocation of consumer data. The main contribution is to show that, under different conditions, both types of data allocation can arise. The seller holds either a partially informative dataset or a fully informative one. In the former case one firm gets an exclusive access to partial consumer data. In the latter case the seller induces both firms to acquire partial but different information about consumers. In terms of revenues, inducing partial data differentiation in equilibrium yields the highest payoff for the seller. No firm is entirely foreclosed and, moreover, the non exclusive allocation is efficient. However, this equilibrium further harms consumers. The implication is that partial exclusivity is not bad *per se*, neither for the uninformed rival nor consumers, but it is rather data differentiation that rises market power. From a competition policy point of view, the regulator should be more concerned about the type of data exchanged in the market rather than exclusive access to data. The analysis of a competitive upstream market structure is left for future research.

Appendix A

Proof of Lemma 1. Define $\Delta p = p_1 - p_2$. The fundamental equation for the analysis is

$$\tilde{y}(x) = (1 - x) - \frac{\Delta p}{2t} \tag{29}$$

In the above expression x is unknown to the players. Consider the extremes of the distribution of x, which is known to firms: the types with the lowest realization of x buy from firm 1 if $y < \tilde{y}(0)$, whereas those with the highest realization of x acquire 1's product if $y < \tilde{y}(1)$, with $\tilde{y}(1) < \tilde{y}(0)$. Since a low realization of x implies a preference for product 1 in the x dimension, relatively more consumers in the y dimension prefer firm 1 when x tends to zero, as captured by the negative unitary slope of $\tilde{y}(x)$. In particular: (i) when $\Delta p > 0$, $\tilde{y}(x)$ lies below $1 - x$, (ii) when $\Delta p < 0$, $\tilde{y}(x)$ lies above $1 - x$, whereas (iii) when $\Delta p = 0$ the indifferent consumers are located along the bisector of the unit square.

Consider firm 1 and fix p_2. Demand of firm 1 is necessarily zero whenever $p_1 \geq p_2 + 2t$ (i.e. $\Delta p \geq 2t$), whereas firm 1 captures the total mass of consumers for any $p_1 \leq p_2 - 2t$ (i.e.

$-\Delta p \geq 2t$). Let us focus on interior cases ($-2t < \Delta p < 2t$) in which both firms have positive demand[16]. Moreover, prices are restricted to be non negative (i.e. above or at least equal to the marginal cost). The sign of Δp gives rise to two distinct segments of the demand function (notice that demand is continuous at $\Delta p = 0$, as shown later). We look for a symmetric equilibrium in uniform prices.

Case I: $\Delta p > 0$. The locus of indifferent consumers lies below the bisector. Thus, the intercepts with the axis are respectively on the left y-axis and the bottom x-axis. The coordinates are:

$$(0, \hat{y}) \;\Leftrightarrow\; \hat{y} = \frac{2t + p_2 - p_1}{2t} = \frac{2t - \Delta p}{2t} \tag{30a}$$

$$(\hat{x}, 0) \;\Leftrightarrow\; \hat{x} = \frac{2t + p_2 - p_1}{2t} = \frac{2t - \Delta p}{2t}. \tag{30b}$$

When firm 1 is pricing above the rival, the demand of firm 1 corresponds to the area of the triangle determined by these coordinates. Therefore

$$D_1^I = \int_0^{\hat{x}} F(y \leq \tilde{y}(x)) f(x) dx = \frac{(2t - \Delta p)^2}{8t^2} \tag{31}$$

whereas firm 2's demand is just the complement to one

$$D_2^I = 1 - D_1^I = \frac{4t^2 + 4t\Delta p - (\Delta p)^2}{8t^2}. \tag{32}$$

It is easy to show that for both firms $\partial D_i / \partial p_i = -\frac{1}{2t}\left(\frac{2t - \Delta p}{2t}\right) < 0$ and that for $\Delta p > 0$

$$\frac{\partial^2 D_1}{\partial p_1 \partial p_1} \geq 0 \qquad\qquad \frac{\partial^2 D_2}{\partial p_2 \partial p_2} < 0. \tag{33}$$

Case II: $\Delta p < 0$. The locus of indifferent consumers lies above $1 - x$. The intercepts with the

[16] Therefore I simply denote with $\Delta p > 0$ cases in which $0 < \Delta p < 2t$ and with $\Delta p < 0$ cases in which $-2t < \Delta p < 0$.

axis are respectively on the right y-axis and the top x-axis. The coordinates are:

$$(1, \bar{y}) \;\Leftrightarrow\; \bar{y} = \frac{p_2 - p_1}{2t} = \frac{-\Delta p}{2t} \tag{34a}$$

$$(\bar{x}, 1) \;\Leftrightarrow\; \bar{x} = \frac{p_2 - p_1}{2t} = \frac{-\Delta p}{2t}. \tag{34b}$$

Notice that in this case it is easier to firstly derive the demand of firm 2 as the area of the triangle

$$D_2^{II} = \int_{\bar{x}}^{1} \left(1 - F(y \leq \tilde{y}(x))\right) f(x)dx = \frac{(2t + \Delta p)^2}{8t^2} \tag{35}$$

and then the demand of firm 1 as

$$D_1^{II} = 1 - D_2^{II} = \frac{4t^2 - 4t\Delta p - (\Delta p)^2}{8t^2}. \tag{36}$$

For both firms we find again that $\partial D_i / \partial p_i = -\frac{1}{2t}\left(\frac{2t+\Delta p}{2t}\right) < 0$, and

$$\frac{\partial^2 D_1}{\partial p_1 \partial p_1} < 0 \qquad \frac{\partial^2 D_2}{\partial p_2 \partial p_2} \geq 0. \tag{37}$$

Demand is continuous at the inflection point $p_1 = p_2$. Consider firm 1 and take the right and left limits of its price over the two demand segments characterized above:

$$\lim_{p_1 \to p_2^+} D_1^I(p_1, p_2) = \frac{1}{2} \qquad \lim_{p_1 \to p_2^-} D_1^{II}(p_1, p_2) = \frac{1}{2}. \tag{38}$$

Therefore, demand of firm i holding fixed the price of the rival j can be written as

$$D_i(p_i, p_j) = \begin{cases} 0 & p_i \geq p_j + 2t & \text{(39a)} \\ \dfrac{(2t + p_j - p_i)^2}{8t^2} & p_j < p_i < p_j + 2t & \text{(39b)} \\ \dfrac{1}{2} & p_i = p_j & \text{(39c)} \\ \dfrac{4t^2 + 4t(p_j - p_i) - (p_j - p_i)^2}{8t^2} & p_j - 2t < p_i < p_j & \text{(39d)} \\ 1 & p_i \leq p_j - 2t. & \text{(39e)} \end{cases}$$

Finally, it remains to show that the solution to firms' maximization problems is the same under both structures of demand. Consider case I (case II is symmetric). First order conditions are quadratic in prices, and taking them equal to zero yields two best replies for each firm. Recalling that prices are restricted to be non-negative and that pricing above the rival's price by more than $2t$ leads firms out the market, we can disregard degenerate best replies that violates these conditions. Consider firm 1. Taking the first order conditions with respect to p_1 yields

$$b_1(p_2) = \frac{p_2 + 2t}{3} \tag{40a}$$

$$b_1(p_2) = p_2 + 2t. \tag{40b}$$

It is immediate to see that the second response leads firm 1 out of the market, given that the firm sets a uniform price such that zero consumers are willing to buy the product for any price of the rival firm. Consequently, this best reply is eliminated.

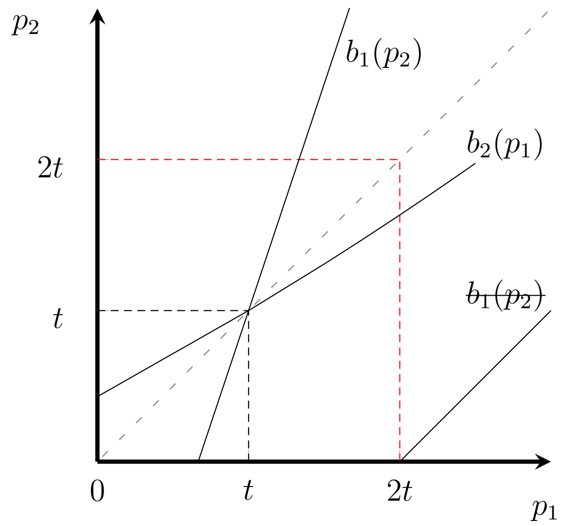

Figure 11: Best responses - not violating conditions on prices - in *Case I* ($\Delta_p > 0$).

Similarly, it is possible to show that the unique best response of firm 2 satisfying the conditions on prices is $b_2(p_1) = (2p_1 - 4t + z)/3$ where $z = \sqrt{p_1^2 - 4tp_1 + 28t^2}$ (notice that the second computed response of firm 2 lies entirely in the quadrant $(p_1(+), p_2(-))$, and it does not even appear in Figure (11). Solving the system of best responses yields a unique equilibrium in positive prices: $p_1^* = p_2^* = t$.

■

Proof of Lemma 2. When firm 1 charges $p_1(x)$ and firm 2 sets $p_2(y)$ the indifference condition writes

$$p_1(x) + tx + ty = p_2(y) + t(1-x) + t(1-y). \tag{41}$$

Given that firms target consumers asymmetrically, firm 1 considers as indifferent consumers those located along

$$y(x) = \frac{p_2(y) - p_1(x)}{2t} + (1-x) \tag{42}$$

and firm 2 considers the line

$$x(y) = \frac{p_2(y) - p_1(x)}{2t} + (1-y). \tag{43}$$

Notice that $y(x)$ and $x(y)$ draw the same line within the unit square, so that $D_1 = y(x)$ and $D_2 = 1 - x(y)$. Existence and uniqueness of a discriminatory price equilibrium is proved in steps.

Part 1. Consider firm 1. For each group x, the rival is setting a continuum of prices in the y dimension. We will show that what matters for the optimization problem of a firm is only the average price of the rival. Therefore, we firstly solve for competition in average prices; the requirement is that in equilibrium the optimal price schedules must be equal to the average price derived in this part of the proof

$$\int_0^1 p_1(\tilde{x})dF(\tilde{x}) = \bar{p}_1 \qquad \int_0^1 p_2(\tilde{y})dF(\tilde{y}) = \bar{p}_2. \tag{44}$$

Firm 1 maximizes

$$\tilde{\pi}_1 = \int_0^1 p_1(\tilde{x})D_1(p_1(\tilde{x}), \bar{p}_2, \tilde{x})dF(\tilde{x}) \tag{45}$$

and firm 2 maximizes

$$\tilde{\pi}_2 = \int_0^1 p_2(\tilde{y})D_2(p_2(\tilde{y}), \bar{p}_1, \tilde{y})dF(\tilde{y}). \tag{46}$$

Taking the first order conditions we get $\bar{p}_1 = \frac{\bar{p}_2+t}{2}$ and $\bar{p}_2 = \frac{\bar{p}_1+t}{2}$. The average price schedules are $\bar{p}_1 = \bar{p}_2 = t$.

Part 2. Now we derive the unique profit maximizing schedule satisfying the above constraint, taking into account the average discriminatory schedule of the rival[17]. Firms' final objective functions therefore are

$$\pi_1 = p_1(x)\left(\frac{\bar{p}_2 - p_1(x)}{2t} + (1-x)\right), \forall x \tag{47}$$

$$\pi_2 = p_2(y)\left(1 - \left(\frac{p_2(y) - \bar{p}_1}{2t} + (1-y)\right)\right), \forall y. \tag{48}$$

Equilibrium prices are $p_1^*(x) = t(\frac{3}{2} - x)$ and $p_2^*(y) = t(\frac{1}{2} + y)$, with average schedules indeed equal to the transportation cost. The market boundary is $y^*(x) = 1 - x$ (or equivalently $x^*(y) = 1 - y$). Profits are

$$\pi_1 = \int_0^{1-x}\int_0^1 p_1^*(x)\ f(x)f(y)dxdy = \frac{7t}{12} \tag{49a}$$

$$\pi_2 = \int_{1-y}^{1}\int_0^1 p_2^*(y)\ f(y)f(x)dydx = \frac{7t}{12}. \tag{49b}$$

Consumer surplus is equal to

$$CS = \int_0^{1-x}\int_0^1 (v - p_1^*(x) - tx - ty)\ f(x)f(y)dxdy + \int_{1-y}^{1}\int_0^1 (v - p_2^*(y) - t(1-x) - t(1-y))\ f(y)f(x)dydx = v - \frac{11t}{6}. \tag{50}$$

■

Proof of Lemma 3. When firm 1 charges a personalized price $p_1(x)$ and firm 2 a uniform price, the expression for the indifference line modifies to

$$y(x) = \frac{p_2 - p_1(x) + 2t(1-x)}{2t}. \tag{51}$$

As long as $|p_1(x) - p_2| < 2t$ for all x, the market boundary is interior, and the payoffs of the

[17] This procedure is equivalent to guessing linear schedules $p_1(x) = a - bx$ and $p_2(y) = \alpha + \beta y$, and plugging them into the optimization problem of the rival firm, taking the integral with respect to the information unobserved to that firm. The equilibrium values of $(a^*, b^*, \alpha^*, \beta^*)$ yields the same schedules derived in the two-step proof.

players are continuous. Players set prices simultaneously but player 2 is able to best reply only "on average" to the personalized price of player 1.

The objective functions of the two players are

$$\pi_1 = p_1(x)\left(\frac{p_2 - p_1(x)}{2t} + (1-x)\right), \forall x \tag{52}$$

and

$$\pi_2 = p_2\left(1 - \int_{x\in[0,1]}\left(\frac{p_2 - \bar{p}}{2t} + (1-x)\right)dx\right) \tag{53}$$

where $\bar{p}$ is the average of $p_1(x)$ over $x \in [0,1]$. Solving for the first order conditions we get

$$b_1(p_2, x) = \frac{p_2 + 2t(1-x)}{2} \qquad \text{and} \qquad b_2(\bar{p}) = \frac{\bar{p} + t}{2} \tag{54}$$

To get the optimal uniform price we plug the integral (i.e. the average) of the informed firm's best response into the above equation

$$p_2 = \frac{1}{2}\left(\int_{x\in[0,1]}\left(\frac{p_2 + 2t(1-x)}{2}\right)dx + t\right)dx \tag{55}$$

which yields $p_2^* = t$. The informed firm is always best responding by setting $p_1^*(x) = t(\frac{3}{2} - x)$ and the uninformed firm is best responding "on average". The equilibrium market boundary is equal to $y^*(x) = \frac{3-2x}{4}$ and final payoffs of the players are

$$\pi_2 = \int_{\frac{3-2x}{4}}^{1}\int_0^1 p_2^*\, f(x)f(y)dxdy = \frac{t}{2} \tag{56a}$$

$$\pi_1 = \int_0^{\frac{3-2x}{4}}\int_0^1 p_1^*(x)\, f(x)f(y)dxdy = \frac{13t}{24}. \tag{56b}$$

Consumer surplus is equal to

$$\begin{aligned} CS = &\int_0^{\frac{3-2x}{4}}\int_0^1 (v - p_1^*(x) - tx - ty)\; f(x)f(y)dxdy + \\ &\int_{\frac{3-2x}{4}}^{1}\int_0^1 (v - p_2^* - t(1-x) - t(1-y))\, f(x)f(y)dxdy = v - \frac{83t}{48}. \end{aligned} \tag{57}$$

■

Proof of Lemma 4. Notice that the difference in transportation costs $t(1-2x)+t(1-2y)$ is negative for $x+y>1$. Firstly we show that $p_2=0$ cannot be the equilibrium uniform price. When $p_2=0$, the price of firm 1 is given by (23) and the informed firm serves all consumers with $x+y\leq 1$, whereas the other half of the market buys from firm 2. The market boundary is $y(x)=1-x$, the bisector of the unit square. However, this cannot be an equilibrium provided that $\pi_2=0$, which implies that firm 2 has an incentive to deviate and to set a price larger than zero, serving less consumers but making a positive profit.

Consider therefore a candidate equilibrium $p_2>0$. Let us firstly provide a geometric argument that simplifies the problem. When $p_2>0$ the market boundary must necessarily lie above the bisector: the informed firm is now able to match the net utility guaranteed by firm 2 also for some consumers with $x+y>1$. The intuition is that in this region a positive price p_2 offsets the negative term $t(1-2x)+t(1-2y)$ that appears in (23), and the more p_2 increases, the more the market boundary switches in the north-east direction. Indeed, the boundary is always identified by $p_1(x,y)=0$ because, as long as the price of the informed firm is nonnegative, firm 1 has the advantage of "winning" the consumers at the margin by just matching the rival's offer. Therefore, the problem of characterizing the optimal uniform price reduces to the characterization of the set of locations (x,y) at which $p_1(x,y)=0$ when $p_2>0$.

Moreover, notice that for each line that is parallel to the market boundary, the coordinates (x,y) are such that the transportation costs are the same along the entire line: these are the isocost lines within the unit square. Given that $p_1(x,y)$ depends only on x, y and p_2, which is the same for all consumers, it must necessarily be that along the isocost lines the price set by the informed firm is the same[18]. We can therefore further simplify the problem by focusing on locations along the curve $y=x$. Thus, set $x=y=z$ where $\frac{1}{2}<z<1$. Then, for all $x+y=2z$, the price of the informed firm will be the same by construction[19]. From the indifference condition we can write

$$p_2=p_1(z,z)+t(4z-2). \tag{58}$$

[18]In a three dimensional space, the personalized price of the informed firm can be represented as a negatively sloped plane with domain $[0,1]^2$: if we cut the plane along the isocost lines, then we find the same price level for all (x,y) along each line.

[19]Basically, condition $x+y=2z$ is equivalent to $x=y=z$.

The marginal consumers are such that $p_1(z,z)=0$ and the demand of firm 2 is then geometrically characterized in Figure (12) as $D(z)=\frac{(2(1-z))^2}{2}$.

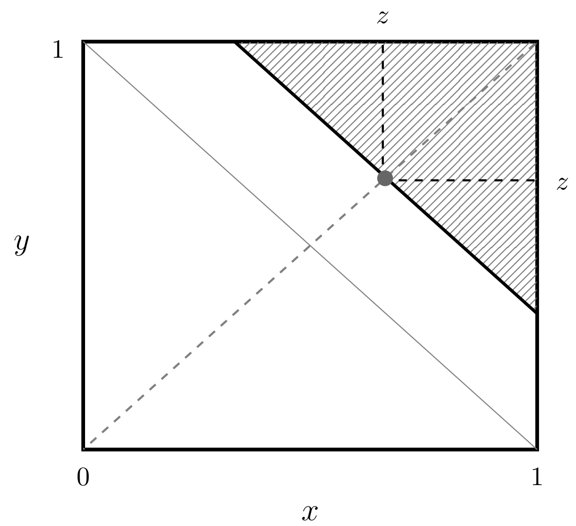

Figure 12: Demand of firm 2 when $p_2 > 0$.

The objective function of the uninformed firm is

$$\pi_2 = t(4z-2)\frac{(2(1-z))^2}{2}. \tag{59}$$

Formally the equilibrium is characterized by the following equations:

$$\begin{cases} x = y = z & (60a) \\ p_1(z,z) = 0 & (60b) \\ 8t(1-z)^2 - 4t(1-z)(2-4z) = 0 & (60c) \end{cases}$$

Solving with respect to z we get $z^* = \frac{2}{3}$. Substituting back into (58) yields $p_2^* = \frac{2t}{3}$. From (23) it follows that $p_1^*(x,y) = 2t\left(\frac{4}{3}-(x+y)\right)$, which is equal to zero along $y^*(x) = \frac{4}{3}-x$. Indeed, it is easy to verify that $x+y = 2\left(\frac{2}{3}\right)$: along this line consumers are indifferent between the two outlets at prices p_2^* and $p_1^*(x,y)$. Profits are equal to

$$\pi_2 = \int_{\frac{4}{3}-x}^{1}\int_{\frac{1}{3}}^{1} p_2^*\,dxdy = \frac{4t}{27} \tag{61a}$$

$$\pi_1 = \int_0^1\int_0^{\frac{1}{3}} p_1^*(x,y)\,dxdy + \int_0^{\frac{4}{3}-x}\int_{\frac{1}{3}}^{1} p_1^*(x,y)\,dxdy = \frac{62t}{81}. \tag{61b}$$

■

Proof of Lemma 5. Both firms are partially informed about the x dimension. Suppose now that the seller can further offer data on y to firm 1, which gets full information whereas the rival remains only partially informed. Consumer (x, y) is always served by firm 1 when

$$p_1(x, y) + tx + ty = p_2(x) + t(1 - x) + t(1 - y) \tag{62}$$

and the informed firm optimally sets its personalized price at each location to match the rival's price

$$p_1(x, y) = p_2(x) + t(1 - 2x) + t(1 - 2y). \tag{63}$$

The objective function of firm 2 is defined for each x as

$$\pi_2 = p_2 \left(1 - \left(\frac{p_2}{2t} + (1 - x)\right)\right) \tag{64}$$

and taking the first order condition yields the personalized price $p^*(x) = tx$. In turn, firm 1 sets $p_1^*(x, y) = t(2 - x - 2y)$. The market boundary is interior ($y^*(x) = \frac{2-x}{2}$) and profits are equal to

$$\pi_1 = \int_0^{\frac{2-x}{2}} \int_0^1 p_1^*(x, y)\ f(x)f(y)dxdy = \frac{7t}{12} \tag{65a}$$

$$\pi_2 = \int_{\frac{2-x}{2}}^1 \int_0^1 p_2^*(x)\ f(x)f(y)dxdy = \frac{t}{6}. \tag{65b}$$

■

References

Acquisti, A., C. Taylor, and L. Wagman (2016): "The economics of privacy," *Journal of Economic Literature*.

Acquisti, A. and H. R. Varian (2005): "Conditioning prices on purchase history," *Marketing Science*, 24, 367–381.

Armstrong, M. (2006): "Price discrimination," *MPRA Paper*.

Baye, I., T. Reiz, and G. Sapi (2018): "Customer Recognition and Mobile Geo-Targeting," *DICE Discussion Paper*.

Belleflamme, P., W. M. W. Lam, and W. Vergote (2017): "Price Discrimination and Dispersion under Asymmetric Profiling of Consumers," *AMSE Working Papers*.

——— (2020): "Competitive Imperfect Price Discrimination and Market Power," *Marketing Science*, 39, 996–1015.

Belleflamme, P. and W. Vergote (2016): "Monopoly price discrimination and privacy: the hidden cost of hiding," *Economics Letters*, 149, 141–144.

Bhaskar, V. and T. To (2004): "Is perfect price discrimination really efficient? An analysis of free entry," *The RAND Journal of Economics*, 35, 762–776.

Bounie, D., A. Dubus, and P. Waelbroeck (2018): "Selling Strategic Information in Digital Competitive Markets," *CESifo Working Paper, No 7078*.

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

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

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

Corts, K. S. (1998): "Third-Degree Price Discrimination in Oligopoly: All-Out Competition and Strategic Commitment," *The RAND Journal of Economics*, 29, 306–323.

Dube, J., Z. Fang, N. Fong, and X. Luo (2017): "Competitive price targeting with smartphone coupons," *Marketing Science*, 36, 944–975.

Dube, J. and S. Misra (2017): "Scalable price targeting," Discussion paper, University of Chicago Booth School of Business.

Duch-Brown, N., B. Martens, and F. Mueller-Lang (2017): "The economics of ownership, access and trade in digital data," Tech. rep., Digital Economy Working Paper 2017-01; JRC Technical Reports.

Fudenberg, D. and J. M. Villas-Boas (2006): "Behavior-based price discrimination and customer recognition," in *Handbook on Economics and Information Systems*, ed. by T. Hendershott, Elsevier.

Ganuza, J. and G. Llobet (2018): "Personalized prices in the digital economy," in *Economic analysis of the digital revolution*, ed. by J. Ganuza and G. Llobet, FUNCAS Social and Economic Studies, 117–142.

Goldfarb, A. and C. Tucker (2012): "Shifts in privacy concerns," *The American Economic Review*, 102, 349–353.

——— (2019): "Digital economics," *The Journal of Economics Literature*, 57, 3–43.

Hannak, A., G. Soeller, D. Lazer, A. Mislove, and C. Wilson (2014): "Measuring price discrimination and steering on e-commerce web sites," in *Proceedings of the 2014 conference on internet measurement conference*, ACM, 305–318.

Hindermann, C. M. (2018): "Price Discrimination in Online Retail," ZBW – Leibniz Information Centre for Economics, Kiel, Hamburg.

Hotelling, H. (1929): "Stability in competition," *The Economic Journal*, 39, 41–57.

Hupperich, T., D. Tatang, N. Wilkop, and T. Holz (2018): "An Empirical Study on Online Price Differentiation," in *Proceedings of the Eight ACM Conference on Data and Application Security and Privacy*, 76–83.

Jehiel, P. and B. Moldovanu (2000): "Auctions with downstream interaction among buyers," *The RAND Journal of Economics*, 31, 768–791.

Jentzsch, N., G. Sapi, and I. Suleymanova (2013): "Targeted Pricing and Customer Data Sharing Among Rivals," *International Journal of Industrial Organization*, 31, 131–144.

Kim, J.-H., L. Wagman, and A. L. Wickelgren (2018): "The Impact of Access to Consumer Data on the Competitive Effects of Horizontal Mergers and Exclusive Dealing," *Journal of Economics & Management Strategy*, 28, 373–391.

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

Liu, Q. and J. Shuai (2013): "Multi-dimensional Price Discrimination," *International Journal of Industrial Organization*, 31, 417–428.

Montes, R., W. Sand-Zantman, and T. Valletti (2018): "The value of personal information in markets with endogenous privacy," *Management Science*, Articles in Advance, 1–21.

Neven, D. and J.-F. Thisse (1987): "Combining horizontal and vertical differentiation: The principle of max-min differentiation," .

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

Shaked, A. and J. Sutton (1982): "Relaxing price competition through product differentiation," *Review of Economic Studies*, 3–13.

Shy, O. and R. Stenbacka (2015): "Customer privacy and competition," *Journal of Economics & Management Strategy*, 25, 539–562.

Stole, L. (2003): "Price discrimination and imperfect competition," .

Stucke, M. E. (2018): "Should we be concerned about Data-opolies?" *Geo. L. Tech. Rev.*, 275.

Taylor, C. (2004): "Consumer privacy and the market for customer information," *The RAND Journal of Economics*, 35, 631–650.

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

Thisse, J.-F. and X. Vives (1988): "On the strategic choice of spatial price policy," *The American Economic Review*, 78, 12–137.

Download ZEW Discussion Papers from our ftp server:

http://ftp.zew.de/pub/zew-docs/dp/

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim
ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de