CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 13, Issue 8, No. 352 – August 2015

The value of personal information in markets with endogenous privacy

Rodrigo Montes, Wilfried Sand-Zantman, and Tommaso Valletti

The value of personal information in markets with endogenous privacy*

Rodrigo Montes[†], Wilfried Sand-Zantman[‡], and Tommaso Valletti[§]

July 21, 2015

Abstract

This paper investigates the effects of price discrimination on prices, profits and consumer surplus, when one or more competing firms can use consumers' private information to price discriminate and consumers can pay a privacy cost to avoid it. While a monopolist always benefits from higher privacy costs, this is not true in the competing duopoly case. In this last case, firms' individual profits are decreasing while consumer surplus is increasing in the privacy cost. Finally, under competition, we show that the optimal selling strategy for the owner of consumer data consists in dealing exclusively with one firm in order to create maximal competition between the winner and the loser of data. This brings inefficiencies, and we show that policy makers should concentrate their attention on exclusivity deals rather than making it easier for consumers to protect their privacy.

1 Introduction

This paper studies how customer information and privacy affect the discriminatory behavior of a firm facing imperfect competition. In particular, we investigate the effects of price discrimination on prices, profits and consumer surplus, when firms can use consumers' private information to discriminate and consumers can potentially avoid being discriminated. This analysis allows us to draw some interesting conclusions on the value of customers' information, that is, the willingness-to-pay of firms to acquire data about customers.

*We thank Florian Baumann, Paul Belleflamme, Steffen Hoernig, Morten Hviid, Bruno Jullien, Marc Lebourges, Yassine Lefouili, Martin Peitz, Olga Slivkó, Oz Shy, Rune Stenbacka, Chia-Yu Tsai, Catherine Tucker, J. Miguel Villas-Boas, Liad Wagman, as well as seminar participants at DICE, Florence, Liège, Mannheim and UEA. Sand-Zantman gratefully acknowledges Orange for financial support under the IDEI/Orange convention. Montes and Sand-Zantman also acknowledge the support from the Agence Nationale de la Recherche, grant ANR- 12-BSH1-0009.

[†]Toulouse School of Economics, rodrigomontesm@gmail.com

[‡]Toulouse School of Economics and ESSEC Business School, wsandz@tse-fr.eu

[§]Imperial College London and University of Rome II, t.valletti@imperial.ac.uk

In the past few years, price discrimination on the Internet has been documented many times.[1] Probably the most (in)famous case occurred already in 2000, when a customer complained that, after erasing his cookies, he observed a lower price for a DVD on *Amazon.com*.[2] More recently, the Wall Street Journal (2012a) reported that the travel agency *Orbitz Worldwide* was showing more expensive hotel offers to Mac users than to PC users. A similar practice has been conducted by *Staples.com*: The Wall Street Journal (2012b) uncovered that this site was displaying different prices once the location of the potential buyer had been identified. Other firms employing similar methods were identified by the newspaper: *Discover Financial Services*, *Rosetta Stone* and *Home Depot*. Another firm, *Office Depot*, admitted to using "customers' browsing history and geolocation to vary the offers and products it displays to a visitor to its site" (Wall Street Journal 2012b).

The widespread use of personal data in marketing has created a large market for consumers' personal information, which is estimated to be a $156-billion-a-year industry, with over 4000 *data brokers*. Some of these firms are large data aggregators, while others deal with specific types of consumers (Pasquale, 2014). Individual attributes are increasingly regarded as business assets that can be used to target services or offers and to provide relevant advertising. The aim of this article is to study the role of information and privacy in markets. More precisely, we would like to understand the value of consumers' information, both for the firms selling this information and for the firms buying this information. We also study whether consumers should be given more or less control on the information that can be used by firms to adjust their commercial offers.

To do so, we use a model where firms can acquire information on consumers' taste and then propose a set of different personalized prices. We assume that there are two broad groups of consumers. For the first group, individual information cannot be extracted, as in the standard Hotelling model with uniform prices. For the second group, information is potentially available, i.e., this information can be bought from a data supplier. One possible interpretation for this setting is that the first group is made of consumers that are not active on-line, or not yet active because they are newcomers. The second group comprises consumers that are active, leave many traces of their activity, and so need to engage *ex post* in costly actions to erase those traces. This means that, if a firm does not know a consumer's type after having bought some information, the firm cannot tell whether this consumer is a less active consumer or an active one who erased his browsing history. The way active consumers can protect their privacy is by paying a cost to *disappear* from the firm's database. We can interpret this *privacy cost* as the effort consumers spend to conceal their actions online. As a benchmark, we consider a

[1] See, e.g., Mikians et al. (2012, 2013) for a systematic data collection and Shaw and Vulkan (2012) for an experimental approach.

[2] `https://en.wikipedia.org/wiki/Amazon.com_controversies#Differential_pricing`. See also Wall Street Journal (2012b)

case where privacy is not allowed, so the information on old consumers is available on a data exchange market. If the firm is a monopolist, it will fully exploit this situation and get all the surplus from every old consumer while behaving as an standard non-informed monopoly for the group of the new consumers. Under duopoly, the outcome depends on the informational structure, that is, on which firms acquire some information about old consumers. We show that, at the equilibrium, the data supplier chooses to sell the data about consumers to only one firm. Then, the uninformed firm will set a price lower than the Hotelling price while the informed firm chooses to be less aggressive for new consumers. As far as the old consumers are concerned, the informed firm will match what these consumers could gain by buying from the other firm, but will not capture the whole market.

Next we study the case where consumers can pay for privacy to avoid being in the database. In this case, there are two markets *de facto*, one with all the new consumers and the old consumers that paid the privacy cost, and another comprising the old consumers that choose not to pay the privacy cost. On the first market, called the *anonymous market*, a monopolist firm chooses to set a price increasing with the number of old consumers in this market, or equivalently decreasing with the cost of privacy. On the second market, this firm captures the whole surplus from consumers. As increasing the privacy cost makes the new consumers on the anonymous market better-off but the old consumers in the market worse off, it is ambiguous from a consumers' point of view. But this always benefits the firm whose profit increases with the privacy cost. In the duopoly case, we show that the data supplier still chooses to deal exclusively with one firm, and that the price of information is U-shaped in the privacy cost. The investigation of the market prices in this case reveals that, as long as the privacy cost is not too large, the prices on the anonymous will be higher than in the Hotelling case, and they decrease with the privacy cost. Indeed, for small privacy cost, most of the old consumers choose to buy on the anonymous market. The uninformed firm can therefore obtain a large profit by focusing on this market and setting a high price. When the privacy cost increases, the size and the taste characteristics of consumers on the anonymous market change, making both firms more aggressive. In this setting, larger privacy costs reduce competition that leads both firms' individual profits to decrease in the privacy cost and consumer's surplus to increase.

We then extend our analysis in two directions. First, we relax the assumption that the data supplier can commit *ex ante* to a particular selling strategy. Relaxing the commitment of the data seller forces consumers to form beliefs on the information structure *ex post* - on which firms will be able to price discriminate. We show that, in spite of this change, the conclusions that the data supplier chooses to deal exclusively with one firm still holds. The second extension discusses the case where the market sizes of old and new consumers differ. We show that varying the relative size of the market changes the pricing strategies of the firms, with a tougher competition when the (relative) mass of new consumers is small. We also show that our result on exclusive dealing on the data market is robust to changing the relative masses of consumers.

Privacy is not a new topic in economics. Already in the 80's, Posner (1981) and Stigler (1980) argued that privacy could decrease efficiency by allowing individuals to hide some characteristics. The disclosure of information, and the induced price discrimination (as shown by Thisse and Vives, 1988), is known to have different effects depending on the perspective chosen. It can increase welfare, given that it allows the firm to sell to consumers with lower valuations (see Tirole, 1988, on the welfare effect of price discrimination). Ex ante however, early information disclosure destroys insurance possibilities hindering efficient risk-sharing (Hirshleifer, 1971). In our paper, we assume that the market is always fully covered, so price discrimination only affects the distribution of surplus among the agents, but not efficiency directly. We can still conduct a meaningful welfare analysis because of the existence of transportation costs and privacy costs in our model, so that the privacy question we are after still generates tangible economic trade-offs.

In the last decade, motivated by the rise of online markets and Big Data, there has been a revival of the research on privacy, with three main directions.[3] The first one is the study of behavior-based price discrimination (see, e.g., Fudenberg and Tirole, 2000; Esteves, 2010; Fudenberg and Villas-Boas, 2012; and Villas-Boas, 1999, 2004).[4] These papers look at dynamic pricing situations where competition is repeated, and firms use the past behavior of consumers to infer their taste and price accordingly. When each consumer is biased toward one particular firm, the first-period choices tell who is the fan of which firm. This leads firms to set high prices to their fans and low prices to the fans of the other firms. Our model is different in that we do not assume a period in which there is no information. By doing so, we abstract from the process of creating information, focusing rather on the privacy actions undertaken by consumers and on the way this information is sold to the firms. The second strand of the literature has tackled the privacy issued more directly. Taylor (2004) and Acquisti and Varian (2005) consider repeated sales by a monopolist at each period, assuming inter-temporal correlation between tastes. Both papers show that, while using past information can benefit the firm when consumers are myopic, it is not the case anymore with rational consumers. Following the Coase (1972) conjecture argument, using past information deters consumers to consume in the first place and hurts the firm's profit. The analysis of information disclosure in a more competitive setting has been more limited so far, with the exception of Taylor and Wagman (2014) who show that full privacy could be detrimental to consumer wel-

[3]For a comprehensive review of the literature on the economics of privacy see Acquisti et al. (2015). See also Lane et al. (2014) for an accessible approach to the policy issues. For empirical work on privacy, see Goldfarb and Tucker (2011 and 2012).

[4]The marketing literature on couponing, market segmentation, and consumer addressability is closely related (see Chen et al., 2001; Chen and Iyer, 2002).

fare, but that this result depends on the details of the model used. We endogenize the privacy choice by the consumers, and also insist on the optimal selling strategy by the data owner, which has not been analyzed before in the context of the privacy literature.[5] The last strand of the literature shares with us the idea that consumers may decide how much information to reveal. In Conitzer et al. (2012) or Koh et al. (2013), consumers face a monopoly and they may chose to reveal who they are. In the first contribution, the monopoly cannot commit not to use past information, as in the Coase conjecture. When the cost of protecting its privacy is high, consumers refrain from paying this cost but also from consuming in the first place. Lowering the cost of anonymity is therefore similar to an increase in the commitment power of the monopolist, and increases the firm's profit. Our paper differs in that the cost of privacy only reduces the monopolist's ability to tailor prices and it does not affect commitment. In the second contribution by Kho et al. (2013), the focus is put on voluntary profiling, where the refusal to participate in such a program leads the monopolist to propose high prices. Voluntary profiling allows to reduce search costs to find the ideal product but generates unsolicited adds. Both papers focus on the monopoly case, and therefore do not discuss the strategy of the seller of information. In Casadesus-Masanell and Hervas-Drane (2015), consumers can also choose the amount of personal data they provide to competing platform firms, this information increasing the quality of the product. Additionally, firms get revenue from selling consumers' data, and consumers value privacy. We instead separate the market for goods and the market for information, and model differently the way consumers choose their privacy. Chen et al. (2001) shed some light on the market for information. In their model, firms compete to attract non-loyal customers. The authors find that a data seller may sell data non-exclusively if targetability (the firm's ability to distinguish loyal from non-loyal consumers) is sufficiently low. This result differs from our own and will be discussed in a subsequent section.

The paper is organized as follows. The next section presents the model. Section 3 solves the case where consumers are not allowed to have privacy. Section 4 introduces endogenous privacy. Section 5 solves for extensions to the model. Results are discussed in section 6 that concludes the paper.

2 The model

There are three different types of agents: consumers, two competing sellers, and a firm - a data supplier - that can collect information about consumers. We start by describing preferences of consumers and the two competing sellers. There is a continuum of customers,

[5]For a related literature on intermediary gatekeepers, see, e.g., Wathieu (2002), Pancras and Sudhir (2007), and Weeds (2015). Recently, Chiou and Tucker (2014) have studied if larger quantities of consumer historical data provide competitive advantages to Internet search firms.

each of them with a unit demand. They receive a utility $v > 0$ if they buy the product and nothing otherwise. Consumers have horizontal preferences θ uniformly distributed on $[0, 1]$. Let p_i be the price a firm i charges for its product ($i = 1, 2$). Each firm's marginal cost is normalized to zero. Additionally, consumers have to pay a linear transportation cost $t > 0$ to buy the good. Therefore, if firm i is located at point $x \in [0, 1]$, a consumer θ makes a net utility $v - p_i - |x - \theta|t$. When two firms are competing on the market, consumers buy from the firm that gives them the highest possible utility. Throughout the paper we assume that $v \geq 2t$. By doing so we ensure that the gross utility is high enough so that all consumers would consume even if the price was set by a monopolist seller.[6]

Let us have two sets of consumers like the one just described, which differ in terms of the information that can be obtained by firms on consumer preferences. We will refer to the first set as the *new consumers*. For them, there is no way to obtain any detailed information, so firms can only offer basic or uniform prices that do not depend on θ. The second set consists of *old consumers*. For these consumers, some information can potentially be obtained by the firms. Therefore firms could make tailored offers than depend on θ in a way that we specify below. Both sets have a total mass of one.[7]

To interpret the two sets, consider an online retailer that supplies two distinct groups. On the one hand, there are newcomers: those for whom there is not enough data yet. These consumers do not necessarily need to be young, only new to the Internet. These consumers alternatively could have preferences very different from what the retailer has to offer and therefore, even if they are active on-line, that information will not help the firm. If, say, the retailer in question sells sporting goods, even the most detailed data about consumers who never visited any sports websites or never shopped anything sport related will not be useful for the firm. On the other hand, there are old consumers, i.e., people who have been actively using the Internet in the past, visiting websites, shopping, leaving reviews, and so forth. These data are very likely to be informative about the consumer preferences for the goods or services the retailer sells.

Old consumers can pay a cost to avoid detection by the firms. We call this the *privacy cost* and denote it by $c \geq 0$. This cost represents the actions consumers take, or the payments they make, to prevent any firm or third party from holding personal data about them. For example, as discussed earlier, this cost could represent how hard it is to erase the cookies after shopping on the Internet or visiting a website. In turn, regulators could make the cost smaller by imposing a full disclosure policy on the use of cookies.[8] Note that this cost could also be made larger if the firms were allowed to trade customers' data, so that consumers would need to request potentially many websites

[6]This assumption simplifies the analysis by reducing the number of cases to be studied without loss of economic insight.

[7]In an extension, we discuss the impact of changing the relative size of the two sets (see section 5.2).

[8]In Europe, see for example Directive 2002/58/EC.

to erase their data.[9] There is also evidence that some consumers incur in a monetary cost to protect their privacy. For example, *Reputation.com* charges individuals \$9.95 per month to remove personal data from on-line data markets. Or, another company, *Private Internet Access*, charges \$3.33 per month for a VPN (virtual private network) connection.

Throughout the paper, we set $c \leq t$, which generates the economically more interesting cases (when $c > t$ we can show that no consumer would ever pay the privacy cost). We assume that c affects the utility directly: if a consumer θ buys the good at p_i from a firm located at x and also pays the privacy cost, his utility is $v - p_i - |x - \theta|t - c$. For the rest of the paper, c will not depend on the consumer type in order to explore the effects of policies that facilitate or impede privacy on the Internet in general.

Firm i can price differently the two groups. A basic price will be offered to all consumers it does not recognize. We call this the *anonymous market*. This group is comprised of every new consumers plus the old consumers who paid the privacy cost and thus cannot be detected. The basic price will be based on the firm's belief about the average willingness to pay in the anonymous market.

For the consumers the firm does recognize, it will make tailored offers, based on each consumer's location θ. The consumers on this *personalized market* is comprised of the subset of the old consumers who did not pay the privacy cost.

The last agent in our model is the firm capable of collecting and selling information about consumer preferences. This information can be thought as personal taste, willingness to pay, brand fidelity, etc. We do not model how the data supplier has acquired its data in the first place. We will rather investigate how it chooses to sell the data, and the way the data is used by firms operating on the good market. On the internet, for instance, a company like Bluekai plays the role of the data supplier, by centralizing information (on consumers' past experience and on their observed characteristics) which is auctioned off among retailers. If a retailer wants to start an advertising campaign, it must go to this *data exchange market* to get access to data on consumers it is interested in, and make a personalized offer.

However, consumers are not passive and can take actions to avoid being in the database. They may have many reasons to do so. First, privacy is a good in itself, and some people consider that it is not right for a company to compile their personal information. Second, some criminal or shameful activities could be actively concealed by some people. Third, consumers might be aware of the possibility that a firm could set different prices for different people, based on personal data. As seen in the introduction, this has been documented and it is not always in favor of consumers. It is this last

[9]Firms could also facilitate privacy voluntarily. For example the *Network Advertising Initiative*, an association of advertisers, encourages the adoption of opting-out measures across its members, so that consumers can more easily avoid being tracked and targeted. However, there is evidence that some participants are not respecting that consumers opted-out (Mayer, 2012).

particular aspect that our model captures.[10]

3 No privacy

As a benchmark, we first study the case where consumers cannot pay for privacy. This is equivalent to saying that the cost c is very large. First we analyze the monopoly case, followed by the case where two firms compete. In this section, the set of new consumers is equivalent to the anonymous market, and the set of old consumers is equivalent to the personalized market. We therefore use the terms interchangeably.[11] In this section and throughout the paper we solve for Perfect Bayesian Equilibria in pure strategies.

3.1 Monopoly

Consider a monopolist seller located at 0, both for the sets of old and new consumers.[12] The timing of the model goes as follows:[13]

1. The data supplier posts a price T for the data.
2. The firm chooses whether or not to purchase the data.
3. The firm chooses the basic price p.
4. The firm can offer tailored prices.
5. Consumers buy and consume.

Let us assume that the monopolist bought the information in period 2 from the data seller. The monopolist's market share among the new consumers would be all the consumers $\theta \in [0, \theta_1]$ such that $v - p - \theta_1 t = 0$. This means that θ_1 represents the consumer indifferent between buying the good and getting the outside option of zero:

$$\theta_1 = \min\{\frac{v-p}{t}, 1\}.$$

For the old consumers, the firm offers tailored prices $p(\theta)$. Given that the outside option is equal to zero, the tailored prices can be as high as to capture all the surplus. Hence:

$$p(\theta) = v - t\theta,$$

[10] We emphasize that the only reason for paying the privacy cost in our model is anticipating a tailored price higher than the basic price for a particular consumer θ. We do not include any intrinsic benefits of privacy. However, c can be thought as a *net* privacy cost that includes an intrinsic value for privacy. In this case, the comparative statics remain the same.

[11] This will not be the case in section 4, where we introduce privacy.

[12] The results of this section would be qualitatively similar if the firm was located in the middle of both markets.

[13] Merging stages 2 and 3 would make no difference in this case, but we keep this distinction for consistency with the competition case that we analyze in Section 3.2.

where $v \geq 2t > t$ guarantees that that $p(\theta) > 0$. Therefore, the monopolist solves

$$\max_{p \geq 0} \int_0^{\theta_1} p d\theta + \int_0^1 p(\theta) d\theta$$

$$\text{s.t.} \quad 0 \leq \theta_1 \leq 1.$$

With the assumption that $v \geq 2t$, we have that $p^* = v - t$ and $\theta_1 = 1$. The price p^* is the equilibrium basic price and it is only offered to the new consumers. Old consumers are offered the tailored price $p(\theta)$ and are left with no surplus. These results imply that both markets are fully covered.

Total profits π are given by the sum of the profits obtained from selling to old consumers π^O and the profits from selling to new consumers π^N. The equilibrium values are

$$\pi^O = \int_0^1 (v - t\theta) d\theta = v - \frac{t}{2} \quad \text{and} \quad \pi^N = \frac{v - p^*}{t} p^* = v - t.$$

Hence total profits are $\pi = 2v - \frac{3t}{2}$. Note that without consumers' information, the monopolist would make a profit $2(v - t)$, i.e., twice π^N. Therefore, the maximum price the firm is willing to pay to acquire the dataset from the data supplier is

$$T^* = \pi - 2(v - t) = \frac{t}{2}.$$

The surplus for old consumers CS^O and for new consumers CS^N are given by

$$CS^O = \int_0^1 (v - (v - t\theta) - t\theta) d\theta = 0 \quad \text{and} \quad CS^N = \int_0^1 (v - (v - t) - t\theta) d\theta = \frac{t}{2}.$$

The above equation implies that total consumers' surplus is simply $CS = CS^N = \frac{t}{2}$.

3.2 Competition

We now study the case where 2 firms, A and B, compete. We assume that in both sets the locations are fixed, with firm A located at $\theta = 0$ and firm B at $\theta = 1$. We look at different cases depending on which firms have information. We are interested in the competitive effects of information when access to it is potentially asymmetric in the market. In such a case, we will assume - without loss of generality - that only A has information about the consumers.

As in the previous section, information is only available for old consumers and no information can be obtained about the new consumers. The timing for the competitive case is as follows:[14]

[14] In this section, sequentiality in the choice of basic and tailored prices is necessary to have an equilibrium in pure strategies. A will tailor a price that is optimally a function of B's basic price, hence B can always reduce its price and gain a large portion of the market. Sequentiality helps us ensure that

1. The data supplier posts a price T for the data.
2. Firms A and B choose whether or not to purchase the data.
3. Firms A and B choose basic prices p_A and p_B.
4. Firms A and B can offer tailored prices.
5. Consumers buy and consume.

In stage 2, firms can buy information from an upstream data supplier. We assume that only then information can be sold, that is, the data supplier commits to selling ex ante, prior to competition between sellers.[15]

Note that the duopoly setting can be seen as one where a firm with customers' information is able to offer a rebate to the consumers on the competitor's turf.[16] For example, this could be understood as the ability of the most informed firm to send targeted coupons to some consumers.

As a benchmark, consider a situation where neither firm has information. This is equivalent to solving a Hotelling model in two identical markets. In each of them, prices p_A and p_B are the same and equal to the transportation cost t, which implies that the market is evenly split between firms.[17] Then, profits for each firm from selling to old and new consumers are $\frac{t}{2}$ while total profits for each firm are $\pi_A = \pi_B = t$.

Finally, consumers' surpluses are given by

$$CS^O = CS^N = \int_0^{1/2} (v - t - t\theta)d\theta + \int_{1/2}^{1} (v - t - t(1-\theta))d\theta = v - \frac{5t}{4}.$$

Hence, total consumers' surplus $CS = 2v - \frac{5t}{2}$.

3.2.1 Both firms have information

The case where both firms have information has been studied by Taylor and Wagman (2014). The only difference with respect to our paper lies in that we consider two sets of consumers and two types of prices.

a pure-strategy equilibrium will exist, because the optimal tailored price will always exist. We prove in the Appendix that a pure-strategy equilibrium does not exist if prices are chosen simultaneously. This methodological point is often encountered but not formalized in the literature, so it is worth making it a stand-alone result.

[15]We will relax this assumption in an extension where we allow for interim sales, that is, firms can buy information at later stages (see section 5.1).

[16]See Fudenberg and Tirole (2000). Note that, in contrast, in the monopoly case studied before there was no poaching and price discrimination simply allowed the firm to fully extract rent from consumers.

[17]See Belleflamme and Peitz (2010).

Proposition 1. *Assume both firms A and B buy the information and consumers cannot pay for privacy. Basic prices are:*

$$p_A = t \quad and \quad p_B = t,$$

while tailored prices are:

$$p_A(\theta) = \max\{t(1-2\theta), 0\} \quad and \quad p_B(\theta) = \max\{t(2\theta-1), 0\}.$$

Profits are $\pi_A = \pi_B = \frac{3t}{4}$ and consumer surplus is given by $CS = 2(v-t)$.

Proof. See Taylor and Wagman (2014). □

Taylor and Wagman (2014) show that if both firms have information and compete simultaneously only for old consumers, competition intensifies to the point that both firms make lower profits than they would by charging basic prices only.

Note that, compared to the case without information, profits decrease while consumers' surplus increases. This is due to the fact that competition reduces the potential for rent extraction when both firms hold information about consumers. It is as if both firms where competing *à la* Bertrand for each consumer. The goods are differentiated so both prices are equal to marginal cost only for the consumer in the middle located at equal distance both firms. For all the other consumers, one of the firm is preferred so it can set a strictly positive price and still sell its products.

3.2.2 Only firm A buys information

In this section we focus on the case where information is only sold to A. Later, we will address the conditions under which that information is indeed acquired. Note that the outside option for the consumers is not zero,[18] but the utility obtained from buying from B instead. Buying from A leads to a utility level given by

$$v - t\theta - p_A,$$

while buying from B leads to

$$v - t(1-\theta) - p_B,$$

where p_A and p_B are A and B's prices respectively. In the anonymous market, market shares $[0, \theta_1]$ and $[\theta_1, 1]$ are determined by the indifferent consumer between buying from A or B:

$$\theta_1(p_A, p_B) = \frac{1}{2} + \frac{p_B - p_A}{2t}.$$

[18]We have guaranteed this by assuming $v \geq 2t$, so that even consumers closer to A would make strictly positive surplus if they were to buy from B at the equilibrium price.

In the personalized market, firm A offers in stage 4 a tailored price $p_A(\theta)$ that makes consumers indifferent between accepting and buying B's product.[19] This implies:

$$p_A(\theta) = p_B + (1 - 2\theta)t.$$

We define the last consumer buying from A by θ_2^0, where θ_2^0 is such that $p_A(\theta) = 0$. Using the above expression for $p_A(\theta)$, this leads to

$$\theta_2^0(p_B) = \frac{p_B + t}{2t}.$$

We now consider the third stage of the game where both firms choose their basic prices. Firm A's profits are given by:

$$\pi_A = \int_0^{\theta_1} p_A d\theta + \int_0^{\min\{\theta_2^0,1\}} p_A(\theta) d\theta,$$

and B's profits are:

$$\pi_B = \int_{\theta_1}^{1} p_B d\theta + \int_{\min\{\theta_2^0,1\}}^{1} p_B d\theta.$$

There are two possible cases. In the former, B operates in the personalized market, which means that prices are such that some old consumers would buy from B and pay p_B. This corresponds to the case where $\theta_2^0 < 1$. When instead $\theta_2^0 \geq 1$, firm B does not operate in the personalized market. Because $p_A(\theta)$ does not depend on p_A, A's reaction function is always the same, regardless of the case. This implies that it is B that decides on the outcome of the game based on its own profits. B can choose either to give away every costumer in the personalized market by setting a high price, or to lower the price to keep some consumers on this market.

In the next proposition, we show that firm B finds it optimal to operate in the personalized market, and therefore chooses a low basic price. In this sense, we say that competition is more aggressive.

Proposition 2. *Assume only firm A buys the information and consumers cannot pay for privacy. Then the equilibrium prices on the anonymous market are:*

$$p_A = \frac{6t}{7} \quad \textit{and} \quad p_B = \frac{5t}{7},$$

profits are:

$$\pi_A = \frac{54t}{49} \quad \textit{and} \quad \pi_B = \frac{25t}{49},$$

and consumer surplus is given by $CS = 2v - \frac{110t}{49}$.

[19] Under indifference, consumers buy from the informed firm. In this case it is firm A.

Proof. We verify that B finds it profitable to set a price such that it operates in the personalized market ($\theta_2^0 < 1$). Remark that A's reaction function is the same in every case and it is obtained from

$$\arg\max_{p_A} \quad \theta_1(p_A, p_B)p_A.$$

The reaction of firm A is then simply $p_A = \frac{p_B+t}{2}$. If B decides to compete in the personalized market, its reaction function will be the solution of

$$\max_{p_B} \quad \left[\left(1-\theta_2^0(p_B)\right) + \left(1-\theta_1(p_A, p_B)\right)\right] p_B,$$

which leads to $p_B(p_A) = \frac{2t+p_A}{4}$. Combining with the reaction function of A leads to $p_A = \frac{6t}{7}$, $p_B = \frac{5t}{7}$ and a profit for A and B of $\pi_A = \frac{54t}{49}$ and $\pi_B = \frac{25t}{49}$. If B does not operate in the personalized market, its program writes as

$$\arg\max_{p_B} \quad [1-\theta_1(p_A, p_B)]\, p_B,$$

Its reaction function is then given by the standard Hotelling reaction function, $p_B(p_A) = \frac{t+p_A}{2}$, resulting in a price $p_A = t$ and profit $\pi_B = \frac{t}{2}$. As the latter profit is smaller than the former, B chooses a price to compete both in the personalized market and in the anonymous market. $\square$

We find that prices are lower than in the no-information case. As mentioned earlier, this is because B chooses a price that allows him to operate effectively in both markets, making thereby competition tougher. As a consequence, consumers' surplus is at its highest when both firms have information and at its lowest when no firm has information. Concerning profits, the informed firm makes the highest profits while the competitor makes the lowest. Furthermore, the informed firm in this case makes higher profits than in any one of the previous cases.

3.2.3 The price of information

We can now analyze how the value of information in the various setups studied in the preceding section should influence the data supplier's selling strategy. We suppose here that the data supplier, denoted as DS, owns the information about old consumers and can post a price T for it and sell it in period 2, and only at that period. DS can choose different selling strategies by setting different prices, which in equilibrium will induce one or two firms to buy the data. We call T_i the price paid by each firm when a number i of firms buys the information. The allocation and the prices are set and payments realized in period 1 and we further assume that this trade is common knowledge.

Assuming that DS chooses the strategy maximizing its profit, this profit is given by

$$\pi_{DS} = \max\{T_1, 2T_2\}.$$

To characterize the prices T_1 and T_2, we assume that DS has all the bargaining power, a natural assumption since both firms A and B are competing. This setting is similar to an auction with externalities as in Jehiel and Moldovanu (2000). Indeed, suppose first that DS sells to only one firm. Then, the maximum price T_1 that DS can set is the profit difference for one firm between winning and losing the auction, that is the profits made by the firm with information minus the profits the firm would make when its rival has it. We obtain in this first case

$$T_1 = \frac{54t}{49} - \frac{25t}{49} = \frac{29t}{49}.$$

Analogously, if DS sells to both firms, T_2 represents the difference in (say) firm A's profit between the case where both firms can use consumers' information and when only firm B has access to this information. Using the profit values derived earlier, we find that

$$T_2 = \frac{3t}{4} - \frac{25t}{49} = \frac{47t}{196}.$$

Therefore, DS chooses to sell the information only to one firm because $T_1 > 2T_2$. The next Proposition presents the result of this section.

Proposition 3. *Information is sold to only one firm at $T = \frac{29t}{49}$ and net profits for both firms are $\pi_A = \pi_B = \frac{25t}{49}$.*

Proof. See text. □

Proposition 3 shows that it is optimal for the owner of information to give exclusive rights. This is because competition is too tough when both firms have information and the potential for rent extraction is minimized. On the other hand, an informed firm can extract more rents from the consumers if its competitor can only use a single basic price. The exclusive allocation can be implemented with an auction where the DS commits to always sell the data only to one firm.

4 Privacy

We now turn to the case where consumers can pay for privacy. As noted in the description of the model, consumers can incur a cost c to be left out of the database that the firm(s) can buy. We assume that consumers observe whether the firm(s) bought the information and form rational expectation on the prices.[20]

[20] In section 5.1, we will investigate the situation where consumers make their privacy decision before information is bought.

4.1 Monopoly

We still assume that firm A is located a $\theta = 0$ in the both sets of new and old consumers. Compared to the timing used before, we add a third stage when consumers can protect their privacy.

1. The data supplier posts a price T for the data.
2. The firm chooses whether or not to purchase the data.
3. Consumers make their privacy choice, i.e., decide whether or not to pay the cost c.
4. The firm chooses the basic price p.
5. The firm can offer tailored prices.
6. Consumers buy and consume.

As in the case with no privacy, any consumer located at θ for whom the firm has information will be charged the tailored price $p(\theta) = v - t\theta > 0$ and make zero utility. This group is called the *personalized market* and it includes every old consumer who did not opt for privacy. The rest of the consumers constitute the anonymous market. Since the choice of privacy is made in period 3 (i.e., before any price has been set) we need to define an anticipated price $p^a \geq 0$. This price is the basic price that old consumers expect to pay if they are not in the firm's database. Effectively, for these consumers the basic price is now $p^a + c$. As this price increases, fewer consumers would avoid detection. If $p^a + c$ was small, however, it would be very attractive for consumers to enforce their privacy in period 3. A consumer located at θ will pay the privacy cost when

$$v - t\theta - p^a - c \geq 0 \iff \theta \leq \theta_2 = \frac{v - p^a - c}{t}.$$

To solve the model, assume the firm buys the information in period 2. In period 4, the monopolist maximizes profits $\pi(p)$ given the anticipated p^a:

$$\pi(p) = \int_0^{\theta_1} p d\theta + \int_0^{\min\{\theta_1,\theta_2\}} p d\theta + \int_{\theta_2}^1 p(\theta) d\theta. \tag{1}$$

The first term on the right hand side of equation (1) represents the profits from selling to new consumers. The second term covers the profits from the old consumers who are not in the firm's database, i.e., consumers who have paid the privacy cost. Together, these two terms represent the profits on the anonymous market. Note that the firm could set a large price p so that the actual market share θ_1 is smaller than the amount of people who pay the privacy cost. This is why the upper bound in the second integral is $\min\{\theta_1, \theta_2\}$.

The last term represents the personalized market. The consumers on this latter market are subject to price discrimination and pay the tailored price $p(\theta)$.

The firm then maximizes the profits expressed in equation (1) for $p \geq 0$ subject to the following constraint:

$$1 \geq \theta_1 \geq 0. \tag{2}$$

Furthermore, in any equilibrium we require that the proportion of people who pay the privacy cost stays bounded:

$$1 \geq \theta_2 \geq 0. \tag{3}$$

Proposition 4. *Assume the firm buys the information.*

- *If* $c \leq \max\{3t - v, 0\}$, *then*

$$p^* = \frac{2v - c}{3} \quad \textit{and} \quad \theta_2^* = \frac{v - 2c}{3t},$$

and profits and consumer surplus are:

$$\pi = \frac{2c^2 + v^2}{6t} + \frac{1}{2}(2v - t) \quad \textit{and} \quad CS = \frac{5c^2 - 2cv + 2v^2}{18t}.$$

- *If* $t \geq c > 3t - v$, *then*

$$p^* = v - t \quad \textit{and} \quad \theta_2^* = 1 - \frac{c}{t},$$

and profits and consumer surplus are:

$$\pi = \frac{c^2}{2t} + 2(v - t) \quad \textit{and} \quad CS = t + \frac{c^2}{2t} - c.$$

Proof. Note that $\theta_2 > \theta_1$ cannot be an equilibrium because consumers in period 3 correctly anticipate the prices chosen by the monopolist in period 4, so $p^a = p$. Then it has to be that, in any equilibrium, $\theta_2 \leq \theta_1$. Under this condition the firm takes the quantity of consumers who pay the privacy cost θ_2 as given. Neglecting conditions (2) and (3), the first order condition $\pi'(p) = 0$ leads to

$$p = \frac{-p^a - c + 2v}{2}.$$

Finally, we apply rational expectations, which imply the condition $p^a = p$ on the first order condition. Note that with $p^a = p$, $\theta_1 \leq 1$ implies $\theta_2 \leq 1$. Then this solution requires only $\theta_1 \leq 1$ to be interior. The second solution is the corner solution when $\theta_1 = 1$. □

The basic price p^* is now obtained considering the average willingness to pay in the anonymous market, which corresponds to the weighted average between the willingness to pay of new consumers and the willingness to pay among old consumers who paid the cost c. Note also that the average willingness to pay from those who paid the privacy cost is larger than the average willingness to pay of new consumers, given that consumers who pay the cost are closer to the firm.

There are two effects at play to understand the effects of an increase in c on the price. The first one is the *size effect*. When c increases, the total number of old consumers who pay the basic price goes down. As a result, the overall weighted average willingness to pay is reduced. This effect leads to a price decrease. The second one is the *valuation effect*. As c increases, consumers who do not value the good as much will not pay the basic price any more and the average valuation in the market will increase. This second effect leads to a price increase. But on the aggregate the size effect dominates, and the price goes down with c.[21] We now look at the impact of privacy costs on profits and consumer surplus.

Corollary 1. *Assume the firm buys the information. Profits are always increasing in c. Additionally, if $c \leq 3t - v$, consumer surplus is U-shaped in c and minimized at $\frac{v}{5}$. If $c > 3t - v$, consumer surplus is always decreasing in c.*

Proof. Direct from Proposition 4. □

Corollary 1 states that profits are increasing in the privacy cost. Profits increase in c because, in spite of having a lower basic price, the share of consumers who pay the privacy cost decreases. Then, higher profits in the personalized market compensate the losses in the anonymous market.

Consider first the interior case $c \leq 3t - v$, where consumers' surplus is U-shaped in c. This happens because new consumers always gain with a larger c, while old consumers always lose. For the former, a larger cost is associated with a larger market share and a lower price. For the latter, it is the opposite. This can be seen by noting that the effective basic price paid by old consumers is $p^* + c$. The corollary shows that for $c \leq \frac{v}{5}$ the losses coming from old consumers dominate, whereas for $c > \frac{v}{5}$ the gains coming from new consumers dominate.

When $c > 3t - v$ instead, increasing c has no effect on the set of new consumers because full coverage is achieved and the firm does not lower its price any more on the anonymous market. Then when c increases the surplus of old consumers decreases as fewer consumers pay for privacy and have to pay their full valuation.[22]

As in the case without privacy, the data supplier can extract any profit the monopolist makes in excess of $2(v - t)$, which are the profits it could make without having any information at all.

[21] This result holds in the limit for any distribution of θ. Indeed, define c^* such that $\theta_2^*(c^*) = 0$. If c^* decreases, some consumers would pay the privacy cost and the price will rise. The uniform distribution assumption only ensures that this effect is true for any c.

[22] When $c = t$, Proposition 4 leads to the No-Privacy result, with $\theta_2^* = 0$.

4.2 Competition

Now we consider the duopoly case where consumers can pay for privacy. This section is similar to 3.2 but it includes a third stage in which consumers can choose to pay for privacy. The timing in this section changes to accommodate for the consumers' actions.

1. The data supplier posts a price T for the data.
2. Firms A and B choose whether or not to purchase the data.
3. Consumers make their privacy choice, i.e., decide whether or not to pay the cost c.
4. Firms A and B choose basic prices p_A and p_B.
5. Firms A and B can offer tailored prices.
6. Consumers buy and consume.

Some of the old consumers might pay the privacy cost to avoid the tailored price. This share is given by those who prefer paying the privacy cost c — and buying from A at the basic price — than buying from A at the tailored price:

$$v - p_A^a - \theta t - c \geq v - p_A^a(\theta) - \theta t = v - t(1-\theta) - p_B^a.$$

As in the previous section, this share must be based on anticipated prices, given that consumers make their privacy choices in period 3 and prices are set in periods 4 and 5. In other words, by the time prices are set, firm A treats the proportion of consumers who paid the privacy cost as fixed. Then, the indifferent consumer type is given by

$$\theta_2 = \frac{1}{2} + \frac{p_B^a - p_A^a - c}{2t}. \tag{4}$$

Finally, if neither firm has information, the fact that consumers can pay for privacy does not change the benchmark result found in section 3.2. Therefore, if consumers observe that no firm will have information, no consumer will want to pay the cost c, and resulting prices are the standard Hotelling prices.

4.2.1 Both firms have information

Consider the case where both firms buy information in period 2. In this case, tough competition leads to tailored prices low enough so that it is not profitable for *any* consumer to pay for privacy.

Lemma 1. *If consumers observe that both firms bought the information, no consumer pays the privacy cost.*

Proof. Take any subset of old consumers who did not pay the privacy cost. Competition in this subset is for every consumer; therefore, tailored prices are equivalent to those in Proposition 1. This implies that any consumer who hides does it to be allowed to pay the basic price of the nearest firm, because the alternative (the tailored price) is already matching the competitor's offer. The rest of the proof relies on two arguments:

1. If a consumer $\hat{\theta} \in [0, \frac{1}{2}]$ pays the privacy cost, every consumer $\theta \in [0, \hat{\theta}]$ pays the privacy cost. By symmetry, a similar statement is true in $(\frac{1}{2}, 1]$. Note that if consumer $\hat{\theta}$ pays, it implies that

$$\frac{t - c - p_A}{2t} \geq \hat{\theta}.$$

2. If a strictly positive mass of consumers pays the privacy cost, then p_A and p_B should be greater than t. Say consumers $\theta \in [0, \theta_A]$ pay the privacy cost to hide from A. Then A solves

$$\max_{p_A} \int_0^{\theta_A} p_A d\theta + \int_0^{\frac{1}{2} + \frac{p_B - p_A}{2t}} p_A d\theta.$$

B solves an analogous problem. Then prices are

$$p_A = t + \frac{2t(2\theta_A + 1 - \theta_B)}{3} \quad \text{and} \quad p_B = t + \frac{2t(\theta_A + 2(1 - \theta_B))}{3}.$$

The first argument implies that consumers to the left of $\frac{1}{2}$ would pay the privacy cost if and only if $p_A + c < t(1 - 2\theta)$. The second argument implies that $p_A > t$. These two conditions cannot hold for $c > 0$. Therefore, no consumer would pay the privacy cost. □

In the personalized market, both firms will compete for every consumer and the tailored prices will be low. Therefore, to observe consumers paying the privacy cost, basic prices should be even lower than tailored prices, after taking the privacy cost c into account, and such basic prices are not profitable.

Note that, from a total welfare point of view, we achieve a first best allocation in this case: transportation costs are minimized (market shares along both Hotelling lines are symmetric) and no consumer engages in wasteful activities (no one pays the privacy cost).

To summarize, when consumers *can* pay for privacy and information is symmetric among firms, in equilibrium consumers *do not* pay for privacy. This means that the fact that consumers are entitled to privacy has no effect over the outcome of the model when both firms have information. Therefore, Proposition 1 still holds with Hotelling prices offered to new consumers and lower personalized prices offered to old consumers.

4.2.2 Only firm A buys information

As before, we assume that firm A buys the information and B does not. The equilibrium is given in the following result.

Proposition 5. *Assume only firm A buys the information and consumers can pay for privacy. There exists a threshold $\overline{c} < t$ such that, if $c \leq \overline{c}$, the equilibrium is:*

$$p_A = t + \frac{t-c}{2}, \quad p_B = t + \frac{t-c}{4} \quad \textit{and} \quad \theta_2 = \frac{3(t-c)}{8t},$$

and if $c > \overline{c}$, the equilibrium is the one in Proposition 2.

Proof. First, we verify that B chooses a price such that it does not sell to old consumers ($\theta_2^0 \geq 1$). Note that A's reaction function is always given by:

$$\arg\max_{p_A} \quad [\theta_1(p_A, p_B) + \theta_2(p_A^a, p_B^a)]\, p_A.$$

If B does not sell to old consumers, its reaction function is:

$$\arg\max_{p_B} \quad (1 - \theta_1(p_A, p_B))\, p_B.$$

These two equations lead to $p_A(\theta_2)$ and $p_B(\theta_2)$. Finally, because in equilibrium $p_A = p_A^a$ and $p_B = p_B^a$ (which implies $\theta_2 < \theta_1$), we can solve for θ_2 and find $\theta_2 = \frac{3(t-c)}{8t}$. Prices lead to $\pi_B = \frac{(5t-c)^2}{32t}$. On the contrary, if B does sell to old consumers, its reaction function is

$$\arg\max_{p_B} \quad \left[\left(1 - \theta_2^0(p_B)\right) + (1 - \theta_1(p_A, p_B))\right] p_B.$$

Note that θ_2^0 is a direct consequence of $p_A(\theta) \geq 0$. Given that $p_A(\theta) = p_B + (1-2\theta)t$ is always chosen by A in period 5, θ_2^0 is a function of p_B and not of p_B^a.
Following the same procedure as above, we obtain $\theta_2 = \frac{6t-7c}{20t}$ and $\pi_B = \frac{(8t-c)^2}{100t}$. Then to keep $\theta_2 \geq 0$ we must have $c \leq \frac{6t}{7}$. For a larger c, no consumer pays for privacy and prices are the same as in Proposition 2. Therefore, it is sufficient to establish for $c \leq \frac{6t}{7}$ that $\frac{(5t-c)^2}{32t} \geq \frac{(8t-c)^2}{100t}$. Then, we can find a $\overline{c} = 5t - \frac{20}{7}\sqrt{2}t \approx 0.96t$ such that, for $c \leq \overline{c}$, $\frac{(5t-c)^2}{32t} \geq \frac{25t}{49}$. □

We find that prices are larger than Hotelling prices for $c \leq \bar{c}$. From A's perspective, the average willingness to pay in the segment $[0, \theta_2]$ is higher than the average willingness to pay of new consumers, because these consumers are closer to A. Then it follows that the average willingness to pay in the anonymous market is higher than the average willingness to pay of a set that consists of the new consumers only.[23] This leads to

[23]To see this, let v^N be the average willingness to pay of new consumers and v^O be the average willingness to pay in the segment $[0, \theta_2]$ of old consumers. Since $v^O > v^N$, then $\frac{v^N + \theta_2 v^O}{1+\theta_2} > v^N$.

$p_A \geq t$. Noteworthy, the fact that $p_A \geq p_B$ is a consequence of the difference in the average willingness to pay for each product, as no consumer in $[0, \theta_2]$ buys from B.

The result that the price $p_B \geq t$ comes from the fact that B does not find it profitable to compete for the old consumers. In other words, A supplies every old consumer. This means that $\theta_2^0 \geq 1$, which in turn implies that $p_B \geq t$. The change of regime at $\bar{c}$ is given by B's decision to sell to old consumers,[24] then B's equilibrium profits are continuous at $\bar{c}$. However, θ_2 is discontinuous at $\bar{c}$ and drops to zero, i.e., no consumer pays for privacy. Hence, we will focus on $c \leq \bar{c}$ to study the effects of the privacy cost on the equilibrium.

Analogous to the monopoly case, a change in c has two effects: a size effect and valuation effect. Then, because prices decrease when c increases, the size effect dominates and A reduces its price as c goes up. In this equilibrium, B supplies only new consumers and its reaction function corresponds to the one in the standard Hotelling model. This implies that prices are strategic complements, then p_B also decreases with c. Furthermore, $|p_A'(c)| > |p_B'(c)|$ given that c affects p_A directly as fewer consumers pay for privacy but also trough p_B. However, c affects p_B only through the competitor's price.

An important implication from Proposition 5 is that price discrimination cannot be avoided even when enforcing privacy is free. Indeed, for some consumers, it is better to accept the tailored price than to pay p_A. Consumers relatively closer to B will accept the tailored price, because such price is decreasing in θ. When c goes to zero, the set of old consumers who pay for privacy and the set made of those who would accept p_A in stage 4 are identical, but there are still some consumers who prefer to go for the tailored offer. When c is large however, some consumers would accept p_A had c been zero but decide to pay the tailored price instead if $p_A(\theta) - p_A$ is lower than the cost c.

Corollary 2. *Assume only firm A buys the information and consumers can pay for privacy. Then, the profits of both firms are decreasing in c while consumer surplus is increasing in c.*

Proof. If $c \leq \bar{c}$ profits and consumer surplus are given by

$$\pi_A = \int_0^{\frac{p_B - p_A + t}{2t}} p_A \, d\theta + \int_0^{\theta_2} p_A \, d\theta + \int_{\theta_2}^1 (p_B + (1 - 2\theta)t) \, d\theta = \frac{11c^2 - 22ct + 107t^2}{64t},$$

$$\pi_B = \int_{\frac{p_B - p_A + t}{2t}}^1 p_B \, d\theta = \frac{(5t - c)^2}{32t},$$

$$CS = \int_0^{\frac{p_B - p_A + t}{2t}} (v - p_A - \theta t) \, d\theta + \int_0^{\theta_2} (v - p_A - \theta t - c) \, d\theta + \int_{\theta_2}^1 (v - (p_B + (1 - 2\theta)t) - \theta t) \, d\theta$$
$$+ \int_{\frac{p_B - p_A + t}{2t}}^1 (v - p_B - (1 - \theta)t) \, d\theta = \frac{10ct + 5c^2 - 103t^2}{32t} + 2v,$$

[24] The details can be found the in proof of Proposition 5.

and, if $c > \overline{c}$, the equilibrium is the same as the one found in Proposition 2. We can show that $\pi_A'(c), \pi_B'(c) < 0$ since $c < t$ and $CS'(c) > 0$. In the proof of Proposition 5 we showed that there is a $\overline{c}$ above which firm B poaches some old consumers which leads to basic prices low enough so that consumers do not pay for privacy and the equilibrium in Proposition 2 applies. Then profits and consumer surplus are linear in $c > \overline{c}$. Thus the results in this proposition hold for any $c \leq t$. □

Note that, as c increases, the proportion of consumers who pay the privacy cost decreases. This lowers the average willingness to pay for A's basic price. As a response, A reduces its price and B does the same to keep up with competition for the new consumers. Because p_B decreases, the tailored price should also decrease as A has to leave more surplus to the consumers to make them indifferent between buying from it or from B. This brings down the profits of both firms, which contrasts sharply with the monopoly case.

In turn, consumers benefit from this competition. From Corollary 2, consumer surplus is always increasing in c. This is because when c increases, all prices go down and the amount of consumers who pay the privacy cost goes down as well. These two effects are more than enough to compensate for the consumers who pay a higher c. Furthermore, it can be checked that consumers' surplus exhibits an upward jump at $\bar{c}$ because firms compete more aggressively when B supplies old consumers.

We can also compute the privacy gains for old consumers close to A as the difference in utility from paying for privacy minus what they get if they do not:

$$(v - p_A - t\theta - c) - (v - p_A(\theta) - t\theta) = \frac{1}{4}(t(3 - 8\theta) - 3c),$$

which is decreasing in c. Indeed, as c increases, these consumers benefit from the decrease in the price but suffers from a direct effect of increasing the privacy cost. And this direct negative effect dominates the positive price-mediated effect. Therefore, consumers who protect their privacy are worse off when privacy is more costly, while the rest of the consumers are better off.

4.2.3 The price of information

The price of information when consumers can pay for privacy is calculated in the same way as in section 3.2.3. The data supplier DS compares what it can obtain from selling to one firm, T_1, to what he can obtain when selling to both firms, $2T_2$. Proposition 6 shows that, even when consumers can pay for privacy, DS finds it optimal to sell the information to one firm only.

Proposition 6. *If $c \leq \overline{c}$, then information is sold to one firm at a price $T = \frac{9c^2 - 2ct + 57t^2}{64t}$ and net profits are $\pi_A = \pi_B = \frac{(5t-c)^2}{32t}$. If $c > \overline{c}$, Proposition 3 holds.*

Proof. The case where $c > \bar{c}$ has been already solved. For $c \leq \bar{c}$ we use the proof of Corollary 2 to check that $T_1 > 2T_2$, where $T_1 = \pi_A - \pi_B$ and $T_2 = \frac{3t}{4} - \pi_B$. Indeed,

$$\pi_A - \pi_B \geq 2\left(\frac{3t}{4} - \pi_B\right),$$

and therefore $T = T_1$. □

The result that the data supplier DS prefers to sell information exclusively to one firm is therefore robust to introducing privacy. Remarkably, the price firm DS can set for the information is higher when the privacy cost is not too high. Indeed, the informed firm (A) knows that, when privacy is allowed (or the privacy cost is not too high), it can charge higher prices because it faces consumers with a larger average valuation (only consumers close to the informed firm will pay for privacy). Therefore, competition is less aggressive and there are larger gains to be made when acquiring information exclusively. Because of this, the possibility of privacy reinforces the result that information is allocated to only one firm.

The previous result is at odds with Chen et al. (2001), who postulate that a non-exclusive allocation can be optimal for the data seller. This happens because price competition is not strong when targetability is sufficiently low, and firms set higher prices the more they mistake loyal customers from searches. The authors call privacy the fact that targetability is imperfect and firms cannot properly recognize its loyal customers. Noteworthy, mistargeting occurs uniformly across consumers. In our model, on the contrary, privacy (and hence targetability) depends on the information structure and on competition. This is why, we do not find a situation with high levels of privacy (low targetability) and non-exclusive sales of data.

Finally, given that T in Proposition 6 is U-shaped in c, we can show that the corner $c = \bar{c}$ yields the highest revenue for firm DS.[25] Furthermore, since T is minimized at $c = \frac{t}{9}$, the value of information is for the most part increasing in c as the outside option of not buying π_B decreases faster than the profits generated from buying the information π_A when $c > \frac{t}{9}$.

Note that, as p_A decreases faster in c than p_B, B's market share among new consumers decreases with c. This means that B must lose profits when c increases. Firm A, on the contrary, faces two effects in the new consumers segment, as its market share increases but its price decreases with c. For old consumers, a larger privacy cost leads to fewer consumers paying for it, which increases the revenue from tailored prices, even though $p_A(\theta)$ decreases. At the same time, less of the old consumers pay the basic price. This makes A's (gross) profits less sensitive to c than B's profit.

[25] Going above $\bar{c}$ results in a change of regime whereby no consumer opts for privacy as it is too costly, but the competitive value of information is lower.

5 Extensions

5.1 Interim sales of information

So far we have assumed that the data supplier DS was selling the information *ex ante* at a price T. We now consider that the sales of data could be made at a later stage of the game. More precisely, we change the timing so that buying information occurs after consumers have decided whether or not to pay the privacy cost, but before the pricing game. Note that if the sale of information occurs after the basic prices are set, all consumers can benefit from the basic price at no cost, and information has no value.[26] In this extension, we concentrate on the range of privacy costs $c \leq \bar{c}$ to avoid any change of regime. The timing is then as follows:

1. Consumers make their privacy choice, i.e., decide whether or not to pay the cost c.
2. The data supplier posts a price T for the data.
3. Firms A and B choose whether or not to purchase the data.
4. Firms A and B choose basic prices p_A and p_B.
5. Firms A and B can offer tailored prices.
6. Consumers buy and consume.

In this section, consumers must decide whether to pay the privacy cost or not before knowing whose firm will be able to make targeted offers. Therefore, not only the firms A and B, but more importantly the data supplier, may want to deviate from the actions anticipated by the consumers. Given these possible deviations, the next lemma shows that information cannot be sold to both firms at the same time but that exclusive sales should have a structure different from the one characterized in the previous section.

Lemma 2. *When consumers choose to pay the privacy cost before knowing which firm will be informed, there is no equilibrium where*

1. *information is sold to the same firm with probability 1,*
2. *information is sold to both firms with probability 1.*

Proof. We first show that there is no equilibrium when information is sold to the same firm with probability one. Consider a candidate equilibrium in which consumers anticipate that only firm A buys the data. In this case, the outcome of the game is the same as

[26] If the basic prices are set before the information is bought by the firm(s) but only publicly announced to the anonymous market after information is bought, the equilibrium is the same as the one we derive in this section.

the one in Proposition 5. However, the data supplier DS can deviate and sell to B. If it does, it can charge a price equivalent to the profits of B in the Old market. Note that after DS's deviation in period 2, both firms choose the same basic prices in period 4 as in Proposition 5, while θ_2 is fixed since period 1 and defined in Proposition 5 as well. Therefore,

$$\pi_B^O = \int_{\frac{3(t-c)}{8t}}^{1} \left(\frac{t-c}{2} + t(2\theta - 1) + t \right) d\theta = \frac{10ct - 21c^2 + 75t^2}{64t}.$$

It can be shown that π_B^O is strictly larger than the price of information as stated in Proposition 6. Therefore, if consumers anticipate that the information will be bought by A (resp., by B), the data supplier has some incentives to sell to B (resp., to A). So there cannot be an equilibrium in which the data seller allocates the data to a single firm with probability one.

We now show that there is no equilibrium when information is sold to both firms with probability 1. Consider a candidate equilibrium where it is the case. Then, as shown earlier in section 3, consumers do not pay the privacy cost. But then, the data supplier has some incentives to sell the information to only one firm (we are back to section 3). So consumers should rationality anticipate this deviation. Therefore, this cannot happen at the equilibrium. □

One way to construct an equilibrium is to allow the data supplier to randomize to whom the information is sold, and charge the price such that the chosen firm is indifferent between buying or not buying.

Proposition 7. *There exists an equilibrium where information is sold exclusively to one firm with probability $\frac{1}{2}$, and the price of information is*

$$T = \begin{cases} 4c - \frac{4c^2}{t} & \text{if } c \leq \frac{t}{2} \\ t & \text{if } c > \frac{t}{2} \end{cases}.$$

Proof. Assume consumers anticipate each firm will have information with probability $\frac{1}{2}$. Furthermore, consider a symmetric equilibrium in which a fraction θ_A pays for privacy (to hide from A) and a fraction $1-\theta_B$ pays for privacy (to hide from B), with $\theta_A = 1-\theta_B$. If a consumer close to A pays for privacy, she anticipates making $-p_A^a - c - t\theta_A + v$. On the other hand, if the same consumer does not pay for privacy, she expects to make:

$$\frac{1}{2}(-p_A^a - t(2\theta_A - 1) - t(1 - \theta_A) + v) + \frac{1}{2}(-p_B^a - t\theta_A - t(1 - 2\theta_A) + v).$$

The first term represents the utility such consumer makes from paying a tailored price (as calculated before) to A. The second term represents utility from paying a tailored price to B. Both terms are multiplied by the probability of such events. The indifferent consumer θ_A is then given by:

$$\theta_A(p_A^a, p_B^a) = \frac{p_B^a - p_A^a + t - 2c}{2t}.$$

Given θ_A, A's profits are given by

$$\frac{1}{2}\left(\theta_1(p_A, p_B)p_A + \theta_2 p_A + \int_{\theta_A}^{\theta_B} p_B + (1-2\theta)tdx\right) + \frac{1}{2}\left(\theta_1(p_A, p_B)p_A + \theta_2 p_A\right),$$

where $\theta_1(p_A, p_B)$ represents A's market share among the new consumers. The first term represents profits when the firm has data and the second term when it does not (but B does). By symmetry, firms A and B solve respectively

$$\arg\max_{p_A} \quad [\theta_1(p_A, p_B) + \theta_A(p_A^a, p_B^a)]\, p_A,$$

$$\arg\max_{p_B} \quad [(1 - \theta_1(p_A, p_B)) + \theta_A(p_A^a, p_B^a)]\, p_B.$$

Equations $p_B = p_B^a$ and $p_A = p_A^a$ (consumers' rational expectations) lead to equilibrium prices $p_A = p_B = 2(t-c)$ and $\theta_A = \frac{t-2c}{2t}$. Then, for such solution to be feasible, the condition $c \leq \frac{t}{2}$ must be satisfied. In this solution, the data supplier can set a price T equal to the profits in the personalized market, i.e., $[\theta_A, \theta_B]$. This is because these profits are additional to the profits the firm would make without data. Hence,

$$T = \int_{\theta_A}^{\theta_B} t(1-2\theta) + 2(t-c)dx = 4c - \frac{4c^2}{t}.$$

When $c > \frac{t}{2}$ consumers do not pay for privacy and the personalized market consists of all the old consumers. Note that, because $\theta_A = 0$, $p_A = p_B = t$. Then, the extra profits information allows to make (i.e., the price of information) is equal to:

$$T = \int_0^1 p_A(\theta)d\theta = t.$$

By construction of the price T, firms do not deviate at stage 3. Finally, it can be shown that the data seller does not deviate to selling to both firms, because more revenue is always raised by selling to one firm only. Indeed, such deviation generates revenue:

$$T^D = 2\int_{\theta_A}^{\frac{1}{2}} t(1-2\theta)d\theta < T.$$

We can rule out price deviations by the firms in period 4 by noting that, because both firms are playing their best response in the anonymous market, changing basic prices would only generate losses. Furthermore, the data supplier has no incentive to change the allocation rule, because it is indifferent between selling the data to A or B; and this is consistent with the consumers' expectations. □

The result on exclusivity is robust to changing the timing of the baseline model, if we consider *ex post* exclusivity rather than *ex ante* exclusivity. Furthermore, we can show that the data supplier would like to commit *ex ante* for small values of c, but would rather sell later for a large c (see Figure 1).

To understand this result, recall that without commitment and for small value of c, almost all old consumers will pay the privacy cost and the data supplier has almost nothing to sell. With commitment, even for c small, some consumers choose not to pay the privacy cost, and there is tough competition for them as the basic price is then high (so their outside option limited). Consider instead the case where the privacy cost is high. Without commitment, no consumer decides to pay this cost because its preferred firm will only have some information on him with probability 1/2, so the competition for information is intense, and the resulting price set by the data supplier is high. With commitment, there may be more consumers paying for privacy and less competition for information.

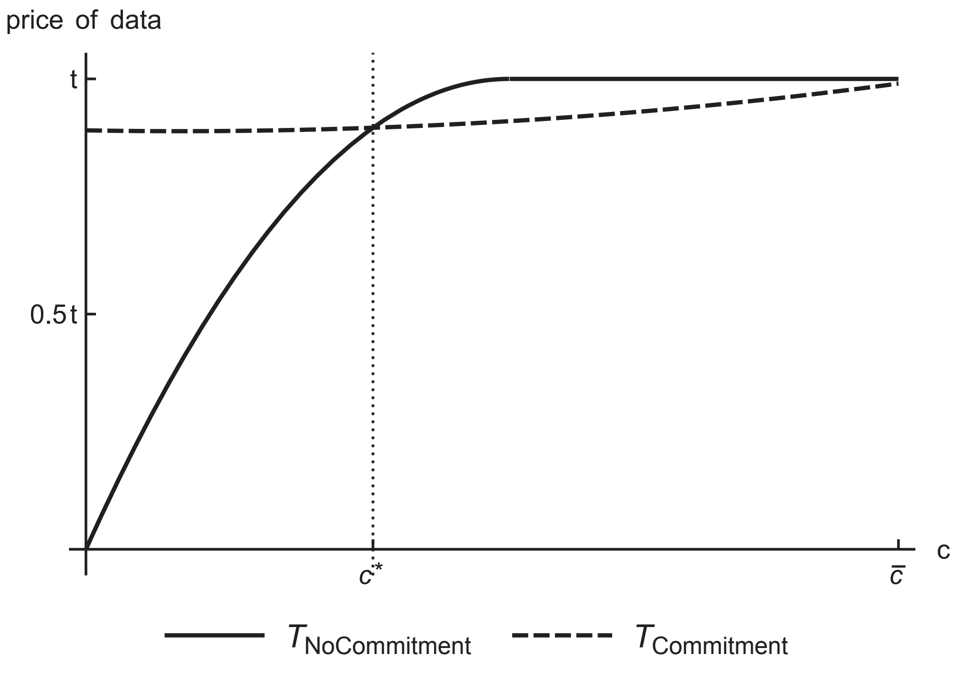

Figure 1: Information prices with and without commitment

5.2 Markets of different sizes

In the preceding sections, we have assumed that there were as many old as new consumers. We now generalize the analysis assuming that the mass of old consumers is still equal to 1 but allowing the mass of new consumers to be of any size $m > 0$. We are interested in the effect of changing the relative market size on the incentives to acquire information in the presence of privacy. Therefore, we assume that $c \leq \frac{t(3m+3)}{3m+4} < t$ so that some consumers pay for privacy in every case ($\theta_2 \geq 0$).

The main effect of varying m on the equilibrium is to change the incentives of firm B to compete for the old consumers.[27] We recall that, for B to do so, it needs to set a low price (smaller than t) because it faces disadvantageous competition from A, who holds information about consumers. Furthermore, B can set only one price for both markets. Therefore, if m is large, it is more profitable for B to keep a higher price and compete only for new consumers. Conversely, if m is small, forsaking old consumers will not be profitable.

Lemma 3. *Assume only firm A buys the information and consumers can pay for privacy. There are two types of equilibria:*

1. *If B does not compete for consumers, then*

$$p_A = t + \frac{2(t-c)}{1+3m}, \quad p_B = t + \frac{t-c}{1+3m} \quad \text{and} \quad \theta_2 = \frac{3m(t-c)}{2t(1+3m)}.$$

2. *If B compete for old consumers, then*

$$p_A = t + \frac{t-2c}{2+3m}, \quad p_B = t - \frac{cm + (1+m)t}{(1+m)(2+3m)} \text{ and } \theta_2 = \frac{3m(m+1)(t-c) - cm}{2(m+1)(3m+2)t}.$$

Proof. In equilibrium we must have $p_A = p_A^a$ and $p_B = p_B^a$, then $\theta_2 < \theta_1$. However, reaction functions ought to be calculated taking the consumers' actions as given. In case 1, B's reaction function is given by

$$\arg\max_{p_B} \quad m\left(1 - \theta_1(p_A, p_B)\right) p_B,$$

while A's reaction function is

$$\arg\max_{p_A} \quad \left[m\theta_1(p_A, p_B) + \theta_2(p_A^a, p_B^a)\right] p_A.$$

These two equations allow to find $p_A(\theta_2)$ and $p_B(\theta_2)$. Finally, because in equilibrium $p_A = p_A^a$ and $p_B = p_B^a$, we can plug these reaction functions into Equation (4). Doing so, leads to the expression for θ_2.

In case 2 we solve in the same manner but we change B's reaction function to account

[27] Note that B's choice is always independent of A's actions.

for selling to old consumers:

$$\arg\max_{p_B} \quad \left[\left(1-\theta_2^0(p_B)\right) + m\left(1-\theta_1(p_A,p_B)\right)\right] p_B.$$

Recall that θ_2^0, B's market share among old consumers, depends on p_B and not on p_B^a. This is because θ_2^0 is set by $p_A(\theta_2^0)=0$. And given that $p_A(\theta)=p_B+(1-2\theta)t$ is always chosen by A in period 5, θ_2^0 is a function of p_B and not of p_B^a. Therefore, B optimizes in both markets at the same time. Note that the assumption $c \leq \frac{t(3m+3)}{3m+4}$ is necessary to ensure that the proportion of consumers who pay the privacy cost is positive. $\square$

Prices are decreasing in m when B does not compete for the old consumers. In fact, as m goes large, prices converge to the Hotelling prices. Indeed, as the market for new consumers becomes relatively more important, firms focus increasingly on this market and the situation is akin to a standard Hotelling competition. When B competes for old consumers, the effect of the relative size m on the prices is more ambiguous.

Lemma 4. *Suppose that only firm A buys the information and that consumers can pay for privacy. There exists $m^*(c)>0$, such that B does not compete for old consumers if and only if $m \geq m^*(c)$.*

Proof. Compute the difference in B's profits when B does not sell to old consumers (case 1 in Lemma 3) and when B does (case 2 in Lemma 3):

$$\Delta\pi = \frac{m(c-(3m+2)t)^2}{2(3m+1)^2 t} - \frac{\left(-cm+3m^2t+4mt+t\right)^2}{2(m+1)(3m+2)^2 t}.$$

The first term on the right hand side represents the profits B makes when it does not sell to old consumers. The second term is the profits when it does. It can be verified that $\Delta\pi$ goes to $\frac{t}{6}$ as m increases to infinity. Moreover, as m goes to zero, $\Delta\pi$ tends to $-\frac{t}{8}$. Furthermore, it can be proved that $\Delta\pi$ is always increasing. Therefore a threshold $m^*>0$ exists. $\square$

We can then study the data supplier's behavior for any $m>0$. Lemma 4 gives us two cases to analyze based on whether m is greater or lower than $m^*(c)$. By Lemma 1 (which can be generalized to any $m>0$), if both firms buy information, consumers do not pay for privacy. Then, following Proposition 1, firms make equal profits from selling to both sets of consumers: $\frac{mt}{2}$ for new consumers and $\frac{t}{4}$ for old consumers.

Proposition 8. *For any $m>0$, it is always more profitable for the data supplier to sell the information only to one firm.*

Proof. For $m \geq m^*(c)$, B does not compete for the old consumers. Then, as in the proof of Proposition 6, it suffices to show $\pi_A - \pi_B \geq 2\left(\frac{t(2m+1)}{4} - \pi_B\right)$, where

$$\pi_A = m\int_0^{\frac{p_B - p_A + t}{2t}} p_A\, d\theta + \int_0^{\theta_2} p_A\, d\theta + \int_{\theta_2}^{1} (p_B + (1-2\theta)t)\, d\theta,$$

$$\pi_B = m\int_{\frac{p_B - p_A + t}{2t}}^{1} p_B\, d\theta,$$

and p_A, p_B and θ_2 are given by Lemma 3, 1.

For $m < m^*(c)$, B supplies old consumers. Again, as in the proof of Proposition 6, it suffices to show $\pi_A - \pi_B \geq 2\left(\frac{t(2m+1)}{4} - \pi_B\right)$, where

$$\pi_A = m\int_0^{\frac{p_B - p_A + t}{2t}} p_A\, d\theta + \int_0^{\theta_2} p_A\, d\theta + \int_{\theta_2}^{\frac{p_B + t}{2t}} (p_B + (1-2\theta)t)\, d\theta,$$

$$\pi_B = m\int_{\frac{p_B - p_A + t}{2t}}^{1} p_B\, d\theta + \int_{\frac{p_B + t}{2t}}^{1} p_B\, d\theta,$$

and p_A, p_B and θ_2 are given by Lemma 3. □

Proposition 8 implies that exclusivity in the allocation of information is robust to any relative market size, provided that information is sold *ex ante*.

6 Main Highlights and Final Remarks

Big data allow firms to make targeted offers to consumers. These offers come in three forms: targeted advertisements, customization of products, and tailored prices. In this paper, we have focused on the latter.[28] What is novel is that we have allowed consumers to respond to the possibility that they can be targeted. We have also studied the original aspect of the selling strategy for the data owner. Using our model, we can offer an answer to two important questions, related first to the value of privacy for consumers and second to the value of information for firms.

On the first issue, one current debate is on the impact arising from restricting the way commercial firms use information. If we model increasing restrictions as a decrease in the privacy cost, it appears from our study that one must distinguish between the case of monopoly and the case of more competitive market structures. In the monopoly case, making it easier for consumers to protect their privacy has an ambiguous effect on consumers, as some consumers gain while others lose. Using an utilitarian social welfare

[28] However, we could have easily modeled product customization by making the firm offer consumer of type θ a good that generates gross utility $v + \theta t$, at a cost θx for the firm, where $0 < x < t$. In this case, the data supplier would appropriate any efficiency gains and still prefer exclusivity. In turn, the uninformed firm's profits and consumer surplus would not change. Furthermore, if one firm is more efficient than its competitor at tailoring goods, the data supplier always sells to the most efficient firm.

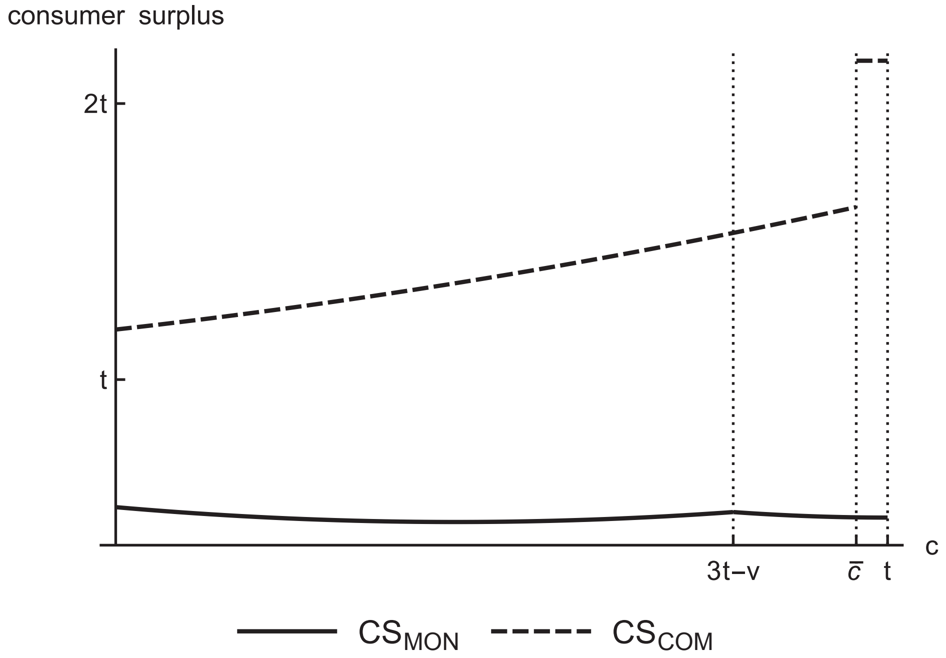

Figure 2: Consumer Surplus for $t = 1$ and $v = 2.2$

function weighting equally consumers' surplus and firms' profits still leads to ambiguous results, even if there is a stronger presumption that society is worse off when privacy costs are low. In the duopoly case, the results are less ambiguous, with total consumers' surplus increasing with privacy cost, and total welfare maximized in the case where consumers cannot protect their data. However, if consumers do protect their data (cases with $c < \bar{c}$), welfare is U-shaped in c. Thus, from a social welfare perspective, it would make sense to promote either a little or a lot of protection, since intermediate cases are inferior.[29] Figures 2 and 3 illustrate these results, where the continuous lines describe the monopoly case and the dashed lines the duopoly case.

On the second issue, one may be interested in computing the value of information for the firm acquiring data, or equivalently for the data supplier. This value depends on the strategy chosen by the data supplier, so the question could rather be phrased in terms of a study on the optimal selling strategy for the data supplier. We show that this optimal selling strategy consists in creating a strong competition between firms for acquiring information. As explained previously, competition for information can be analyzed as an auction with externalities where the payoff of each firm is influenced by the final informational structure. The best way to generate a large value of information is to increase the difference between the gains when one firm has the information and when it has no information. This is achieved by selling the information in an exclusive

[29]We recall that, in our model, efficiency is reached when market shares along the Hotelling lines are shared equally by the firms (under competition), and consumers do not spend resources on privacy. The other variables in the model (such as the prices to consumers and the price of data) are transfers that affect individual payoffs but not overall welfare.

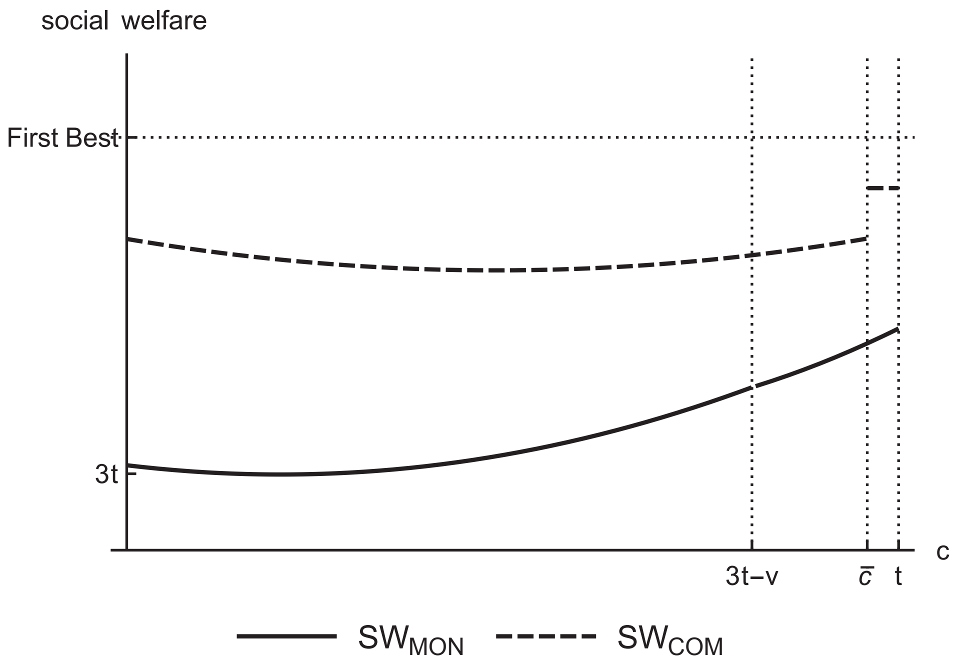

Figure 3: Social Welfare for $t = 1$ and $v = 2.2$

way. Figure 4 represents the price for data in the monopoly case (when the exclusivity strategy is not available) and in the duopoly case where the data supplier excludes one of the competing firms. Even if the profit in the monopoly case is greater than the profits in the duopoly case, for every c, T_{COM}, the price of data in duopoly, is larger than T_{MON}, the price of data in the monopoly case. These prices are computed considering the differential impact of information on profit and this impact is greater in the duopoly case as shown in Figure 5. This illustrates the idea that, from the data supplier's point of view, creating a maximal competition between firms and creating a winner-takes-all situation is optimal.

A first investigation of the strategies of firms in the market for data seems to go in the direction of our paper. Indeed, the largest platform to acquire consumers data worldwide, Bluekai (Oracle), auctions off information about consumers. The Bluekai Exchange offers access to "(...) data on more than 300 million users offering more than 30000 data attributes; it processes more than 750 million data events and transacts over 75 million auctions for personal information a day" (OECD, 2013, p. 15). Bidders are typically firms currently engaged in a marketing campaign that require information to tailor offers. According to their website, given how "audiences" are segmented, it is unlikely that two firms get data on the same consumers, even if they belong to the same industry (Bluekai, 2015).

Exclusivity arises quite often for premium rights on media platforms (e.g., sport rights or premium movies). We highlight in this paper that there is an analogy in the context of Big Data, even without network effects. It would be interesting to know more about

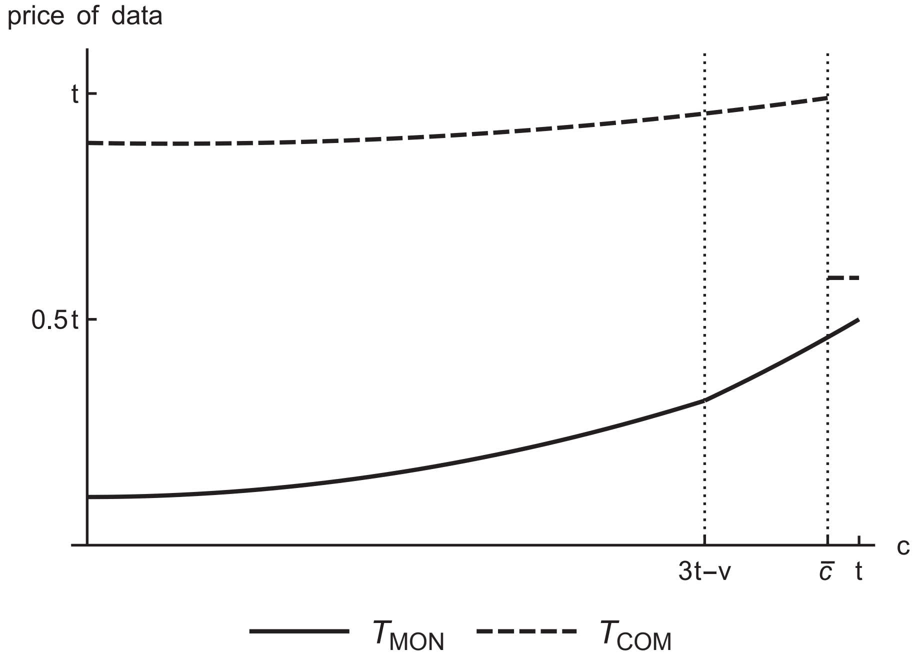

Figure 4: Price of Data for $t = 1$ and $v = 2.2$

exclusivity deals offered by data suppliers, though many details about data transactions are less visible to researchers. Our model predicts that exclusivity should be offered to one firm in a competing market, e.g., one local restaurant, or one manufacturer of training shoes, and so forth. This empirical implication of our results is testable with suitable data.

Finally, we draw a clear policy recommendation concerning privacy and price discrimination. Policy makers can have two tools in our model: a more standard one which makes privacy more or less costly, and the oversight of exclusivity data arrangements. We find that regulators should promote a symmetric allocation of consumers' data across competing firms, but should possibly not advocate for easier privacy when only one firm has consumers' data.[30] Furthermore, an asymmetric allocation of information is doubly inefficient, both because it induces inefficient consumption patterns (oversupply from firm A in both markets in our model) and wasteful privacy expenditures by consumers. Hence, having started with price discrimination concerns in mind, regulation should redirect its focus on the way information is transacted rather than on facilitating privacy for consumers.

[30]For example, it is currently discussed whether bankrupt *Radioshack* should be allowed to sell its customer data collected, potentially, many years ago (Thielman, 2015). This sheds light on the fact that consumers may be myopic when assessing the disclosure of personal data. This issue, related to Gabaix and Laibson (2006), is left for future research.

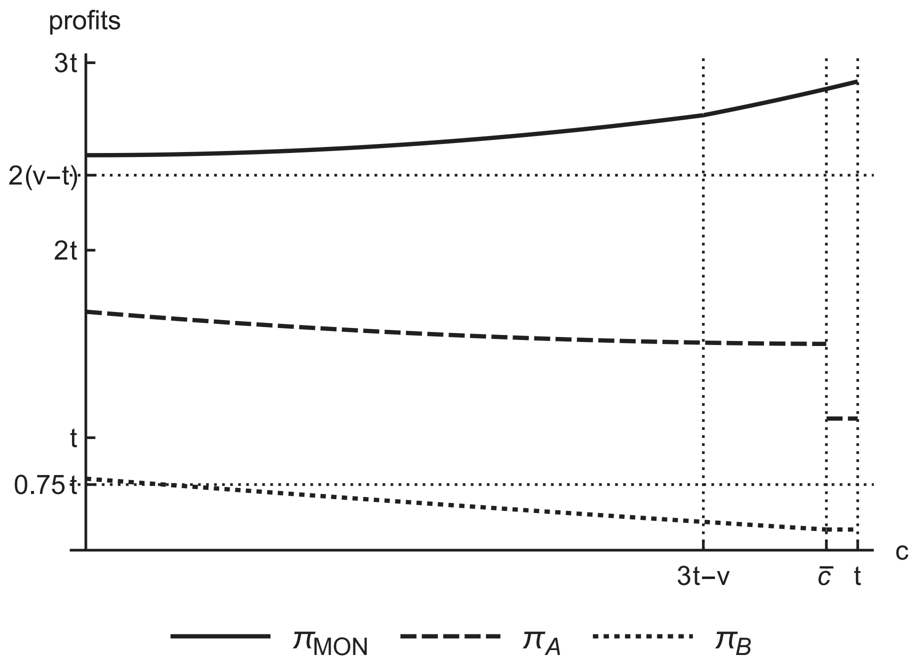

Figure 5: Profits for $t = 1$ and $v = 2.2$

References

[1] Acquisti, A. and H. Varian (2005). "Conditioning Prices on Purchases History", *Marketing Science*, Vol. 14(3), 367-381.

[2] Acquisti, A., C. Taylor and L. Wagman (2015). "The Economics of Privacy", *Forthcoming in the Journal of Economic Literature*.

[3] Belleflamme, P and M. Peitz (2010). *Industrial Organization, Markets and Strategies*, Cambridge University Press.

[4] BlueKai, Oracle (2015). "Frequently Asked Questions", retrieved on 16/04/2015, `http://www.bluekai.com`

[5] Casadesus-Masanell, R. and A. Hervas-Drane (2015). "Competing with Privacy", *Management Science*, Vol. 61(1), 229-246.

[6] Chen, Y. and G. Iyer (2002). "Consumer addressability and customized pricing", *Marketing Science*, Vol. 22(2), 197-208.

[7] Chen, Y., C. Narasimhan, and Z. J. Zhang (2001). "Individual marketing with imperfect targetability", *Marketing Science*, Vol. 20(1), 23-41.

[8] Chiou, L. and C. Tucker (2014). "Search engines and data retention: Implications for privacy and antitrust", *Mimeo*.

[9] Coase R. (1972). "Durability and Monopoly", *Journal of Law and Economics*, Vol. 15(1), 143-149.

[10] Conitzer, V., C. Taylor and L. Wagman (2012). "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases", *Marketing Science*, Vol. 31(2): 277-292.

[11] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector, 2002 O.J. L 201/37.

[12] Esteves, R. (2010). "Pricing with customer recognition", *International Journal of Industrial Organization*, Vol. 28, 669-681.

[13] Fudenberg, D. and J. Tirole (2000). "Customer Poaching and Brand Switching", *RAND Journal of Economics*, Vol. 31(4), 634-657.

[14] Fudenberg, D. and J. Villas-Boas (2012). "Price Discrimination in the Digital Economy", in M. Peitz and J. Waldgofel (Eds.), *The Oxford Handbook of the Digital Economy.*

[15] Gabaix, X. and D. Laibson (2006). "Shrouded attributes, consumer myopia, and information suppression in competitive markets", *The Quarterly Journal of Economics*, Vol. 121 (2), 505-540.

[16] Goldfarb, A. and C. Tucker (2011). "Privacy regulation and online advertising", *Management Science*, Vol. 57(1), 57-71.

[17] Goldfarb, A. and C. Tucker (2012). "Shifts in privacy concerns", *American Economic Review: Papers and Proceedings*, Vol. 102(3), 349-353.

[18] Hirshleifer, J. (1971). "The Private and Social Value of Information and the Reward to Inventive Activity", *American Economic Review*, Vol. 61(4), 561-574.

[19] Jehiel, P. and B. Moldovanu (2000). "Auctions with Downstream Interaction among Buyers", *RAND Journal of Economics*, Vol. 31(4), 768-791.

[20] Koh, B., S. Raghunathan and B. Nault (2013). "Is voluntary profiling welfare enhancing?", *Working Paper.*

[21] Lane, J., V. Stodden, S. Bender, and H. Nissenbaum (2014). *Privacy, Big Data, and the Public Good: Frameworks for Engagement.* Cambridge University Press.

[22] Mayer, J. (2012). "Tracking the Trackers", The Center for Internet and Society at Stanford Law School, retrieved on 16/04/2015, `http://cyberlaw.stanford.edu/blog/2011/07/tracking-trackers-early-results`

[23] Mikians, J., L. Gyarmati, V. Erramilli and N. Laoutaris (2012). "Detecting price and search discrimination on the internet", *Hotnets'12, October 29-30, 2012, Seattle, WA, USA.*, 79-84.

[24] Mikians, J., L. Gyarmati, V. Erramilli and N. Laoutaris (2013). "Crowd-assisted Search for Price Discrimination in E-Commerce: First results", *Mimeo.*

[25] OECD (2013). "Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value", OECD Digital Economy Papers, No. 220, OECD Publishing.

[26] Pancras, J. and K. Sudhir (2007). "Optimal marketing strategies for a customer data intermediary", *Journal of Marketing Research*, Vol. 44(4), 560-578.

[27] Pasquale, F. (2014). "The Dark Market for Personal Data", *The New York Times*, 16/10/14, retrieved on 06/05/15, `http://www.nytimes.com/2014/10/17/opinion/the-dark-market-for-personal-data.html?_r=1`

[28] Posner, R. (1981). "The Economics of Privacy", *American Economic Review*, Vol. 71(2), 405-409.

[29] Shy, O. and R. Stenbacka (2014). "Customer Privacy and Competition", *Mimeo.*

[30] Stigler, G. (1980). "An Introduction to Privacy in Economics and Politics". *The Journal of Legal Studies*, Vol 9(4), 623-644.

[31] Taylor, C. (2004). "Consumer Privacy and the Market for Customer Information", *RAND Journal of Economics*, Vol. 35(4), 631-650.

[32] Taylor, C. and L. Wagman (2014). "Customer privacy in oligopolistic markets: winners, losers, and welfare", *International Journal of Industrial Organization*, Vol. 34, 80-84.

[33] Thielman, S. (2015). "Bankrupt RadioShack aims to sell 67m customer names and addresses", *The Guardian*, 18/05/15, retrieved on 21/07/15, `http://www.theguardian.com/business/2015/may/18/bankrupt-radioshack-selling-customer-data-names-addresses`

[34] Tirole, J. (1988). *The Theory of Industrial Organization*, MIT Press.

[35] Thisse J. and X. Vives (1988). "On The Strategic Choice of Spatial Price Policy", *American Economic Review*, Vol. 78(1), 122-137.

[36] Villas-Boas, J. (1999). "Dynamic competition with customer recognition", *RAND Journal of Economics*, Vol. 30(4), 604-631.

[37] Villas-Boas, J. (2004). "Price cycles in markets with customer recognition", *RAND Journal of Economics*, Vol. 35(3), 486-501.

[38] Wall Street Journal (2012a). "On Orbitz, Mac Users Steered to Pricier Hotels", 23/08/2012, retrieved on 17/05/2015, `http://online.wsj.com/news/articles/SB10001424052702304458604577488822667325882?mg=reno64-wsj\&url=http\%3A\%2F\%2Fonline.wsj.com\%2Farticle\%2FSB10001424052702304458604577488822667325882.html`

[39] Wall Street Journal (2012b). “Websites Vary Prices, Deals Based on Users’ Information”, 24/12/2012, retrieved on 17/05/2015, http://online.wsj.com/news/articles/SB10001424127887323777204578189391813881534?mg=reno64-wsj&url=http\%3A\%2F\%2Fonline.wsj.com\%2Farticle\%2FSB10001424127887323777204578189391813881534.html

[40] Wathieu, L. (2002). “Privacy, exposure and price discrimination”, Technical Report 03-015, Harvard Business School.

[41] Weeds, H. (2015). “TV Wars: Exclusive Content and Platform Competition in Pay TV”, *Economic Journal* (forthcoming).

7 Appendix

In this section we present an interesting stand-alone result that justifies our sequential setting in the choice of basic and tailored prices. The alternative timing of simultaneous choices has the disadvantage of not yielding equilibria in pure strategies.

Lemma A1: *If only one firm has information, and all prices (basic and tailored) are chosen simultaneously, there is no equilibrium in pure strategies.*

Proof. Assume (wlog) that A has information about some consumers. Note that for any p_B, A can tailor a price $p_A(\theta) \in [0, \max\{0, p^I(\theta, p_B)\}]$, where $p^I(\theta, p_B)$ is the price that makes the consumer θ indifferent between buying from A or from B, at p_B. Then, it is a dominant strategy for A to set $p_A(\theta) = \max\{0, p^I(\theta, p_B)\}$. Consider any equilibrium candidate with $p_B > 0$. For any $p^I(\theta, p_B)$, B can deviate and set $p_B - \epsilon$. Then, by assumption, old consumers are not indifferent any more and buy from B. ϵ needs to be small enough so that B's losses from selling to new consumers do not offset the gains from the deviations, and such ϵ always exists. Consider now a candidate equilibrium with $p_B = 0$, which implies that B's profit is null. In this case, $p_A = (2t - c)/3$. But this cannot be an equilibrium since B should then rather choose $p_B > 0$ to maximize its profit on the anonymous market only. Its best reaction is then to set $p_B = (5t - c)/6$ and makes a strictly positive profit. Therefore, for any possible candidate pure-strategy equilibrium, there is a profitable deviation for firm B. We conclude that, when prices are chosen simultaneously, there is no equilibrium in pure strategies. □

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Second Homes: Households' Life Dream or (Wrong) Investment?
Marianna Brunetti and Costanza Torricelli
CEIS Research Paper, 351 ,August 2015

Insider's Dilemma: a General Solution in a Repeated Game
Berardino Cesi and Walter Ferrarese
CEIS Research Paper, 350 ,July 2015

The Life Saving Effects of Hospital Proximity
Paola Bertoli and Veronica Grembi
CEIS Research Paper, 349 ,July 2015

The (W)health of Nations: The Impact of Health Expenditure on The Number of Chronic Diseases
Leonardo Becchetti, Pierluigi Conzo and Francesco Salustri
CEIS Research Paper, 348 ,June 2015

Seasonal Changes in Central England Temperatures
Tommaso Proietti and Eric Hillebrand
CEIS Research Paper, 347 ,June 2015

Experience and Gender Effects in an Acquiring-a-Company Experiment Allowing for Value Messages
Daniela Di Cagno, Arianna Galliera, Werner Güth, Noemi Pace and Luca Panaccione
CEIS Research Paper, 346 ,June 2015

Contract and Procurement Design for PPPs in Highways: the Road Ahead
Elisabetta Iossa
CEIS Research Paper, 345 ,June 2015

Generalised Partial Autocorrelations and the Mutual Information Between Past and Future
Alessandra Luati and Tommaso Proietti
CEIS Research Paper, 344 ,June 2015

A Hybrid Public Good Experiment Eliciting Multi-Dimensional Choice Data
Daniela Di Cagno, Arianna Galliera, Werner Güth and Luca Panaccione
CEIS Research Paper, 343 ,May 2015

Does Experience Rating Improve Obstetric Practices? Evidence From Geographical Discontinuities in Italy
Sofia Amaral-Garcia, Paola Bertoli and Veronica Grembi
CEIS Research Paper, 342 ,April 2015

DISTRIBUTION

Our publications are available online at www.ceistorvergata.it

DISCLAIMER

COPYRIGHT

MEDIA INQUIRIES AND INFORMATION

For media inquiries, please contact Barbara Piazzi at +39 06 72595652/01 or by e-mail at ufficiostampa@ceis.uniroma2.it. Our web site, www.ceistorvergata.it, contains more information about Center's events, publications, and staff.

DEVELOPMENT AND SUPPORT

For information about contributing to CEIS Tor Vergata, please contact Carmen Tata at +39 06 72595615/01 or by e-mail at segr.ceis@economia.uniroma2.it