EconomiX

# Competition, Privacy, and Multi-Homing

Jean-Marc Zogheib

2023-34 Document de Travail/ Working Paper

EconomiX - UMR 7235 Bâtiment Maurice Allais
Université Paris Nanterre 200, Avenue de la République
92001 Nanterre Cedex

Site Web : economix.fr
Contact : secreteriat@economix.fr
Twitter : @EconomixU

Université Paris Nanterre

# Competition,Privacy, and Multi-Homing

Jean-Marc Zogheib*†

**Abstract**

Two digital firms compete in prices and information disclosure levels. A consumer signing up to one firm's service decides how much personal information to provide. We find that firms essentially trade-off between consumer valuations and disclosure levels to determine their business strategies when consumers single-home. Under multi-homing, business strategies are more complex to assess and may completely shift compared to single-homing. All things being equal, implementing a strict privacy regime with no data disclosure can be optimal under single-homing, while a soft privacy regime with data disclosure may be preferred under multi-homing.

**Keywords**: competition, online privacy, information disclosure, multi-homing.
**JEL codes**: D11; D40; L21; L41.

## 1 Introduction

Online privacy has turned into a critical variable of competition between digital firms. They serve online users subject to growing privacy concerns regarding the collection and use of their data. In a survey conducted June 3-17 2019 by the Pew Research Center on 4,272 U.S. adults, it was found that 79% of them are concerned about how much data companies collect about them while 52% said they decided recently not to use a product or service because they were worried about how much personal information would be collected about them.[1] Consequently, how a firm determines the privacy level of its service can have critical implications on profitability and market position.

Competition in digital markets is shaped by the exploitation of consumer data. A digital firm can offer consumers low-price or even free services, thereby attracting consumer attention through higher consumer data collection. Then, it can monetise consumer data on a data market, e.g., to data intermediaries. This type of business model is driven by indirect network externalities whereby more consumers increase the value of joining a firm's service for data intermediaries. As a result, consumer information disclosure

*Economix, UPL, University Paris Nanterre, CNRS, University Paris Nanterre and EconomiX-CNRS, Nanterre F92000, France. Email: jmzogheib@parisnanterre.fr

†This is an updated version of the paper "Privacy, Competition, and Multi-Homing.

[1] Pew Research Center (June 3, 2019) (1) and (2).

and the pricing framework are strongly interlinked and affect competition dynamics. For instance, video streaming platforms such as Disney+, Netflix, or Prime Video are now developing plans with ads at a lower price in addition to their ad-free premium plan due in response to competitive pressure.[2] Online newspapers usually propose paying ad-free plans and non-paying plans with ads. In the online shopping sector, cashback firms such as Pogo, Dosh, or Drop even allow consumers to be paid for providing their data.[3]

Consumers' online behaviours are of first importance in the competitive framework. First, digital users may provide a lower amount of personal information to firms due to privacy concerns, e.g., by using privacy-enhancing technologies such as Adblock or Ghostery, as well as opt-out behaviours towards consent.[4] Moreover, consumers have the ability to sign up for the service of only one firm (*single-home*) or to sign up for the service of two or more firms (*multi-home*). In other words, a consumer may buy different variants of horizontally differentiated services, depending on the extent to which service functionalities overlap. The latter applies well to the abovementioned examples and can further alter competition dynamics: what about the pricing framework and consumer information disclosure in a context where digital firms have *overlapping* consumers? How is consumer data monetised onto a data intermediary? To what extent firms' business stratégies are changed in the presence of multi-homing?

In this paper, we examine the impact of competition between two digital firms in prices and information disclosure levels. Firms supply a horizontally differentiated online service to consumers and can disclose data from these consumers to a monopoly data intermediary. Consumers observe the level of disclosure to which firms engage and their price before deciding which service to patronise and how much personal information to provide. The level of information disclosure is an inverse measure of privacy. The perceived quality of the firm's service for each consumer thus increases with information provision and decreases with the firm's disclosure level. Firms thus derive revenue from (i) purchase revenues from the prices charged to consumers and (ii) information disclosure revenues obtained from the data intermediary.[5]

We study a single-homing (SH) framework and find that the trade-off between consumer valuations and disclosure levels is a crucial determinant of firms' business strategies. Higher consumer valuations are associated with a higher level of privacy (i.e., a lower disclosure

[2]Disney+ charges \$7.99 per month for its plan with ads whereas the ad-free plan costs \$13.99 per month. Disney to raise price on ad-free Disney+ to \$13.99 per month starting October 12, CNBC, 09 August 2023.

[3]See Pogo's, Dosh's, and Drop's presentation pages.

[4]Opting out means a user takes action to withdraw their consent. In contrast, opting in means that a user will take an affirmative action to offer their consent. In this respect, the recent Never-Consent option on Ghostery (2023) is quite interesting.

[5]To some extent, and as in Casadesus-Masanell and Hervas-Drane (2015) and Anderson, Foros, and Kind (2019), this framework can be called "two-sided" in that digital firms can derive revenues not only from the demand side (consumers) but also from the supply side (data intermediary).

level) and, in turn, a higher price and less consumer information provision. We then distinct between two privacy regime. First, a strict privacy regime if consumer valuations are sufficiently high in which firms do not engage in the disclosure of consumer data and rely exclusively on purchase revenues. Second, a soft privacy regime where firms engage in consumer data disclosure and rely more on disclosure revenues as consumer valuations decreases. The price charged to consumers represents a trade-off between purchase and disclosure revenues per consumer, and above some thresholds, consumer subsidisation becomes optimal. Firms represent bottlenecks on the data from their users and charge the data intermediary a price that leaves her with zero surplus.

We then examine the possibility of multi-homing (MH), where it is more difficult for firms to monetise MH consumer data as it may be sold twice to the data intermediary. We find that firms' business strategies rely on a complex trade-off between different economic parameters; on top of consumer valuations, the value of MH, product differentiation and MH consumer data valuation now impact information disclosure levels. A strict or soft privacy regime may then be implemented depending on a combination of these parameters. We distinguish three possible scenarios. First, if MH consumer data is not too valuable, consumer valuations and the value of MH negatively impact the firm's disclosure level, whereas it is the contrary for product differentiation. It entails that with a lower competition intensity, there is less privacy, price decreases, and consumers are less willing to provide information. Second, if MH consumer data is sufficiently valuable, we find the same type of results as in the first scenario, provided that consumer marginal disutility from a lower privacy level offsets the firm's marginal disclosure revenue. When the latter condition does not hold, we shift to a third scenario where the firm charges a maximal disclosure level.

MH consumer data valuation is a decisive component of our equilibrium analysis. Notably, the positive relation between MH consumer data valuation and disclosure levels is complex. It depends on how MH consumer data is monetised onto the data intermediary, the marginal benefit from a higher disclosure level, and the privacy-sensitiveness of consumers. If the disclosure level decreases with MH consumer data valuation, consumers provide more information, while it is unclear whether a firm compensates for the lower disclosure by increasing its price. Note that the price charged to consumers represents a trade-off between purchase and disclosure revenues per MH consumer. Consumer subsidisation becomes optimal above some thresholds, provided that MH consumer data is not valueless.

In an extension, we study the situation of a merger to a multi-product monopoly. Under SH and by specifying the consumer utility function, we find that a merger increases market power and has no effect on privacy, ultimately harming consumers. With the possibility of MH, a merger increases market power and positively affects privacy if MH consumer data valuation is sufficiently high. Otherwise, the effect on privacy is ambiguous.

In the next section, we present the related literature. Section 3 presents the model framework. In Section 4, we solve for firms' decisions when consumers single-home. In Section 5, we examine the case where consumers can also multi-home. Section 6 explores some extensions. All proofs are in the appendix.

## 2 Relation to literature

Our paper is related to the literature on the economics of privacy. A strand of the literature investigates the link of privacy with allocation efficiencies and externalities (Stigler, 1980; Posner, 1981; Hermalin et Katz, 2006; Calzolari and Pavan, 2006; Hui and Png., 2006). Calzolari and Pavan (2006) consider information disclosure between two firms (principals) interested in discovering consumers' willingness to pay. In a model of sequential contracting, a common agent strategically decides whether to report her true type. They find that the transmission of personal data from one company to another may in some cases reduce information distortions and enhance social welfare.[6] We contribute to this literature by setting a framework where consumers strategically decide to provide some of their data to the firm they patronise. Consumers have privacy concerns over the disclosure of their data while providing more information to firms is beneficial to consumers through more personalised service.

Another strand of the literature studies the link between privacy and competition (Noam, 1995a,1995b; Spulber, 2009; Taylor and Wagman, 2014; Casadesus-Masanell and Hervas-Drane, 2015; Shy et al., 2016; Montes et al., 2018; Choi et al., 2019; Lefouili and Toh, 2020, Kim, 2020; Argenziano and Bonatti, 2020; Ichihashi, 2020a). Casadesus-Masanell and Hervas-Drane (2015) analyse the effect of competition on consumer privacy in a model of vertical differentiation. Consumers voluntarily provide personal information to firms in order to obtain higher-quality products. Firms can disclose and sell some of this information to an outside firm, thereby harming consumer utility. The authors show that (i) one firm positions itself as a high-quality (low-disclosure) provider, and the other firm as a low-quality (high-disclosure) provider, and (ii) moving from monopoly to duopoly implies more disclosure by the firms, but it does not necessarily harm consumer welfare. We build on the framework of Casadesus-Masanell and Hervas-Drane (2015). Our paper differs in that we examine a model of horizontal differentiation with a general utility function where consumers either single or multi-home. We find that moving from monopoly to duopoly has no effect on information disclosure under SH whereas it is the contrary under MH. Moreover, we model the interaction between digital firms and the data intermediary, where the former intends to monetise SH and MH consumer data by disclosing it to the latter.

[6]See also Acquisti and Wagman (2016) for a comprehensive literature review on the economics of privacy.

Our paper has connections with a strand of the digital economics literature dealing with the impact of collection and use of consumer data on the competition between digital firms (Prufer and Schottmuller, 2017; Prat and Valletti, 2019; Kim et al., 2019; Belleflamme et al., 2020; De Cornière and Taylor, 2020; Ichihasi, 2020b).[7] Prat and Valletti (2019) study digital platforms as attention brokers with proprietary information about their users' product preferences and sell targeted ad space to retail product industries. They show that increased concentration among attention brokers may reduce entry into retail product industries, which ends up harming consumer welfare. A monopolistic attention broker has an incentive to create an attention bottleneck by reducing the supply of targeted advertising. The authors finally evaluate that a merger between platforms can increase market power in the retail industry to the detriment of consumers. In our framework, digital firms are attention brokers because they are bottlenecks for access to consumer data. However, this bottleneck position is weakened because of consumer multi-homing and the data intermediary is left with a positive surplus. In such a case, firms may find reducing information disclosure profitable and even entering a strict privacy regime.

Our paper is finally related to the seminal literature digital economics literature on two-sided markets and the impact of consumer single and multi-homing (Rochet and Tirole, 2003; Caillaud and Jullien, 2003; Armstrong, 2006; Choi, 2010; Belleflamme and Peitz, 2019; Anderson et al., 2019). Anderson et al. (2019) study the impact of consumer multi-homing on market equilibrium and performance. The authors show that equilibrium consumer prices are independent of the number of platforms when some but not all consumers multi-home. On the contrary, advertising prices decrease as the fraction of multi-homers increases. They conclude that compared to single-homing, multi-homing flips the side of the market on which platforms compete. In our paper, we find that due to monetisation issues, multi-homing tends to decrease the selling price of consumer data, which negatively affects disclosure revenues. In a companion paper, Anderson et al. (2017) examine the characteristics of single-homing and multi-homing equilibrium in a one-sided market. The authors assume that each product has its own specific part, while a common overlapping part belongs to both products. Without product overlap, allowing multi-homing should be better for the firms because they have greater demand and less fierce competition. With overlap, each firm cannot charge for the common part, as they compete à la Bertrand. Therefore, allowing multi-homing could make the firms worse off because of overlap and horizontal differentiation. As Anderson et al. (2017), we assume that consumers choose to multi-home by considering the overlapping between the two digital services and how it affects price and disclosure revenues.

[7] Closer to targeted advertising, see Tag (2009), Anderson and Gans (2011), and De Cornière and Nijs (2014, 2016).

# 3 Model

**Setting**. We study a market where two digital firms, denoted by $A$ and $B$, compete in prices and information disclosure levels. They are located on a Hotelling line and supply a service to consumers at zero marginal cost. Firms are located at the two extremes of a line of length 1, with firm $A$ located at point 0 and firm $B$ at point 1. Consumers are uniformly distributed along the line and choose to sign up for a service from firm $A$, firm $B$, or both if possible.

**Single-homing (SH)**. A consumer purchasing the service of firm $i = A, B$ decides on the level of personal information $y_i \geq 0$ to provide to this firm. The net utility of the consumer, located at $x \in [0, 1]$, when purchasing service $i$, for given price $p_i$ and information disclosure $d_i \geq 0$, is

$$U_i = vq(y_i, d_i) - p_i - tx_i, \tag{1}$$

where $v > 0$ is a parameter which reflects the intrinsic benefit of the service, and $tx_i = tx$ is the transportation cost incurred when buying for firm $i = A$ and $tx_i = t(1 - x)$ the transportation cost when buying from firm $i = B$.

We interpret term $q(y_i, d_i)$ in equation (1) as the *informational* quality of firm $i$'s service,[8] where quality is assumed to be increasing and concave in the level of information provision $y_i$ (i.e., $\partial q/\partial y_i \geq 0$ and $\partial^2 q/\partial y_i^2 \leq 0$). It means that consumers benefit from providing information because it allows the firm to provide a personalised, higher-quality service; but this benefit is decreasing at the margin. Moreover, quality is assumed to be decreasing and concave in the level of information disclosure $d_i$, meaning that consumers incur disutility from the disclosure of their personal information (i.e., $\partial q/\partial d_i \leq 0$ and $\partial q^2/\partial d_i^2 \leq 0$ ). Finally, as $d_i$ increases, providing more information negatively affects consumer utility ($\partial q/\partial y_i \partial d_i \leq 0$). For the sake of exposition, let $q(y_i, d_i) \equiv q_i$.

**Possibility of multi-homing (MH)**. Let us characterise how consumers value MH over SH. Let $\sigma \in [0, 1]$ represent the incremental value of signing up for a second service. If $\sigma = 1$, there is no overlap between the two services and consumers derive the full utility from the second service. On the other hand, if $\sigma = 0$, there is a significant overlap and the consumption of the second service brings no incremental gross utility.

The utility of a consumer who purchases service $j$ in addition to service $i$ is

$$U_{i,j} = U_i + \{\sigma vq(y_j, d_j) - p_j - tx_j\} \tag{2}$$

If $\sigma = 0$, the consumer does not benefit from the consumption of a second service and we assume that there is no MH (i.e., $vq(y_i, d_i) - p_A - p_B - t < 0$). As $\sigma$ increases, the consumer increasingly values the consumption of a second service, which implies that MH is more

[8] We use the same term as in Casadesus-Masanell and Hervas-Drane (2015).

valuable. If $\sigma = 1$, there is no overlap between services $A$ and $B$ and $U_{A,B} = U_{B,A}$.[9]

**Payoffs**. Firm $i$ decides on a price $p_i \in \mathbb{R}$ and a disclosure level $d_i \in \mathbb{R}^+$. Firms derive revenues from purchases, and we allow them to set negative prices (i.e., to subsidise consumers). Firms can also derive revenues by disclosing consumer data on a data market. More precisely, firm $i$ sells consumer data to a monopoly data intermediary $D$, at a price $r_i$ per user and piece of information. It means that firm $i$'s revenue from disclosing the information of one consumer is $d_i y_i \times r_i$. Let $D_i(.)$ be the demand of firm $i$. The profit of firm $i$ is then $\Pi_i = (p_i + d_i y_i r_i)\, D_i$.

We now characterise the payoff function of the data intermediary. Let $s^d > 0$ be the value the data intermediary places on the information of each user signing up to service $i$. When acquiring firm $i$'s consumer data, the data intermediary $D$ payoff is thus given by $\left(s^d - r_i\right) D_i$. For simplicity, we assume that $s^d = 1$.

We finally determine how the data intermediary $D$ values MH consumer data. Since the information of a MH consumer is in possession of both firms, it can be sold twice and becomes less valuable to $D$. Let $\alpha s^d = \alpha$ be the value of the information of each user when sold twice to $D$, where $\alpha \in (0, 1]$.[10] In other words, MH consumer data are less valuable if $\alpha < 1$, thereby negatively affecting disclosure revenues.[11]

We make the following stability assumption.

**Assumption 1.**

$$D_i \left| \frac{\partial^2 D_i}{\partial d_i^2} \right| > \left( \frac{\partial D_i}{\partial d_i} \right)^2 .$$

**Timing.** We consider the following sequence of events. At the first stage, firms simultaneously set their disclosure level.[12] At the second stage, firms simultaneously set the price of their service to consumers and the price of consumer data charged to the data intermediary $D$. At the third stage, having observed prices and disclosure levels, consumers choose to purchase firm $A$'s, firm $B$'s service, or both if possible, or to stay out of the market. At the fourth stage, consumers decide on the level of information provision to the firm they have patronised.

We look for the subgame perfect equilibrium of this game.

[9]Note that Equation (2) could be rewritten $U_{i,j} = U_i + U_j - (1-\sigma)vq(y_j, d_j)$, where $(1-\sigma)vq(y_j, d_j)$ can be interpreted as the MH disutility from purchasing service $j$ in addition to service $i$; this disutility increases as $\sigma$ decreases.

[10]Anderson et al. (2019) use this type of modelling with $\alpha \in [0, 1]$.

[11]Notably, $\alpha$ could be interpreted as an "expected" value placed by the data intermediary when consumer data is sold twice. Another possible interpretation is that each piece of information of MH users takes only two extreme values, 0 with probability $1 - \alpha$ (i.e., the data is not valuable) or 1 with probability $\alpha$ (i.e., the data is fully valuable). In this case, the expected value of data per MH user for the data intermediary would be $\alpha \times 1 + (1 - \alpha) \times 0 = \alpha$.

[12]As in Casadesus-Masanell and Hervas-Drane (2015), firms commit to the information disclosure levels announced in the first stage.

# 4 Single-homing

We start by considering single-homing, a situation where consumers purchase one service or none. In what follows, we determine the equilibrium under duopoly and restrict our attention to parameter values such that the market is covered in equilibrium.

At the fourth stage, a consumer decides on the level of information to the firm she has patronised, let us say firm $i$. For a given disclosure level $d_i$ and price $p_i$, the consumer chooses $y_i$ by maximizing (1), which gives

$$y_i^c(d_i) \equiv \arg\max_{y_i} vq(y_i, d_i) - p_i - tx_i. \tag{3}$$

Using the implicit function theorem, we show that the higher the firm $i$'s disclosure level is, the less a consumer is willing to provide information

$$\frac{\partial y_i}{\partial d_i}_{|y_i = y_i^c(d_i)} = -\frac{\partial^2 q_i / \partial y_i \partial d_i}{\partial^2 q_i / \partial y_i^2}_{|y_i = y_i^c(d_i)} \leq 0.$$

Moreover, let us assume that $\partial^2 y_i^c(d_i)/\partial d_i^2 \leq 0$. To save notations, let $y_i^c(d_i) = y_i^c$.

At the third stage, we compute firms' demands by determining the location of the consumer who is indifferent between $A$ and $B$. Replacing for $y_i^c$ into (1), we find that the indifferent consumer is located at

$$x^* = \frac{1}{2} - \frac{p_A - p_B}{2t} + \frac{v}{2t}(q_A - q_B). \quad \text{where} \quad q_i \equiv q(y_i^c, d_i). \tag{4}$$

The demand of firm $A$ is therefore $D_A(p_A, p_B) = x^*$, while the demand of firm $B$ is $D_B(p_A, p_B) = 1 - x^*$.

At the second stage, firm $i$ sets the price of consumer data to the data intermediary $D$. Each firm makes a take-it-or-leave-it offer to the data intermediary because they represent a bottleneck for access to consumer data. The data intermediary is willing to buy firm $i$'s consumer data as long as $r_i \leq s^d = 1$. Firm $i$ therefore charges the data intermediary $r_i = 1$, and D is left with zero surplus.

Firm $i$'s profit is then given by

$$\Pi_i(p_A, p_B) = (p_i + d_i y_i^c) D_i(p_A, p_B). \tag{5}$$

From (5), we observe that firm $i$ derives profits from selling its service at a price $p_i$, i.e., through purchase revenues, and from consumer data $y_i^c$ at a price $r_i = 1$ and a disclosure level $d_i$, i.e., through disclosure revenues.

We now write the following lemma on firm $i$'s disclosure revenue per consumer, $d_i y_i^c$.

**Lemma 1.** *Firm $i$'s disclosure revenues per consumer $d_i y_i^c$ are concave and increasing in $d_i$ if and only if $d_i \in (0, \hat{d}]$, where $\hat{d} \equiv \arg\max_{d_i} d_i y_i^c$ and $\hat{d} > 0$.*

From Lemma 1, firm $i$'s disclosures revenues per consumer are maximized at $d_i = \hat{d}$, where $\hat{d} \equiv \arg\max_{d_i} d_i y_i^c$. It implies that it does not benefit from setting a disclosure level $d_i$ higher than $\hat{d}$.

To ensure that an interior solution is obtained when firm $i$ sets the disclosure level $d_i$, we make the following assumption:

**Assumption 2.** A disclosure level $d_i$ is such that $d_i \in (0, \hat{d}]$.

A higher disclosure level $d_i$ affects firm $i$'s disclosure revenues per consumer through two opposite effects. On the one hand, there is a positive effect coming from the consumer providing personal information to firm $i$ ($y_i^c > 0$). On the other hand, the consumer suffers from higher information disclosure and then provides less information to firm $i$ at the margin ($d_i \partial y_i^c / \partial d_i \leq 0$). By Assumption 2, $\partial(d_i y_i^c)/\partial d_i \geq 0$.

Each firm sets its price $p_i$ to maximise its profit, which is given by equation (5), taking the rival's price $p_j$ as given. Solving for the price reaction functions, we obtain the equilibrium prices[13]

$$p_i^c(d_i, d_j) = t + \frac{v(q_i - q_j) - 2d_i y_i^c - d_j y_j^c}{3}.$$

Let us study how disclosure levels $d_i$ and $d_j$ affect firm $i$'s price.

$$\frac{\partial p_i^c}{\partial d_i}(d_i) = \frac{v}{3}\frac{\partial q_i}{\partial d_i} - \frac{2}{3}\frac{\partial(d_i y_i^c)}{\partial d_i} \leq 0; \quad \frac{\partial p_i^c}{\partial d_j}(d_j) = -\frac{v}{3}\frac{\partial q_j}{\partial d_j} - \frac{1}{3}\frac{\partial(d_j y_j^c)}{\partial d_j}. \tag{6}$$

The first expression in (6) establishes that a higher disclosure level $d_i$ affects negatively firm $i$'s price because of a lower perceived quality of service and higher disclosure revenues. The second expression in (6) is more complex. On the one hand, a higher $d_j$ positively affects firm $i$'s price since consumers who subscribe to firm $i$'s service benefit from a relatively better quality, which induces firm $i$ to increase its price. On the other hand, a higher $d_j$ negatively affects $p_i^c$ since it negatively affects firm $j$'s price $p_j^c$, due to strategic complementarity.

Plugging equilibrium prices into the profit function (5), we now solve for firms' optimal disclosure levels at Stage 1. Firm $i$'s profit can be rewrited

$$\Pi_i(d_i, d_j) = 2t\left(D_i(d_i, d_j)\right)^2 \quad \text{where} \quad D_i(d_i, d_j) = \frac{1}{2} + \frac{v(q_i - q_j) + d_i y_i^c - d_j y_j^c}{6t}.$$

Each firm sets its disclosure level $d_i$ to maximise its profit, taking the rival's disclosure level $d_j$ as given. Solving for the disclosure reaction functions, we obtain the equilibrium disclosure levels, $d_A^c$ and $d_B^c$.

[13] Price reaction function are denoted by $p_i(p_j) = (t + v(q_i - q_j) - d_i y_i^c + p_j)/2$ ($i = A, B$). Moreover, the second-order condition is always satisfied, as $\partial^2 \Pi_i / \partial p_i^2 = -1/t < 0$.

The following proposition characterises the duopoly equilibrium:

**Proposition 1.** *(i) If $0 < t \leq vq(y^c, d^c) + d^c y^c/2$, consumers provide information $y_i = y^c$. Firms' optimal prices and disclosure levels are $p_i = p^c = t - d^c y^c$ and $d_i = d^c$.*

*(ii) Firms' decision to charge positive disclosure levels depends only on consumer valuations $v$, where $\partial d^c/\partial v \leq 0$.*

*(iii) The data intermediary is left with zero surplus.*

From Proposition 1, a central result is that disclosure levels are determined by consumer valuations $v$. Indeed, $d^c$ decreases with $v$, which implies that higher consumer valuations are associated with a higher privacy level (lower $d^c$). In turn, disclosure levels affect firms' pricing and consumer information provision. To clarify intuitions, suppose that there exists $v^c > 0$ such that $d^c > 0$ if $v < v^c$ and $d^c = 0$ otherwise.

If consumer valuations for the service are sufficiently high ($v \geq v^c$), firms do not engage in information disclosure ($d^c = 0$) and charge positive prices ($p^c = t$). Firms rely exclusively on purchase revenues and thus adopt a *strict privacy regime*.

If consumers have low valuations ($v < v^c$), firms engage in information disclosure ($d^c > 0$). As $v$ increases, the level of information disclosure decreases ($\partial d^c/\partial v \leq 0$) whereas consumers are willing to provide more information ($\partial y^c/\partial v \geq 0$) and prices increase ($\partial p^c/\partial v \geq 0$). As $v$ decreases, firms rely comparatively more on revenues from the disclosure of consumer data than purchase revenues. In this respect, the equilibrium price $p^c$ represents firm $i$'s trade-off between purchase revenues ($t$) and disclosure revenues ($d^c y^c$) per user. As consumer valuations $v$ decrease, $d^c$ increases and firm $i$ charges a lower price $p^c$, which explains why it relies more on disclosure revenues. In sum, if $v < v^c$, firms adopt a *soft privacy regime*.

**Corollary 1.** *Negative pricing exists under a soft privacy regime if, for $d^c > 0$, product differentiation $t$ or consumer valuations $v$ are sufficiently low.*

From Corollary 1, firm $i$ may charge a negative price, that is, subsidising consumers. It may happen when competition intensifies (lower $t$) or consumer valuations ($v$) for the service are low, as this can be seen on Figure 1 (left).

## Numerical application

We solve the model in a horizontal differentiation framework using the utility function in Casadesus-Masanell and Hervas-Drane (2015). The net utility of the consumer has the following form:

$$U_i = vy_i(1 - y_i - d_i) - p_i - tx_i,$$

At Stage 4, we find that a consumer chooses the level of information provision to the firm she has patronised $y_i^c = (1 - d_i)/2$. At Stage 3, the indifferent consumer is given by

$$x^* = \frac{1}{2} - \frac{p_A - p_B}{2t} + \frac{v}{8t}\left((1-d_A)^2 - (1-d_B)^2\right),$$

At Stage 2, the selling price of consumer data is the monopoly price, that is, $r_i = 1$. Firm $i$ chooses the price $p_i$ that maximizes its profit and its equilibrium price is $p_i^c(d_A, d_B)$ (the SOCs are satisfied: $\partial^2\Pi_i/\partial p_i^2 = -1/t < 0$). At Stage 1, we find that the equilibrium disclosure levels (such that the second-order conditions are satisfied) are symmetric and given by $d_i^c = (1-v)/(2-v)$.

**Result 1.** *In the duopoly equilibrium, the market is covered and consumers provide information* $y^c = (1 - d^c)/2$.

*If* $\begin{cases} 0 < v < 1, \\ 0 < t \leq 1/6(2-v), \end{cases}$ *firms' optimal prices and disclosure levels are* $\begin{cases} p^c = \frac{v-1}{2(v-2)^2} + t, \\ d^c = \frac{1-v}{2-v}; \end{cases}$

*If* $\begin{cases} v \geq 1, \\ 0 < t \leq v/6, \end{cases}$ *firms' optimal prices and disclosure levels are* $\begin{cases} p^c = t, \\ d^c = 0. \end{cases}$

Figure 1 (right) illustrates the duopoly equilibrium. We observe that firms face a trade-off between the disclosure of consumer data ($d^c$), the price ($p^c$), and the level of information provision ($y^c$), depending on consumer valuations ($v$). A higher disclosure level decreases the level of information provision and prices. Consumers with low valuations provide less information, but this information is more exploited to generate larger disclosure revenues, while these consumers may be subsidised (soft privacy regime). Consumers with high valuations provide more information, but firms generate more value through purchase revenues (strict privacy regime).

# 5 Multi-homing

We now consider that some consumers may purchase both services (i.e., multi-home).

Figure 2 depicts a possible market outcome, where consumers located to the left of point $x_B^0$ purchase service $A$ only, those on the right of $x_A^0$ purchase service $B$ only, and finally consumers between $x_B^0$ and $x_A^0$ purchase both.

We solve for the equilibrium in the duopoly case.

At Stage 4, each consumer decides on the level of information to provide to the firm(s) he has patronised. If the consumer signs up for service $i$ only, he chooses $y_i$ by maximising (1), and the level of information provision is then given by (3). If the consumer signs up for both services, he chooses $y_A$ and $y_B$ to maximise $U_{i,j}$, which gives $y_i^c(d_i) \equiv \arg\max_{y_i} U_{i,j}$.

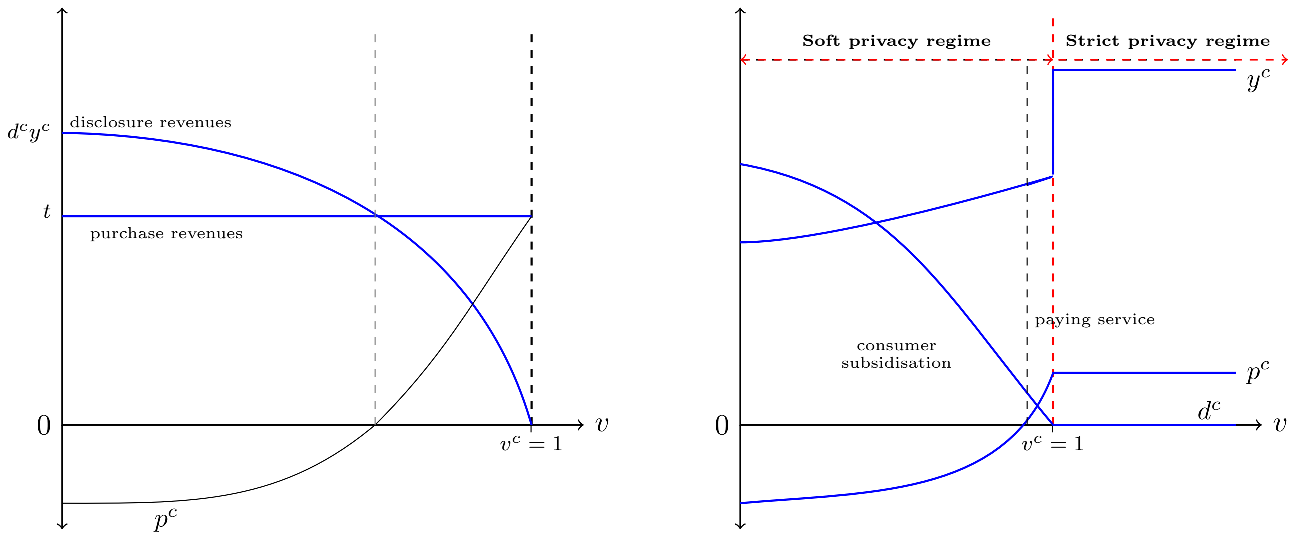


Figure 1: *Left*: Pricing trade-off in soft privacy regime (t=0.04);
*Right*: Equilibrium

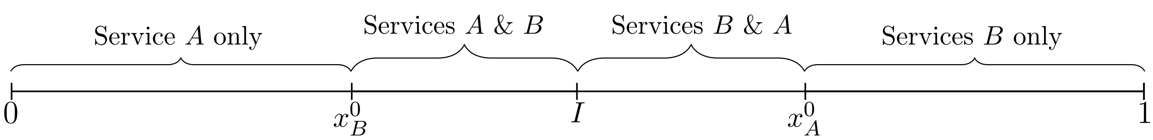


Figure 2: Demand composition with multi-homing

Therefore, MH and SH consumers provide the same level of information when using service $i$.

At Stage 3, consumers choose which service(s) to patronise.

With our specification, a consumer who purchases service $i$ in addition to service $j$ does not necessarily derive the same utility as if she purchases $j$ in addition to $i$, i.e., we can have either $U_{AB} > U_{BA}$ or $U_{AB} < U_{BA}$.

We therefore make the following simplifying assumption.

**Assumption 3.** MH consumers to the left of $I = 1/2$ sign up for service $B$ in addition to service $A$ and the ones to the right of $I$ sign up for service $A$ in addition to service $B$

We justify this assumption insofar as MH consumers incur disutility from firstly signing up for a service far from their location; MH consumers to the left of 1/2 (to the right of 1/2) are therefore assumed to primarily sign up for service $A$ (service $B$). This disutility from location is exogenous to this framework and is simply assumed (Figure 2).[14]

The location of the consumer who is indifferent between purchasing service $A$ and purchasing both services is then given by $U_A = U_{A,B}$ (location $x_B^0$ on Figure 2). Similarly,

[14] In a framework focused on product functionalities and location, Anderson et al. (2017) specify a utility function where a consumer who purchases service $i$ for its incremental value over service $j$ does so depending on its location $x$.

the location of the consumer who is indifferent between purchasing service $B$ and purchasing both services is given by $U_B = U_{B,A}$ (location $x_A^0$ on Figure 2).

Let $D_i^{sh}(.)$ be the SH demand of firm $i$, $D^{mh}(.)$ be the MH demand, common to firms $A$ and $B$, and $D_i(.)$ be firm $i$'s total demand. We then have

$$\begin{cases} D_i^{sh}(p_j) = 1 - \frac{\sigma v q_j - p_j}{t}, \quad i \neq j, & \text{where} \quad \begin{cases} x_B^0 \equiv D_A^{sh}(p_B), \\ 1 - x_A^0 \equiv D_B^{sh}(p_A), \end{cases} \\ D^{mh}(p_A, p_B) = \frac{\sigma v(q_A+q_B)-(p_A+p_B)}{t} - 1, & \text{where} \quad x_A^0 - x_B^0 \equiv D^{mh}(p_A, p_B). \end{cases}$$

$$\text{and} \quad D_i(p_i) = \frac{\sigma v q_i - p_i}{t}, \quad i = A, B, \quad \text{where} \quad D_i = D_i^{sh} + D^{mh}. \tag{7}$$

Firm $i$'s total demand therefore only depends on its own price and disclosure level. However, the composition of firm $i$'s demand between SH and MH consumers depends on the prices and disclosure levels of the two firms.

At Stage 2, each firm determines the selling price of consumer data, charged to the data intermediary $D$. Firms are bottlenecks on SH consumer data and then charge the monopoly price $r_i = 1$ to $D$. However, MH consumer data may be sold twice (i.e., by both firms) to the data intermediary, meaning each firm $i$ would charge $r_i = \alpha$ to $D$, where $\alpha \in (0, 1]$.

We write the following lemma on the selling of consumer data to $D$ under MH.

**Lemma 2.** *The data intermediary $D$ obtains SH consumer data from firm $i$ at a price $r_i = 1$ whereas she obtains MH consumer data from both firms at a price $r_i = \alpha$.*

From Lemma 2, we show that selling MH consumer data is a Nash equilibrium for both firms. Firms are therefore unable to "fully" value MH consumer data since $\alpha \leq 1$, which will ultimately affect disclosure revenues.

Let $D_i^{sh}(p_j) + \alpha D^{mh}(p_i, p_j) = (1-\alpha) D_i^{sh}(p_j) + \alpha D_i(p_i)$. Firm $i$'s profit can then be written as:

$$\Pi_i(p_i) = p_i D_i(p_i) + d_i y_i^c \left( (1-\alpha) D_i^{sh}(p_j) + \alpha D_i(p_i) \right), \quad \alpha \in (0, 1]. \tag{8}$$

The first term in (8) represents firm $i$'s purchase revenues. The second term represents firm $i$'s disclosure revenues, increasing in $\alpha$. Note that the expression $(1-\alpha) D_i^{sh}(p_j) + \alpha D_i(p_i)$ can be interpreted as the expected number of users from whom firm $i$ can raise disclosure revenues. With probability $1 - \alpha$, it generates disclosure revenues from SH users only, whereas with probability $\alpha$, it generates them from all users, i.e., SH and MH users.

Each firm $i$ sets its price $p_i$ to maximise its profit given by (8), taking its rival's price $p_j$ as given. Solving for the first-order conditions of profit maximisation, we obtain the equilibrium price[15]

$$p_i^c(d_i) = \frac{\sigma v q_i - \alpha d_i y_i^c}{2}, \quad i = A, B. \tag{9}$$

We can divide firm $i$'s price $p_i^c(d_i)$ in (9) into two terms. The first term ($\sigma v q_i$) represents consumers' willingness to pay for service $i$, decreasing in the disclosure level $d_i$. The second term ($\alpha d_i y_i^c$) represents firm $i$'s disclosure revenues per MH user, increasing in $d_i$. We see that as $\alpha$ gets closer to 1, i.e., MH consumer data is more valuable, firm $i$'s price tends to decrease. Firm $i$'s pricing trade-off is therefore affected by the extent of MH consumer data monetisation.

Firm $i$'s demand now writes:

$$D_i(d_i) = \frac{\sigma v q_i + \alpha d_i y_i^c}{2t}, \quad D_i^{sh}(d_j) = 1 - D_j(d_j), \quad D^{mh}(d_i, d_j) = D_i(d_i) + D_j(d_j) - 1.$$

Plugging the equilibrium prices into the profit function (8), let us examine how firm $i$'s profit $\Pi_i(p_i^c(d_i), p_j^c(d_j), d_i, d_j)$ is affected by a variation of $d_i$. Using the envelope theorem, we observe that the impact of a variation of $d_i$ on firm $i$'s profit is such that

$$\frac{d\Pi_i}{dd_i} = (p_i + \alpha d_i y_i^c)\frac{\partial D_i}{\partial d_i}(p_i) + \frac{\partial (d_i y_i^c)}{\partial d_i}(\alpha D_i(d_i) + (1-\alpha) D_i^{sh}(d_j)). \tag{10}$$

The first term on the right-hand side of (10) represents the negative impact of a higher disclosure $d_i$ on firm $i$'s demand $D_i(p_i)$, which ultimately affects its total revenue per consumer ($p_i + \alpha d_i y_i^c$). Firm $i$'s demand decreases since a higher $d_i$ has a negative effect on the quality of service. The second term on the right-hand side of (10) represents the marginal disclosure revenues raised by firm $i$ on the expected number of users over whom data is disclosed.

We now solve for the firms' equilibrium disclosure levels at Stage 1. Firm $i$'s profit can be rewrited

$$\Pi_i(d_i) = t(D_i(d_i))^2 + (1-\alpha) d_i y_i^c D_i^{sh}(d_j)$$

Each firm sets its disclosure level $d_i$ to maximise its profit, taking the rival disclosure level $d_j$ as given. We then obtain equilibrium disclosure levels, $d_A^c$ and $d_B^c$.

The following proposition summarises the results:

**Proposition 2.** *(i) If $(\sigma v q^c + \alpha d^c y^c)/2 < t < \sigma v q^c + \alpha d^c y^c$, consumers provide information $y_i = y^c$. Firms' optimal prices and disclosure levels are $p_i = p^c = (\sigma v q^c - \alpha d^c y^c)/2$ and $d_i = d^c$.*

*(ii) Firms' choice to charge positive disclosure levels depends on consumer valuations $v$, the value of MH $\sigma$, MH consumer data valuation $\alpha$, and product differentiation $t$.*

[15] The second-order conditions are always satisfied: $\partial^2 \Pi_i / \partial p_i^2 = -2/t < 0$.

(a) *If* $\alpha \leq 1/2$*,* $\partial d^c/\partial v \leq 0$*,* $\partial d^c/\partial \sigma \leq 0$*,* $\partial d^c/\partial t \geq 0$*, and* $\partial d^c/\partial \alpha \gtreqless 0$*.*

(b) *If* $\alpha > 1/2$ *and* $\sigma v|\partial q_i/\partial d_i| \geq (2\alpha - 1)\partial(d_i y_i^c)/\partial d_i$*, we find the same results as in (a);*

(c) *If* $\alpha > 1/2$ *and* $\sigma v|\partial q_i/\partial d_i| < (2\alpha - 1)\partial(d_i y_i^c)/\partial d_i$*,* $d^c = d^{max}$*.*

From Proposition 2, we observe that disclosure levels are determined not only by consumer valuations $v$, but also by the value of MH $\sigma$, product differentiation $t$, and MH consumer data valuation $\alpha$. Consequently, we may end up in a soft or strict privacy regime depending on the latter parameters, and not only on $v$ as in Proposition 1. To clarify intuitions, suppose that there exist a $v^{c'} > 0$ such that $d^c > 0$ if $v < v^{c'}$, and $d^c = 0$ otherwise. We proceed similarly for $\sigma$, $t$, and $\alpha$.

Proposition 2(ii) describes three possible scenarios.

• *MH consumer data valuation is low (*$\alpha \leq 1/2$*).* If consumer valuations are sufficiently high ($v \geq v^{c'}$), and for given $\sigma$, $t$, and $\alpha$, firms do not engage in information disclosure ($d^c = 0$) and charge strictly positive prices ($p^c = \sigma v q^c/2$). Firms rely exclusively on purchase revenues and adopt a strict privacy regime. We adopt a similar reasoning for the value of MH $\sigma$. A higher $\sigma$ means that consumers value more the consumption of a second service. Since this parameter interacts with $v$, we obtain the same implications regarding the privacy regime. Conversely, a low $\sigma$ or a low $v$, i.e., lower consumer valuations, leads firms to adopt a soft privacy regime ($d^c > 0$). Indeed, in such a case, firms make comparatively more profit with disclosure revenues than purchase revenues, as in Proposition 1.

In contrast, higher product differentiation $t$ drives up firm $i$'s disclosure level. The intuition is that more product differentiation negatively impacts the number of MH consumers and hence increases the number of SH consumers, as MH is costlier. Firm $i$ is thus better able to monetise consumer data through SH consumers and earns higher disclosure revenues. In other words, a lower competitive intensity implies less privacy ($\partial d^c/\partial t \geq 0$), lower prices ($\partial p^c/\partial t \leq 0$), and less information provision ($\partial y^c/\partial t \leq 0$). Higher values of $t$ incentivise firms to enter more in a soft privacy regime, and possibly to a "minimal" privacy regime if for some $t > t'$, $d = d^{max}$.

• *MH consumer data valuation is high (*$\alpha > 1/2$*).* Provided that consumers' marginal disutility from higher disclosure level ($\sigma v|\partial q_i/\partial d_i|$) is superior to firm $i$'s marginal disclosure revenue ($(2\alpha - 1)\partial(d_i y_i^c)/\partial d_i$),[16] we obtain the same type of results as in Proposition 2(a).

• *MH consumer data valuation is high (*$\alpha > 1/2$*).* Provided that consumers' marginal disutility from a higher disclosure level is inferior to firm $i$'s marginal disclosure revenue, firm $i$ charges a maximal disclosure level $d^{max}$. Indeed, the negative effect of a higher $d_i$

[16]Note that the marginal disclosure revenue can be rewrited $(2(\alpha - 1) + 1)\partial(d_i y_i^c)/\partial d_i$. The term $2(\alpha - 1) \leq 1$ represents the limited ability of firm $i$ to value MH consumer data while the term 1 represents the same thing but on SH consumer data.

on quality is more than compensated by higher disclosure revenues at the margin. If MH consumer data is sufficiently valuable ($\alpha > 1/2$), we obtain a corner solution where firm $i$ charges $d^c = d^{max}$.

The study of the variation of $d^c$ with $\alpha$ is more complex. By equation (A.1), a higher $\alpha$ has two opposite effects on firm $i$' marginal profit from disclosure. On the one hand, a higher $\alpha$ has a positive impact on firm $i$'s demand ($d_i y_i^c/2t \geq 0$) since the firm has a better ability to monetise MH consumer data and, in turn, decreases its price. However, due to the limited ability to monetise MH consumer data and a lower service quality when $d_i$ is higher ($\sigma v \partial q_i/\partial d_i + (2\alpha - 1)\partial(d_i y_i^c)/\partial d_i \leq 0$), we obtain a negative effect. On the other hand, a higher $\alpha$ impacts positively the marginal revenue from MH consumer data ($\partial(d_i y_i^c)/\partial d_i \times D^{mh}) \geq 0$.

Even if the sign of $\partial d^c/\partial\alpha$ depends on the value of the parameters, we can think of the intuition behind a positive or a negative sign. First, a disclosure level $d^c$ increasing with $\alpha$ (i.e., $\partial d^c/\partial\alpha \geq 0$) is not counterintuitive. Indeed, if firm $i$ has more valuable MH consumer data, it is direct to see that it will be able to earn more disclosure revenues, which encourages it to increase $d^c$. It is more likely to happen provided consumers are not too privacy sensitive (e.g., with low consumer valuations) and $\alpha > 1/2$. Second, a disclosure level $d^c$ decreasing with $\alpha$ (i.e., $\partial d^c/\partial\alpha \leq 0$) is less obvious in terms of intuition. Indeed, a higher $\alpha$ makes MH consumer data more valuable, but this positive effect on disclosure revenues appears insufficient to induce firm $i$ to increase $d^c$. One important reason is the privacy-sensitiveness of consumers which deters the firm from charging a higher disclosure level. For instance, a higher $v$ (or $\sigma$) increases the magnitude of quality degradation and may induce a firm to ultimately decrease $d^c$ when $\alpha$ increases.

We now study the pricing of firm $i$ and write the following corollary.

**Corollary 2.** *(i) If $\partial d^c/\partial\alpha \leq 0$, $\partial p^c/\partial\alpha \gtreqless 0$, whereas if $\partial d^c/\partial\alpha > 0$, $\partial p^c/\partial\alpha < 0$.*

*(ii) There is price subsidisation if for $d^c > 0$, consumers' willingness to pay ($\sigma v q^c$) are sufficiently low compared to disclosure revenues per MH user ($\alpha d^c y^c$).*

From Corollary 2(i), if the disclosure level $d^c$ decreases as MH consumer data valuation $\alpha$ increases ($\partial d^c/\partial\alpha \leq 0$), firm $i$'s price $p^c$ has an ambiguous variation with respect to $\alpha$. On the one hand, firm $i$ relies less on disclosure revenues and comparatively more on purchase revenues, which induces it to increase its price. On top of that, consumers' willingness to pay for the service increases since a lower $d^c$ means better-perceived quality. On the other hand, a higher $\alpha$ entails that firm $i$ has a better ability to generate disclosure revenues from MH consumers, which induces it to lower its price $p^c$. However, if $d^c$ increases as $\alpha$ increases ($\partial d^c/\partial\alpha > 0$), firm $i$ reacts by decreasing its price ($\partial p^c/\partial\alpha < 0$). The intuition is that a higher value of MH consumer data $\alpha$ increases disclosure revenues. In turn, a higher disclosure decreases consumers' willingness to pay for the service, which is why firm $i$ has an incentive to decrease $p^c$. It then rely comparatively more on disclosure revenues.

From Corollary 2(ii), we observe that firm $i$ can subsidise its consumers by trading-off between consumers' willingness to pay and disclosure revenues per MH user. Notably, consumer subsidisation is possible as long as $\alpha > 0$.

## Numerical application

We solve the model using the same utility function $U_i = vy_i(1 - y_i - d_i) - p_i - tx_i$. In the end, we obtain the following results. We focus on equilibria with interior solutions where some consumers single-home and some other multi-home.[17]

**Result 2.** *In the duopoly equilibrium, an interior solution is obtained and consumers provide information $y^c = (1 - d^c)/2$.*

*(i) For given $\sigma$, $v$, $t$, and $\alpha$, an equilibrium where $d^c > 0$ may exist, i.e., we have $p^c = (\sigma v q^c - \alpha d^c y^c)/2$ and $d^c = d^c(\sigma, v, t, \alpha)$ if $((1 - d^c)/8)(\sigma v + d^c(2\alpha - \sigma v)) < t < ((1 - d^c)/4)(\sigma v + d^c(2\alpha - \sigma v))$.*

*(ii) An equilibrium where $d^c = 0$ may exist where $p^c = \sigma v/8$ and $d^c = 0$ if $(\sigma v)/8 < t \leq (\sigma v)/4$. This outcome may occur (a) if $v > v^{\cdot}c'$ for given ($\sigma$,$t$,$\alpha$), (b) or if $\sigma \geq \sigma^c$ for given ($v$,$t$,$\alpha$), or (c) if $t \leq t^c$ for given ($\sigma$,$v$,$\alpha$), or (d) if $\alpha \gtreqless \alpha^c$, depending on the sign of $\partial d^c / \partial \alpha$.*

On top of consumer valuations, the value of MH, as well as the differentiation parameter or the valuation of MH consumer data, may change the extent of implementing a strict or soft privacy regime. On Figure 3 in Appendix C, for given $\sigma = 1$, $t = 0.04$, and $\alpha = 0.25$, firm $i$ implements a soft privacy regime if $v \in [0.125, 0.3]$ and a strict privacy regime if $v \in [0.3, 0.32]$. Consumers are subsidised if $v$ is sufficiently low and pay the service otherwise.

On Figure 4 when $\alpha = 1/2, \sigma = 0.2$ and $t = 0.04$, we find that a soft privacy regime is always implemented if $v \in (0, 0.21]$ while there is no scope for a strict privacy regime. The disclosure level remains high and consumers are always subsidised. It suggests that when MH consumer data is more valuable, it can be better monetised and firm $i$ has incentives to rely more heavily, and even exclusively on disclosure revenues.

On Figure 6, we observe that when $\sigma = 0.2, v = 1.2$, $t = 0.04$, there is only a scope for a soft privacy regime, which occurs if $\alpha \in [0.18, 0.48]$. Low value of MH ($\sigma = 0.2$) exerts a downward pressure on consumers' willingness to pay for the service, which induces firm $i$ to set a high disclosure level and to rely exclusively on disclosure revenues and subsidise consumers. As $\alpha$ increases, $d^c$ decreases and then increases. However, if $\sigma = 1$ and $v = 0.3$, firm $i$ implements a strict privacy regime if $\alpha \in (0, 0.25]$ and then switch to soft privacy

[17] Given the specification on $U_i$, we focus on cases where $\alpha \leq 1/2$, to ensure valid and interior solutions.

regime if $\alpha \in (0.25, 0.4]$. A higher $\sigma$ therefore induces firm $i$ to set a strict privacy regime when MH consumer data valuation is not too high. Note that firm $i$ no longer subsidises consumers in the soft privacy regime and the disclosure level $d^c$ always increases with $\alpha$, contrary to when $\sigma = 0.2$ and $v = 1.2$.

**Economic analysis: SH only vs. presence of MH**

A firm's business strategy can be strongly altered in the presence of MH compared to SH only.

A firm's business model with SH is well-defined: set a high privacy level if consumer valuations are high, which can result in a strict privacy regime, and set a lower privacy level if valuations are lower (soft privacy regime). In this respect, firms' privacy choices are unifactorial.

Under MH, the analysis complexifies since what determines the setting of a high or low privacy level is multi-factorial: it depends not only on consumer valuations, but also on the value of MH for consumers, product differentiation, and the value of MH consumer data. Let us take some examples with numerical applications on Table 1.

When $v = 0.34$, a firm implements a soft privacy regime ($d^c = 0.397$) and subsidises consumers ($p^c < 0$) under SH. Under MH, we also have a soft privacy regime but with a lower disclosure level (0.394) and the firm never subsidises consumers ($p^c > 0$). A potential explanation is that the value of MH $\sigma$ shifts upward consumers' willingness to pay ($\sigma v$) and MH consumer data valuation is not too high ($\alpha = 0.25$), which induces the firm to set a lower $d^c$. It may explain why there is no consumer subsidisation under MH compared to SH only.

When $v = 1.1$, we have a strict privacy regime and a positive price under SH whereas it is the complete opposite under MH. Even if consumer valuations are high, the value of MH $\sigma = 0.2$ exerts a downward pressure on consumers' willingness to pay. The firm is therefore induced to rely exclusively on disclosure revenues. However, if $\sigma$ goes up to 0.28, there is a shift in the firm's strategy: it relies exclusively on price revenues and implements a strict privacy regime; we then have qualitatively the same situation under MH and with SH only.

| | v | $\sigma$ | t | $\alpha$ | $d^c$ | $p^c$ | $y^c$ |
|---|---|---|---|---|---|---|---|
| SH MH | 0.34 | 1 | 0.06 | 0.25 | 0.394 | $6 \times 10^{-4}$ | 0.303 |
| SH | 0.34 | | 0.06 | | 0.397 | -0.059 | 0.301 |
| SH MH | 1.1 | 0.28 | 0.04 | 0.25 | 0 | 0.07 | 0.5 |
| SH MH | 1.1 | 0.2 | 0.04 | 0.25 | 0.409 | -0.005 | 0.295 |
| SH | 1.1 | | 0.04 | | 0 | 0.04 | 0.5 |

Table 1: Comparison between SH and SH+MH

# 6 Extension: multi-product monopoly

## 6.1 Single-homing

We now study the equilibrium when $A$ and $B$ act as a monopoly, for example, after a merger. We consider a multi-product monopoly, which supplies services $A$ and $B$ at prices $p_A$ and $p_B$, with disclosure levels $d_A$ and $d_B$, respectively.

At Stage 4, the level of information provision $y_i$ for service $i$ is as given in (3): for a given $d_i$, $y_i = y_i^m(d_i) = y_i^c(d_i)$.

At Stage 3, proceeding similarly as in the duopoly, we find that the demand for service $A$ is $D_A(p_A, p_B) = x^*$ and the demand for service $B$ is $D_B(p_A, p_B) = 1 - x^*$, where $x^*$ is given by (4).

At Stage 2, we determine the price of consumer data. As in the duopoly, the monopolist has monopoly power on consumer data as it represents a bottleneck on consumer data. Thus, the data intermediary $D$ obtains this data at a price $r_A = r_B = 1$.

We now consider the pricing problem of the monopoly. The monopoly profit is given by:

$$\Pi^m(p_A, p_B) = \sum_{i=A,B} (p_i + d_i y_i^m) D_i(p_A, p_B). \tag{11}$$

At the optimum of the monopoly with a covered market, the indifferent consumer receives zero surplus, i.e., $vq_A - p_A - tx^* = 0$. Substituting for $x^*$ in (4), we obtain the relation between prices that ensures market coverage: $p_B = v(q_A + q_B) - p_A - t$. We can thus express $\Pi^m$ as a function of only $p_A$.[18] We solve for the first-order condition $\partial \Pi^m / \partial p_A = 0$ (the SOC holds as $\partial^2 \Pi^m / \partial p_A^2 = -4/t < 0$). We therefore obtain the equilibrium prices denoted by

$$\begin{cases} p_i^m(d_i, d_j) = \frac{v(q_i+q_j)}{2} - \frac{d_i y_i^m - d_j y_j^m}{4} - \frac{t}{2}; \quad i \neq j = A, B, \\ \text{with} \quad D_i(d_i, d_j) = \frac{1}{2} + \frac{v(q_i - q_j)}{2t} + \frac{d_i y_i^m - d_j y_j^m}{4t}. \end{cases}$$

Consider now the first stage where the monopolist chooses the disclosure levels.

The monopolist sets the disclosure levels to maximise its profit $\Pi^m$. Solving for the first-order conditions, we obtain the equilibrium disclosure levels, $d_A^m$ and $d_B^m$.

The following proposition characterises the monopoly outcome:

**Proposition 3.** *(i) If $0 < t < vq(y^m, d^m) + d^m y^m$, consumers provide information $y_i = y^m$. Firm's optimal prices and disclosure levels are $p_i = p^m = vq^m - t/2$ and $d_i = d^m$.*

*(ii) The monopolist's decision to charge positive disclosure levels depends only on consumer valuations $v$, where $\partial d^m / \partial v \leq 0$.*

[18] Therefore, $D_A(p_A) = (vq_A - p_A)/t$ and $D_B(p_A) = 1 - (vq_A - p_A)/t$.

As in Proposition 1, disclosure levels are determined and impacted negatively by consumer valuations $v$. Assume that there exists $v^m > 0$ such that $d^m > 0$ if $v < v^m$ and $d^m = 0$ otherwise. First, if consumer valuations are sufficiently high ($v \geq v^m$), the monopolist does not engage in the disclosure of consumer data ($d^m = 0$) and charges positive prices ($p^m = vq(y^m(0), 0) - t/2 > 0$). The monopolist therefore relies exclusively on purchase revenues and adopts a strict privacy regime. Second, if consumer valuations are sufficiently low, ($v < v^m$), the monopolist engages in the disclosure of consumer data ($d^m > 0$). As consumer valuations ($v$) increase, the level of information disclosure decreases ($\partial d^m/\partial v \leq 0$) while consumers provide more information ($\partial y^m/\partial v \geq 0$) and prices increase ($\partial p^m/\partial v > 0$). The firm here adopts a soft privacy regime.

The monopolist is a bottleneck on consumer data and extracts all surplus from the data intermediary $D$.

**Corollary 3.** *There is negative pricing under a soft privacy regime if consumer valuations $v$ are sufficiently low (for a given $t$) or if product differentiation $t$ is sufficiently high (for a given $v$).*

From Corollary 3, we observe that the monopolist subsidises consumers if product differentiation is sufficiently high and consumer valuations are sufficiently low. It differs from the duopoly where a sufficiently low $t$, that is, a higher competitive intensity, drives prices below zero. The price maximisation of the monopolist explains this result as it chooses the maximum price that extracts the surplus of the indifferent consumer.

**Numerical application**

We find the following results for the multi-product monopoly with SH only.[19]

**Result 3.** *In a multi-product duopoly, the market is covered and consumers provide information $y^m = (1 - d^m)/2$.*

*If* $\begin{cases} 0 < v < 1, \\ 0 < t < 1/4(2-v), \end{cases}$ *firms' optimal prices and disclosure levels are* $\begin{cases} p^m = \frac{v}{4(2-v)^2} - \frac{t}{2}, \\ d^m = \frac{1-v}{2-v}; \end{cases}$

*If* $\begin{cases} v \geq 1, \\ 0 < t < v/4, \end{cases}$ *firms' optimal prices and disclosure levels are* $\begin{cases} p^m = \frac{v}{4} - \frac{t}{2}, \\ d^m = 0. \end{cases}$

We therefore have that $p^c < p^m$, $d^c = d^m$, and $y^c = y^m$. An implication is that a merger to a multi-product monopoly under SH increases market power and decreases consumer surplus.

[19] To determine the range of values for which the monopolist covers the market, we solve $\partial\Pi^u/\partial p_i|_{\{p_i=p_i^m, d_i=d_i^m\}} < 0$, where $\Pi^u$ is the profit if the market is not covered.

### 6.2 Multi-homing

At Stage 4, the level of information $y_i$ is determined in the same way as the previous sub-section: for SH consumers, it is given by (3) while for MH consumers, it is given $y_i^m(d_i) \equiv \arg\max_{y_i} U_{ij}$. As before, MH and SH consumers provide the same level of information when using service $i$.

At Stage 3, the consumer chooses which service(s) to patronise. Proceeding in a similar way as in the previous sub-section, the demand functions of the monopoly for services $A$ and $B$ are given by equation (7).

At Stage 2, the monopolist determines the selling price of consumer data, charged to the data intermediary $D$. The monopolist is a bottleneck on SH consumer data and thus charges $r_i = 1$ to $D$. An issue is that MH consumer data may be sold twice, thereby becoming less valuable since $\alpha \leq 1$.

Two scenarios can be considered regarding the selling of MH consumer data. First, the case where the monopolist is able to identify precisely the location of each MH user so that it can manage to sell MH consumer data at once. For instance, it would sell the data of MH users to the left (right) of 1/2 at a disclosure level $d_A$ ($d_B$). In other words, it is able to internalise the double-selling issue of the duopoly. Second, the alternative case where the monopolist is not able to do such identification and will not be able to monetise MH consumer data fully. We argue that the first scenario may not be the most realistic. Indeed, it involves consumer targeting, which can imply a lower level of privacy, in addition to information disclosure to the data intermediary.[20] On top of that, data protection laws (e.g., GDPR in EU) impose severe restrictions on the exploitation of consumer data; it may be unlikely that further data processing to monetise such data is compliant with data privacy.[21] Here, we focus on the second scenario where the issue of double-selling of MH consumer data remains. The first scenario is solved in Appendix B.

The monopoly profit is given by

$$\Pi^m(p_i, p_j) = \sum_{i=A,B} (p_i + \alpha d_i y_i^c) D_i(p_i) + d_i y_i^c (1-\alpha) D_i^{sh}(p_j). \tag{12}$$

The monopoly chooses $p_A$ and $p_B$ to maximise (12). Solving for the first-order conditions, we obtain the equilibrium price (the SOC holds as $\partial^2 \Pi^m / \partial p_i^2 \leq 0$ and $\partial^2 \Pi^m / \partial p_i \partial p_j = 0$)

$$p_i^m(d_i, d_j) = \frac{\sigma v q_i - \alpha d_i y_i^m + (1-\alpha) d_j y_j^m}{2}, \quad i = A, B. \tag{13}$$

As in equation (9), we observe that the monopolist prices service $i$ by trading off between

[20]Consumer targeting can also involve personalised pricing, which is outside the scope of this paper.

[21]Even if it were compliant with data protection law, questions around user consent to be more targeted may arise. We do not deal with this issue in the paper.

consumers' willingness to pay ($\sigma v q_i$) and disclosure revenues per MH user ($\alpha d_i y_i^m$). On top of that, the firm now takes into account potential disclosure revenues that can be raised on users of service $j$: $p_i^m(d_i, d_j)$ depends (i) negatively on disclosure revenues per MH user ($-\alpha d_j y_j^m$), i.e., those consumers signing up to service $i$ in addition to service $j$, and (ii) positively on disclosure revenues per SH user of service $j$ ($d_j y_j^m$), and the overall effect is positive.

The monopoly demand now writes,

$$D_i(d_i, d_j) = \frac{\sigma v q_i + \alpha d_i y_i^m - (1-\alpha) d_j y_j^m}{2t}, \quad D_i^{sh}(d_i, d_j) = 1 - D_i(d_i, d_j).$$

Plugging the equilibrium prices into the profit function (12), we now solve for the optimal disclosure levels of the monopoly at Stage 1. The monopoly can be written

$$\Pi^m(d_i, d_j) = t\left((D_i(d_i, d_j))^2 + (D_j(d_i, d_j))^2\right) + (1-\alpha)(d_i y_i^m + d_j y_j^m).$$

The monopoly sets its disclosure levels to maximise its profit $\Pi^m(d_i, d_j)$. We obtain the equilibrium disclosure levels $d_A^m$ and $d_B^m$.

The following proposition summarises the analysis.

**Proposition 4.** *(i) If $(\sigma v q^m + (2\alpha-1)d^m y^m)/2 < t < v q^m + (2\alpha-1)d^m y^m$, consumers provide information $y_i = y^m$. Optimal prices and disclosure levels are $p_i = p^m = (\sigma v q^m - (2\alpha-1)d^m y^m)/2$ and $d_i = d^m$.*

*(ii) The monopolist's choice to charge positive disclosure levels depends on consumer valuations $v$, the value of MH $\sigma$, MH consumer data valuation $\alpha$, and the level of product differentiation $t$.*

*(a) If $\alpha \leq 1/2$, $\partial d^m/\partial v \leq 0$, $\partial d^m/\partial \sigma \leq 0$, $\partial d^m/\partial t \geq 0$, and $\partial d^m/\partial \alpha \gtreqless 0$.*

*(b) If $\alpha > 1/2$ and $\sigma v|\partial q_i/\partial d_i| \geq (2\alpha-1)\partial(d_i y_i^m)/\partial d_i$, we find the same results as in (a);*

*(c) If $\alpha > 1/2$ and and $\sigma v|\partial q_i/\partial d_i| < (2\alpha-1)\partial(d_i y_i^c)/\partial d_i$, we obtain a corner solution where $d^m = d^{max}$.*

From Proposition 4, as in the duopoly under MH, the monopolist's choice to charge disclosure levels depends on the value of MH $\sigma$, product differentiation $t$, the value of MH consumer data $\alpha$, and consumer valuations $v$. We obtain qualitatively similar results and intuitions, as examined in Proposition 2.

We now analyse the variations of the monopolist's price $p^m$ with respect to $v$, $\sigma$, $t$, and $\alpha$, and the possibility of consumer subsidisation.

**Corollary 4.** *(i) If $\alpha \leq 1/2$, the variations of the monopolist price $p^m$ with respect to $v$, $\sigma$, $t$, and $\alpha$ are ambiguous, whereas if $\alpha > 1/2$, $\partial p^m/\partial v \geq 0$, $\partial p^m/\partial \sigma \geq 0$, $\partial p^m/\partial t \leq 0$, and $\partial p^m/\partial \alpha \gtreqless 0$.*

*(ii) There is price subsidisation if and only if $\alpha > 1/2$ and consumers willingness to pay ($\sigma v q^m$) are lower than expected disclosure revenues ($(2\alpha - 1)d^m y^m$).*

From Corollary 4(i), we observe that the variations of the monopoly price $p^m$ can be analysed by distinguishing between two cases: either the value of MH consumer data is low ($\alpha \leq 1/2$) or it is high ($\alpha > 1/2$). Rewriting $p^m$ such as $p^m = \sigma(vq^m + d^m y^m - 2\alpha d^m y^m)/2$, we see that the monopolist trades-off between consumers' willingness to pay ($\sigma v q^m$), disclosure revenues per SH user ($1 \times d^m y^m$), and disclosure revenues per MH user ($\alpha(d^m y^m + d^m y^m)$), i.e., consumers purchasing service $A$ ($B$) in addition to service $B$ ($A$). We have seen in equation (13) that this trade-off comes from the monopolist, which prices service $i$ by internalising the impact on the consumption of service $j$. The monopolist decreases its price for service $i$ when it earns higher disclosure revenues from MH users, whereas it increases it when it earns higher disclosure revenues from service $j$'s SH users.

If $\alpha \leq 1/2$, disclosure revenues per SH user (of service $j$) are comparatively higher than disclosure revenues per MH user: consumers' willingness to pay ($\sigma v q^m$) and total disclosure revenues ($(2\alpha - 1)d^m y^m$) varies in opposite signs as $d^m$ increases, which is why the variations of $p^m$ are unclear. However, if $\alpha > 1/2$, disclosure revenues per MH user prevails: $p^m$ increases with consumer valuations $v$ and the value of MH $\sigma$, while it decreases with product differentiation $t$.

It is interesting to analyze how $p^m$ varies with $\alpha$:

$$\frac{\partial p^m}{\partial \alpha} = \frac{1}{2}\left(\sigma v \frac{dq^m}{d\alpha} - (2\alpha - 1)\frac{d(d^m y^m)}{d\alpha} - 2d^m y^m\right). \tag{14}$$

If $\partial d^m/\partial \alpha \leq 0$, there are three effects at work. First, $p^m$ increases since a higher $\alpha$ increases consumers' willingness to pay (first term in (14)). Second, a higher $\alpha$ lowers disclosure revenues at the margin but the impact on price $p^m$ depends on $\alpha$: if $\alpha \leq 1/2$, the impact is positive, whereas if $\alpha > 1/2$, it is negative (second term in (14)). Third, a higher $\alpha$ impacts positively disclosure revenues and drives the price $p^m$ down (third term in (14)). The sign of $\partial p^m/\partial \alpha$ is therefore unclear. However, if $\partial d^m/\partial \alpha > 0$ and $\alpha > 1/2$, $p^m$ decreases as $\alpha$ increases ($\partial p^m/\partial \alpha > 0$); this is due to (i) the negative effect of $\alpha$ on consumers' willingness to pay, (ii) the positive effect of $\alpha$ disclosure revenues at the margin, and (iii) higher disclosure revenues.

From Corollary 4(ii), we find that a necessary condition for the monopolist to subsidise consumers is that the value of MH consumer data is sufficiently high ($\alpha > 1/2$), which is stricter condition than in MH duopoly. It then trades-off between consumers' willingness to pay ($\sigma v q^m$) and disclosure revenues raised on its consumers ($(2\alpha - 1)d^m y^m$).

**Proposition 5.** *(i) A multi-product monopoly charges a higher price than in the duopoly, i.e.,* $p^c \leq p^m$*.*

*(ii) If* $\alpha < 1/2$*,* $d^m \gtreqless d^c$ *whereas if* $\alpha \geq 1/2$*,* $d^c \geq d^m$*.*

From Proposition 5(i), we find that moving from duopoly to multi-product monopoly increases market power ($p^c \leq p^m$). The intuition is that when pricing service $i$, the monopolist internalises that it also generates revenues from service $j$'s consumers. In particular, higher disclosure revenues from service $j$'s consumers (i.e., SH and MH users) induce the monopolist to increase service $i$'s price, which is not the case in the duopoly.

Proposition 5(ii) states that if MH consumer data is sufficiently valuable ($\alpha \geq 1/2$), information disclosure levels in duopoly are greater in multi-product monopoly ($d^c \geq d^m$). The idea is that when $\alpha \geq 1/2$, the expected number of users from whom it is possible to earn disclosure revenues is higher in duopoly than in monopoly, which induces duopoly firms to charge higher disclosure levels. This effect originates from the monopolist which internalises the externality of the pricing of service $i$ on service $j$; it relies relatively more on purchase revenues and may subsidise consumers only to the extent that $\alpha \geq 1/2$.

The overall effect on welfare of moving from a duopoly to a multi-product monopoly is therefore ambiguous.

### Numerical application

We solve the model using the same utility function $U_i = vy_i(1 - y_i - d_i) - p_i - tx_i$. In the end, we obtain the following results. We focus on equilibria with interior solutions where some consumers single-home and some other multi-home.[22]

**Result 4.** *In the duopoly equilibrium, an interior solution is obtained and consumers provide information* $y^m = (1 - d^m)/2$*.*

*(i) For given* $\sigma$*,* $v$*,* $t$*, and* $\alpha$*, an equilibrium where* $d^c > 0$ *may exist, i.e., we have* $p^m = (\sigma v q^m - (2\alpha - 1)d^m y^m)/2$ *and* $d^m = d^m(\sigma, v, t, \alpha)$ *if* $((1 - d^m)/8)(\sigma v + d^m(2(2\alpha - 1) - \sigma v)) < t < ((1 - d^m)/4)(\sigma v + d^m(2(2\alpha - 1) - \sigma v))$*.*

*(ii) An equilibrium where* $d^m = 0$ *may exist where* $p^m = \sigma v/8$ *and* $d^c = 0$ *if* $(\sigma v)/8 < t \leq (\sigma v)/4$*. This outcome may occur under conditions similar to the duopoly in the presence of MH.*

# 7 Conclusion

In a framework of horizontal differentiation investigates competition in prices and information disclosure levels. We study how privacy and consumer behaviours impact competition

[22] As discussed in the proof of Proposition 4, an interior solution on disclosure levels is much more likely to be found if $\alpha > 1/2$. The numerical application confirms this result. Therefore, it is almost impossible to compare monopoly and duopoly numerically with MH given the reduced form specification of $U_i$.

between two digital firms. The latter supply a service to consumers and earn revenues that originate from two sources: purchase revenues from consumers purchasing the service and disclosure revenues from a data intermediary.

Under SH, a decisive component of firms' business strategies is the negative relation between consumer valuations and disclosure levels. Two privacy regimes are then possible: a strict privacy regime where firms rely only on purchase revenues and consumers provide many information, and a soft privacy regime where firms rely on both sources of revenues; consumers provide less information and may be subsidised. Firms are bottlenecks on consumer data and the data intermediary is left with zero surplus.

We develop a MH framework where it is more challenging to monetise MH consumer data as it may be sold twice to the data intermediary. We find that firms' business strategies become more complex to assess as they rely on a combination of economic parameters: consumer valuations, the value of MH, product differentiation and MH consumer data valuation. Notably, higher MH consumer data valuation is not necessarily linked to a higher disclosure level, as a firm also has to consider the marginal disclosure revenue and the privacy-sensitiveness of consumers. If disclosure levels decrease with MH consumer data valuation, it is not direct to conclude that firms compensate for lower disclosure levels with higher prices.

We believe our analysis is relevant for digital businesses in a market where privacy and consumer behaviours are essential competition parameters. As we have seen, a firm's business strategy tends to complexify under MH, to the point where a shift in the firms' business strategy may be necessary; what is clear under SH deserves more scrutiny under MH. For instance, it may be optimal to set a strict (soft) privacy regime under SH and a soft (strict) one under MH. Note that our framework focuses on partial MH where consumers do not fully enjoy the consumption of an additional product ($\sigma \leq 1$), as it may be the case in reality. We also highlight the importance of the valuation of MH consumer data and how it strongly impacts firms' business models and the profitability of a data intermediary. MH data valuation ($\alpha$) of the data intermediary could (exogenously) depend on the quality of the data firms have on MH consumers or on how much they invest in data collection.

This analysis may also be relevant to competition authorities whishing to evaluate competition dynamics and the potential impact of a merger in a digital market where privacy is considered as a proxy of quality. It might eventually be relevant to data protection authorities that can set ex-ante rules on how personal data should or should not be exploited, thereby protecting individuals' online privacy.

# References


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

Anderson, Simon P, Oystein Foros, and Hans Jarle Kind (2017) "Product functionality, competition, and multipurchasing," *International Economic Review*, Vol. 58, pp. 183–210.

——— (2019) "The importance of consumer multihoming (joint purchases) for market performance: Mergers and entry in media markets," *Journal of Economics & Management Strategy*, Vol. 28, pp. 125–137.

Anderson, Simon P and Joshua S Gans (2011) "Platform siphoning: Ad-avoidance and media content," *American Economic Journal: Microeconomics*, Vol. 3, pp. 1–34.

Argenziano, Rossella and Alessandro Bonatti (2020) "Information Revelation and Privacy Protection."

Armstrong, Mark (2006) "Competition in two-sided markets," *The RAND Journal of Economics*, Vol. 37, pp. 668–691.

Belleflamme, Paul, Wing Man Wynne Lam, and Wouter Vergote (2020) "Competitive imperfect price discrimination and market power," *Marketing Science*, Vol. 39, pp. 996–1015.

Belleflamme, Paul and Martin Peitz (2019) "Platform competition: Who benefits from multihoming?" *International Journal of Industrial Organization*, Vol. 64, pp. 1–26.

Caillaud, Bernard and Bruno Jullien (2003) "Chicken & egg: Competition among intermediation service providers," *RAND journal of Economics*, pp. 309–328.

Calzolari, Giacomo and Alessandro Pavan (2006) "On the optimality of privacy in sequential contracting," *Journal of Economic theory*, Vol. 130, pp. 168–204.

Casadesus-Masanell, Ramon and Andres Hervas-Drane (2015) "Competing with privacy," *Management Science*, Vol. 61, pp. 229–246.

Choi, Jay Pil (2010) "Tying in two-sided markets with multi-homing," *The Journal of Industrial Economics*, Vol. 58, pp. 607–626.

Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim (2019) "Privacy and personal data collection with information externalities," *Journal of Public Economics*, Vol. 173, pp. 113–124.

De Corniere, Alexandre and Romain De Nijs (2016) "Online advertising and privacy," *The RAND Journal of Economics*, Vol. 47, pp. 48–72.

De Corniere, Alexandre and Greg Taylor (2020) "Data and competition: a general framework with applications to mergers, market structure, and privacy policy."

Hermalin, Benjamin E and Michael L Katz (2006) "Privacy, property rights and efficiency: The economics of privacy as secrecy," *Quantitative Marketing and Economics*, Vol. 4, pp. 209–239.

Hui, Kai-Lung and PL Ivan (2006) "Png. 2006," *The Economics of Privacy*, pp. 471–498.

Ichihashi, Shota (2020a) "Online privacy and information disclosure by consumers," *American Economic Review*, Vol. 110, pp. 569–95.

——— (2020b) "Non-competing data intermediaries,"Technical report, Bank of Canada.

Kim, Jin-Hyuk, Liad Wagman, and Abraham L Wickelgren (2019) "The impact of access to consumer data on the competitive effects of horizontal mergers and exclusive dealing," *Journal of Economics & Management Strategy*, Vol. 28, pp. 373–391.

Kim, Soo Jin (2020) "Privacy, Information Acquisition, and Market Competition," *Working Paper.*

Lefouili, Yassine and Ying Lei Toh (2020) "Privacy and quality," *TSE Working Paper 17 (795).*

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

Noam, Eli (1995a) "Privacy in telecommunications, part I," *New Telecommunications Quarterly*, pp. 52–59.

——— (1995b) "Privacy in telecommunications, part III," *New Telecommunications Quarterly.*

Posner, Richard A (1981) "The economics of privacy," *The American economic review*, Vol. 71, pp. 405–409.

Prat, Andrea and Tommaso M Valletti (2019) "Attention oligopoly," *Working paper.*

Prufer, Jens and Christoph Schottmuller (2017) "Competing with big data," *TILEC Discussion Paper No. 2017-006.*

Rochet, Jean-Charles and Jean Tirole (2003) "Platform competition in two-sided markets," *Journal of the european economic association*, Vol. 1, pp. 990–1029.

Shy, Oz and Rune Stenbacka (2016) "Customer privacy and competition," *Journal of Economics & Management Strategy*, Vol. 25, pp. 539–562.

Spulber, Daniel F (2009) "The map of commerce: Internet search, competition, and the circular flow of information," *Journal of Competition Law and Economics*, Vol. 5, pp. 633–682.

Stigler, George J (1980) "An introduction to privacy in economics and politics," *The Journal of Legal Studies*, Vol. 9, pp. 623–644.

Tag, Joacim (2009) "Paying to remove advertisements," *Information Economics and Policy*, Vol. 21, pp. 245–252.

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

# Appendix

## Appendix A: Proofs

### Lemma 1

First, if $d_i y_i^c(d_i)$ is concave in $d_i$, it means that

$$\frac{\partial^2(d_i y_i^c)}{\partial d_i^2}(d_i) = 2\frac{\partial y_i^c}{\partial d_i} + d_i \frac{\partial^2 y_i^c}{\partial d_i^2} \leq 0,$$
$$\text{and} \quad y_i^c(0) > y_i^c(1) + \left.\frac{\partial y_i^c}{\partial d_i}\right|_{d'>0}.$$

Second, let $F(d_i) \equiv \partial(d_i y_i^c)/\partial d_i$. Assuming continuity and monotonicity of $F(d_i)$ on $\mathbb{R}^+$, by the intermediary values theorem, there exists a $d_i$ such that $F(d_i) = 0$. Consequently, $d_i y_i^c$ is concave in $d_i$.

### Proposition 1

Note that

$$\frac{\partial D_i}{\partial d_i}(d_i) = \frac{1}{6t}\left(v\frac{\partial q_i}{\partial d_i} + \frac{\partial(d_i y_i^c)}{\partial d_i}\right),$$

The first-order conditions of firm $i$ are given by

$$\frac{\partial \Pi_i}{\partial d_i} = 0 \Leftrightarrow 2t\left(2D_i \frac{\partial D_i}{\partial d_i}\right) = 0.$$

We look for an interior solution and check that the second-order conditions are satisfied:

$$\frac{\partial^2 \Pi_i}{\partial d_i^2} = 4t \left( \left( \frac{\partial D_i}{\partial d_i} \right)^2 + D_i \frac{\partial^2 D_i}{\partial d_i^2} \right) \leq 0, \quad \text{by} \quad \text{Assumption 1}.$$

We now check the variations of $d^c$ with respect to $v$.

$$\frac{\partial^2 \Pi_i}{\partial d_i \partial v}_{|\{d_i = d^c, d_j = d^c\}} = 4t \left( D_i \frac{\partial^2 D_i}{\partial d_i \partial v} + \underbrace{\frac{\partial D_i}{\partial v}}_{=0} \frac{\partial D_i}{\partial d_i} \right) \leq 0.$$

From the implicit function theorem, we obtain

$$\frac{\partial d_i}{\partial v}_{|\{d_i = d^c, d_j = d^c\}} = -\frac{\partial^2 \Pi_i / \partial d_i \partial v}{\partial^2 \Pi_i / \partial d_i^2} \leq 0.$$

If $d^c > 0$, we have the following variations:

$$\frac{\partial y^c}{\partial v} = \frac{\partial y^c}{\partial d_i} \frac{\partial d^c}{\partial v} \geq 0; \quad \frac{\partial p^c}{\partial v} = -\left( \frac{\partial d^c}{\partial v} \left( y^c + \frac{\partial y^c}{\partial d_i} \right) \right) \geq 0; \quad \frac{\partial \left( d^c y^c \right)}{\partial v} \leq 0.$$

Note that the optimal disclosure level $d^c$ does not depend on $t$:

$$\frac{\partial^2 \Pi_i}{\partial d_i \partial t}_{|\{d_i = d^c, d_j = d^c\}} = 0.$$

All consumers are served if the marginal consumer derives a positive utility in equilibrium, that is, if $vq(y^c, d^c) - p^c - tx^* \geq 0$, i.e., if $0 < t \leq vq(y^c, d^c) + d^c y^c / 2$.

**Lemma 2**

Firm $i$ is a bottleneck on SH consumer data and therefore charges the data intermediary $r_i = 1$. By contrast, if firms $A$ and $B$ compete to sell MH consumer data, the price of this data depends on how the data intermediary values it, that is, $r_i = \alpha$. If firms $A$ and $B$ had to choose between selling and not selling MH consumer data to $D$, it is direct to show that "selling" would be a Nash equilibrium for both firms.

| | | Firm B | |
|---|---|---|---|
| | | Not selling | Selling |
| Firm A | Not selling | $0, 0$ | $0, 1$ |
| | Selling | $1, 0$ | $\alpha, \alpha$ |

Table 2: Data selling game

**Proposition 2**

The first-order conditions write as follows,

$$\frac{\partial \Pi_i}{\partial d_i} = 0 \Leftrightarrow 2tD_i(d_i)\frac{\partial D_i}{\partial d_i} + (1-\alpha)\frac{\partial (d_i y_i^c)}{\partial d_i} D_i^{sh}(d_j) = 0.$$

We check that the second-order conditions are satisfied:

$$\frac{\partial^2 \Pi_i}{\partial d_i^2} = 2t\left(\left(\frac{\partial D_i}{\partial d_i}\right)^2 + D_i \frac{\partial^2 D_i}{\partial d_i^2}\right) + (1-\alpha)\frac{\partial^2 (d_i y_i^c)}{\partial d_i^2} D_i^{sh}(d_j) \leq 0 \quad \text{by Assumption 1.}$$

In equilibrium, all consumers are served and there is multi-homing if the following set of conditions is satisfied:

$$\begin{aligned}
&U_i^c = vq(y^c, d^c) - p^c - tx^* \geq 0,\\
&U_{ij}^c = (1+\sigma)vq(y^c, d^c) - 2p^c - t \geq 0,\\
&0 < \underbrace{\frac{\sigma v q^c + \alpha d^c y^c}{t} - 1}_{D^{mh}} < 1.
\end{aligned}$$

The above conditions are satisfied if $(\sigma v q^c + \alpha d^c y^c)/2 < t < \sigma v q^c + \alpha d^c y^c$.

We now check the variations of $d^c$. Let us inspect how $d^c$ varies with consumer valuations $v$.

$$\begin{cases}
\frac{\partial D_i}{\partial v}(d_i) = \frac{\sigma q_i}{2t} > 0; \quad \frac{\partial^2 D_i}{\partial d_i \partial v}(d_i) = \frac{\sigma}{2t}\frac{\partial q_i}{\partial d_i} \leq 0;\\
\frac{\partial D_i^{sh}}{\partial v}(d_j) = -\frac{\sigma q_j}{2t} < 0; \quad \frac{\partial D_i}{\partial d_i}(d_i) = \frac{1}{2t}\left(\sigma v \frac{\partial q_i}{\partial d_i} + \alpha \frac{\partial (d_i y_i^c)}{\partial d_i}\right).
\end{cases}$$

At the symmetric equilibrium where $d_i = d_j = d^c$, $\partial D_i^{sh}/\partial v \equiv -\partial D_i/\partial v$.

$$\begin{aligned}
\frac{\partial^2 \Pi_i}{\partial d_i \partial v}_{|\{d_i = d^c, d_j = d^c\}} &= 2t\left(\frac{\partial D_i}{\partial v}\frac{\partial D_i}{\partial d_i} + D_i \frac{\partial^2 D_i}{\partial d_i \partial v}\right) + (1-\alpha)\frac{\partial (d_i y_i^c(d_i))}{\partial d_i}\frac{\partial D_i^{sh}}{\partial v},\\
&= \frac{\partial D_i}{\partial v}\left(2t\frac{\partial D_i}{\partial d_i} - (1-\alpha)\frac{\partial (d_i y_i^c(d_i))}{\partial d_i}\right) + 2tD_i\frac{\partial^2 D_i}{\partial d_i \partial v}\\
&= \frac{\partial D_i}{\partial v}\left(\underbrace{\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha - 1)\frac{\partial (d_i y_i^c(d_i))}{\partial d_i}}_{\gtreqless 0}\right) + \sigma D_i \underbrace{\frac{\partial q_i}{\partial d_i}}_{\leq 0}.
\end{aligned}$$

If $\alpha \leq 1/2$, the above equation is negative whereas the sign is unclear if $\alpha > 1/2$. In fact, we have

$$\frac{\partial^2 \Pi_i}{\partial d_i \partial v}_{|\{d_i=d^c, d_j=d^c\}} \begin{cases} \leq 0 & \text{if} \quad \sigma v \left|\frac{\partial q_i}{\partial d_i}\right| \geq (2\alpha-1)\frac{\partial(d_i y_i^c(d_i))}{\partial d_i} \quad \text{and} \quad \alpha > \frac{1}{2}, \\ \gtreqless 0 & \text{if} \quad \sigma v \left|\frac{\partial q_i}{\partial d_i}\right| < (2\alpha-1)\frac{\partial(d_i y_i^c(d_i))}{\partial d_i} \quad \text{and} \quad \alpha > \frac{1}{2}. \end{cases}$$

However if $\sigma v \left|\frac{\partial q_i}{\partial d_i}\right| < (2\alpha-1)\frac{\partial(d_i y_i^c(d_i))}{\partial d_i}$ and $\alpha > \frac{1}{2}$,

we have $\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha-1)\frac{\partial(d_i y_i^c(d_i))}{\partial d_i} > 0 \Rightarrow \sigma v \frac{\partial q_i}{\partial d_i} + \alpha \frac{\partial(d_i y_i^c(d_i))}{\partial d_i} > 0$;

then $\frac{\partial D_i}{\partial d_i} => 0 \Rightarrow \frac{\partial \Pi_i}{\partial d_i} > 0 \Rightarrow d^c = d^{max}$,

which means that we obtain a corner solution with a maximal disclosure level $d^{max}$.

From the implicit function theorem, we therefore obtain

$$\left.\frac{\partial d_i}{\partial v}\right|_{\{d_i=d^c, d_j=d^c\}} = -\frac{\partial^2 \Pi_i / \partial d_i \partial v}{\partial^2 \Pi_i / \partial d_i^2} \leq 0 \quad \text{if} \begin{cases} \alpha \leq \frac{1}{2}, \\ \alpha > \frac{1}{2} \quad \text{and} \quad \sigma v \left|\frac{\partial q_i}{\partial d_i}\right| > (2\alpha-1)\frac{\partial(d_i y_i^c(d_i))}{\partial d_i}. \end{cases}$$

We obtain similar results when studying how $d^c$ varies with $\sigma$.

We now study how $d^c$ varies with $t$.

$$\frac{\partial D_i}{\partial t} = -\frac{\sigma v q_i + \alpha d_i y_i^c}{2t^2} < 0; \quad \frac{\partial^2 D_i}{\partial d_i \partial t} = -\frac{\sigma v \partial q_i / \partial d_i + \alpha \partial(d_i y_i^c)/\partial d_i}{2t^2}; \quad \frac{\partial D_i^{sh}}{\partial t} = \frac{\sigma v q_j + \alpha d_j y_j}{2t^2}.$$

Note that at the symmetric equilibrium where $d_i = d_j = d^c$, $D_i \equiv D_j$.

$$\begin{aligned} \frac{\partial^2 \Pi_i}{\partial d_i \partial t}_{|\{d_i=d^c, d_j=d^c\}} &= 2D_i \frac{\partial D_i}{\partial d_i} + 2t\left(\frac{\partial D_i}{\partial t}\frac{\partial D_i}{\partial d_i} + D_i \frac{\partial^2 D_i}{\partial d_i \partial t}\right) + (1-\alpha)\frac{\partial(d_i y_i^c)}{\partial d_i}\frac{\partial D_i^{sh}}{\partial t} \\ &= 2D_i \frac{\partial D_i}{\partial d_i} + 2t(-\frac{D_i}{t}\frac{\partial D_i}{\partial d_i} - \frac{D_i}{t}\frac{\partial D_i}{\partial d_i}) + (1-\alpha)\frac{\partial(d_i y_i^c)}{\partial d_i}\frac{D_j}{t} \\ &= -2D_i \frac{\partial D_i}{\partial d_i} + (1-\alpha)\frac{\partial(d_i y_i^c)}{\partial d_i}\frac{D_j}{t} \\ &= D_i\left(\frac{1-\alpha}{t}\frac{\partial(d_i y_i^c)}{\partial d_i} - 2\frac{\partial D_i}{\partial d_i}\right) \\ &= -\frac{D_i}{t}\left(\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha-1)\frac{\partial(d_i y_i^c)}{\partial d_i}\right). \end{aligned}$$

Using the implicit function theorem, we find that

$$\frac{\partial d_i}{\partial t}_{|\{d_i=d^c, d_j=d^c\}} \geq 0 \begin{cases} \text{if} \quad \alpha \leq \frac{1}{2}, \\ \text{if} \quad \alpha > \frac{1}{2} \quad \text{and} \quad \sigma v \left|\frac{\partial q_i}{\partial d_i}\right| > (2\alpha-1)\frac{\partial(d_i y_i^c(d_i))}{\partial d_i}. \end{cases}$$

We now study how $d^c$ varies with $\alpha$.

$$\frac{\partial D_i}{\partial \alpha} = \frac{d_i y_i^c}{2t} > 0; \quad \frac{\partial^2 D_i}{\partial d_i \partial \alpha} = \frac{\partial (d_i y_i^c)/\partial d_i}{2t} > 0; \quad \frac{\partial D_i^{sh}}{\partial \alpha} = -\frac{d_j y_j}{2t} < 0.$$

Note that at the symmetric equilibrium where $d_i = d_j = d^c$, let $\partial D_i^{sh}/\partial \alpha \equiv -\partial D_i/\partial \alpha$.

$$\begin{aligned}
\frac{\partial^2 \Pi_i}{\partial d_i \partial \alpha}_{|\{d_i=d^c, d_j=d^c\}} &= 2t\left(\frac{\partial D_i}{\partial \alpha}\frac{\partial D_i}{\partial d_i} + D_i \frac{\partial^2 D_i}{\partial d_i \partial \alpha}\right) + \frac{\partial (d_i y_i^c)}{\partial d_i}\left((1-\alpha)\frac{\partial D_i^{sh}}{\partial \alpha} - D_i^{sh}\right) \\
&= \frac{\partial D_i}{\partial \alpha}\left(2t\frac{\partial D_i}{\partial d_i} - (1-\alpha)\frac{\partial (d_i y_i^c)}{\partial d_i}\right) + 2tD_i \frac{\partial^2 D_i}{\partial d_i \partial \alpha} - \frac{\partial (d_i y_i^c)}{\partial d_i} D_i^{sh} \\
&= \frac{\partial D_i}{\partial \alpha}\left(\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha - 1)\frac{\partial (d_i y_i^c)}{\partial d_i}\right) + \frac{\partial (d_i y_i^c)}{\partial d_i}(D_i + D_j - 1) \\
&= \frac{d_i y_i^c}{2t}\left(\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha - 1)\frac{\partial (d_i y_i^c)}{\partial d_i}\right) + \frac{\partial (d_i y_i^c)}{\partial d_i} D^{mh}. \qquad \text{(A.1)}
\end{aligned}$$

Therefore, the sign of $\partial d_i/\partial \alpha$ at equilibrium is ambiguous.

**Corollary 2**

At equilibrium, the variation of $p^c$ with respect to $\alpha$ is given by:

$$\begin{aligned}
\frac{\partial p^c}{\partial \alpha} =& \frac{1}{2}\left(\sigma v \frac{dq^c}{d\alpha} - \alpha \frac{d(d^c y^c)}{d\alpha} - d^c y^c\right) \\
=& \frac{1}{2}\left(\sigma v \frac{\partial d^c}{\partial \alpha}\left(\frac{\partial q^c}{\partial d_i} + \frac{\partial q^c}{\partial y_i}\frac{\partial y^c}{\partial d_i}\right) + \alpha \frac{\partial d^c}{\partial \alpha}\left(y^c + d^c \frac{\partial y^c}{\partial d_i}\right) - d^c y^c\right).
\end{aligned}$$

**Proposition 3**

$$\begin{cases}
\frac{\partial D_i}{\partial d_i}(d_i, d_j) = \frac{v}{2t}\frac{\partial q_i}{\partial d_i} + \frac{1}{4t}\frac{\partial (d_i y_i^m)}{\partial d_i}; \quad \frac{\partial D_j}{\partial d_i}(d_i, d_j) = -\frac{\partial D_i}{\partial d_i}. \\
\frac{\partial p_i}{\partial d_i}(d_i, d_j) = \frac{v}{2}\frac{\partial q_i}{\partial d_i} - \frac{1}{4}\frac{\partial (d_i y_i^m)}{\partial d_i} < 0; \quad \frac{\partial p_j}{\partial d_i}(d_i, d_j) = t\frac{\partial D_i}{\partial d_i}; \quad \frac{\partial^2 D_i}{\partial d_i \partial t}(d_i, d_j) = -\frac{1}{t}\frac{\partial D_i}{\partial d_i}.
\end{cases}$$

We determine the first-order conditions for the maximization of $d_i$ (similarly for $d_j$):

$$\begin{aligned}
\frac{\partial \Pi^m}{\partial d_i} &= \left(\frac{\partial p_i}{\partial d_i} + \frac{\partial (d_i y_i^m)}{\partial d_i}\right) D_i + (p_i + d_i y_i^m)\frac{\partial D_i}{\partial d_i} + (p_j + d_j y_j)\frac{\partial D_j}{\partial d_i} + \frac{\partial p_j}{\partial d_i} D_j = 0 \\
&= \left(\frac{v}{2}\frac{\partial q_i}{\partial d_i} + \frac{1}{4}\frac{\partial (d_i y_i^m)}{\partial d_i} + \frac{1}{2}\frac{\partial (d_i y_i^m)}{\partial d_i}\right) D_i + (p_i + d_i y_i^m)\frac{\partial D_i}{\partial d_i} + t\frac{\partial D_i}{\partial d_i} D_j - (p_j + d_j y_j^m)\frac{\partial D_i}{\partial d_i} = 0 \\
&= t\frac{\partial D_i}{\partial d_i} + \frac{D_i}{2}\frac{\partial (d_i y_i^m)}{\partial d_i} + \frac{d_i y_i^m - d_j y_j^m}{2}\frac{\partial D_i}{\partial d_i} = 0
\end{aligned}$$

We now check that the second-order conditions are satisfied when $d^c = d_i = d_j$, that is

$$\begin{cases} \frac{\partial^2 \Pi^m}{\partial d_i^2} = \frac{v}{2}\left(\frac{\partial^2 q_i}{\partial d_i^2} + \frac{1}{t}\frac{\partial q_i}{\partial d_i}\right) + \frac{1}{4}\left(3\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2} + \frac{\partial (d_i y_i^m)}{\partial d_i}\right) \leq 0 \quad \text{if} \quad \left|\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2}\right| > \frac{1}{3}\frac{\partial (d_i y_i^m)}{\partial d_i}; \\ \frac{\partial^2 \Pi^m}{\partial d_i \partial d_j} = \frac{1}{2}\frac{\partial (d_i y_i^m)}{\partial d_i}\frac{\partial D_i}{\partial d_j} - \frac{1}{2}\frac{\partial (d_j y_j^m)}{\partial d_j}\frac{\partial D_i}{\partial d_i}; \\ \text{Then assume that} \quad \frac{\partial^2 \Pi^m}{\partial d_i^2}\frac{\partial^2 \Pi^m}{\partial d_j^2} - \frac{\partial^2 \Pi^m}{\partial d_i \partial d_j}\frac{\partial^2 \Pi^m}{\partial d_j \partial d_i} > 0. \end{cases}$$

We now check the variations of $d^c$.

$$\frac{\partial^2 \Pi^m}{\partial d_i \partial v}_{|\{d_i = d^m, d_j = d^m\}} = t\frac{\partial^2 D_i}{\partial d_i \partial v} \leq 0.$$

We therefore use the implicit function theorem to find that

$$\frac{\partial d_i}{\partial v}_{|\{d_i = d^m, d_j = d^m\}} = -\frac{\partial^2 \Pi^m / \partial d_i \partial v}{\partial^2 \Pi^m / \partial d_i^2}_{|\{d_i = d^m, d_j = d^m\}} \leq 0.$$

If $d^m > 0$, we then have that

$$\frac{\partial p^m}{\partial v} = q(y^m, d^m) + v\left(\frac{\partial q^m}{\partial y}\frac{\partial y^m}{\partial d}\frac{\partial d^m}{\partial v} + \frac{\partial q^m}{\partial d}\frac{\partial d^m}{\partial v}\right) \geq 0 \quad \frac{\partial y^m}{\partial v} \geq 0.$$

Note that the optimal disclosure level $d^c$ does not depend on $t$:

$$\frac{\partial^2 \Pi^m}{\partial d_i \partial t}_{|\{d_i = d^m, d_j = d^m\}} = \frac{\partial D_i}{\partial d_i} + t\frac{\partial^2 D_i}{\partial d_i \partial t} = 0.$$

Since we wish to restrict our attention to parameter values such that the market is covered, we check under which conditions the monopoly chooses to cover the market. The monopolist covers the market if $\partial \Pi^u / \partial p_i|_{\{p_i = p_i^m, d_i = d_i^m\}} < 0$, which holds if $0 < t < vq(y^m, d^m) + d^m y^m$.

**Proposition 4**

$$\begin{cases} \frac{\partial D_i}{\partial d_i}(d_i, d_j) = \frac{1}{2t}\left(\sigma v \frac{\partial q_i}{\partial d_i} + \alpha \frac{\partial (d_i y_i^m)}{\partial d_i}\right); \quad \frac{\partial D_j}{\partial d_i}(d_i, d_j) = -\frac{(1-\alpha)}{2t}\frac{\partial (d_i y_i^m(d_i))}{\partial d_i}; \\ \frac{\partial^2 D_j}{\partial d_i^2}(d_i, d_j) = -\frac{(1-\alpha)}{2t}\frac{\partial^2 (d_i y_i^m(d_i))}{\partial d_i^2}; \quad \frac{\partial^2 D_i}{\partial d_i^2}(d_i, d_j) = \frac{1}{2t}\left(\sigma v \frac{\partial^2 q_i}{\partial d_i^2} + \alpha \frac{\partial^2 (d_i y_i^m(d_i))}{\partial d_i^2}\right). \end{cases}$$

The first-order conditions write as follows (similarly for $d_j$),

$$\frac{\partial \Pi^m}{\partial d_i} = 0 \Leftrightarrow 2t\left(\frac{\partial D_i}{\partial d_i}D_i + \frac{\partial D_j}{\partial d_i}D_j\right) + (1-\alpha)\frac{\partial (d_i y_i^m)}{\partial d_i} = 0$$

We check that the second-order conditions for an interior solution are satisfied.

$$\begin{cases} \frac{\partial^2 \Pi^m}{\partial d_i^2} = 2t\left(\frac{\partial^2 D_i}{\partial (d_i)^2}D_i + \left(\frac{\partial D_i}{\partial d_i}\right)^2 + \frac{\partial^2 D_j}{\partial (d_i)^2}D_j + \left(\frac{\partial D_j}{\partial d_i}\right)^2\right) + (1-\alpha)\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2} \\ = 2t\left(D_i\left(\frac{\partial^2 D_i}{\partial (d_i)^2}\right) + D_j\left(\frac{\partial^2 D_j}{\partial (d_i)^2}\right) + \left(\frac{\partial D_i}{\partial d_i}\right)^2 + \left(\frac{\partial D_j}{\partial d_i}\right)^2\right) + (1-\alpha)\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2} \\ = 2t\left(D_i \frac{\sigma v}{2t}\frac{\partial^2 q_i}{\partial d_i^2} + \frac{\partial^2 (d_i y_i^m)}{\partial d_i^2}\frac{\alpha(D_i+D_j)-D_j}{2t} + \left(\frac{\partial D_i}{\partial d_i}\right)^2 + \left(\frac{\partial D_j}{\partial d_i}\right)^2\right) + (1-\alpha)\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2} \\ \text{In the symmetric equilibrium where } d_i = d_j \text{, we then have} \\ = 2t\left(\frac{D_i}{2t}\left(\underbrace{\sigma v\frac{\partial^2 q_i}{\partial d_i^2}}_{<0} + \underbrace{(2\alpha-1)\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2}}_{\gtreqless 0}\right) + \left(\frac{\partial D_i}{\partial d_i}\right)^2 + \left(\frac{\partial D_j}{\partial d_i}\right)^2\right) + \underbrace{(1-\alpha)\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2}}_{\leq 0}; \\ \frac{\partial^2 \Pi^m}{\partial d_i \partial d_j} = 2t\left(\frac{\partial D_i}{\partial d_i}\frac{\partial D_i}{\partial d_j} + \frac{\partial D_j}{\partial d_i}\frac{\partial D_j}{\partial d_j}\right). \end{cases}$$

Let $A \equiv \sigma v \partial^2 q_i/\partial d_i^2 + (2\alpha-1)\partial^2(d_i y_i^m)/\partial d_i^2$ and $B \equiv (\partial D_i/\partial d_i)^2 + (\partial D_j/\partial d_i)^2$. We assume that $A < 0$ and $|A| > B$, then $\partial^2 \Pi^m/\partial d_i^2 \leq 0$. Notably, if $\alpha > 1/2$, the assumption is more likely to be valid compared to if $\alpha \leq 1/2$.

$$\text{Assume that} \quad \frac{\partial^2 \Pi^m}{\partial d_i^2}\frac{\partial^2 \Pi^m}{\partial d_j^2} - \frac{\partial^2 \Pi^m}{\partial d_i \partial d_j}\frac{\partial^2 \Pi^m}{\partial d_j \partial d_i} > 0,$$

to ensure the validity of the second-order conditions.

All consumers are served and there is multi-homing if the following set of conditions is satisfied:

$$\begin{aligned} U_i^m &= vq(y^m, d^m) - p^m - tx^* \geq 0, \\ U_{ij}^m &= (1+\sigma)vq(y^m, d^m) - 2p^m - t \geq 0, \\ 0 &< \underbrace{\frac{\sigma v q^m + (2\alpha-1)d^m y^m}{t} - 1}_{D^{mh}} < 1. \end{aligned}$$

The above conditions are satisfied if $(\sigma v q^m + (2\alpha-1)d^m y^m)/2 < t < vq^m + (2\alpha-1)d^m y^m$.

We now check the variations of which $d^m$ . Let us examine how $d^m$ varies with consumer valuations $v$.

$$\frac{\partial D_i}{\partial v}(d_i, d_j) = \frac{\sigma q_i}{2t} > 0; \quad \frac{\partial^2 D_i}{\partial d_i \partial v}(d_i, d_j) = \frac{\sigma}{2t}\frac{\partial q_i}{\partial d_i} \leq 0, \quad \frac{\partial^2 D_j}{\partial d_i \partial v}(d_i, d_j) = 0.$$

Note that at the symmetric equilibrium where $d_i = d_j = d^m$, $\partial D_i/\partial v \equiv \partial D_j/\partial v$.

$$\begin{aligned}
\frac{\partial^2 \Pi^m}{\partial d_i \partial v}_{|\{d_i=d^m, d_j=d^m\}} &= 2t\left(\frac{\partial^2 D_i}{\partial d_i \partial v} D_i + \frac{\partial D_i}{\partial d_i}\frac{\partial D_i}{\partial v} + \frac{\partial^2 D_j}{\partial d_i \partial v} D_j + \frac{\partial D_j}{\partial d_i}\frac{\partial D_j}{\partial v}\right) \\
&= 2t\left(\frac{\partial^2 D_i}{\partial d_i \partial v} D_i + \frac{\partial D_i}{\partial v}\left(\frac{\partial D_i}{\partial d_i} + \frac{\partial D_j}{\partial d_i}\right)\right) \\
&= \sigma \frac{\partial q_i}{\partial d_i} D_i + \frac{\partial D_i}{\partial v}\left(\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha - 1)\frac{\partial (d_i y_i^m)}{\partial d_i}\right).
\end{aligned}$$

From the implicit function theorem, we therefore obtain

$$\left.\frac{\partial d_i}{\partial v}\right|_{\{d_i=d^m, d_j=d^m\}} = -\frac{\partial^2 \Pi^m / \partial d_i \partial v}{\partial^2 \Pi^m / \partial d_i^2} \quad \text{if} \begin{cases} \alpha \leq \frac{1}{2}, \\ \alpha > \frac{1}{2} \quad \text{and} \quad \sigma v \left|\frac{\partial q_i}{\partial d_i}\right| > (2\alpha - 1)\frac{\partial (d_i y_i^m)}{\partial d_i}. \end{cases}$$

Notably,

$$\text{If} \quad \sigma v \left|\frac{\partial q_i}{\partial d_i}\right| < (2\alpha - 1)\frac{\partial (d_i y_i^m)}{\partial d_i} \quad \text{and} \quad \alpha > \frac{1}{2} \quad \Rightarrow \frac{\partial \Pi^m}{\partial d_i} > 0 \quad \Rightarrow d^m = d^{max}.$$

We obtain a similar result when studying the variations of $d^m$ with respect to $\sigma$.

We also study how $d^m$ varies with $t$.

$$\frac{\partial D_i}{\partial t} = -\frac{\sigma v q_i + \alpha d_i y_i^m - (1-\alpha) d_j y_j^m}{2t^2} < 0; \quad \frac{\partial^2 D_i}{\partial d_i \partial t} = -\frac{\sigma v \partial q_i/\partial d_i + \alpha \partial (d_i y_i^m)/\partial d_i}{2t^2};$$
$$\frac{\partial^2 D_j}{\partial d_i \partial t} = \frac{(1-\alpha)}{2t^2}\frac{\partial (d_i y_i^m)}{\partial d_i} > 0.$$

At the symmetric equilibrium where $d_i = d_j = d^m$, $D_i \equiv D_j$.

$$\begin{aligned}
\frac{\partial^2 \Pi^m}{\partial d_i \partial t}_{|\{d_i=d^m, d_j=d^m\}} =& 2\left(\frac{\partial D_i}{\partial d_i} D_i + \frac{\partial D_j}{\partial d_i} D_j\right) \\
&+ 2t\left(\frac{\partial^2 D_i}{\partial d_i \partial t} D_i + \frac{\partial D_i}{\partial d_i}\frac{\partial D_i}{\partial t} + \frac{\partial^2 D_j}{\partial d_i \partial t} D_j + \frac{\partial D_j}{\partial d_i}\frac{\partial D_j}{\partial t}\right) \\
=& 2\frac{\partial D_i}{\partial d_i} D_i + 2\frac{\partial D_j}{\partial d_i} D_j + 2t\left(-\frac{\partial D_i}{\partial d_i}\frac{D_i}{t} - \frac{\partial D_i}{\partial d_i}\frac{D_i}{t} - \frac{\partial D_j}{\partial d_i}\frac{D_j}{t} - \frac{\partial D_j}{\partial d_i}\frac{D_j}{t}\right) \\
=& -2\left(\frac{\partial D_i}{\partial d_i} D_i + \frac{\partial D_j}{\partial d_i} D_j\right) \\
=& -\frac{D_i}{t}\left(\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha - 1)\frac{\partial d_i y_i^m}{\partial d_i}\right).
\end{aligned}$$

Using the implicit function theorem, we find that

$$\left.\frac{\partial d_i}{\partial t}\right|_{\{d_i=d^m,d_j=d^m\}} \geq 0 \begin{cases} \text{if} \quad \alpha \leq \frac{1}{2}, \\ \text{if} \quad \alpha > \frac{1}{2} \quad \text{and} \quad \sigma v \left|\frac{\partial q_i}{\partial d_i}\right| < (2\alpha-1)\frac{\partial(d_i y_i^m)}{\partial d_i}. \end{cases}$$

We finally study how $d^m$ varies with $\alpha$.

$$\begin{cases} \frac{\partial D_i}{\partial \alpha}(d_i,d_j) = \frac{\partial D_j}{\partial \alpha}(d_i,d_j) = \frac{d_i y_i^m + d_j y_j^m}{2t} > 0; \\ \frac{\partial^2 D_i}{\partial d_i \partial \alpha}(d_i,d_j) = \frac{\partial(d_i y_i^m)/\partial d_i}{2t} = \frac{\partial^2 D_j}{\partial d_i \partial \alpha}(d_i,d_j) \geq 0. \end{cases}$$

$$\begin{aligned} \frac{\partial^2 \Pi^m}{\partial d_i \partial \alpha}_{|\{d_i=d^m,d_j=d^m\}} &= 2t\left(\frac{\partial^2 D_i}{\partial d_i \partial \alpha}D_i + \frac{\partial D_i}{\partial d_i}\frac{\partial D_i}{\partial \alpha} + \frac{\partial^2 D_j}{\partial d_i \partial \alpha}D_j + \frac{\partial D_j}{\partial d_i}\frac{\partial D_j}{\partial \alpha}\right) - \frac{\partial(d_i y_i^m)}{\partial d_i} \\ &= 2t\left(\frac{\partial D_i}{\partial \alpha}\left(\frac{\partial D_i}{\partial d_i} + \frac{\partial D_j}{\partial d_i}\right)\right) + 2t\frac{\partial^2 D_i}{\partial d_i \partial \alpha}(D_i + D_j) - \frac{\partial(d_i y_i^m)}{\partial d_i} \\ &= \frac{\partial D_i}{\partial \alpha}\left(\sigma v \frac{\partial q_i}{\partial d_i} + (2\alpha-1)\frac{\partial(d_i y_i^m)}{\partial d_i}\right) + \frac{\partial(d_i y_i^m)}{\partial d_i}(\underbrace{D_i + D_j - 1}_{D^m}). \end{aligned} \tag{A.2}$$

Therefore, the sign of $\partial d^m/\partial \alpha$ is ambiguous.

**Proposition 5**

-Proposition 5(i): for all $d_i$ and $d_j$, comparing $p^c(d_i)$ and $p^m(d_i,d_j)$, we have $p_i^c(d_i) \leq p_i^m(d_i,d_j)$. It implies that $p^c \leq p^m$.

-Proposition 5(ii): for all $y_i^m(d_i)$, $y_j^m(d_j)$, $p_i^m(d_i,d_j)$, and $p_j^m(d_i,d_j)$,

$$\begin{aligned} \frac{d\Pi^m}{dd_i}(d_i,d_j) &= \frac{\partial \Pi^m}{\partial d_i} + \frac{\partial \Pi^m}{\partial p_i^m}\frac{\partial p_i^m}{\partial d_i} + \frac{\partial \Pi^m}{\partial p_j^m}\frac{\partial p_j^m}{\partial d_i} \\ &= \left(\underbrace{p_i^m + \alpha d_i y_i^m - (1-\alpha)d_j y_j^m}_{A^m}\right)\frac{\partial D_i^m}{\partial d_i} + \frac{\partial d_i y_i^m}{\partial d_i}\left(\underbrace{\alpha D_i^m + (1-\alpha)D_i^{sh,m}}_{B^m}\right). \end{aligned}$$

For all $y_i^c(d_i)$, $y_j^c(d_j)$, $p_i^c(d_i)$, and $p_j^c(d_i)$,

$$\begin{aligned} \frac{d\Pi_i^c}{dd_i}(d_i,d_j) &= \frac{\partial \Pi^c}{\partial d_i} + \frac{\partial \Pi^c}{\partial p_i^c}\frac{\partial p_i^c}{\partial d_i} \\ &= \left(\underbrace{p_i^c + \alpha d_i y_i^c}_{A^c}\right)\frac{\partial D_i^c}{\partial d_i} + \frac{\partial d_i y_i^c}{\partial d_i}\left(\underbrace{\alpha D_i^c + (1-\alpha)D_i^{sh,c}}_{B^c}\right). \end{aligned}$$

First, $\partial D_i^c/\partial d_i = \partial D_i^m/\partial d_i$ and $A^m < A^c$ for all $d_i$. Second, since $\partial d_i y_i^c/\partial d_i = \partial d_i y_i^m/\partial d_i$

for all $d_i$, we compare $B^m$ and $B^c$ (reminding that $y_i^c(d_i) = y_i^m(d_i)$, we remove the superscript)

$$\begin{aligned} B^m - B^c &= \alpha\,(D_i^m - D_i^c) + (1-\alpha)(D_i^{sh,m} - D_i^{sh,c}) \\ &= \alpha(\underbrace{D_i^m - D_i^c}_{<0}) + (1-\alpha)(\underbrace{D_j^c - D_j^m}_{>0}) \\ B^m - B^c \geq 0 &\Leftrightarrow \alpha \leq \frac{d_i y_i}{d_i y_i + d_j y_j}. \end{aligned}$$

The duopoly equilibrium as well the multi-product monopoly one are symmetric, i.e., we obtained $d_A = d_B = d^c$ and $d_A = d_B = d^m$. Therefore, for all $d_i = d_j$, we have $B^m - B^c \geq 0 \Leftrightarrow \alpha \leq 1/2$. Therefore,

$$\begin{cases} \left.\frac{d\Pi^m}{dd_i}(d_i,d_j) - \frac{d\Pi_i^c}{dd_i}(d_i,d_j)\right|_{d_i=d_j} \gtreqless 0 & \text{if} \quad \alpha < \frac{1}{2} \\ \left.\frac{d\Pi^m}{dd_i}(d_i,d_j) - \frac{d\Pi_i^c}{dd_i}(d_i,d_j)\right|_{d_i=d_j} \leq 0 & \text{if} \quad \alpha \geq \frac{1}{2}. \end{cases}$$

Given that $d^m$ is such that $d\Pi^m/dd_i|_{d_i=d_j=d^m} = 0$ and $d\Pi_i^c/dd_i|_{d_i=d_j=d^c} = 0$, if $\alpha \geq 1/2$,

$$\begin{cases} \left.\frac{d\Pi^m}{dd_i}(d_i,d_j)\right|_{d_i=d_j=d^m} = 0 \leq \left.\frac{d\Pi_i^c}{dd_i}(d_i,d_j)\right|_{d_i=d_j=d^m}, \\ \left.\frac{d\Pi^m}{dd_i}(d_i,d_j)\right|_{d_i=d_j=d^c} \leq \left.\frac{d\Pi_i^c}{dd_i}(d_i,d_j)\right|_{d_i=d_j=d^c} = 0 \end{cases} \Rightarrow \left.\frac{d\Pi^m}{dd_i}(d_i,d_j)\right|_{d_i=d_j=d^c} \leq \left.\frac{d\Pi_i^c}{dd_i}(d_i,d_j)\right|_{d_i=d_j=d^m},$$

which implies that $d^c \geq d^m$ if $\alpha \geq 1/2$.

**Proposition 4'**

$$\frac{\partial D_i}{\partial v} = \frac{\sigma q_i}{2t} \geq 0; \quad \frac{\partial D_i}{\partial d_i} = \frac{\sigma v}{2t}\frac{\partial q_i}{\partial d_i} \leq 0; \quad \frac{\partial^2 D_i}{\partial d_i \partial v} = \frac{\sigma}{2t}\frac{\partial q_i}{\partial d_i} \leq 0.$$

The first-order conditions write as follows,

$$\frac{\partial \Pi^m}{\partial d_i} = 0 \Leftrightarrow 2t\frac{\partial D_i}{\partial d_i} D_i + \frac{1}{2}\frac{\partial (d_i y_i^m)}{\partial d_i} = 0$$

We check that the second-order conditions for an interior solution are satisfied:

$$\begin{cases} \frac{\partial^2 \Pi^m}{\partial d_i^2} = 2t\left(\frac{\partial^2 D_i}{\partial (d_i)^2} D_i + \left(\frac{\partial D_i}{\partial d_i}\right)^2\right) + \frac{1}{2}\frac{\partial^2 (d_i y_i^m)}{\partial d_i^2} \leq 0, & \text{by Assumption 1}, \\ \frac{\partial^2 \Pi^m}{\partial d_i \partial d_j} = 0. \end{cases}$$

We now check how $d^m$ varies with $v$.

$$\frac{\partial^2 \Pi^m}{\partial d_i \partial v}_{|\{d_i=d^m, d_j=d^m\}} = 2t\left(\frac{\partial D_i}{\partial v}\frac{\partial D_i}{\partial d_i} + D_i \frac{\partial^2 D_i}{\partial d_i \partial v}\right) \leq 0$$

From the implicit function theorem, we therefore obtain

$$\frac{\partial d_i}{\partial v}_{|\{d_i=d^m, d_j=d^m\}} = -\frac{\partial^2 \Pi^m / \partial d_i \partial v}{\partial^2 \Pi_i / \partial d_i^2} \leq 0.$$

We obtain a similar result regarding how $d^m$ varies with $\sigma$.

We also check how $d^m$ varies with $t$.

$$\frac{\partial D_i}{\partial t} = -\frac{\sigma v q_i}{2t^2} = -\frac{D_i}{t} < 0; \quad \frac{\partial^2 D_i}{\partial d_i \partial t} = -\frac{\sigma v}{2t^2}\frac{\partial q_i}{\partial d_i} = -\frac{1}{t}\frac{\partial D_i}{\partial d_i} \geq 0.$$

$$\begin{aligned}
\frac{\partial^2 \Pi^m}{\partial d_i \partial t}_{|\{d_i=d^m, d_j=d^m\}} &= 2\frac{\partial D_i}{\partial d_i} D_i + 2t\left(\frac{\partial D_i}{\partial t}\frac{\partial D_i}{\partial d_i} + D_i \frac{\partial^2 D_i}{\partial d_i \partial t}\right) \\
&= 2\frac{\partial D_i}{\partial d_i}\left(D_i + t\frac{\partial D_i}{\partial t}\right) + 2tD_i \frac{\partial^2 D_i}{\partial d_i \partial t} \\
&= 2tD_i \frac{\partial^2 D_i}{\partial d_i \partial t} \geq 0.
\end{aligned}$$

Using the implicit function theorem, we find that

$$\left.\frac{\partial d_i}{\partial t}\right|_{\{d_i=d^m, d_j=d^m\}} \geq 0.$$

All consumers are served and there is multi-homing if the following set of conditions is satisfied:

$$\begin{aligned}
&U_i^m = vq(y^m, d^m) - p^m - tx^* \geq 0, \\
&U_{ij}^m = (1+\sigma)vq(y^m, d^m) - 2p^m - t \geq 0, \\
&0 < \underbrace{\frac{\sigma v q^m}{t} - 1}_{D^m} < 1.
\end{aligned}$$

The above conditions are satisfied if $\sigma v q^m/2 < t < vq^m$.

## Appendix B

**The monopolist internalized the double-selling issue on MH consumer data**

The monopoly profit is given by

$$\Pi^m(p_i, p_j) = \sum_{i=A,B} p_i D_i(p_i) + \frac{1}{2} d_i y_i^m. \tag{A.3}$$

The monopoly chooses $p_A$ and $p_B$ to maximize (12) by setting $\partial \Pi^m / \partial p_i = 0$. Solving for the first-order conditions, we obtain the equilibrium price

$$p_i^m(d_i) = \frac{\sigma v q_i}{2}, \quad i = A, B.$$

Notice that here, the monopolist always charges a positive price ($p^m(d_i) > 0$) and this price is decreasing in $d_i$ ($\partial p_i^m(d_i)/\partial d_i \leq 0$).

The monopoly demand now writes,

$$D_i(d_i) = \frac{\sigma v q_i}{2t}, \quad D_i^{sh}(d_i) = 1 - D_i.$$

Plugging the equilibrium prices into the profit function (12), we now solve for the optimal disclosure levels of the monopoly at Stage 1. The monopoly can be written as

$$\Pi^m(d_i, d_j) = t\left((D_i(d_i))^2 + (D_j(d_j))^2\right) + \frac{1}{2}(d_i y_i^c + d_j y_j^m)$$

The monopoly sets its disclosure levels to maximize its profit $\Pi^m(d_i, d_j)$. We obtain the equilibrium disclosure levels $d_A^m$ and $d_B^m$.

The following proposition summarizes the analysis.

**Proposition 4′.** *(i) In the monopoly equilibrium, if $\sigma v q^m/2 < t \leq v q^m$, consumers provide information $y_i = y^m$ ($i = A, B$). Optimal prices and disclosure levels are $p_i = p^m = \sigma v q^m/2$ and $d_i = d^m$.*

*(ii) The monopolist's choice to charge positive disclosure levels depends on consumer valuations $v$, the value of MH $\sigma$, and the level of product differentiation $t$, where $\partial d^m/\partial v \leq 0$, $\partial d^m/\partial \sigma \leq 0$, and $\partial d^m/\partial t \geq 0$.*

# Appendix C: Figures

## MH duopoly

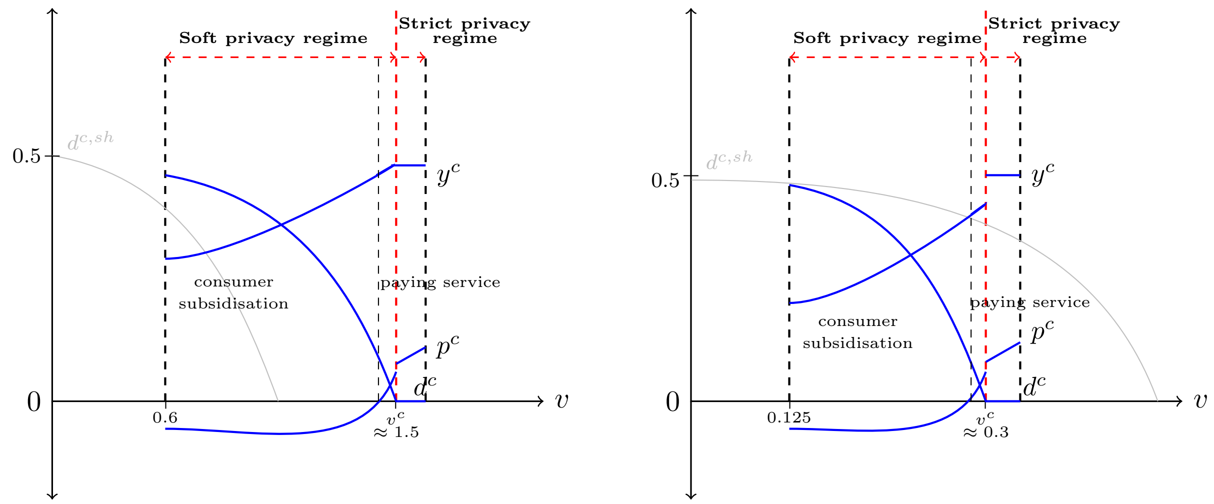


Figure 3: *Left*: $\sigma = \underline{0.2}, \alpha = 0.25, t = 0.04$ ; *Right*: $\sigma = \underline{1}, \alpha = 0.25, t = 0.04$.

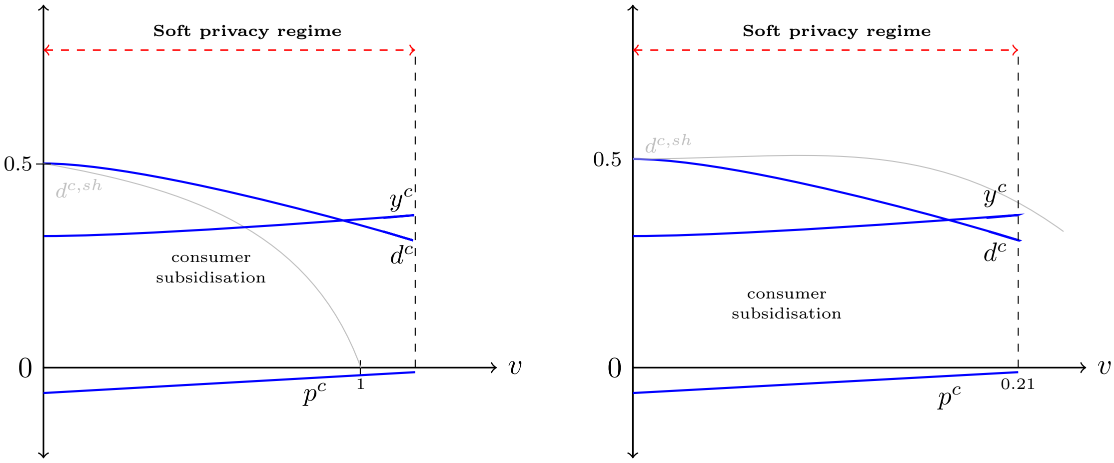


Figure 4: *Left*: $\sigma = \underline{0.2}, \alpha = 0.5, t = 0.04$; *Right*: $\sigma = \underline{1}, \alpha = 0.5, t = 0.04$.

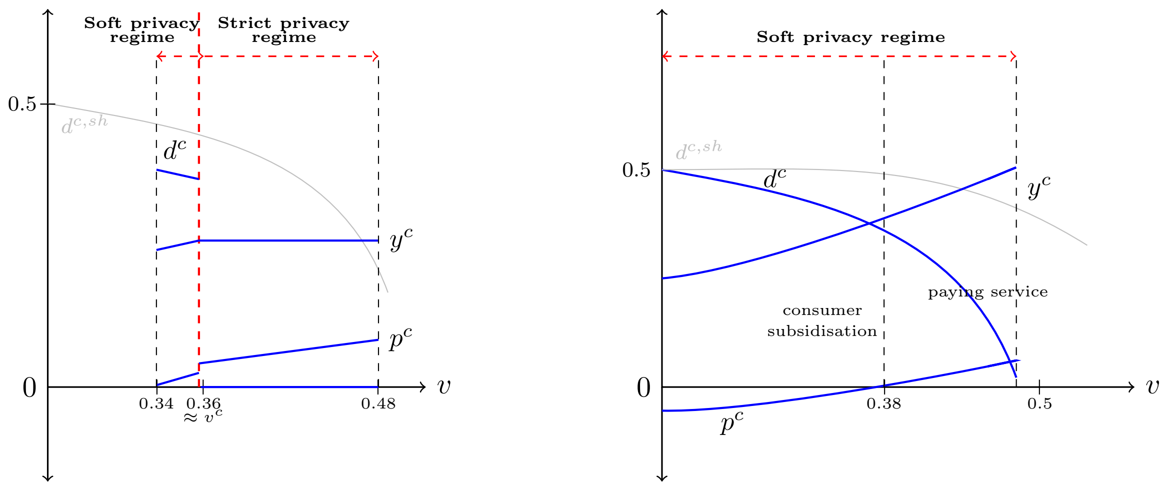


Figure 5: *Left*: $\sigma = 1, \alpha = \underline{0.25}, t = 0.06$; *Right*: $\sigma = 1, \alpha = \underline{0.5}, t = 0.06$.

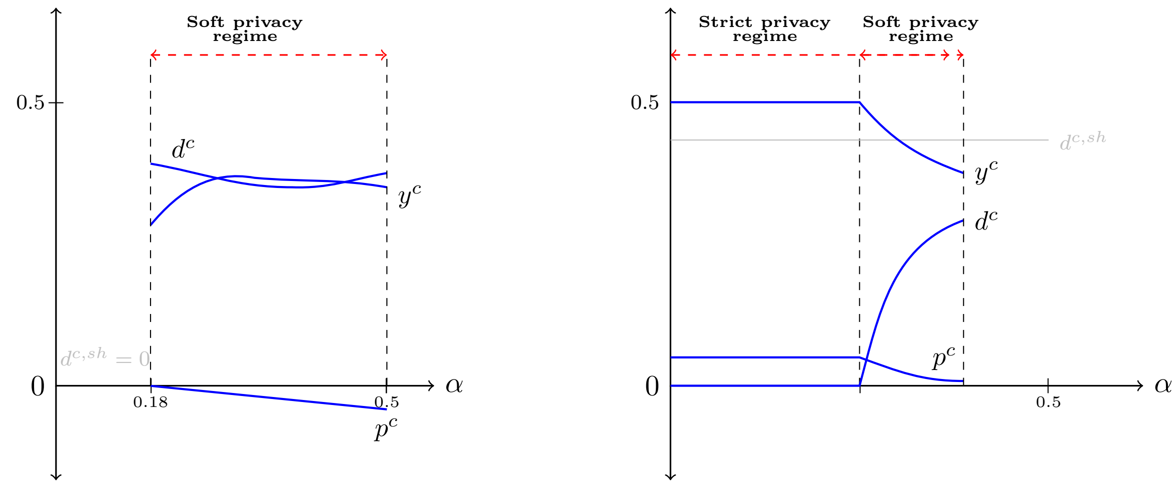


Figure 6: *Left*: $\sigma = \underline{0.2}, v = \underline{1.2}, t = 0.04$; *Right*: $\sigma = \underline{1}, v = \underline{0.3}, t = 0.04$.