Privacy and Quality[*]

Yassine Lefouili[†] Ying Lei Toh[‡]

April 20, 2017

Abstract

This paper analyzes the effects of a privacy regulation that caps the level of data disclosure on investment in quality and social welfare. We develop a model in which a monopolist offers a service for free to consumers with heterogeneous privacy preferences, and derives revenues from disclosing consumer data to third parties. We assume that the users of the service choose how much information they provide to the firm. In this setting, regulating the disclosure level can alter both the extensive margin effect and the intensive margin effect of an investment in quality. In the case where the market is fully covered, a welfare-maximizing regulator who can commit *ex ante* to a cap on disclosure level finds it optimal to do so when the complementarity between quality and information is not too strong. In the case where the market is partially covered, such an *ex ante* disclosure cap may be socially desirable even when quality and information are strongly complementary. Finally, we extend our analysis to the case where the regulator maximizes consumer surplus, and to a scenario where the regulator sets a disclosure cap *ex post*.

Keywords: Privacy Regulation, Data Disclosure, Quality.

JEL Classification: D83, L15, L4, M21.

[*]We are grateful to Romain De Nijs, Andres Hervas-Drane, Bruno Jullien and Mike Riordan for very useful comments and discussions. We also thank the participants at the tenth IDEI-TSE-IAST Conference on the Economics of Intellectual Property, Software, and the Internet and the Conference on the Economics of Information and Communication Technologies, and the anonymous reviewers at at Workshop of the Economics of Information Security for helpful suggestions. The financial support of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 670494) is gratefully acknowledged.

[†]Toulouse School of Economics, University of Toulouse Capitole. E-mail: yassine.lefouili@tse-fr.eu

[‡]Toulouse School of Economics, University of Toulouse Capitole. E-mail: yinglei.toh@tse-fr.eu

1 Introduction

In today's digital economy where most services are offered essentially for free, consumer data has emerged as a new form of currency. Many firms, including Internet giants such as Google and Facebook, have chosen to monetize the user data that they have collected, rather than charge a positive price for the content or services that they provide. Consumer data could be either sold directly to third parties (e.g. direct marketing companies and data brokers) and/or used to tailor the advertisements shown to consumers (Lambrecht *et al.*, 2014). As firms' ability to collect and analyze information continue to improve with the advancement of technology, many have begun to voice their concerns over the data practices of firms. Some have argued that the existing notice and consent requirement with regards to a firm's data practices is inadequate for protecting consumer privacy, because only "in some fantasy world, users read these notices, understand their implications ... then click to indicate their consent." (The White House, 2014, p. 38). Others further questioned if consumers have any meaningful alternatives to consenting at all, given the dominant positions of many of these firms. Facebook, for example, has recently been investigated by the German Federal Cartel Office for abusing its dominant position through the terms and conditions it imposes on the use of personal data.[1]

One means of protecting consumer privacy is to regulate data use/disclosure by firms. For example, a regulator could restrict the type and quantity of information that a firm is allowed to disclose to third parties or limit the duration for which consumer data could be stored. Although such a policy would help to protect consumer privacy, it may hurt the incentives of firms to innovate or to invest in the quality of their services. These restrictions has been likened to the act of "killing the goose that lays the Internets golden eggs".[2] Athey (2014) points out, for instance, how privacy regulation may reduce the efficiency of online advertising, making it more difficult for a firm to build up its initial user base and to derive revenues from it, and hence, lowering the firm's incentives to innovate or to create new content. It is unclear *a priori* whether a regulation on data use/disclosure is socially desirable; this paper attempts to shed some light on this issue.

We develop a model in which a monopolist offers a service to consumers with heterogeneous privacy preferences. The monopolist does not charge a price but instead derives revenues from disclosing consumer information to third parties (e.g., for the purpose of targeted advertising). Further, it may invest in the quality of its service. The gross utility that a consumer derives from using the service is increasing in both the quality level of its service and the amount of information input that the consumer provides. This information can take the form of purchase and browsing histories, location information, personal information shared on a social media profile, reviews and comments and so on.[3] When providing

[1] Boot, Nuria and Petropoulos, Georgios, "German Facebook probe links data protection and competition policy," *Bruegel*, 14 March 2016, `http://bruegel.org/2016/03/german-facebook-probe-links-data-protection-and-competition-policy`.

[2] Thierer, Adam, "Privacy regulation and the free Internet", *Reuters*, 23 December 2010, `http://blogs.reuters.com/mediafile/2010/12/23/privacy-regulation-and-the-free-internet/`

[3] More generally, one could also consider the level of usage as the amount of information input provided by a consumer. By increasing her level of usage, the consumer provides the firm with more opportunities to gather information about herself (e.g., the firm could track the consumer's activities by means of cookies).

information to the firm, the consumer also incurs a privacy cost. This privacy cost depends on the strength of the consumer's preference for privacy, her level of information provision and the monopolist's level of disclosure.

An important modelling innovation of our paper lies in the introduction of a quality dimension to the firm's service. The service's quality level and the consumer's information provision level jointly determine the gross utility that the consumer derives from using the service. More specifically, we consider the case where quality and information are complements from the consumer's perspective. In other words, the marginal benefit of information provision, and hence the amount of information provided, is increasing in the quality level of the service. As an example, consider the services provided by social media platforms such as Facebook and Instagram. By innovating and creating new and better sharing tools (which corresponds to a higher quality of service), a social media site enhances the users' sharing experience. The marginal utility of sharing and therefore the amount of content shared by users is likely to increase as a result. The service provided by a product recommendation site (or, more generally, a matching site) serves as another example. The higher the quality of a firm's recommendation/matching algorithm, the better the match the consumer obtains when she reveals her preferences (personal information) to the firm.

We first examine the participation and information provision decisions of a consumer. For given levels of quality and disclosure, a consumer patronizes the firm if her idiosyncratic privacy cost from information disclosure is sufficiently small. Further, a larger share of consumers participate when the quality level is higher and the disclosure level is lower. Conditional on participation, a consumer always provides less information when disclosure level is higher, when the quality level is lower and when her preference for privacy is stronger. We then compare the privately and socially optimal choice of quality and disclosure levels. We show that the monopolist under-provides quality for a given level of disclosure and over-discloses information for a given level of quality relative to the social optimum. These intuitive results stem from the fact that the monopolist fails to take into account the benefit of quality and the cost of disclosure to the consumers when deciding on its levels of quality and disclosure.

Next, we analyze the effect of a privacy regulation taking the form of a cap on information disclosure on quality and social welfare. We first consider the scenario of an *ex ante* regulation, in which a social-welfare-maximizing regulator sets the level of the disclosure cap before the monopolist decides on its level of quality.[4] We further distinguish between two sub-cases - full and partial market coverage. Full market coverage arises when the opportunity cost of using the firm's service is sufficiently low.[5] When this is the case, quality and disclosure levels only affect the amount of information provided by consumers (the intensive margin) but not their participation decisions (the extensive margin).[6] Under full market coverage, a cap on disclosure weakly decreases the quality level chosen by

[4]This implicitly assumes that the regulator is able to commit to the level of the cap.

[5]This may correspond to the scenario where consumers do not have meaningful alternatives to the service offered by the monopolist.

[6]Our results under that scenario would also hold in a more general setting where the market is not fully covered but quality and disclosure levels do not affect consumers' participation (while they affect the amount of information provided by consumers to the firm).

the monopolist — there is a trade-off between more consumer privacy and higher service quality. This reduction in quality level lowers social welfare. However, because the cap has a positive (direct) effect on reducing the consumers' privacy costs, the overall impact on social welfare is still positive when quality and information are sufficiently weak complements. When the market is only partially covered, a disclosure cap may lead to an increase or decrease in the quality of service, depending on the shape of the distribution of the idiosyncratic privacy cost. If the distribution exhibits a weakly increasing density (which implies that a relatively large share of consumers have high idiosyncratic privacy costs), a cap leads to an increase in quality. In this case, there is no trade-off between privacy and quality, and the cap unambiguously increases social welfare. However, if the distribution exhibits a decreasing density function, the effect of a cap on disclosure on social welfare is ambiguous.

As an extension, we explore the case where the regulator is a consumer protection agency (and therefore, cares only about maximizing consumer surplus) in the *ex ante* regulation setting. While the results in this case do not qualitatively differ from those derived under a welfare-maximizing regulator, we show that the consumer protection agency will set a (weakly) lower disclosure cap than a social-welfare-maximizing regulator. In another extension, we examine the scenario of an *ex post* regulation. Under *ex post* regulation, the regulator imposes a disclosure cap after the firm has invested in quality. This could either correspond to the case where the regulator is unable to commit to the level of the disclosure cap, or to the scenario where a regulator attempts to prevent a firm from abusing its dominant position after it has acquired a high degree of market power (as in the case of Facebook described earlier). As compared to the case of *ex ante* regulation, the impact of an *ex post* regulation depends also on how the value of information compares with the average privacy cost of the consumers.

Related literature. Our work contributes to the growing pool of economic literature on the use and protection of consumer data by firms (see Acquisti et al. 2016 for a survey). Closely related to our work are papers that consider the collection and exploitation of consumer data by firms. Bloch and Demange (2017) and Bourreau et al. (2017) examine the taxation and regulation of digital monopoly platforms that derive (part of) their revenues from the exploitation of consumer data. Like in our paper, they consider models in which: (i) a monopolist offers a service to consumers who have homogeneous valuation for the service but heterogeneous privacy costs, and (ii) the exploitation of consumer data raises the monopolist's revenues and the value of the service to the consumers. The amount of data that a consumer has to provide is set by the monopolist in Bloch and Demange (2017), while it is chosen by the consumer in Bourreau et. al (2017) and in our work. One key difference between our model and those in Bloch and Demange (2017) and Bourreau et. al (2017) stems from the monopolist information disclosure/exploitation policy. In both of these papers, the monopolist discloses all of the data that is used in the provision of its service whereas we allow for partial disclosure. More specifically, all of the data that a consumer shares is used to provide the service but only a share of that data is disclosed to third parties (or exploited by the monopolist). This enables us to analyze how a firm's data disclosure/exploitation policy affects a consumer's decision to share information, which has not been considered in the aforementioned papers. In this regard, our model bears much resemblance with that of Casadesus-Masanell and Hervas-Drane (2015), who also distinguish between data provision by consumers and data disclosure by the firm. That said, we differ from them in several ways. First, in our model, the firm derives revenues solely from disclosure, whereas in Casadesus-Masanell and Hervas-Drane (2015), the firm has two revenue streams - disclosure and price - and faces a trade-off between the two. Second, Casadesus-Masanell and Hervas-Drane (2015) focus their analysis on the effect of competition on the disclosure levels chosen by firms while we focus on the regulatory impact of a cap on disclosure level.

An important distinction between our work and the three papers cited above is in how the value of the firm's service depends additionally on an endogenous quality component in our model. Because the firm's only source of revenues comes from data disclosure, its incentives to invest in quality depends on the impact of quality on disclosure revenues. This implies that a regulation which affects the disclosure revenues of the firm will have implications on its investment in quality. The interlinkage between privacy regulation and firms' incentives to innovate has been examined by Goldfarb and Tucker (2012). Drawing on empirical studies in the healthcare (see Tucker and Miller 2009, 2011a and 2011b) and the online advertising (see Goldfarb and Tucker 2011) sectors, they highlight how privacy regulations may raise the costs and/or lower the benefits associated with data-driven innovation, hence creating adverse consequences for such innovation. Our work shows, in addition, that privacy regulation (in the form of a disclosure cap) can affect the level of service innovation (quality) even when data is not a direct input for innovation.

The rest of the paper is structured as follows. In section 2, we introduce the model setup. In section 3, we examine the consumers' participation and information provision decisions. We then compare, in section 4, the privately and socially choice of quality and disclosure levels, before presenting our analysis of the *ex ante* regulation in section 5. In section 6, we present two extensions to our model - the analysis of an *ex ante* regulation with a consumer-surplus-maximizing regulator and that of an *ex post* regulation. In Section 7, we provide an alternative interpretation of our model and discuss the way our results would be modified if two assumptions of our model were relaxed. Section 8 concludes.

2 Setup

Consider a firm that offers a service to a mass of consumers for free, and derives revenues from selling (some of) its customers' personal information to third parties (e.g., advertisers). The firm can choose the quality $q \geq 0$ of its service and a disclosure level $d \in [0,1]$ of personal information. A lower disclosure level can be interpreted either as more restrictions on the type of data that can be sold to third parties or more restrictions on the set of third parties that the data can be sold to.[7]

Consumers' utility. Consumers benefit from a better service when they provide the

[7] Alternatively, the disclosure level could also be interpreted as the degree of *informativeness* of the data. A disclosure level $d < 1$ would then mean that the firm engages in some (potentially imperfect) form of anonymization. This would however require to interpret differently the parameter r capturing the value of information.

firm with (more) personal information, but incur a utility loss from having (part of) that information disclosed to third parties. More specifically, we assume that the utility of a consumer who patronizes the firm and provides it with an amount of information $x \in [0,1]$ is

$$U(x, \theta, q, d) \equiv V(x, q) - (\alpha + \theta d) x - K$$

where α is the cost of providing personal information[8] (which we assume to be the same for all consumers), and θ is an idiosyncratic privacy cost parameter distributed over an interval $\left[\underline{\theta}, \bar{\theta}\right]$ according to a differentiable density function $f(.)$. This parameter captures the intensity of a consumer's preference for privacy; the higher the value of θ, the stronger the consumer's preference for privacy. The parameter $K \geq 0$ is a fixed opportunity cost of using the service, which we assume to be the same for all consumers.[9] Moreover, the gross utility $V(x, q)$ is bounded, and twice continuously differentiable, increasing and concave in both its arguments.

Value of personal information. We suppose that all the buyers of personal data have the same willingness to pay $r > 0$ for a unit of (personal) information, and call r the *value of information*. Moreover, the firm is a monopolist in the market for personal information. These simplifying assumptions have two straightforward implications. First, the firm always sets the unit price for personal information to r independently of its other strategic choices. Second, the price set by the monopolist leaves no surplus to third parties buying information, which simplifies our welfare analysis.[10]

Firm's profit and social welfare. Denoting $\mathbf{x}$ the function mapping each $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$ to the amount of information $x(\theta) \in [0,1]$ provided by a consumer of type θ,[11] the firm's profit is

$$\Pi(\mathbf{x}, q, d) = rd \int_{\underline{\theta}}^{\bar{\theta}} x(\theta) f(\theta) d\theta - C(q)$$

where $C(q)$ is the (fixed) cost of producing a service of quality q. Assume that $C(.)$ is twice differentiable, and such that $C(0) = 0$, $C(q) \underset{q \to +\infty}{\longrightarrow} +\infty$, $C'(0) = 0$, $C'(q) > 0$ for any $q > 0$, and $C''(q) > 0$ for any $q \geq 0$.

Social welfare is defined as the sum of the firm's profit and the consumers' utility.

Interdependency between quality and information. We capture the (local) interdependency between quality and information from consumers' perspective through the following parameter:

$$\gamma(x, q) \equiv -\frac{\frac{\partial^2 V}{\partial x \partial q}}{\frac{\partial^2 V}{\partial x^2}}.$$

For our analysis, we assume that $\gamma(x, q) \geq 0$, which implies that quality and information

[8]This could be for instance the cost of uploading content on a social network or the cost of filling out an electronic form on a website.

[9]For notational convenience we do not include K in the arguments of the utility function.

[10]In Section 7 we discuss the scenario in which third parties buying personal data pay a price lower than r, thus making a positive surplus.

[11]The amount of information $x(\theta)$ can be equal to zero either because the consumer of type θ decides not to use the service, or because she uses the service but decides to provide no personal information at all.

are complements.[12] Further, the larger the value of $\gamma(x,q)$, the stronger the extent of complementarity (in relative terms).

Timing. We consider the following two-stage game:

1. The firm chooses a level of quality q and commits to a disclosure level d.

2. Consumers observe the levels of quality and disclosure. They decide then whether to patronize the firm and, if they do, how much personal information they provide.

3 Consumers' choice

We begin our analysis with the consumers' problem. Having observed the firm's quality and disclosure levels, a consumer has to decide whether or not to patronize the firm and how much information to provide the firm with if she patronizes it.

Let us first examine a consumer's optimal level of information provision. Conditional on patronizing the firm, a consumer chooses the amount of information she provides the firm with so as to maximize her utility $U(x,\theta,q,d)$. Denote[13]

$$\tilde{x}(\theta,q,d) \equiv \underset{x\in[0,1]}{\arg\max} U(x,\theta,q,d).$$

To ease the exposition, we assume throughout the paper that

$$\alpha > \sup_{q\geq 0} \frac{\partial V}{\partial x}(1,q)$$

and

$$\bar{\theta} < \inf_{q\geq 0} \frac{\partial V}{\partial x}(0,q)$$

which ensures that, conditional on using the service, the amount of information that a consumer provides to the firm is always interior (i.e., $x \in (0,1)$).

The following lemma shows the effect of quality and disclosure levels on the amount of personal information provided by consumers.

Lemma 1 *(Comparative statics - Information amount) The amount of information provided by a consumer to the firm is decreasing in the disclosure level and the idiosyncratic privacy cost parameter, and is weakly increasing in the quality level. More precisely,*

$$\frac{\partial \tilde{x}}{\partial d}(\theta,q,d) = \frac{\theta}{\frac{\partial^2 V}{\partial x^2}(\tilde{x}(\theta,q,d),q)} < 0,$$

[12]This is likely to be the relevant case when we consider social media platforms. For instance, Facebook has recently developed and introduced new sharing tools (including Facebook Live video and the "On this day" feature) in an attempt to boost the sharing of original personal content (for full story, visit `https://www.bloomberg.com/news/articles/2016-04-07/facebook-said-to-face-decline-in-people-posting-personal-content`). This provides support for our assumption.

[13]The existence and uniqueness of $\tilde{x}(\theta,q,d)$ follows from the fact that $U(x,\theta,q,d)$ is concave in x over the compact set $[0,1]$.

$$\frac{\partial \tilde{x}}{\partial \theta}(\theta, q, d) = \frac{d}{\frac{\partial^2 V}{\partial x^2}(\tilde{x}(\theta, q, d), q)} < 0,$$

and

$$\frac{\partial \tilde{x}}{\partial q}(\theta, q, d) = \gamma(\tilde{x}(\theta, q, d), q) \geq 0.$$

Proof. See Appendix A. ■

It is intuitive that the amount of information provided by consumers is decreasing in the level of disclosure and in the value of her (idiosyncratic) privacy cost parameter. Both an increase in the level of disclosure and an increase in the privacy cost parameter have the same effect of raising the marginal privacy cost associated with information provision. Thus, for any given level of quality, it is optimal for a consumer to reduce the amount of information she provides. An increase in quality level raises the consumer's marginal gross utility from providing information. Holding the cost of disclosure fixed (i.e., at a given level of disclosure), it is therefore optimal for the consumer to provide more information at higher levels of quality.

We now consider the participation decision of a consumer. Denoting

$$\tilde{U}(\theta, q, d) \equiv U(\tilde{x}(\theta, q, d); \theta, q, d),$$

a consumer of type θ chooses to patronize the firm if and only if[14]

$$\tilde{U}(\theta, q, d) > 0.$$

The following lemma characterizes the demand for the service offered by the firm and shows how it is affected by the levels of quality and disclosure.

Lemma 2 *(Comparative statics - Demand). There exists $\tilde{\theta}(q, d) \in [\underline{\theta}, \overline{\theta}]$ such that a consumer patronizes the firm if and only if*

$$\theta < \tilde{\theta}(q, d).$$

Moreover, $\tilde{\theta}(q, d)$ is weakly increasing in the level quality, and weakly decreasing in the disclosure level. More precisely, the following expressions hold whenever $\tilde{\theta}(q, d) \in (\underline{\theta}, \overline{\theta})$:

$$\frac{\partial \tilde{\theta}}{\partial q}(q, d) = \frac{\frac{\partial V}{\partial q}\left(\tilde{x}\left(\tilde{\theta}(q, d), q, d\right), q\right)}{d\tilde{x}\left(\tilde{\theta}(q, d), q, d\right)} > 0$$

and

$$\frac{\partial \tilde{\theta}}{\partial d}(q, d) = -\frac{\tilde{\theta}(q, d)}{d} < 0.$$

Proof. See Appendix A. ■

The above lemma tells us that consumers with (sufficiently) weak preference for privacy patronize the firm while those with strong preference for privacy stay out of the market.

[14] For technical reasons, we assume that a consumer who is indifferent between patronizing the firm does not use the service.

Conditional on participation, consumers with stronger preference for privacy (higher θ) incur higher privacy costs and derive lower gross utility from using the firm's service. This implies that the consumer's utility from participation is decreasing in θ. Therefore, consumers with (sufficiently) strong preference for privacy find it optimal not to patronize the firm. It is straightforward to see why the demand (or the level of participation) is increasing in the firm's quality level but decreasing in its disclosure level. The overall utility from patronizing the firm is increasing in its quality level (higher gross utility level of the service) and decreasing in its disclosure level (higher privacy costs). Since a consumer only patronizes a firm when she obtains a sufficiently high level of utility, the firm's demand is increasing in its quality level but decreasing in its disclosure level.

4 Private versus social incentives

We next examine the privately and socially optimal choice of quality and disclosure levels. Let us first consider the private incentives to invest in quality and to disclose personal information. The firm's profit when consumers make their participation and information provision decision optimally is

$$\tilde{\Pi}(q,d) = rd \int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \tilde{x}(\theta,q,d) f(\theta) \, d\theta - C(q). \tag{1}$$

From (1) it follows that the firm's marginal net benefit from investing in quality is

$$\frac{\partial \tilde{\Pi}}{\partial q} = \underbrace{rd \int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \frac{\partial \tilde{x}}{\partial q}(\theta,q,d) f(\theta) \, d\theta}_{\text{intensive margin effect}} + \underbrace{rd \frac{\partial \tilde{\theta}}{\partial q} \tilde{x}\left(\tilde{\theta}(q,d),q,d\right) f\left(\tilde{\theta}(q,d)\right)}_{\text{extensive margin effect}} - C'(q). \tag{2}$$

Lemma 1 tells us that the sign of the intensive margin effect is positive and Lemma 2 shows that the extensive margin effect is weakly positive. The intensive margin effect here captures the impact of a change in quality level on the firm's revenue via the change in the amount of information provided by the consumers. A higher level of quality induces consumers to provide more information, which increases the firm's disclosure revenues. Therefore, the intensive margin effect is positive. The extensive margin effect captures the impact of the change in demand resulting from a change in quality on the firm's profit. Since the consumers' utility is always (weakly) increasing in the level of quality (all else equal), the firm's demand is also always (weakly) increasing.

The marginal net benefit from increasing the disclosure level is

$$\frac{\partial\tilde{\Pi}}{\partial d} = \underbrace{r\int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \tilde{x}(\theta,q,d) f(\theta)\,d\theta + rd\int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \frac{\partial\tilde{x}}{\partial d}(\theta,q,d) f(\theta)\,d\theta}_{\text{effect on the intensive margin}}$$
$$+\underbrace{rd\frac{\partial\tilde{\theta}}{\partial d}\tilde{x}\left(\tilde{\theta}(q,d),q,d\right) f\left(\tilde{\theta}(q,d)\right)}_{\text{effect on the extensive margin}}.$$

Lemma 1 implies that the sign of the intensive margin effect is ambiguous, while Lemma 2 shows that the extensive margin effect is weakly negative. The sign of the intensive margin effect is ambiguous because a change in the disclosure level generates two opposing effects. An increase in the level of disclosure raises the firm's disclosure revenues per unit of information provided by consumers but reduces the amount of information provided by the consumers at the same time.

Let us assume that $\tilde{\Pi}(.,.)$ is concave in both its arguments.[15] The privately optimal level of quality for a given level of disclosure and the privately optimal level of disclosure for a given level of quality are defined as follows:[16]

$$q^M(d) \equiv \arg\max_{q\geq 0} \tilde{\Pi}(q,d),$$

$$d^M(q) \equiv \arg\max_{d\in[0,1]} \tilde{\Pi}(q,d).$$

The privately optimal levels of quality and disclosure is given by

$$\left(\tilde{q}^M, \tilde{d}^M\right) = \arg\max_{q\geq 0,\ d\in[0,1]} \tilde{\Pi}(q,d).$$

Let us now consider the social incentives for quality and information disclosure. The social planner's objective function is given by

$$\tilde{W}(q,d) \equiv \tilde{\Pi}(q,d) + \widetilde{CS}(q,d)$$

where

$$\widetilde{CS}(q,d) = \int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \tilde{U}(\theta,q,d) f(\theta)\,d\theta.$$

[15]For a large part of the subsequent analysis we only need $\tilde{\Pi}(.,.)$ to be concave in *each* of its arguments.

[16]The existence and uniqueness of $q^M(d)$ follows from the fact that $\tilde{\Pi}(q,d)$ is continuous and concave in q and $\tilde{\Pi}(q,d) \underset{q\longrightarrow+\infty}{\longrightarrow} -\infty$, while the existence and uniqueness of $d^M(q)$ follows from the fact that $\tilde{\Pi}(q,d)$ is continuous and concave in d over the compact set $[0,1]$.

Hence, the social marginal net benefit from investing in quality is

$$\begin{aligned} \frac{\partial \tilde{W}}{\partial q} &= \frac{\partial \tilde{\Pi}}{\partial q} + \int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \frac{\partial \tilde{U}}{\partial q} (\theta, q, d) f(\theta) d\theta + \frac{\partial \tilde{\theta}}{\partial q} \tilde{U} \left(\tilde{\theta}(q,d), q, d \right) f \left(\tilde{\theta}(q,d) \right) \\ &= \frac{\partial \tilde{\Pi}}{\partial q} + \underbrace{\int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \frac{\partial V}{\partial q} \left(\tilde{x}(\theta, q, d), q \right) f(\theta) d\theta}_{>0} + \underbrace{\frac{\partial \tilde{\theta}}{\partial q} \tilde{U} \left(\tilde{\theta}(q,d), q, d \right) f \left(\tilde{\theta}(q,d) \right)}_{=0 \text{ by definition of } \tilde{\theta}(q,d)}. \end{aligned} \tag{3}$$

As compared to the firm's marginal benefit of investing in quality, the social marginal benefit contains an additional term that captures the positive effect of an increase in quality on consumers. This implies that the social marginal net benefit from investing in quality is greater than the corresponding private benefit.

Likewise, the social marginal net benefit from information disclosure is

$$\begin{aligned} \frac{\partial \tilde{W}}{\partial d} &= \frac{\partial \tilde{\Pi}}{\partial d} + \int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \frac{\partial \tilde{U}}{\partial d} (\theta, q, d) f(\theta) d\theta + \frac{\partial \tilde{\theta}}{\partial d} \tilde{U} \left(\tilde{\theta}(q,d), q, d \right) f \left(\tilde{\theta}(q,d) \right) \\ &= \frac{\partial \tilde{\Pi}}{\partial d} \underbrace{- \int_{\underline{\theta}}^{\tilde{\theta}(q,d)} \theta \tilde{x}(\theta, q, d) f(\theta) d\theta}_{<0} + \underbrace{\frac{\partial \tilde{\theta}}{\partial d} \tilde{U} \left(\tilde{\theta}(q,d), q, d \right) f \left(\tilde{\theta}(q,d) \right)}_{=0}. \end{aligned}$$

It is lower than the corresponding private marginal benefit of disclosure because the social planner takes into account the negative effect of disclosure on the consumers' utility.

Similarly, we assume that $\tilde{W}(.,.)$ is concave in each of its arguments. The socially optimal quality level for a given level of disclosure and the socially optimal disclosure level for a given level of quality are defined as follows:[17]

$$q^W(d) \equiv \arg\max_{q \geq 0} \tilde{W}(q,d)$$

$$d^W(q) \equiv \arg\max_{d \in [0,1]} \tilde{W}(q,d).$$

We now compare the socially and privately optimal level of quality for a given level of disclosure and the socially and privately optimal level of disclosure for a given level of quality. The comparison of $q^M(d)$ and $q^W(d)$ is useful for the analysis of the *ex ante* regulation of information disclosure that we consider in the next section, while the comparison of $d^M(q)$ and $d^W(q)$ is key to understanding the effect of an *ex post* regulation of disclosure in the corresponding extension.[18] The two aforementioned comparisons are provided in the

[17] The existence and uniqueness of $q^W(d)$ follows from the fact that $\tilde{W}(q,d)$ is continuous and concave in q and $\tilde{W}(q,d) \underset{q \longrightarrow +\infty}{\longrightarrow} -\infty$, while the existence and uniqueness of $d^W(q)$ follows from the fact that $\tilde{W}(q,d)$ is continuous and concave in d over the compact set $[0,1]$.

[18] The comparison of the socially and privately optimal pair of quality and disclosure levels is not necessary for the policy analysis and is therefore omitted.

following two lemmas. In both cases, the results stem from the fact that the monopolist does not internalize the effects of its choices on consumers.

Lemma 3 *(Under-provision of quality) For a given disclosure level, the monopolist under-provides quality from a social welfare perspective:* $q^M(d) \leq q^W(d)$.

Proof. See Appendix A. ■

Lemma 4 *(Over-disclosure of information) For a given quality level, the monopolist over-discloses information from a social welfare perspective:* $d^M(q) \geq d^W(q)$.

Proof. See Appendix A. ■

5 *Ex ante* regulation

In this section we investigate the social desirability of a policy whereby an authority regulates the information disclosure level *ex ante* (i.e., before investment in quality is decided by the firm). More specifically, we study the decision of a social-welfare-maximizing regulator whose only instrument is a cap on the disclosure level.

Consider the following game:

- First, the regulator decides whether to impose a cap for the disclosure level, and sets the value of that cap $\bar{d}$ if it does so.

- Second, the firm decides on its disclosure and quality levels.

- Third, consumers decide whether to patronize the firm and how much information to provide if they do.

Let us first analyze the firm's behavior for a given regulator's choice. The firm's optimal disclosure level maximizes $\tilde{\Pi}\left(q^M(d), d\right)$ subject to the constraint $d \leq \bar{d}$. If $\bar{d} \geq \tilde{d}^M$, then the constraint is not binding, which means that the firm's decision will be the same as in the unregulated scenario. If $\bar{d} < \tilde{d}^M$, however, the constraint is binding. From the concavity of $\tilde{\Pi}\left(q^M(d), d\right)$ with respect to d, it then follows that the firm will choose $d = \bar{d}$.
In the first stage, the regulator seeks to maximize

$$\hat{W}(d) \equiv \tilde{W}\left(q^M(d), d\right) = \tilde{\Pi}\left(q^M(d), d\right) + \int_{\underline{\theta}}^{\tilde{\theta}\left(q^M(d), d\right)} \tilde{U}\left(\theta, q^M(d), d\right) f(\theta)\, d\theta.$$

Assume that $\hat{W}(.)$ is concave and denote

$$\hat{d}^W \equiv \underset{d \in [0,1]}{\arg\max} \hat{W}(d).$$

Given the firm's behavior (and the concavity of $\hat{W}(.)$), the regulator finds it optimal to set a cap (and the optimal cap is then $\bar{d} = \hat{d}^W$) if and only if $\hat{d}^W \leq \tilde{d}^M$. This is equivalent to

$$\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M} \leq 0$$

whenever $\hat{d}^W < 1$, which we assume to hold hereafter. Using the Envelope Theorem, we have

$$\begin{aligned}
\frac{\partial \hat{W}}{\partial d} &= \frac{\partial \tilde{\Pi}}{\partial d}\left(q^M(d), d\right) + \int_{\underline{\theta}}^{\tilde{\theta}\left(q^M(d),d\right)} \left[\frac{\partial \tilde{U}}{\partial q}\left(\theta, q^M(d), d\right)\frac{\partial q^M}{\partial d} + \frac{\partial \tilde{U}}{\partial d}\left(\theta, q^M(d), d\right)\right] f(\theta)\, d\theta \\
&\quad + \left[\frac{\partial \tilde{\theta}}{\partial q}\frac{\partial q^M}{\partial d} + \frac{\partial \tilde{\theta}}{\partial d}\right] \underbrace{\tilde{U}\left(\tilde{\theta}\left(q^M(d), d\right), q^M(d), d\right)}_{=0} f\left(\tilde{\theta}(q^M(d), d)\right) \\
&= \frac{\partial \tilde{\Pi}}{\partial d}\left(q^M(d), d\right) + \int_{\underline{\theta}}^{\tilde{\theta}\left(q^M(d),d\right)} \left[\frac{\partial V}{\partial q}\left(\tilde{x}\left(\theta, q^M(d), d\right), q^M(d)\right)\frac{\partial q^M}{\partial d} - \theta \tilde{x}\left(\theta, q^M(d), d\right)\right] f(\theta)\, d\theta.
\end{aligned}$$

Evaluating this at $d = \tilde{d}^M$ and using the fact that $\tilde{q}^M = q^M(\tilde{d}^M)$ yields

$$\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M} = \underbrace{\frac{\partial \tilde{\Pi}}{\partial d}\left(\tilde{q}^M, \tilde{d}^M\right)}_{=0} + \int_{\underline{\theta}}^{\tilde{\theta}\left(\tilde{q}^M,\tilde{d}^M\right)} \left[\frac{\partial V}{\partial q}\left(\tilde{x}\left(\theta, \tilde{q}^M, \tilde{d}^M\right), \tilde{q}^M\right)\left.\frac{\partial q^M}{\partial d}\right|_{d=\tilde{d}^M} - \theta \tilde{x}\left(\theta, \tilde{q}^M, \tilde{d}^M\right)\right] f(\theta)\, d\theta$$

or, equivalently,

$$-\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M} = \underbrace{\int_{\underline{\theta}}^{\tilde{\theta}\left(\tilde{q}^M,\tilde{d}^M\right)} \theta \tilde{x}\left(\theta, \tilde{q}^M, \tilde{d}^M\right) f(\theta)\, d\theta}_{\text{direct effect}} - \underbrace{\int_{\underline{\theta}}^{\tilde{\theta}\left(\tilde{q}^M,\tilde{d}^M\right)} \frac{\partial V}{\partial q}\left(\tilde{x}\left(\theta, \tilde{q}^M, \tilde{d}^M\right), \tilde{q}^M\right)\left.\frac{\partial q^M}{\partial d}\right|_{d=\tilde{d}^M} f(\theta)\, d\theta.}_{\text{indirect effect}}$$

This shows that a (marginal) decrease in the disclosure level starting from the unregulated level has two effects: a *direct* effect on the privacy costs incurred by consumers (keeping the quality level constant), and an *indirect* effect capturing how a decrease in the disclosure level alters the firm's quality choice. The direct effect is always positive because there is over-disclosure by the firm from a social perspective, while the sign of the indirect effect depends on whether the firm's quality choice increases or decreases in response to the reduction in disclosure level. Since the firm under-provides quality from a social perspective,social welfare is increasing in quality level at the unregulated equilibrium. Therefore, the indirect effect is weakly positive if the firm weakly increases its quality level when disclosure level is decreased; i.e., if

$$\left.\frac{\partial q^M}{\partial d}\right|_{d=\tilde{d}^M} \leq 0.$$

In this case, the overall effect of setting a disclosure cap on social welfare is unambiguously positive. However, if

$$\left.\frac{\partial q^M}{\partial d}\right|_{d=\tilde{d}^M} > 0$$

then the indirect effect is negative and, therefore, the overall effect of a disclosure cap is *a priori* ambiguous. The following lemma relates the effect of a change in disclosure level on the firm's optimal quality level to the cross-effect of quality and disclosure on the firm's profit.

Lemma 5 *(Effect of the disclosure level on quality) If $q^M(d) \neq 0$, then*

$$\frac{\partial q^M}{\partial d} = -\frac{\frac{\partial^2 \tilde{\Pi}}{\partial q \partial d}\left(q^M(d), d\right)}{\frac{\partial^2 \tilde{\Pi}}{\partial q^2}\left(q^M(d), d\right)}.$$

Proof. See Appendix A. ■

From Lemma 5, we see that the effect of a change in disclosure level on the firm's choice of quality has the same sign as the cross-effect of quality and disclosure on the firm's profit. The intuition behind this result is straightforward. If the marginal benefit of investing in quality is increasing in the level of disclosure, the firm will invest more at higher levels of disclosures.

We now study the sign of the effect of a change in disclosure level on quality, which in turn determines the sign of the indirect effect of a disclosure cap. To simplify the analysis, we assume that the function $\gamma(x, q)$, that captures the complementarity between information and quality, is constant:

$$\gamma(x, q) = \gamma$$

for any $(x, q) \in [0, 1] \times [0, +\infty)$.[19]

We first focus on the scenario in which the market is fully covered for any levels of quality and disclosure (and hence, there is no extensive margin effect from a change in quality level). We then extend the analysis to the scenario in which the market is not fully covered.

5.1 Full market coverage

Suppose that $K < V(0, 0)$. Under this assumption, the consumer's utility level when she is providing information optimally, $\tilde{U}(\theta, q, d)$, is strictly positive for all types $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$ at any given levels of quality and disclosure.[20] Therefore, the market is fully covered for all levels of quality and disclosure.[21]

Consider the firm's optimal choice of quality for a given level of disclosure. The firm's marginal benefit from investing in quality is given by (2), with the term capturing the effect of disclosure on the extensive margin effect equal to zero (due to full market coverage).

[19]This amounts to restricting our attention to the class of gross utility functions $V(x, q)$ for which there exists a real number γ, a twice continuously differentiable function $g(.)$, and a continuously differentiable function $h(.)$ such that

$$V(x, q) = g(q) + \int_0^x h(u - \gamma q) du.$$

[20]This follows from the fact that $\tilde{U}(\theta, q, d) \geq V(0, 0) - K$ for any $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$ and any levels of quality and disclosure.

[21]The assumption that $K < V(0, 0)$ is sufficient but not necessary for a large part of our analysis. For many of results that we derive, it suffices that the unregulated market (i.e., when $(q, d) = (\tilde{q}^M, \tilde{d}^M)$) is fully covered. The necessary condition for full coverage at $(\tilde{q}^M, \tilde{d}^M)$ is $K < V(\tilde{x}(\bar{\theta}, \tilde{q}^M, \tilde{d}^M), \tilde{q}^M) - \bar{\theta}\tilde{d}^M \tilde{x}(\bar{\theta}, \tilde{q}^M, \tilde{d}^M)$.

Substituting $\frac{\partial \tilde{x}}{\partial q}$ by its expression in Lemma 1, we obtain

$$\frac{\partial \tilde{\Pi}}{\partial q} = rd\gamma - C'(q).$$

The marginal benefit of investing in quality is weakly positive when evaluated at $q = 0$:

$$\left.\frac{\partial \tilde{\Pi}}{\partial q}\right|_{q=0} = rd\gamma - \underbrace{C'(0)}_{=0} \geq 0.$$

This implies that the firm's optimal choice of quality level for a given level of disclosure, $q^M(d)$, satisfies the first-order condition

$$\frac{\partial \tilde{\Pi}}{\partial q}(q^M(d), d) = 0$$

for all levels of disclosure. From Lemma 5, we know that the sign of the effect of a change in disclosure level on the firm's choice of quality level is the same as the sign of the cross-effect of quality and disclosure on the firm's profit. Under full market coverage,

$$\frac{\partial^2 \tilde{\Pi}}{\partial q \partial d}\left(q^M(d), d\right) = r\gamma,$$

which is weakly positive since $\gamma \geq 0$. Therefore, the firm's optimal level of quality is weakly increasing in the level of disclosure. The next lemma summarizes the analysis above.

Lemma 6 *(Effect of the disclosure level on quality under full market coverage) The firm's optimal choice of quality for a given level of disclosure, $q^M(d)$, is weakly positive and weakly increasing in the level of disclosure; more precisely, $\frac{\partial q^M}{\partial d} = \frac{r\gamma}{C''(q^M(d))}$ for all $d \in [0,1]$.*

The intuition behind the above lemma is as follows. The marginal benefit from investing in quality arises from the complementarity between quality and information. Consumers provide more information when the firm's service is of better quality; consequently, the firm obtains higher revenues from disclosure. Further, the firm benefits more from this increase in information provision when it is disclosing a larger share of consumer information (i.e., at higher disclosure levels). The firm's investment in quality is therefore increasing in the level of disclosure.

One direct implication of Lemma 6 is that a disclosure cap would lower the level of quality chosen by the firm. This means that there exists a trade-off between privacy and quality (at the unregulated equilibrium) — an increase in the level of privacy (lower d) results in decrease in level of quality. Moreover, observe from Lemma 6 that this reduction in quality level is larger the stronger the complementarity between quality and information provision.

Consider now the regulator's problem. Recall from the preceding discussion that it is optimal for the regulator to set a disclosure cap if and only if $\hat{d}^W < \tilde{d}^M$, or equivalently, $-\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M} > 0$. From Lemma 6, we have $\left.\frac{\partial q^M}{\partial d}\right|_{d=\tilde{d}^M} = \frac{r\gamma}{C''(\tilde{q}^M)}$. This implies that the social marginal benefit of decreasing the disclosure level, evaluated at $d = \tilde{d}^M$, can thus be expressed as follows:

$$-\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M} = \int_{\underline{\theta}}^{\bar{\theta}} \theta \tilde{x}(\theta, \tilde{q}^M, \tilde{d}^M) f(\theta) d\theta - \frac{r\gamma}{C''(\tilde{q}^M)} \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial V}{\partial q}(\tilde{x}(\theta, \tilde{q}^M, \tilde{d}^M), \tilde{q}^M) f(\theta) d\theta.$$

Since $\tilde{q}^M$, $\tilde{d}^M$ and $\tilde{x}(\theta, ., .)$ also depend on γ, it is unclear how the expression above depends on γ. However, denoting

$$\bar{\gamma}(r) = \sup \left\{ \gamma' \geq 0 | \left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M} \leq 0 \text{ for all } \gamma \leq \gamma' \right\},$$

we obtain that a cap on the disclosure level is socially desirable when $\gamma \leq \bar{\gamma}(r)$. Using the fact that $\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M}$ is continuous in γ and (strictly) negative for $\gamma = 0$, we have $\bar{\gamma}(r) > 0$. Thus, a cap on the disclosure level is socially desirable whenever the complementarity between quality and information is not too strong. When quality and information are sufficiently strong complements (i.e., $\gamma > \bar{\gamma}(r)$) , the sign of $-\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\tilde{d}^M}$ is ambiguous and, therefore, so is the impact of a disclosure cap on social welfare.[22] Intuitively, the stronger the level of complementarity between quality and information, the larger the trade-off between privacy and quality. The decrease in social welfare due to the reduction in quality (the indirect effect) may offset the increase in welfare resulting from the reduction in privacy costs (the direct effect). Consequently, the impact of the cap on social welfare may be ambiguous. The following proposition summarizes the above discussion.

Proposition 1 *(Social desirability of an ex ante disclosure cap under full market coverage) When the market is fully covered, an ex ante regulation taking the form of a disclosure cap is socially desirable if quality and information are sufficiently weak complements (i.e., $\gamma \leq \bar{\gamma}(r)$), and has an ambiguous effect on social welfare otherwise.*

We now examine the level of the disclosure cap chosen by the regulator, provided that a cap is indeed socially desirable. The regulator finds it optimal to set a disclosure cap at zero whenever the marginal social benefit of disclosure is weakly negative when evaluated at $d = 0$. We have that

$$\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=0} = \underbrace{\left.\frac{\partial \tilde{\Pi}}{\partial d}\left(q^M(d), d\right)\right|_{d=0}}_{>0} + \int_{\underline{\theta}}^{\bar{\theta}} \left[\frac{\partial \tilde{U}}{\partial q}\left(\theta, q^M(0), 0\right) \left.\frac{\partial q^M}{\partial d}\right|_{d=0} + \underbrace{\frac{\partial \tilde{U}}{\partial d}\left(\theta, q^M(0), 0\right)}_{<0} \right] f(\theta)\, d\theta.$$

Using the results from Lemma 6 and the fact that $q^M(0) = 0$,

$$\left.\frac{\partial \hat{W}}{\partial d}\right|_{d=0} = r \int_{\underline{\theta}}^{\bar{\theta}} \tilde{x}(\theta, 0, 0) f(\theta) d\theta + \frac{r\gamma}{C''(0)} \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial V}{\partial q}(\tilde{x}(\theta, 0, 0), 0) f(\theta) d\theta - \int_{\underline{\theta}}^{\bar{\theta}} \theta \tilde{x}(\theta, 0, 0) f(\theta) d\theta.$$

[22]Note that it is *a priori* possible that $\bar{\gamma}(r)$ takes an infinite value. In this case, *ex ante* regulation would be socially desirable whatever the level of complementarity between quality and information.

Since the second term in the above expression is always positive, the marginal social (net) benefit of disclosure is positive at $d = 0$ (and therefore it is socially optimal to allow for a positive level of disclosure) when r exceeds the "welfare-neutral" value of information at $(q, d) = (0, 0)$,[23] which is given by

$$\hat{r}(0,0) = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \theta \tilde{x}(\theta, 0, 0) f(\theta) d\theta}{\int_{\underline{\theta}}^{\bar{\theta}} \tilde{x}(\theta, 0, 0) f(\theta) d\theta}.$$

When $r \leq \hat{r}(0,0)$, the regulator only finds it optimal to set a positive level of disclosure if γ is sufficiently high, such that the marginal social benefit of disclosure at $d = 0$ is positive; i.e.,[24]

$$\gamma > \frac{C''(0)}{r} \frac{\int_{\underline{\theta}}^{\bar{\theta}} (\theta - r) \tilde{x}(\theta, 0, 0) f(\theta) d\theta}{\int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial V}{\partial q}(\tilde{x}(\theta, 0, 0), 0) f(\theta) d\theta} = \underline{\gamma}(r).$$

Figure 1 provides a graphical illustration of the above discussion.

The following proposition sums up our results on the socially optimal value of the disclosure cap.

Proposition 2 *(Socially optimal disclosure cap under full market coverage) When the market is fully covered, the socially optimal value of the disclosure cap (when a cap is desirable) is*

- always positive when $r > \hat{r}(0,0)$;

[23] Unlike in traditional welfare analysis where the price of product/service represents a (welfare-neutral) transfer of surplus from the consumers to the firm, the "price" (i.e., the privacy costs) paid by the consumers and the "price" received by the firm (i.e., r) in our model may not coincide. The social welfare in our model can be expressed as follows:

$$\tilde{W}(q,d) = \int_{\underline{\theta}}^{\bar{\theta}} V(\tilde{x}(\theta, q, d), q) f(\theta) d\theta - C(q) + d \int_{\underline{\theta}}^{\bar{\theta}} [(r - \theta) \tilde{x}(\theta, q, d)] f(\theta) d\theta,$$

where the third term in the expression captures the potential non-neutrality of the transfer of surplus between the consumers and the firm. For any given $(q, d) \in [0, +\infty) \times [0, 1]$, the value of information r is "welfare-neutral" if it sets this term to zero. Let $\hat{r}(q, d)$ denote the "welfare-neutral" value of r. It is given by

$$\hat{r}(q,d) = \frac{\int_{\underline{\theta}}^{\bar{\theta}} [\theta \tilde{x}(\theta, q, d)] f(\theta) d\theta}{\int_{\underline{\theta}}^{\bar{\theta}} \tilde{x}(\theta, q, d) f(\theta) d\theta}.$$

[24] Note that the RHS of this inequality does not depend on γ. To see why, recall from footnote 5 that the gross utility functions $V(.,.)$ for which $\gamma(.,.)$ is constant are such that $V(x, q) = g(q) + \int_0^x h(u - \gamma q) du$. It is then straightforward to show that $\tilde{x}(\theta, q, d) = \gamma q + h^{-1}(\theta d)$and, in particular, that $\tilde{x}(\theta, 0, 0) = h^{-1}(0)$. Moreover, $\frac{\partial V}{\partial q}(x, q) = g'(q) - \gamma \int_0^x h'(u - \gamma q) du$, which implies that $\frac{\partial V}{\partial q}(\tilde{x}(\theta, 0, 0), 0) = g'(0) - \gamma h(h^{-1}(0)) = g'(0)$.

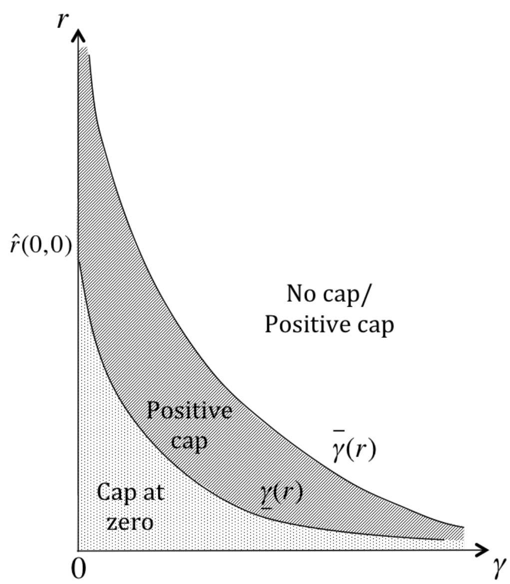

Figure 1: Optimal cap on the disclosure level under full market coverage

- *positive if and only if* $\gamma > \underline{\gamma}(r)$ *when* $r \leq \hat{r}(0,0)$*;*
- *zero otherwise.*

5.2 Partial market coverage

We now consider the case where the demand for the service is positive but the market is not fully covered when disclosure is not regulated, i.e., $\underline{\theta} < \tilde{\theta}\left(\tilde{q}^M, \tilde{d}^M\right) < \bar{\theta}$.

From expression (2) and Lemmas 1 and 2, it follows that the firm's marginal (net) benefit from investing in quality is

$$\frac{\partial \tilde{\Pi}}{\partial q} = \underbrace{rd\gamma F\left(\tilde{\theta}\left(q,d\right)\right)}_{\text{intensive margin effect}} + \underbrace{r\frac{\partial V}{\partial q}\left(\tilde{x}\left(\tilde{\theta}\left(q,d\right),q,d\right),q\right) f\left(\tilde{\theta}\left(q,d\right)\right)}_{\text{extensive margin effect}} - C'(q), \tag{4}$$

where $F(.)$ is the cumulative distribution function of θ. This implies in particular that

$$\left.\frac{\partial \tilde{\Pi}}{\partial q}\right|_{q=0} = \underbrace{rd\gamma F\left(\tilde{\theta}\left(0,d\right)\right)}_{\geq 0} + \underbrace{r\frac{\partial V}{\partial q}\left(\tilde{x}\left(\tilde{\theta}\left(0,d\right),q,d\right),q\right) f\left(\tilde{\theta}\left(0,d\right)\right)}_{>0} - \underbrace{C'(0)}_{=0} > 0.$$

Let us now consider the effect of the disclosure level on the firm's optimal choice of quality, which determines the sign of the indirect effect of a disclosure cap. Differentiating expression (4) with respect to d yields[25]

$$\begin{aligned}\frac{\partial^2\tilde{\Pi}}{\partial q\partial d} &= r\gamma F\left(\tilde{\theta}(q,d)\right) + rd\gamma\frac{\partial\tilde{\theta}}{\partial d}f\left(\tilde{\theta}(q,d)\right) + r\frac{\partial V}{\partial q}\frac{\partial\tilde{\theta}}{\partial d}f'\left(\tilde{\theta}(q,d)\right) \\ &\quad + rf\left(\tilde{\theta}(q,d)\right)\frac{\partial^2 V}{\partial q\partial x}\left[\frac{\partial\tilde{x}}{\partial\theta}\frac{\partial\tilde{\theta}}{\partial d} + \frac{\partial\tilde{x}}{\partial d}\right].\end{aligned}$$

From the expressions of $\frac{\partial\tilde{x}}{\partial\theta}$, $\frac{\partial\tilde{x}}{\partial d}$ and $\frac{\partial\tilde{\theta}}{\partial d}$ provided by Lemmas 1 and 2, it follows that

$$\frac{\partial\tilde{x}}{\partial\theta}\frac{\partial\tilde{\theta}}{\partial d} + \frac{\partial\tilde{x}}{\partial d} = 0$$

and, therefore

$$\frac{\partial^2\tilde{\Pi}}{\partial q\partial d} = \underbrace{r\gamma F\left(\tilde{\theta}(q,d)\right)}_{\equiv A} + \underbrace{rd\gamma\frac{\partial\tilde{\theta}}{\partial d}f\left(\tilde{\theta}(q,d)\right)}_{\equiv B} + \underbrace{r\frac{\partial V}{\partial q}\frac{\partial\tilde{\theta}}{\partial d}f'\left(\tilde{\theta}(q,d)\right)}_{\equiv C}. \tag{5}$$

Term $A+B$ shows how the intensive margin effect of investment in quality depends on the disclosure level. More specifically, term A captures the effect of an increase in the disclosure level on the marginal benefit from investing in quality for a *given* demand for the service. This is the only term that appears in our analysis under full market coverage (where $\tilde{\theta}(q,d) = \bar{\theta}$), and its sign is positive whenever $\gamma \neq 0$. Under partial market coverage, the magnitude of the intensive margin effect is also affected by the drop in demand resulting from an increase in the disclosure level. This effect is captured by term B which is negative whenever $\gamma \neq 0$. Finally, term C shows how the extensive margin effect of investment in quality depends on the disclosure level. Since the extensive margin effect is proportional to the density of consumers at the margin, and because an increase in the disclosure level leads to a decrease in demand, term C is positive (negative) if the density function is locally decreasing (increasing). Using Lemmas 1 and 2 again, we can rewrite (5) as

$$\frac{\partial^2\tilde{\Pi}}{\partial q\partial d} = \underbrace{r\gamma F\left(\tilde{\theta}(q,d)\right)\left[1 - \frac{\tilde{\theta}(q,d)\,f\left(\tilde{\theta}(q,d)\right)}{F\left(\tilde{\theta}(q,d)\right)}\right]}_{A+B}\underbrace{-r\frac{\partial V}{\partial q}\frac{\tilde{\theta}(q,d)}{d}f'\left(\tilde{\theta}(q,d)\right)}_{C}. \tag{6}$$

This shows that the impact of the disclosure level on the intensive margin effect of investment in quality depends on the elasticity of the cumulative distribution $F(.)$, while its impact on the extensive margin effect depends on the convexity/concavity of $F(.)$. More precisely, we have the following result:

Lemma 7 *(Impact of the disclosure level on the intensive and extensive margin effects)*

- If $F(.)$ is relatively inelastic (elastic), i.e., $\frac{\theta f(\theta)}{F(\theta)} < 1$ (> 1), then the impact of the disclosure level on the intensive margin effect of investment in quality is positive (negative).

[25]For notational convenience we drop the arguments of $\frac{\partial V}{\partial q}\left(q, \tilde{x}\left(\tilde{\theta}(q,d), q, d\right)\right)$ and $\frac{\partial^2 V}{\partial q\partial x}\left(q, \tilde{x}\left(\tilde{\theta}(q,d), q, d\right)\right)$.

- *If $F(.)$ is convex (concave), then the impact of the disclosure level on the extensive margin effect of investment in quality is negative (positive).*

The elasticity of $F(.)$ can be related to the shape of the demand for the service, while the second derivative of $F(.)$ can be related to the shape of the derivative of demand with respect to quality. More precisely, in Appendix B, we show that the elasticity of $F(.)$ is equal to the elasticity of the demand for the service with respect to the disclosure level, holding the amount of information constant. We also show that the convexity/concavity of $F(.)$ can be related to the elasticity of the marginal effect of quality on demand with respect to the disclosure level (again holding the amount of information constant): the latter is greater (less) than 1 if $F(.)$ is convex (concave).

Lemma 7 suggests that we should distinguish between four scenarios. However, there are only three possible scenarios because the elasticity of $F(.)$ is always greater than 1 if $F(.)$ is convex. To see why, notice that

$$\frac{\theta f(\theta)}{F(\theta)} = 1 + \frac{\underline{\theta} f(\theta) - \int_{\underline{\theta}}^{\theta} [f(u) - f(\theta)]\, du}{F(\theta)} > 1$$

if $f(.)$ is increasing. Before stating the main result of this section, let us consider two special cases for which we can sign easily the (overall) effect of the disclosure level on the marginal effect of investment in quality. The first one is when the distribution of the privacy cost parameter is uniform (i.e., $F(.)$ is linear). In that case, the disclosure level has no impact on the extensive margin effect, and the intensive margin effect increases in the disclosure level whenever $\gamma \neq 0$. The second scenario of interest is the limiting case in which quality and information are independent (i.e., $\gamma = 0$). In that case, the intensive margin effect is zero whatever the disclosure level, and the extensive margin effect is increasing (decreasing) in the disclosure level if $F(.)$ is concave (convex).

Using Lemmas 5 and 7 and the observations above, we get the following results in the (interesting) scenario where $\tilde{q}^M$ is positive.[26] First, if $F(.)$ is (weakly) convex then the effect of the disclosure level on quality is negative (whenever $\gamma \neq 0$). This implies that, in this case, the regulator can induce a higher quality effect by setting a disclosure cap. Second, if $F(.)$ is concave and relatively inelastic then the effect of the disclosure level on quality is positive. Third, if $F(.)$ is concave and relatively inelastic, then the effect of the disclosure level on quality is positive if quality and information are weak complements (by continuity at $\gamma = 0$), and is ambiguous otherwise.

The following proposition sums up the above discussion.

Proposition 3 *(Effect of an ex ante disclosure cap on quality under partial market coverage) Assume that the market is partially covered and quality is positive when disclosure is not regulated.*

[26] If the firm does not invest at all in quality when the disclosure level is not regulated, i.e., $\tilde{q}^M = 0$, then the effect of a disclosure cap on quality can only be weakly positive, which implies that the effect of a disclosure cap on social welfare is positive.

- If $F(.)$ is (weakly) convex then the effect of a disclosure cap on quality is positive (whenever $\gamma \neq 0$).

- If $F(.)$ is concave and relatively inelastic then the effect of a disclosure cap on quality is negative.

- If $F(.)$ is concave and relatively elastic then the effect of a disclosure cap on quality is negative, if quality and information are weak complements, and is ambiguous otherwise.

This proposition shows, unlike in the case where the market is fully covered, that a disclosure cap can have a positive effect on quality when the market is partially covered. Intuitively, this is because the reduction in disclosure level has an additional effect of boosting demand when the market is not fully covered. This expansion in demand may lead to an overall increase in the marginal benefit of investing in quality, and hence raises the quality level chosen by the firm. In other words, there may be cases (as presented in the proposition) where there is no trade-off between more privacy and higher quality. Combining Proposition 3 with the fact that the direct effect of a disclosure cap on social welfare is always positive,[27] leads to the following result about the social desirability of a cap on the disclosure level.

Proposition 4 *(Social desirability of an ex ante disclosure cap under partial market coverage) Assume that the market is partially covered when disclosure is not regulated.*

- If the distribution of the idiosyncratic privacy cost exhibits a weakly increasing density function (i.e., $F(.)$ is weakly convex) then a disclosure cap is socially desirable.

- If the distribution of the idiosyncratic privacy cost exhibits a decreasing density function (i.e, $F(.)$ is concave) then a disclosure cap has an ambiguous effect on social welfare.

Thus, when the market is partially covered, a disclosure cap may be socially desirable even if quality and information are very strong complements. In particular, this is always the case if the density of the idiosyncratic cost is increasing, as this ensures that both the intensive and extensive margins of investment in quality are positively affected by a disclosure cap.

6 Extensions

6.1 Consumer-surplus-maximizing regulator

In the analysis of the *ex ante* regulation presented previously, we considered the decision problem of a regulator who seeks to maximize social welfare, which is given by the sum of the firm's profit and consumer surplus. We now examine the case where the regulator is a consumer protection agency, whose objective is to maximize consumer surplus. Suppose that we are in the scenario of full market coverage. The agency seeks to maximize

$$\widetilde{CS}(d) = \widetilde{CS}(q^M(d), d) = \int_{\underline{\theta}}^{\bar{\theta}} \tilde{U}(\theta, q^M(d), d) f(\theta)\, d\theta.$$

[27] This implies in particular that a disclosure cap is always socially desirable if $\tilde{q}^M = 0$.

which we assume to be concave in d. Let $\hat{d}^C$ denote the consumer-surplus-maximizing level of disclosure under *ex ante* regulation. As in the case of a social-welfare-maximizing regulator, a cap on disclosure is desirable from the perspective of the consumer protection agency if and only if $\hat{d}^C \leq \tilde{d}^M$. When a cap is desirable, the optimal value of the cap is $\bar{d}^C = \hat{d}^C$.

Let us first investigate the desirability of imposing a cap on disclosure from the perspective of the consumer protection agency. A cap on disclosure is (strictly) desirable if the marginal benefit of disclosure to the consumers is negative when evaluated at $d = \tilde{d}^M$; i.e.,

$$\left.\frac{\partial \widehat{CS}}{\partial d}\right|_{d=\tilde{d}^M} = \int_{\underline{\theta}}^{\bar{\theta}} \left[\frac{\partial \tilde{U}}{\partial q}\left(\theta, \tilde{q}^M, \tilde{d}^M\right) \left.\frac{\partial q^M}{\partial d}\right|_{d=\tilde{d}^M} + \underbrace{\frac{\partial \tilde{U}}{\partial d}\left(\theta, \tilde{q}^M, \tilde{d}^M\right)}_{<0} \right] f(\theta)\, d\theta < 0.$$

Observe that this condition is identical to the one under which a disclosure cap is desirable for a social-welfare-maximizing regulator. This is due to the fact that $\tilde{d}^M$ maximizes the profit of the firm and hence the marginal benefit of disclosure to the firm at $d = \tilde{d}^M$ is zero. Consequently, setting a cap on disclosure is desirable from the consumer protection agency's perspective if and only if it is socially desirable to do so.

Let us now assume that setting a disclosure cap is socially desirable and compare the optimal cap $\hat{d}^W$ for a social-welfare-maximizing regulator and the optimal cap $\hat{d}^C$ for a consumer-surplus-maximizing regulator. Since the disclosure cap is socially desirable, i.e., $\hat{d}^W < \tilde{d}^M$, and $\tilde{\Pi}\left(q^M(d), d\right)$ is concave with respect to d then

$$\frac{\partial \tilde{\Pi}}{\partial d}\left(q^M(\hat{d}^W), \hat{d}^W\right) > 0$$

Therefore,

$$\left.\frac{\partial \widehat{CS}}{\partial d}\right|_{d=\hat{d}^W} < \left.\frac{\partial \hat{W}}{\partial d}\right|_{d=\hat{d}^W} \leq 0,$$

where the second inequality is strict if and only if $\hat{d}^W = 0$. From the concavity of $\widehat{CS}(.)$ it then follows that

$$\hat{d}^C \leq \hat{d}^W.$$

The next proposition summarizes the above results.

Proposition 5 *(Consumer-surplus-maximizing vs socially optimal ex ante disclosure cap)*

(i) An ex ante regulation taking the form of a disclosure cap is desirable for a consumer protection agency if and only if it is socially desirable.

(ii) When such a regulation is desirable, the optimal disclosure cap from a consumer protection agency's perspective is (weakly) lower than the socially optimal disclosure cap.

6.2 *Ex post* regulation

We now investigate the effect of an *ex post* regulation of the disclosure level on investment in quality. More precisely, we consider the following game:

- First, the firm chooses its quality level.

- Second, the regulator decides whether to set a cap on the disclosure level.

- Third, the firm sets its disclosure level.

- Fourth, consumers decide whether to patronize the firm and, if they do, how much information to provide.

Note that in this scenario, the regulator will always find it optimal to set a cap $\bar{d} = d^W(q)$ on the disclosure level because $d^W(q) \leq d^M(q)$ for any quality level q. Moreover, the firm will always choose to set its disclosure level equal to the cap (because its profit is concave). Let us now consider the firm's choice of quality at the first stage of the game. The firm maximizes

$$\hat{\Pi}(q) \equiv \tilde{\Pi}\left(q, d^W(q)\right).$$

Assume that $\hat{\Pi}(.)$ is concave and denote

$$\hat{q}^M \equiv \underset{q \geq 0}{\arg\max} \hat{\Pi}(q)$$

the firm's optimal choice of quality. We focus on the case where the firm invests in quality when the disclosure level is not regulated (i.e., $\tilde{q}^M \neq 0$). *Ex post* regulation has a negative (positive) effect on quality if $\hat{q}^M$ is less (greater) than $\tilde{q}^M$. Since $\hat{\Pi}(.)$ is concave and $\left.\frac{\partial \hat{\Pi}}{\partial q}\right|_{q=\hat{q}^M} = 0$, we have the following equivalence:

$$\hat{q}^M \lesseqgtr \tilde{q}^M \Longleftrightarrow \left.\frac{\partial \hat{\Pi}}{\partial q}\right|_{q=\tilde{q}^M} \lesseqgtr 0.$$

Differentiating $\hat{\Pi}$ with respect to q and evaluating it at $q = \tilde{q}^M$ yields

$$\left.\frac{\partial \hat{\Pi}}{\partial q}\right|_{q=\tilde{q}^M} = \underbrace{\frac{\partial \tilde{\Pi}}{\partial q}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right)}_{\text{direct effect}} + \underbrace{\frac{\partial \tilde{\Pi}}{\partial d}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right) \left.\frac{\partial d^W}{\partial q}\right|_{q=\tilde{q}^M}}_{\text{strategic effect}}.$$

The *direct* effect is the only effect that would exist if the regulator committed *ex ante* to a disclosure cap equal to $d^W(\tilde{q}^M)$. The *strategic* effect captures how the firm adjusts its quality level to favorably influence the regulator's decision when the disclosure cap is decided *ex post*.

Let us consider first the term $\frac{\partial \tilde{\Pi}}{\partial q}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right)$ capturing the *direct* effect. We have

$$\frac{\partial \tilde{\Pi}}{\partial q}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right) = \frac{\partial \tilde{\Pi}}{\partial q}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right) - \underbrace{\frac{\partial \tilde{\Pi}}{\partial q}\left(\tilde{q}^M, d^M(\tilde{q}^M)\right)}_{=0}$$

and, therefore,

$$\frac{\partial \tilde{\Pi}}{\partial q}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right) = -\int_{d^W(\tilde{q}^M)}^{d^M(\tilde{q}^M)} \frac{\partial^2 \tilde{\Pi}}{\partial q \partial d}\left(\tilde{q}^M, u\right) du. \tag{7}$$

Hence, the sign of the direct effect depends on the sign of the cross-effect $\frac{\partial^2 \tilde{\Pi}}{\partial q \partial d}$. If this cross-effect is positive (negative) then the direct effect is negative (positive).

Let us now examine the term $\frac{\partial\tilde{\Pi}}{\partial d}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right) \left.\frac{\partial d^W}{\partial q}\right|_{q=\tilde{q}^M}$ which captures the *strategic* effect. Since $d^W(\tilde{q}^M) < d^M(\tilde{q}^M)$ and $\frac{\partial^2\tilde{\Pi}}{\partial d^2} < 0$ then

$$\frac{\partial\tilde{\Pi}}{\partial d}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right) > 0$$

and, therefore, the sign of $\frac{\partial\tilde{\Pi}}{\partial d}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right) \left.\frac{\partial d^W}{\partial q}\right|_{q=\tilde{q}^M}$ is the same as the sign of $\left.\frac{\partial d^W}{\partial q}\right|_{q=\tilde{q}^M}$. Differentiating

$$\frac{\partial\tilde{W}}{\partial d}(q, d^W(q)) = 0$$

with respect to q yields

$$\left.\frac{\partial d^W}{\partial q}\right|_{q=\tilde{q}^M} = -\frac{\frac{\partial^2\tilde{W}}{\partial q\partial d}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right)}{\frac{\partial^2\tilde{W}}{\partial d^2}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right)}. \tag{8}$$

Since $\frac{\partial^2\tilde{W}}{\partial d^2} < 0$, the sign of the strategic effect is the same as the sign of $\frac{\partial^2\tilde{W}}{\partial q\partial d}\left(\tilde{q}^M, d^W(\tilde{q}^M)\right)$. Hence, the effect of *ex post* regulation on quality depends not only on the cross-effect of quality and disclosure on the firm's profit (as in the case of *ex ante* regulation) but also on their cross-effect on social welfare.

Let us focus hereafter on the full market coverage scenario. In that case, we know that the firm's optimal quality choice $\tilde{q}^M$ when disclosure is not regulated is weakly positive. We have already shown that, under full market coverage, $\frac{\partial^2\tilde{\Pi}}{\partial q\partial d}$ is weakly positive. Therefore, from (7), we know that the direct effect of *ex post* regulation is weakly negative. Moreover, the cross-effect of quality and disclosure on social welfare is given by

$$\begin{aligned}
\frac{\partial^2\tilde{W}}{\partial q\partial d} &= \frac{\partial^2\tilde{\Pi}}{\partial q\partial d}(q,d) + \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial^2\tilde{U}}{\partial q\partial d}(\theta,q,d)\, f(\theta)\, d\theta \\
&= r\gamma - \int_{\underline{\theta}}^{\bar{\theta}} \gamma\theta f(\theta)\, d\theta \\
&= \gamma(r - E(\theta)).
\end{aligned}$$

Combining this with (8) implies that the effect of quality on the disclosure cap $\left.\frac{\partial d^W}{\partial q}\right|_{q=\tilde{q}^M}$ has the same (opposite) sign as γ when $r > E(\theta)$ ($r < E(\theta)$). Thus, the strategic effect of *ex post* regulation is negative (positive) if r is below (above) $E(\theta)$. We can then conclude that the overall effect of *ex post* regulation on quality is negative (ambiguous) if r is below (above) $E(\theta)$.

Proposition 6 *(Effect of an ex post disclosure cap on quality under full market coverage) The effect of an ex post disclosure cap on quality is negative if the value of information r is less than the average privacy cost $E(\theta)$, and is ambiguous otherwise.*

7 Discussion

7.1 Alternative interpretation of the model

In our analysis, we interpreted the consumer's input, x, as the amount of information that a consumer provides to the firm. This interpretation is appropriate when considering a matching website, where each user decides on the preferences she reveals, or services involving user-generated content, where each user decides on the amount of content to share (for example, on a social media platform). This interpretation may, however, be less suited to other forms of Internet services, such as email and search. In these contexts, consumers do not directly supply information to the firm; instead, information is generated as a result of their use of the firm's service. The relevant consumer choice variable (and hence the appropriate interpretation of x) is therefore *usage intensity*, rather than information provision level. Correspondingly, α, which formerly captured the cost of providing information, can be interpreted more generally as the *variable* opportunity cost of using the firm's service.

7.2 Positive surplus for buyers of personal information

We assumed in our baseline model that the firm was able to appropriate all the surplus generated by a transaction with a buyer of personal information. Let us relax this assumption by allowing third parties buying personal information to capture a positive share of that surplus. More precisely, let us assume that the unit price of personal information is βr, where $\beta \in (0, 1]$.

The firm's profit function in this variant of our model can be derived from the one in our baseline setting by replacing r with βr. This implies in particular that the effect of a disclosure cap on firm's choice of quality is qualitatively the same as in the baseline model. More precisely, Lemma 6 holds if we replace r with βr, and Proposition 3 holds in its current formulation (as it does not depend on r).

That said, when it comes to the effect of a disclosure cap on social welfare - defined as the sum of the firm's profit, the third parties' surplus and consumer surplus - it becomes more complicated to derive unambiguous results in the current setting. To see why, consider first the *direct effect* of a marginal decrease in the disclosure level (starting from the firm's optimal disclosure level); that is, the effect for a fixed quality level and a fixed amount of information provided by consumers. In our baseline model (i.e., $\beta = 1$) this effect is always positive. However, when the surplus of third parties buying personal information is positive (i.e., $\beta < 1$), this need not be true: a decrease in the disclosure level still leads to an increase in consumer surplus (through a decrease in privacy costs), but it also leads to a decrease in the surplus of third parties buying personal information from the firm. Let us consider now the *indirect effect* of a marginal decrease in the disclosure level. In our baseline model, this indirect effect was solely driven by the effect of a decrease in the disclosure level on quality. In particular, the sign of the indirect effect of a disclosure cap is the same as its effect on quality. This may not be true when the surplus of third parties is positive, which substantially complicates the analysis. The reason is due to the presence of a new indirect effect, which is associated to the (positive) effect of the disclosure cap on the amount of information provided by consumers. This implies that when the two indirect effects do not have the same sign (which is the the case when a disclosure cap has a negative effect on quality), and the sign of the overall indirect effect depends on their relative *magnitudes*.

Finally, notice that if the regulator maximizes consumer surplus, allowing buyers of personal information to make a positive surplus does not affect the desirability of a disclosure cap, as long as r is replaced with βr in our baseline setting.

7.3 Substitutability between quality and information

For many Internet services, it is natural to think of the firm's quality level and the consumer information provision level as complements (i.e., $\gamma \geq 0$): the better the quality of a firm's service, the higher the level of usage or information provision by the consumers. There may also be cases, however, where the firm's quality level and the consumer's information exhibit substitutability (i.e., $\gamma < 0$). User authentication is one such scenario. Interpreting the firm's quality level as its ability to verify its user identity without the use of personal information provided by the consumer,[28] the higher the firm's quality level, the lower the consumer's utility from providing additional pieces of personal information (phone number, secondary email address, etc.) for authentication purposes.

Note that when quality and information are substitutes, an *ex ante* cap on the level of disclosure is always socially desirable under the full market coverage scenario. There is no trade-off between privacy protection and quality provision in this case because the firm never invests in quality. When the market is partially covered, however, a trade-off between privacy and quality may also exist in the substitutes case.

8 Conclusion

In this paper, we study the impact of privacy regulation - specifically, a cap on data disclosure - on a monopoly's incentives to invest in the quality of its service and on social welfare. We find that the impact of a reduction in disclosure level on the monopolist's optimal choice of quality is negative when the market is fully covered, and depends on the distribution of the consumers' idiosyncratic privacy cost when the market is partially covered. Under full market coverage, a cap on the disclosure level is socially desirable when the degree of complementarity between quality and information is not too strong. Under partial market coverage, a cap is desirable when the distribution of the consumers' privacy cost exhibits a weakly increasing density. As extensions, we also analyzed the case where the regulator's objective is consumer surplus maximization and the scenario of an *ex post* regulation.

Our analysis suggests that it is important for regulators to determine if the market is fully covered when deciding on whether to set a disclosure cap, since the considerations to be taken into account differ under the two settings. The full market coverage scenario is likely to apply for more mature markets. These markets are typically characterized by high rates of penetration, implying that any demand expansion effect that may arise from an increase in quality or decrease in disclosure levels would probably be insignificant. The partial market coverage scenario, by contrast, would be more relevant when considering

[28] For example, the firm could make use of the IP address or geographical location to assess if a login attempt is potentially fraudulent.

younger markets (e.g. for new services). Demand is likely to be responsive to changes in quality and disclosure in these markets. Therefore, one practical implication of our analysis is that regulators should pay attention to the degree of complementarity between quality and information in mature markets and to the distribution of privacy costs in younger markets when deciding on whether or not to impose a disclosure cap.

In addition to a disclosure cap, another privacy regulation that can be explored in our framework is the taxation of disclosure revenues. The taxation of digital monopoly platforms have been studied by Bloch and Demange (2016) and Bourreau et al. (2016); however, neither papers consider the impact of taxation on the quality level of the firm. A (unit) tax on the monopolist's disclosure revenues would translate to a reduction in the value of information in our model. Since this reduction affects both the optimal quality and disclosure levels of the firm, the desirability of a tax on disclosure revenues is *a priori* ambiguous.

Finally, our model may also be interpreted more generally than one of privacy and quality. For example, the value of information can be thought of (more broadly) as the value that the firm derives from the exploitation of consumer data.[29] Correspondingly, the level of disclosure could instead be interpreted as the degree of data exploitation and the privacy cost parameter as a more general parameter reflecting the cost of sharing information. One can even take a step further and consider other types of inputs (besides quality level and personal information) that a firm and the consumers may provide. For instance, the firm could invest in the security level of its service and the consumers could provide time or attention rather than personal information. The interpretation of the consumers' cost parameter (which captured the intensity of privacy preferences in the case of information provision) would then change depending on the input that we are considering.

References

[1] Acquisti, A., Taylor, C., and L. Wagman (2016), "The Economics of Privacy," *Journal of Economic Literature* 54 (2), 442-492.

[2] Athey, S. (2014), "Information, Privacy and the Internet: An Economic Perspective." CPB Lecture.

[3] Bloch, F., and G. Demange (2017), "Taxation and Privacy Protection on Internet platforms," *Journal of Public Economic Theory*, forthcoming.

[4] Bourreau, M., Caillaud, B., and R. De Nijs (2017), "Taxation of a Digital Monopoly Platform," *Journal of Public Economic Theory*, forthcoming.

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

[6] Goldfarb, A., and C. Tucker (2011), "Privacy Regulation and Online Advertising", *Management Science*, 57 (1), 57-71.

[7] Goldfarb, A., and C. Tucker (2012), Privacy and Innovation, in *Innovation Policy and the Economy*, vol 12, Eds J. Lerner and S. Stern, NBER.

[29] Bloch and Demange (2016) provide several interpretations for the degree of data exploitation.

[8] Lambrecht, A., Goldfarb, A., Bonatti, A., Ghose, A., Goldstein, D. G., Lewis, R., Rao, A., Sahni, N., and S. Yao. (2014) "How do Firms Make Money Selling Digital Goods Online?", *Marketing Letters* 25 (3), 331–341.

[9] Miller, A., and C. Tucker (2009), "Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records," *Management Science*, 55 (7), 1077-1093.

[10] Miller, A., and C. Tucker (2011a), "Can Health Care Information Technology Save Babies?," *Journal of Political Economy*, 119 (2), 289-324.

[11] Miller, A., and C. Tucker (2011b), "Encryption and the Loss of Patient Data," *Journal of Policy Analysis and Management*, 30 (3), 534-556.

[12] The White House (2014), "Big Data and Privacy: A Technological Perspective," Available at `https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf`

A Appendix: Proofs

Proof of Lemma 1

Differentiating

$$\frac{\partial U}{\partial x}\left(\tilde{x}\left(\theta,q,d\right),\theta,q,d\right)=\frac{\partial V}{\partial x}\left(\tilde{x}\left(\theta,q,d\right),q\right)-\left(\alpha+\theta d\right)=0 \tag{9}$$

with respect to d yields

$$\frac{\partial^2 V}{\partial x^2}\left(\tilde{x}\left(\theta,q,d\right),q\right)\frac{\partial \tilde{x}}{\partial d}\left(\theta,q,d\right)-\theta=0$$

and, therefore,

$$\frac{\partial \tilde{x}}{\partial d}\left(\theta,q,d\right)=\frac{\theta}{\frac{\partial^2 V}{\partial x^2}\left(\tilde{x}\left(\theta,q,d\right),q\right)}<0$$

Differentiating (9) with respect to θ and q leads to

$$\frac{\partial \tilde{x}}{\partial \theta}\left(\theta,q,d\right)=\frac{d}{\frac{\partial^2 V}{\partial x^2}\left(\tilde{x}\left(\theta,q,d\right),q\right)}<0$$

$$\frac{\partial \tilde{x}}{\partial q}\left(\theta,q,d\right)=-\frac{\frac{\partial^2 V}{\partial x \partial q}\left(\tilde{x}\left(\theta,q,d\right),q\right)}{\frac{\partial^2 V}{\partial x^2}\left(\tilde{x}\left(\theta,q,d\right),q\right)}=\gamma\left(\tilde{x}\left(\theta,q,d\right),q\right).$$

Proof of Lemma 2

Since $U\left(x,\theta,q,d\right)$ is decreasing in θ then $\tilde{U}\left(\theta,q,d\right)=\max_{x\in[0,1]}U\left(x,\theta,q,d\right)$ is decreasing in θ (by the Envelope Theorem). Therefore, there exists $\tilde{\theta}\left(q,d\right)\in\left[\underline{\theta},\overline{\theta}\right]$ such that

$$\tilde{U}\left(\theta,q,d\right)\geq 0\Longleftrightarrow\theta\geq\tilde{\theta}\left(q,d\right)$$

Moreover, differentiating

$$\tilde{U}\left(\tilde{\theta}\left(q,d\right),q,d\right)=0$$

with respect to q and d, and using the Envelope Theorem, we get that

$$\frac{\partial \tilde{\theta}}{\partial q} = -\frac{\frac{\partial \tilde{U}}{\partial q}}{\frac{\partial \tilde{U}}{\partial \theta}} = -\frac{\frac{\partial U}{\partial q}}{\frac{\partial U}{\partial \theta}} = \frac{\frac{\partial V}{\partial q}}{d\tilde{x}\left(\tilde{\theta}\left(q,d\right),q,d\right)} > 0;$$

$$\frac{\partial \tilde{\theta}}{\partial d} = -\frac{\frac{\partial \tilde{U}}{\partial d}}{\frac{\partial \tilde{U}}{\partial \theta}} = -\frac{\frac{\partial U}{\partial d}}{\frac{\partial U}{\partial \theta}} = -\frac{\tilde{\theta}\left(q,d\right)}{d} < 0;$$

Proof of Lemma 3

The result follows directly from the fact that

$$\frac{\partial \tilde{\Pi}}{\partial q} < \frac{\partial \tilde{W}}{\partial q}$$

and the concavity of $\tilde{\Pi}$ and $\tilde{W}$ with respect to q.

Proof of Lemma 4

The result follows directly from the fact that

$$\frac{\partial \tilde{\Pi}}{\partial d} > \frac{\partial \tilde{W}}{\partial d}$$

and the concavity of $\tilde{\Pi}$ and $\tilde{W}$ with respect to d.

Proof of Lemma 5

Assume that $q^M\left(d\right) \neq 0$. Then, by continuity, $q^M(d') \neq 0$ for d' sufficiently close to d. Therefore, or d' sufficiently close to d, $q^M(d')$ is an interior solution given by the first-order condition

$$\frac{\partial \tilde{\Pi}}{\partial q}\left(q^M(d'), d'\right) = 0$$

Differentiating this with respect to d' and evaluating it at $d' = d$ yields

$$\frac{\partial^2 \tilde{\Pi}}{\partial q^2}\left(q^M(d), d\right)\frac{\partial q^M}{\partial d} + \frac{\partial^2 \tilde{\Pi}}{\partial q \partial d}\left(q^M(d), d\right) = 0$$

which leads to the result.

B Appendix: Elasticities

Denote

$$\check{V}(x,q) \equiv V\left(x,q\right) - \alpha x$$

The demand addressed to the firm when the amount of information is chosen optimally by consumers is

$$\tilde{D}\left(q,d\right) = F\left(\tilde{\theta}\left(q,d\right)\right) = F\left(\min\left(\frac{\check{V}\left(\tilde{x}\left(\tilde{\theta}\left(q,d\right),q,d\right),q\right) - K}{d\tilde{x}\left(\tilde{\theta}\left(q,d\right),q,d\right)}, \bar{\theta}\right)\right).$$

Consider the following function:

$$D(x,q,d) = F\left(\min\left(\frac{\check{V}(x,q)-K}{dx}, \bar{\theta}\right)\right),$$

which can be interpreted as the demand addressed to the firm if all consumers using the service (are required to) provide the same amount of information x. Note that

$$\tilde{D}(q,d) = D(\tilde{x}\left(\tilde{\theta}(q,d), q, d\right), q, d).$$

The elasticity of $D(x,q,d)$ with respect to d is (in absolute value)

$$-d\frac{\frac{\partial D}{\partial d}}{D} = \frac{d\frac{1}{d^2}\frac{\check{V}(x,q)-K}{x} f\left(\frac{\check{V}(x,q)-K}{dx}\right)}{F\left(\frac{\check{V}(x,q)-K}{dx}\right)} = \frac{\frac{\check{V}(x,q)-K}{dx} f\left(\frac{\check{V}(x,q)-K}{dx}\right)}{F\left(\frac{\check{V}(x,q)-K}{dx}\right)}$$

whenever $D(x,q,d) \in (0,1)$. In particular, under partial (positive) market coverage,

$$-d\left.\frac{\frac{\partial D}{\partial d}}{D}\right|_{x=\tilde{x}\left(\tilde{\theta}(q,d),q,d\right)} = \frac{\tilde{\theta}(q,d) f\left(\tilde{\theta}(q,d)\right)}{F\left(\tilde{\theta}(q,d)\right)}.$$

which shows that the elasticity of demand holding the amount of information constant (at the level of the marginal consumer) is the same as the elasticity of the cumulative distribution function (computed for the marginal type).
Similarly, straightforward algebraic manipulations show that the elasticity of $\frac{\partial D}{\partial q}$ with respect to d (which can be either positive or negative depending on whether D is supermodular or submodular in (q,d)) is given by:

$$d\frac{\frac{\partial^2 D}{\partial q \partial d}}{\frac{\partial D}{\partial q}} = -1 - \frac{\frac{\check{V}(x,q)-K}{dx} f'\left(\frac{\check{V}(x,q)-K}{dx}\right)}{f\left(\frac{\check{V}(x,q)-K}{dx}\right)}$$

whenever $D(x,q,d) \in (0,1)$. In particular, under partial (positive) market coverage,

$$d\left.\frac{\frac{\partial^2 D}{\partial q \partial d}}{\frac{\partial D}{\partial q}}\right|_{x=\tilde{x}\left(\tilde{\theta}(q,d),q,d\right)} = -1 - \frac{\tilde{\theta}(q,d) f'\left(\tilde{\theta}(q,d)\right)}{f\left(\tilde{\theta}(q,d)\right)} = -1 - \underbrace{\frac{\tilde{\theta}(q,d) F''\left(\tilde{\theta}(q,d)\right)}{F'\left(\tilde{\theta}(q,d)\right)}}_{\text{curvature of } F(.)}$$

This implies that the elasticity of $\frac{\partial D}{\partial q}$ with respect to d holding the amount of information constant is related to the curvature of $F(.)$. It is greater (less) than -1 if $f'(.)$ is positive (negative), that is, if $F(.)$ is convex (concave).