

# The Bright Side of the GDPR: Welfare-improving Privacy Management\*

Chongwoo Choe<sup>†</sup>    Noriaki Matsushima<sup>‡</sup>    Shiva Shekhar<sup>§</sup>

September 23, 2023

## Abstract

We study the GDPR's opt-in requirement in a model with a firm that provides a digital service and consumers who are heterogeneous in their valuations of the firm's service as well as the privacy costs incurred when sharing personal data with the firm. We show that the GDPR boosts demand for the service by allowing consumers with high privacy costs to buy the service without sharing data. The increased demand leads to a higher price but a smaller quantity of shared data. If the firm's revenue is largely usage-based rather than data-based, then both the firm's profit and consumer surplus increase after the GDPR, implying that the GDPR can be welfare-improving. But if the firm's revenue is largely from data monetization, then the GDPR can reduce the firm's profit and consumer surplus.

**JEL Classification:** D18, D61, K24, L12, L51, L86.

**Keywords:** GDPR, opt-in, opt-out, privacy management, welfare

---

\*We thank Joshua Gans, Hooman Hidaji, Ulrich Laitenberger, Barry R. Nault, Markus Reisinger, Tat-How Teh, Mark Tremblay, Vaarun Vijairaghavan, Chengsi Wang, and participants at the Monash-USC Digital Conference and the seminar at Haskayne School of Business for helpful comments and suggestions. We gratefully acknowledge financial support from the Australian Research Council (grant number DP210102015) and the Japan Society for the Promotion of Science (KAKENHI grant numbers JP20H05631, JP21H00702, JP21K01452, JP21K18430, and JP23H00818). Although Matsushima serves as a member of the Competition Policy Research Center at the Japan Fair Trade Commission (JFTC), the views expressed in this paper are solely ours and should not be attributed to the JFTC. The usual disclaimer applies.

<sup>†</sup>Department of Economics, Monash University. Email: [chongwoo.choe@monash.edu](mailto:chongwoo.choe@monash.edu).

<sup>‡</sup>Institute of Social and Economic Research, Osaka University. Email: [nmatsush@iser.osaka-u.ac.jp](mailto:nmatsush@iser.osaka-u.ac.jp).

<sup>§</sup>The corresponding author. Tilburg School of Economics and Management (TiSEM), CESifo Research affiliate. Email: [s.shekhar.1@tilburguniversity.edu](mailto:s.shekhar.1@tilburguniversity.edu), [shiva.shekhar.g@gmail.com](mailto:shiva.shekhar.g@gmail.com).

# 1 Introduction

The European Union (EU)’s General Data Protection Regulation (GDPR) is generally considered to be the most stringent law that governs personal data protection, and is setting a global standard for privacy regulations.<sup>1</sup> Before the GDPR went into effect in 2018, many consumers had little control over their data in the digital space. For example, privacy regulations such as the EU’s Data Protection Directive (Directive 95/46/EC) or Australia’s Privacy Act 1988 were based on principles and guidelines, which were non-binding and not specific enough; nor were they entirely fit for the digital age. As a result, online businesses collected and processed personal data without providing clear information to consumers as to what data is collected, how it is used, and with whom it is shared. Data collection was often set as the default option, from which consumers could not easily opt out. The GDPR superseded the EU’s Data Protection Directive with more specific data protection requirements, stiffer enforcements, and penalties for non-compliance. In addition, to rectify the opaque data collection process, the GDPR (Article 7, Recital 32) requires businesses to allow consumers to make a “freely given, specific, informed and unambiguous” consent to the processing of their personal data, with significant penalties for non-compliance.<sup>2</sup> This requirement essentially mandates opt-in consent to data collection. Evidence shows that opt-in results in much lower levels of participation than opt-out (Johnson et al., 2002, 2020), which may explain why opt-out was a predominant way businesses obtained consent before the GDPR.

Although the European Commission has been positive about the overall success of the GDPR,<sup>3</sup> the media’s assessment has been mixed. Positive assessment was based on more user control over personal data, transparency, and accountability requirements on data processors, while negative assessment pointed out the complexity and compliance costs, especially for smaller businesses, adverse effects on investment in technology ventures, as well as the GDPR’s insufficient data protection measures.<sup>4</sup>

The academic literature has been generally more critical. A number of empirical studies report adverse effects of the GDPR on businesses such as a decrease in the amount of data collected (Schmitt et al., 2021; Congiu et al., 2022; Aridor et al., forthcoming) which adversely affect AI startups and data-based innovation (Bessen et al., 2020; Batikas et al., 2023), dampened incentives to invest in technology ventures (Jia et al., 2021), a decrease in the number of mobile apps (Janßen et al., 2022), reduced website pageviews and e-commerce revenue (Goldberg et al., 2022), and increased market concentration in websites and web technology services (Schmitt et al., 2021; Peukert et al., 2022; Johnson et al., forthcoming).<sup>5</sup> In addition, several

---

<sup>1</sup>Following the GDPR, many countries enacted similar privacy laws or amended existing ones in line with the GDPR. Examples include the California Consumer Privacy Act (CCPA) and Brazil’s General Data Protection Law (LGPD) both of which came into effect in 2020, Singapore’s Personal Data Protection Act (PDPA) amended in 2019, and South Korea’s Personal Information Protection Act (PIPA) amended in 2020, etc.

<sup>2</sup>Some of the largest non-compliance penalties are 746 million euros Luxembourg’s privacy watchdog (CNPD) imposed on Amazon in 2021, 150 million euros on Google levied by France’s privacy watchdog (CNIL) in 2022 and, more recently, 390 million euros Ireland’s Data Protection Commission imposed on Meta in 2023.

<sup>3</sup>[https://ec.europa.eu/commission/presscorner/detail/en/qanda\\_20\\_1166](https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1166)

<sup>4</sup>See, for example, “America should borrow from Europe’s data-privacy law,” *The Economist*, April 5, 2018; “One year on, GDPR needs a reality check,” *Financial Times*; June 30, 2019, “How GDPR is failing,” *Wired*, May 23, 2022.

<sup>5</sup>As potential beneficial effects of the GDPR, Godinho de Matos and Adjerid (2022) provide experimental

theoretical studies show that the GDPR’s opt-in requirement in its current form can be ineffective in allowing consumers to effectively manage their privacy (Choi et al., 2019; Chen, 2022; Chaudhury and Choe, 2023).

Given the negative assessment summarized above, one may wonder if there is any salutary effect to the GDPR. After all, if consumers are able to manage their privacy more effectively thanks to the GDPR, shouldn’t they benefit from it? Or would the potential consumer benefits be skimmed away by the digital business with market power? How does the GDPR change total surplus? Answering these questions is the primary purpose of this paper.

We examine the GDPR’s opt-in requirement for data collection in an environment that we believe incorporates consumer heterogeneity in a richer and more realistic way than in the existing literature on digital privacy.<sup>6</sup> Specifically, we assume that consumers differ in two dimensions, in their preference for the online service under consideration and in the privacy costs incurred while sharing personal data with the online business. Adding consumer heterogeneity in privacy costs to otherwise a standard model of privacy regulation is motivated by two factors. First, growing evidence reports that consumers are highly heterogeneous in their privacy preferences (e.g., Goldfarb and Tucker, 2012; Ghose, 2017; Lin, 2022; Macha et al., forthcoming). For example, Acquisti et al. (2015) review studies that classify individuals into three privacy groups: privacy fundamentalists, pragmatists, and unconcerned. Second, given that one of the stated purposes of the GDPR is to enable consumers to better manage their privacy, a model where all consumers have identical privacy preference is unlikely to yield meaningful insight. Incorporating such multidimensional consumer heterogeneity allows us to study consumer choice in a way that is richer than in the existing literature, which allows us to identify conditions under which the GDPR can increase or decrease welfare.

We briefly sketch our model. A firm sells an online service to consumers who are heterogeneous along two dimensions, valuation for the service and privacy cost, the latter incurred if a consumer shares her personal data with the firm when using the service. If a consumer shares her data with the firm, which we call opt-in as opposed to opt-out, she can enjoy additional opt-in benefits offered by the firm. The firm can monetize the data, so that it earns revenue from two sources: sales of its service that depends on the total demand the firm faces and shared consumer data that depends only on the opt-in demand. We call the former usage-based revenue and the latter data-based revenue. The firm chooses price for its service to maximize profit, and consumers make decisions to buy the firm’s service and whether or not to share data with the firm. We compare the equilibria before and after the GDPR.

Our basic assumption is that the GDPR enables consumers to proactively manage their digital privacy by requiring the firm to allow consumers to buy its service without sharing their personal data, i.e., opt-out purchase. That is, before the GDPR, consumers must share data with the firm when buying its service but, after the GDPR, consumers can choose between opt-in and opt-out purchase. This is clearly a simplifying assumption. Even before the GDPR, some privacy-conscious consumers may bypass the firm’s data collection, for example, by deleting cookies, using VPNs or privacy-oriented browsers. Likewise, even after the GDPR, some

---

evidence that suggests that the GDPR’s enhanced opt-in requirements may improve welfare by providing better privacy protection to consumers and allowing firms to better target their marketing communications.

<sup>6</sup>This literature is reviewed in Section 2.

consumers may routinely opt in without thinking through the ramifications of their decisions. However, insofar as more consumers rationally make opt-in decisions after the GDPR, which we believe is plausible, our main insight should continue to hold.<sup>7</sup> In addition, the absence of strict enforcement before the GDPR implies that firms could not credibly commit not to gather data from consumers who choose opt-out.<sup>8</sup> The GDPR lends credibility to such a commitment. In short, a key change brought about by the GDPR in our model is privacy management made available to consumers, who make opt-in decisions by comparing additional benefits from sharing data with possible privacy costs.<sup>9</sup>

First, we show that the equilibrium price for the firm’s service increases after the GDPR. The intuition is as follows. Before the GDPR, some consumers with high privacy costs may not buy the service because they do not have an option to buy the service without sharing personal data. After the GDPR, they can participate by choosing opt-out purchase, which expands demand for the service. This allows the firm to raise price. The additional choice of opt-out purchase and the higher price will inevitably decrease the opt-in demand, hence the amount of data shared, which will erode the firm’s data-based revenue. But the effect of higher price on data-based revenue is smaller after the GDPR. It is because the firm earns data-based revenue only from opt-in demand after the GDPR in contrast to total demand before the GDPR. That is, the firm’s choice of price is less constrained by data-based revenue after the GDPR. This implies that the firm charges a higher price for its service after the GDPR regardless of the size of its data-based revenue.

Our next result relates to how the GDPR changes the equilibrium demand for the firm’s service. As discussed above, the GDPR expands demand by providing consumers with an additional choice of opt-out purchase, which we call the demand expansion effect. But it leads to a higher price for the service, which counteracts the demand expansion effect. If the demand expansion effect dominates the price effect, then the equilibrium demand increases after the GDPR; otherwise, it decreases. We show that the effect of the GDPR on the equilibrium demand hinges on the size of data-based revenue. When the firm’s data-based revenue is not significant, the firm chooses price for its service close to the monopoly price. As a result, the market is not sufficiently covered before the GDPR, which leaves significant room for demand expansion after the GDPR. In this case, the demand expansion effect dominates the price effect, hence the equilibrium demand increases after the GDPR. However, if the firm earns large data-based revenue, it chooses a low price for its service, resulting in the market sufficiently covered before the GDPR. In this case, the price effect dominates the demand expansion effect, implying that the demand increase following the GDPR is limited. Moreover, if the price increase after

---

<sup>7</sup>For example, we can consider a model where a fraction of tech-savvy, privacy-conscious consumers make optimal opt-in decisions, and the rest always opt in. We can think of the GDPR as increasing the fraction of the first group of consumers.

<sup>8</sup>In the absence of penalty for renegeing by the firm, it is a dominant strategy for the firm to monetize data even if a consumer chooses opt-out. Anticipating this, consumers face only two effective choices: opt-in purchase or no purchase. That is, the firm’s commitment to making opt-out purchase available is not credible.

<sup>9</sup>Another interpretation is that the GDPR allowed firms to credibly exercise second-degree price discrimination by offering a menu of {(price for the service, no opt-in benefits), (price for the service, positive opt-in benefits)}. For example, many German news websites offer free access to its content subject to advertising and tracking or a paid subscription without advertising and tracking. Some examples are <https://abo.spiegel.de/de/c/microsites/werbefreilesen/abo> or <https://www.faz.net/faz-net-services/>.

the GDPR is sufficiently large, then consumers with high privacy costs who purchased the service before the GDPR drop out of the market after the GDPR. As a result, the equilibrium demand can decrease after the GDPR. Although the effect of the GDPR on the total equilibrium demand depends on the firm's data-based revenue, the amount of data collected by the firm unambiguously decreases after the GDPR, consistent with empirical evidence cited earlier. That is, the equilibrium opt-in demand after the GDPR is smaller than the total equilibrium demand before the GDPR. This is a straightforward consequence of a higher price and more purchase options made available to consumers after the GDPR.

Next, we show how the firm's profit changes after the GDPR. As previously alluded to, the effect of the GDPR on profit also depends on the size of data-based revenue. When the firm's data-based revenue is not significant, the higher price and larger demand imply that the firm earns larger sales revenue after the GDPR. Although its data-based revenue decreases because the quantity of shared data decreases, the firm earns larger profit because the data-based revenue is a small part of its profit. On the other hand, the GDPR can erode the firm's profitability if its revenue is largely data-based, mainly because the amount of data collected by the firm is smaller after the GDPR. The implication is that the GDPR's effect on the firm's profit depends on the firm's business model. The GDPR is likely to benefit usage-driven businesses such as ride-hailing platforms like Uber or streaming services like Netflix, but hurt data-driven businesses such as various navigation and weather apps and, to some extent, businesses whose main revenue source is advertising that leverages user data.

The GDPR's effect on consumer surplus also hinges on whether the firm's revenue is largely usage-based or data-based. In the former case, the GDPR's demand expansion effect dominates the price effect. Given that more consumers can participate in the market after the GDPR while benefiting from more effective privacy management, consumer surplus increases in this case. However, when the firm earns significant data-based revenue, the price effect dominates the demand expansion effect, and the equilibrium demand can decrease. Thus consumer surplus decreases in this case. Once again, the implication is that the GDPR can benefit consumers in markets that are largely service-based, but can hurt consumers when data-based revenue is a significant part of digital business.

Although the GDPR's opt-in requirement is the main focus of this paper, the way we describe privacy management made available after the GDPR applies more generally to other online businesses as well. In many e-retailers such as Target, Walmart, Zara, etc., consumers can choose to create an account for transaction or check out as a guest without creating an account. By creating an account, consumers share various personal data such as name, residential address, email address, payment information, etc. with the retailer. The benefits of creating an account include faster checkout, order tracking, and exclusive deals and promotions. On the other hand, sharing data with e-retailers can pose privacy risks such as identity theft, data breaches, cyberstalking, etc. The option to manage privacy is available in other platforms such as Google and YouTube. Users can choose the incognito mode or delete usage history after using services provided by these platforms. Once again, privacy management trades off additional benefits from sharing data against possible privacy risks. The general conclusion from our analysis is that, for digital businesses that do not rely heavily on data-based revenue, em-

powering consumers to proactively manage their digital privacy can improve firm profitability, increase consumer surplus, and hence is welfare-improving. But effective privacy management by consumers can reduce welfare when businesses rely heavily on data-based revenue.<sup>10</sup> We stress that this conclusion applies where consumers can effectively manage their privacy. But this may not necessarily be the case even after the GDPR, as we discuss in Section 6.4.

The rest of the paper is organized as follows. In Section 2, we review the related literature. Section 3 presents a simple example that explains some of our key points, followed by the baseline model. Section 4 analyzes the baseline model and characterizes the equilibria before and after the GDPR. Section 5 provides our key welfare results. In Section 6, we discuss several extensions and variations of the baseline model including the correlation between consumer valuations and privacy costs (Section 6.1), data externalities (Section 6.2), the case where consumer heterogeneity is only in valuations (Section 6.3), testable hypotheses from our analysis (Section 6.4), and the implications for management (Section 6.5). Section 7 concludes the paper. The proofs not provided in the main text are relegated to the Appendix.

## 2 Related Literature

Our work is related to a number of theoretical studies that model some of the key features of the GDPR and examine the digital business's optimal strategy and welfare implications. [Choi et al. \(2019\)](#) focus on the GDPR's opt-in requirement when there are data externalities, and show that the monopolistic firm can render opt-in regulations ineffective if it can exercise price discrimination. In deriving this result, they assume full market coverage, whereas we consider the case where the firm uses uniform pricing and the market is not fully covered. [Fainmesser et al. \(2023\)](#) study the digital business's choice of data collection and data security, and show how the optimal strategy depends on the digital business's revenue model, i.e., whether it is more usage-driven or data-driven. But the digital business's optimal pricing decision is absent in their model, nor do consumers explicitly make opt-in decisions. [Ke and Sudhir \(2023\)](#) analyze a model of behavior-based pricing that endogenizes the consumer's decision to exercise the rights to opt-in, data erasure and data portability, and firms can offer personalized products to opt-in consumers. In their model, data erasure and portability imply that only opt-in decisions arise in equilibrium.<sup>11</sup> [Chen \(2022\)](#) studies socially optimal data collection policy when consumers differ in privacy costs. [Chaudhury and Choe \(2023\)](#) provide an illustrative example of optimal data policy when consumers differ in privacy costs and there are multiple data types. Although not directly related to the GDPR, [Campbell et al. \(2015\)](#) show that a consent requirement in privacy regulation can tilt the playing field in favor of large generalist firms at the cost of small, specialist firms, because it imposes a one-off cost on both but consumer data has multiple use in the former. In addition to the differences already explained above, none of these studies models multidimensional consumer heterogeneity as we do in this paper.

A large and growing body of literature studies consumers' privacy choice and the attendant

---

<sup>10</sup>[Gopal et al. \(forthcoming\)](#) also derive a similar conclusion although welfare-reducing effect of privacy management stems from increased third-party data sharing.

<sup>11</sup>[Cong and Matsushima \(2023\)](#) also consider data management by consumers in a two-dimensional Hotelling model where data collected in one market can be used in another market.

data externalities (Choi et al., 2019; Ichihashi, 2021, 2023; Acemoglu et al., 2022; Bergemann et al., 2022, etc.).<sup>12</sup> A general conclusion from these studies is that opt-in consumers allow the data controller to infer more information about opt-out consumers, which can lead to excessive data sharing and raise the cost from privacy breach.<sup>13</sup> Once again, the main difference between these studies and ours is that we explicitly model multidimensional consumer heterogeneity whereas consumers differ only in one dimension in these studies, e.g., in their valuation of the digital business’s service (Choi et al., 2019; Bergemann et al., 2022) or in their valuation of privacy (Acemoglu et al., 2022; Ichihashi, 2023).

Several studies model multidimensional consumer types as we do in this paper, although their focus is different from ours (Anderson and Gans, 2011; Jentzsch et al., 2013; Miklós-Thal et al., 2023). Anderson and Gans (2011) model consumer heterogeneity in preference and disutility from advertising in the media market. In the oligopoly model of product differentiation in Jentzsch et al. (2013), consumers differ in brand preference and transportation costs. Miklós-Thal et al. (2023) study consumers’ privacy choice when each consumer is represented by two data types, privacy-sensitive data and innocuous data. Methodologically, the way we model opt-out is close to the choice of ad avoidance in Anderson and Gans (2011). That is, consumers in our model can choose opt-out purchase to avoid privacy costs although they lose the opt-in benefits by doing so.

### 3 Model

#### 3.1 Illustrative example

We start with a simple example that illustrates some of our key results. Consider a firm that produces its service at zero marginal cost and sells it at price  $p$ . There are four consumers denoted by  $(v_\theta, c_\theta)$ ,  $\theta = H, L$ , where  $v_\theta$  is the consumer’s valuation of the service and  $c_\theta$  is her privacy cost incurred when sharing her data with the firm. Let  $v_H = 3, v_L = 2, c_H \in (1/2, 1)$ , and  $c_L = 0$ . Then we have  $v_H - c_L > v_H - c_H > v_L - c_L > v_L - c_H > 0$ . Thus efficiency dictates that all consumers buy the service. Each consumer has two options when buying the service: she can buy the service without sharing her data, which we call opt-out purchase; or she can buy the service while sharing her data, which we call opt-in purchase. When the consumer chooses opt-in purchase, the firm offers opt-in benefits denoted by  $\epsilon > 0$ . We assume  $\epsilon$  is arbitrarily small and exogenous given,<sup>14</sup> and that consumers buy the service when indifferent.

First, suppose the consumer must share her data when she buys the service, i.e., only opt-in purchase is available. Then there are four possible prices the firm chooses to maximize profit. If it chooses  $p_1 = v_H - c_L = 3$ , then only type- $(v_H, c_L)$  consumer buys the service and the firm’s profit is  $\pi_1 = 3 - \epsilon$ . If the firm chooses  $p_2 = v_H - c_H = 3 - c_H$ , then type- $(v_H, c_L)$  and type- $(v_H, c_H)$  consumers buy the service. The firm’s profit is then  $\pi_2 = 2(3 - c_H - \epsilon) > \pi_1$ . If

<sup>12</sup>For a comprehensive review of the earlier literature on privacy, see Acquisti et al. (2016).

<sup>13</sup>In addition, based on opt-in consumers’ data, sellers can infer preference of opt-out consumers, which can lead to higher prices and create negative pecuniary externalities (Belleflamme and Vergote, 2016; Braghieri, 2019).

<sup>14</sup>For clarity of exposition, we assume that the firm cannot extract opt-in benefits through price. Our example can be easily modified to allow that possibility without changing the main insight. The firm has incentives to offer opt-in benefits because it can monetize the data. We do not consider data monetization in this illustrative example. But we incorporate the firm’s data-based revenue in the formal model.

the firm chooses  $p_3 = v_L - c_L = 2$ , then all but type- $(v_L, c_H)$  consumer buy the service, yielding profit  $\pi_3 = 6 - 3\epsilon > \pi_2$  for the firm. Finally, if the firm sets  $p_4 = v_L - c_H = 2 - c_H$ , then all consumers buy the service and the firm earns profit  $\pi_4 = 4(2 - c_H - \epsilon) < \pi_3$ . Thus, the firm optimally chooses price  $\hat{p}^* = 2$ , earns profit  $\hat{\pi}^* = 6 - 3\epsilon$ , and excludes type- $(v_L, c_H)$  consumer. In this case, total consumer surplus is  $2 - c_H + 3\epsilon$ .

Next, suppose that the consumer has available both options of opt-in and opt-out purchase. Given the price discount  $\epsilon$ , consumers with low privacy cost ( $c_L$ ) choose opt-in purchase when buying the service. There are two possible prices the firm can charge consumers who choose opt-out purchase. First, if it charges  $p_1 = v_H = 3$ , then only high-valuation consumers buy the service: type- $(v_H, c_L)$  consumer chooses opt-in purchase and type- $(v_H, c_H)$  consumer chooses opt-out purchase. In this case, the firm's profit is  $6 - \epsilon$ . Second, if the firm chooses  $p_2 = 2$ , then all consumers buy the service: type- $(v_H, c_L)$  and type- $(v_L, c_L)$  consumers choose opt-in purchase while type- $(v_H, c_H)$  and type- $(v_L, c_H)$  consumers choose opt-out purchase. The firm's profit in this case is  $8 - 2\epsilon$ . Given that  $\epsilon$  is arbitrarily small, the firm optimally chooses  $\tilde{p}^* = 2$ , and earns profit  $\tilde{\pi}^* = 8 - 2\epsilon > \hat{\pi}^* = 6 - 3\epsilon$ . No consumer is excluded in this case and total consumer surplus is  $2 + 2\epsilon$ , which is larger than that when consumers do not have an option to choose opt-out purchase.

The main takeaway from the above example can be summarized as follows. The option to purchase the service without sharing data allows consumers to better manager their privacy. In the absence of such an option, consumers with high-privacy cost choose not to buy the service in order to protect privacy. But these consumers can choose to buy the service without sharing data when they have such an option. Thus, the option to buy the service without sharing data expands consumer demand, which benefits the firm. The ability to better manage privacy also benefits consumers as a whole. On the other hand, the total amount of data shared with the firm decreases when consumers have the option not to share data. In what follows, we present a model to formally examine these observations and to explore additional implications.

### 3.2 Baseline model

We consider a digital business that provides online service at zero marginal cost and charges price  $p$ . Consumers are heterogeneous along two dimensions: valuation for the service ( $v$ ) and privacy cost ( $c$ ). As discussed previously, consumer heterogeneity in privacy preference is well-documented and, as we show in this paper, it allows us to portray a realistic and rich model of how the GDPR can help boost demand by expanding consumer choice. We assume  $v$  and  $c$  are independent and are uniformly distributed on the interval  $[0, 1]$ .<sup>15</sup> As in the illustrative example, consumers face three possible choices: buying the service while sharing data, called opt-in purchase; buying the service without sharing data, called opt-out purchase; and not buying the service.<sup>16</sup> Denote the mass of consumers choosing opt-in purchase by  $D_S$ , the mass of consumers choosing opt-out purchase by  $D_N$ , and the total demand for the firm's service by  $D = D_S + D_N$ .

<sup>15</sup>In Section 6.1, we discuss the case where  $v$  and  $c$  are correlated.

<sup>16</sup>By opt-in purchase, we mean more than sharing data with the firm; it also allows the firm to share or sell data to third parties. Likewise, opt-out purchase prevents the firm from selling or sharing personal information with third parties, as stipulated in the CCPA Section 120.



The firm can further enhance its service by offering additional benefits to consumers who choose opt-in purchase. We call this opt-in benefits denoted by  $x \in [0, 1]$ .<sup>17</sup> The opt-in benefits include, among others, time saved thanks to faster checkout, exclusive deals and promotions, improved and personalized service, etc. For example, Gerbes, Missouri-based supermarkets, offer opt-in benefits such as exclusive specials, fuel points, personalized coupons, and online or in-app ordering for pickup or delivery.<sup>18</sup> To focus squarely on the GDPR's effect on the demand for the firm's service, we assume  $x$  is exogenously given.

A consumer of type  $(v, c) \in [0, 1] \times [0, 1]$  who chooses opt-in purchase receives utility

$$U_S(v, c) := v + \underbrace{(x - c)}_{\text{net opt-in benefits}} - p \quad (1)$$

where  $p$  is the price charged by the firm and  $x - c$  is the opt-in benefits less privacy cost. If a type- $(v, c)$  consumer chooses opt-out purchase, then her utility is

$$U_N(v, c) := v - p. \quad (2)$$

Clearly, the main difference between opt-in and opt-out purchase is that a consumer choosing opt-in purchase expects the additional opt-in benefits as well as privacy costs. As discussed in Section 1, a key change brought about by the GDPR is that firms need explicit opt-in consent from consumers in order to collect data. Our interpretation of this in our model is that consumers can choose between opt-in and opt-out purchase after the GDPR, although they cannot choose opt-out purchase before the GDPR. Based on this, we can summarize a consumer's decision as follows. Before the GDPR, a consumer buys the firm's service if  $U_S \geq 0$ . After the GDPR, she chooses opt-in purchase if  $U_S \geq \max\{U_N, 0\}$ , opt-out purchase if  $U_N \geq \max\{U_S, 0\}$ , and no purchase otherwise.

The firm's revenue comes from two sources. First, the firm earns revenue from sales of its service. Given the total demand for the firm's service  $D(\cdot)$ , the sales revenue is  $pD(\cdot)$ . Second, the firm can monetize the data it collects, for example, through selling data for targeted advertising (Bergemann and Bonatti, 2015), using data for personalized pricing (Choe et al., 2018; Chen et al., 2020), or through product improvement by data-enabled learning (Hagi and Wright, forthcoming). We do not model how the firm monetizes its data and simply assume that the revenue is exogenously given by  $\alpha \geq 0$  per unit of consumer data. Given that the firm collects data from all opt-in consumers whose mass is denoted by  $D_S(\cdot)$ , the firm's revenue from data is  $\alpha D_S(\cdot)$ . Noting that we assume away all costs for the firm, the firm's profit is given by

$$\Pi = \underbrace{pD(\cdot)}_{\text{sales revenue}} + \underbrace{\alpha D_S(\cdot)}_{\text{data-based revenue}}. \quad (3)$$

Before the GDPR, we have  $D_N = 0$ , hence  $D = D_S$ . Therefore, the main difference in the firm's profit after the GDPR is the possibility that  $D > D_S$ . In the baseline model, we assume

<sup>17</sup>Given that the highest possible privacy cost is 1, we consider the nontrivial case where  $x \leq 1$ .

<sup>18</sup><https://www.gerbes.com/i/privacy-policy>

$x$  is exogenously given so that the firm chooses only  $p$  in order to maximize profit given in (3).

The game proceeds as follows. First, consumers learn their types privately. Next, the firm posts price, which is followed by consumers' decisions: before the GDPR, consumers make only purchase decisions; after the GDPR, they make additional data-sharing decisions. We focus on the equilibrium where the market is not fully covered. Given the monopolistic firm, full market coverage seems not only unrealistic but also makes the comparison of pre- and post-GDPR regimes less interesting. The precise conditions that ensure this will be spelt out as we go on.

## 4 Analysis

### 4.1 Equilibrium before the GDPR

Since consumers cannot choose opt-out purchase in this case, any consumer who buys the firm's service also shares her data with the firm. Thus, consumer utility relevant to her purchase decision is  $U_S(v, c)$ .

A consumer of type  $(v, c)$  buys the service if and only if

$$U_S(v, c) \geq 0 \implies c \leq c(v, p) := v + x - p.$$

In the above,  $c(v, p)$  is the privacy cost that makes the consumer indifferent between opt-in purchase and no purchase given her valuation  $v$  and price  $p$ . Then,  $c(1, p) = 1 + x - p$  is the threshold value of privacy cost above which no consumers buy the service. Likewise,  $v(p) := p - x$  is the threshold value of consumer valuation for the service, below which no consumers buy the service. As discussed previously, we focus on equilibrium where the market is not fully covered. Depending on  $v(p) > 0$  or  $v(p) \leq 0$ , there are two possible cases, as shown in Figure 1. First, the demand for the case where  $0 < v(p) < 1$ , i.e.,  $x < p < 1 + x$ , is illustrated in Figure 1a, which we call case 1a. This is the case where some consumers with low valuation for the service do not purchase the service regardless of their privacy costs. This case arises when the price is higher than the opt-in benefits. Second, Figure 1b illustrates the demand for the case where  $-1 < v(p) \leq 0$ , or  $x - 1 < p \leq x$ , which we call case 1b.<sup>19</sup> This is the case where the opt-in benefits are large enough ( $x \geq p$ ) so that all consumers with relatively low privacy costs purchase the service, regardless of their valuation for the service. But the opt-in benefits are not too large ( $x < 1 + p$ ) so that some consumers with high privacy costs and low valuation do not purchase the service. In case 1b, a subset of consumers along one dimension of consumer types is fully served, but not along the other dimension. Thus, a case like this is possible only when consumer types are multidimensional.

---

<sup>19</sup>Note that the restriction  $x - 1 < p$  allows the possibility that the firm may choose a negative price since  $x \leq 1$  by assumption. This may happen when the firm's revenue is largely data-based, i.e., large  $\alpha$ , in which case the firm may subsidize consumers' purchase of its service with a view to monetizing their data.

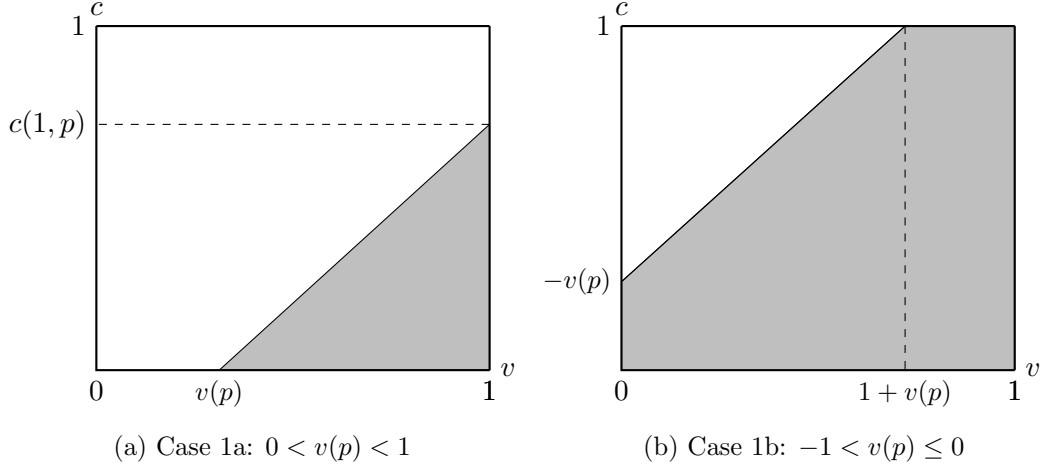


Figure 1: Demand before the GDPR

From the above, we can calculate the demand for the firm's service as<sup>20</sup>

$$\widehat{D}(\widehat{p}) = \begin{cases} \int_{v(\widehat{p})}^1 \int_0^{c(v, \widehat{p})} dc dv = \frac{(1+x-\widehat{p})^2}{2} & \text{if } x < \widehat{p} < 1+x, \\ \int_0^{1+v(\widehat{p})} \int_0^{c(v, \widehat{p})} dc dv + \int_{1+v(\widehat{p})}^1 \int_0^1 dc dv = 1 - \frac{(1-x+\widehat{p})^2}{2} & \text{if } x-1 < \widehat{p} \leq x. \end{cases} \quad (4)$$

The firm chooses  $\widehat{p}$  to maximize  $\widehat{\Pi}(\widehat{p}) = (\widehat{p} + \alpha)\widehat{D}(\widehat{p})$ , leading to the following first-order condition.

$$\frac{\partial \widehat{\Pi}(\widehat{p})}{\partial \widehat{p}} = \underbrace{\widehat{D}(\widehat{p})}_{\text{Margin effect (+)}} + \underbrace{p \frac{\partial \widehat{D}(\widehat{p})}{\partial \widehat{p}}}_{\text{Volume effect (-)}} + \underbrace{\alpha \frac{\partial \widehat{D}(\widehat{p})}{\partial \widehat{p}}}_{\text{Data Volume effect (-)}} = 0.$$

Apart from the trade-off between classical margin and volume effects the firm faces in choosing price, the presence of data-based revenue reduces the firm's incentive to increase price. Thus the firm will set price lower than the case without the data-based revenue.

Solving the first-order condition for  $\widehat{p}$ , we obtain

$$\widehat{p}^* = \begin{cases} \frac{1+x-2\alpha}{3} & \text{if } x < \widehat{p}^* < 1+x, \\ \frac{2(x-1)-\alpha + \sqrt{6+(x+\alpha-1)^2}}{3} & \text{if } x-1 < \widehat{p}^* \leq x. \end{cases}$$

In the first case, one can verify that  $\widehat{p}^* > x$  if and only if  $\alpha < \alpha_1(x)$  and  $\widehat{p}^* < 1+x$  always holds where

$$\alpha_1(x) := \max \left\{ \frac{1-2x}{2}, 0 \right\}. \quad (5)$$

In the second case, we have  $\widehat{p}^* \leq x$  if and only if  $\alpha \geq \alpha_1(x)$  and  $\widehat{p}^* > x-1$  always holds.<sup>21</sup> In addition, given  $\alpha \geq 0$ , both cases are possible if  $(1-2x)/2 \geq 0$  or  $x \leq 1/2$ ; otherwise, only case 1b is possible. The two cases highlight how the firm's pricing strategy depends on its business model. When the firm earns large data-based revenue in the sense that  $\alpha \geq \alpha_1(x)$ ,

<sup>20</sup>We use  $\widehat{\cdot}$  to indicate the regime before the GDPR and  $\widetilde{\cdot}$  for the regime after the GDPR.

<sup>21</sup>It is straightforward to check that the second-order condition is satisfied under these conditions.

the firm chooses a low price for its service ( $\widehat{p}^* \leq x$ ) to induce more consumers to purchase its service, which translates into more data before the GDPR. But, when its data-based revenue is not significant enough ( $\alpha < \alpha_1(x)$ ), the firm chooses a higher price ( $\widehat{p}^* > x$ ) to increase its usage-based revenue.

Substituting  $\widehat{p}^*$  into the demand, we obtain the equilibrium price and demand as

$$(\widehat{p}^*, \widehat{D}^*) = \begin{cases} \left( \frac{1+x-2\alpha}{3}, \frac{2(1+x+\alpha)^2}{9} \right) & \text{if } \alpha < \alpha_1(x), \\ \left( \frac{2(x-1)-\alpha+\beta}{3}, \frac{6-(x+\alpha-1)^2+(x+\alpha-1)\beta}{9} \right) & \text{if } \alpha \geq \alpha_1(x), \end{cases} \quad (6)$$

where  $\beta := \sqrt{6+(x+\alpha-1)^2}$ . The corresponding profit  $\widehat{\Pi}^*$  and consumer surplus  $\widehat{CS}^*$  can be calculated as

$$\widehat{\Pi}^* = \begin{cases} \widehat{\Pi}_{1a}^* = \frac{2(1+x+\alpha)^3}{27} & \text{if } \alpha < \alpha_1(x), \\ \widehat{\Pi}_{1b}^* = \frac{(2(x+\alpha-1)+\beta)(6-(x+\alpha-1)^2+(x+\alpha-1)\beta)}{27} & \text{if } \alpha \geq \alpha_1(x). \end{cases} \quad (7)$$

$$\widehat{CS}^* = \begin{cases} \widehat{CS}_{1a}^* = \frac{4(1+x+\alpha)^3}{81} & \text{if } \alpha < \alpha_1(x), \\ \widehat{CS}_{1b}^* = \frac{81+18(x+\alpha-1)-2(x+\alpha-1)^3-2(12-(x+\alpha-1)^2)\beta}{81} & \text{if } \alpha \geq \alpha_1(x). \end{cases} \quad (8)$$

The following lemma is a direct consequence of the above discussion.

**Lemma 1** *In equilibrium before the GDPR, the price and demand are as in (6), and the corresponding profit and consumer surplus are as in (7) and (8), respectively.*

- *The equilibrium price increases in the opt-in benefits  $x$  and decreases in the marginal data-based revenue  $\alpha$ .*
- *The equilibrium demand, profit, and consumer surplus are all increasing in  $x$  and  $\alpha$ .*

**Proof:** See the appendix.

The comparative statics results in the above lemma are quite intuitive. First, an increase in the opt-in benefits makes the firm's service more valuable when consumers choose opt-in purchase. Given that opt-in purchase is the only choice consumers have when buying the service before the GDPR, an increase in  $x$  expands the demand for the service. This can be seen clearly from Figure 1: an increase in  $x$  expands the shaded areas since  $v(p) = p - x$  and  $c(1, p) = 1 + x - p$ . This allows the firm to charge a higher price and earn larger profit. Consumer surplus also increases since, given the elastic demand, the firm can extract only part of the additional opt-in benefits through higher price. Next, when  $\alpha$  increases, the firm can benefit more by increasing the volume of data, from which it earns the data-based revenue. Thus it reduces price to attract more opt-in consumers. As a result, an increase in  $\alpha$  decreases the equilibrium price and increases the equilibrium demand. This implies that consumer surplus increases when  $\alpha$  increases. In addition, since the firm's equilibrium profit changes in response

to a change in  $\alpha$  exactly by the amount of its equilibrium demand, a direct consequence of the envelope theorem, the firm's profit also increases when  $\alpha$  increases.

## 4.2 Equilibrium after the GDPR

After the GDPR, a consumer of type  $(v, c)$  chooses opt-in purchase if  $U_S \geq \max\{U_N, 0\}$ , opt-out purchase if  $U_N \geq \max\{U_S, 0\}$ , and no purchase if  $\max\{U_S, U_N\} \leq 0$ . Thus, the set of consumers choosing opt-in purchase is given by

$$\{(v, c) \in [0, 1] \times [0, 1] \mid c \leq v + x - p, x \geq c\},$$

and the set of consumers choosing opt-out purchase is given by

$$\{(v, c) \in [0, 1] \times [0, 1] \mid v \geq p, x < c\}.$$

As before,  $U_S \geq 0$  leads to the threshold condition,  $c \leq c(v, p) = v + x - p$ . Then,  $v(p) = p - x$  is the threshold value of consumer valuation for the service, below which no consumers choose opt-in purchase. Next,  $U_N \geq 0$  if  $v \geq p$ , which leads to  $v_N(p) := p$ , the threshold value of consumer valuation, below which no consumers choose opt-out purchase. Clearly, we have  $v(p) \leq v_N(p)$ . In addition,  $U_S \geq U_N$  holds if and only if  $x \geq c$ . Once again, we have two possible cases where the market is not fully covered, depending on  $v(p) > 0$  or  $v(p) \leq 0$ . These are depicted in Figure 2, where the area labeled opt-in represents the demand from consumers who choose opt-in purchase and the area labeled opt-out is the demand from consumers who choose opt-out purchase. Figure 2a illustrates the demand for case 2a where  $0 < v(p) < 1$ , i.e.,  $x < p < 1 + x$ , and Figure 2b shows the demand for case 2b where  $-1 < v(p) \leq 0$ , i.e.,  $x - 1 < p \leq x$ . As expected, the additional opt-out purchase made available by the GDPR allows consumers to self-select themselves: those with low privacy costs choose opt-in purchase in order to enjoy the opt-in benefits; those with high privacy costs choose opt-out purchase to avoid the privacy costs.

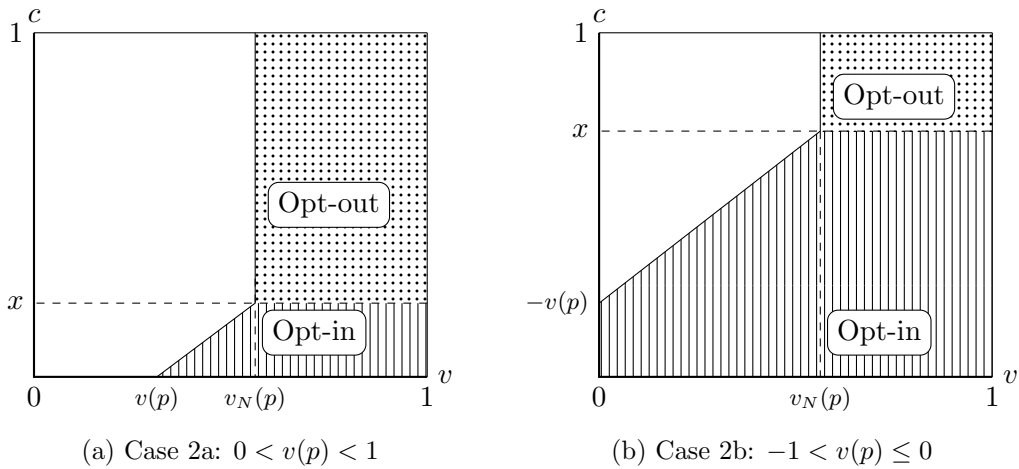


Figure 2: Demand after the GDPR

From the above, we can calculate the opt-in, opt-out, and total demands as follows. First,

the opt-in demand is

$$\tilde{D}_S(\tilde{p}) = \begin{cases} \int_{v(\tilde{p})}^{v_N(\tilde{p})} \int_0^{c(v,\tilde{p})} dcdv + \int_{v_N(\tilde{p})}^1 \int_0^x dcdv = \frac{x^2}{2} + x(1 - \tilde{p}) & \text{if } x < \tilde{p} < 1 + x, \\ \int_0^{v_N(\tilde{p})} \int_0^{c(v,\tilde{p})} dcdv + \int_{v_N(\tilde{p})}^1 \int_0^x dcdv = x - \frac{\tilde{p}^2}{2} & \text{if } x - 1 < \tilde{p} \leq x. \end{cases} \quad (9)$$

The expression for the opt-out demand is the same in both cases and is given by

$$\tilde{D}_N(\tilde{p}) := \int_{v_N(\tilde{p})}^1 \int_x^1 dcdv = (1 - x)(1 - \tilde{p}). \quad (10)$$

The total demand is then

$$\tilde{D}(\tilde{p}) = \tilde{D}_S(\tilde{p}) + \tilde{D}_N(\tilde{p}) = \begin{cases} \frac{x^2}{2} + 1 - \tilde{p} & \text{if } x < \tilde{p} < 1 + x, \\ 1 - (1 - x)\tilde{p} - \frac{\tilde{p}^2}{2} & \text{if } x - 1 < \tilde{p} \leq x. \end{cases} \quad (11)$$

The firm chooses  $\tilde{p}$  to maximize  $\tilde{\Pi}(\tilde{p}) = \tilde{p}\tilde{D}(\tilde{p}) + \alpha\tilde{D}_S(\tilde{p})$ . Compared to the problem before the GDPR, the firm's profit comes from two different sources. Its revenue from sales depends on the total demand but the data-based revenue depends only on the opt-in demand. Solving the firm's problem, we obtain  $\tilde{p}^* = (2 - 2\alpha x + x^2)/4$  if  $x < \tilde{p}^* < 1 + x$  and  $\tilde{p}^* = (\Gamma - 2 - \alpha + 2x)/3$  if  $x - 1 < \tilde{p}^* \leq x$  where  $\Gamma := \sqrt{6 + (2 - 2x + \alpha)^2}$ . In the first case, one can verify that  $\tilde{p}^* < 1 + x$  always holds and  $\tilde{p}^* > x$  if and only if  $\alpha < \alpha_2(x)$  where

$$\alpha_2(x) := \max \left\{ \frac{2 - 4x + x^2}{2x}, 0 \right\}. \quad (12)$$

In the second case, we have  $\tilde{p}^* \leq x$  if and only if  $\alpha \geq \alpha_2(x)$  and  $\tilde{p}^* > x - 1$  always holds.<sup>22</sup> Also, given  $\alpha \geq 0$ , both cases are possible if  $(2 - 4x + x^2)/(2x) \geq 0$  or  $x \leq 2 - \sqrt{2} \approx 0.585$ ; otherwise, only case 2b is possible. Put together, we can write the equilibrium price after the GDPR as

$$\tilde{p}^* = \begin{cases} \frac{2(1 - \alpha x) + x^2}{4} & \text{if } \alpha < \alpha_2(x), \\ \frac{\Gamma - (2 - 2x + \alpha)}{3} & \text{if } \alpha \geq \alpha_2(x). \end{cases} \quad (13)$$

Substituting  $\tilde{p}^*$  into the opt-in demand, we obtain

$$\tilde{D}_S^* = \begin{cases} \frac{x(2 + 2\alpha x + 2x - x^2)}{4} & \text{if } \alpha < \alpha_2(x), \\ x - \frac{3 + (2 - 2x + \alpha)^2 - (2 - 2x + \alpha)\Gamma}{9} & \text{if } \alpha \geq \alpha_2(x). \end{cases} \quad (14)$$

<sup>22</sup>It is straightforward to check that the second-order condition is satisfied under these conditions.

Similarly, the opt-out and total demands can be calculated as

$$(\tilde{D}_N^*, \tilde{D}^*) = \begin{cases} \left( \frac{(1-x)(2+2\alpha x-x^2)}{4}, \frac{2+2\alpha x+x^2}{4} \right) & \text{if } \alpha < \alpha_2(x), \\ \left( \frac{(1-x)(5+\alpha-2x-\Gamma)}{3}, \frac{6-(2+\alpha-2x-\Gamma)(x+\alpha-1)}{9} \right) & \text{if } \alpha \geq \alpha_2(x). \end{cases} \quad (15)$$

Then, the firm's profit  $\tilde{\Pi}^*$  and consumer surplus  $\widetilde{CS}^*$  are

$$\tilde{\Pi}^* = \begin{cases} \tilde{\Pi}_{2a}^* = \frac{4+4x(x+2\alpha+2\alpha x)+x^2(x-2\alpha)^2}{16}, & \text{if } \alpha < \alpha_2(x), \\ \tilde{\Pi}_{2b}^* = \frac{-27(2-\alpha x)+36(1+2x-\alpha)-9(1+2x-\alpha)^2+(1+2x-\alpha)^3}{27} \\ \quad + \frac{(15-(5-2x+\alpha)(1+2x-\alpha))\Gamma}{27} & \text{if } \alpha \geq \alpha_2(x), \end{cases} \quad (16)$$

$$\widetilde{CS}^* = \begin{cases} \widetilde{CS}_{2a}^* = \frac{12+24\alpha x+12(1+\alpha^2)x^2+4(4+3\alpha)x^3-9x^4}{96} & \text{if } \alpha < \alpha_2(x), \\ \widetilde{CS}_{2b}^* = \frac{\delta(\delta^2-9x\delta+27x^2)-9(10-4x+3x^2+2\alpha)\delta+27(5+3x^2)}{162} & \text{if } \alpha \geq \alpha_2(x), \end{cases} \quad (17)$$

where  $\delta := \Gamma - (2 - 2x + \alpha)$ . The above analysis leads to the following lemma.

**Lemma 2** *In equilibrium after the GDPR, the price is as in (13), the opt-in demand, opt-out demand, and total demand are as in (14) and (15), and the corresponding profit and consumer surplus are as in (16) and (17), respectively.*

- *The equilibrium price decreases in  $x$  if and only if  $\alpha_2(x) > \alpha \geq x$ . The equilibrium price decreases in  $\alpha$ .*
- *Both the equilibrium opt-in demand and total demand increase in  $x$  and  $\alpha$ . The equilibrium opt-out demand increases in  $x$  if and only if  $\alpha_2(x) > \alpha > \alpha_3(x) := \frac{2+2x-3x^2}{2-4x}$ , and always increases in  $\alpha$ .*
- *Both the equilibrium profit and consumer surplus increase in  $x$  and  $\alpha$ .*

**Proof:** See the appendix.

The above comparative statics results are similar to those from the case before the GDPR. For example, an increase in  $\alpha$  decreases the equilibrium price and increases the equilibrium demands, profit, and consumer surplus. One main difference is how the equilibrium price responds to a change in the opt-in benefits. Unlike the case before the GDPR, the equilibrium price can decrease in  $x$  when  $\alpha < \alpha_2(x)$ . This highlights the difference in the firm's pricing incentives. Before the GDPR, an increase in  $x$  increases total demand, all of which generates data-based revenue. Therefore, the firm increases price when  $x$  increases. After the GDPR, an increase in  $x$  increases opt-in demand but decreases opt-out demand if price stays the same, as shown in (9) and (10). This can be also seen from Figure 2. Thus, if the firm raises price in

response to an increase in  $x$ , there will be a further decrease in opt-out demand, which erodes the firm's total revenue. Consequently, when the firm increases price in response to an increase in  $x$ , the firm weighs increased data-based revenue thanks to increased opt-in demand against reduced usage-based revenue due to decreased opt-out demand. This implies that the firm is less likely to increase price when its data-based revenue is not significant, or  $\alpha$  is small.

## 5 Comparing the Equilibria before and after the GDPR

Our analysis in the previous section shows that different equilibria exist for different values of  $(x, \alpha)$ . For example, a set of  $(x, \alpha)$  may lead to case 1a equilibrium before the GDPR but case 2b equilibrium after the GDPR. Therefore, in comparing the equilibria before and after the GDPR, we need to be precise about which pair of equilibria exists for given values of  $(x, \alpha)$ . Before the GDPR, we have case 1a equilibrium if  $\alpha \leq \alpha_1(x)$  and case 1b equilibrium if  $\alpha > \alpha_1(x)$  where  $\alpha_1(x)$  is given in (5). Since  $\alpha \geq 0$ , case 1a equilibrium exists only if  $x \leq 1/2$  because  $\alpha_1(x) = 0$  if  $x > 1/2$ . This implies that case 1a equilibrium exists for  $(x, \alpha) \in \{(x, \alpha) | \alpha \leq \alpha_1(x), x \leq 1/2\}$ , and case 1b equilibrium exists for  $(x, \alpha) \in \{(x, \alpha) | \alpha > \alpha_1(x), x \leq 1/2\}$  or  $(x, \alpha) \in \{(x, \alpha) | \alpha > \alpha_1(x), x > 1/2\}$ . Applying a similar argument to the case after the GDPR, we can conclude that case 2a equilibrium exists for  $(x, \alpha) \in \{(x, \alpha) | \alpha \leq \alpha_2(x), x \leq 2 - \sqrt{2}\}$  and case 2b equilibrium exists for  $(x, \alpha) \in \{(x, \alpha) | \alpha > \alpha_2(x)\}$  where  $\alpha_2(x)$  is given in (12). In addition, we can verify  $\alpha_2(x) \geq \alpha_1(x)$ , which implies that we cannot have case 1a equilibrium before the GDPR co-existing with case 2b equilibrium after the GDPR. Based on these discussions, we can identify a pair of equilibria corresponding to all possible values of  $(x, \alpha)$ , which we summarize below.

**Lemma 3** *A pair of equilibria before and after the GDPR can be fully described as follows:*

- *If  $x \leq 1/2$  and  $\alpha < \alpha_1(x)$ , then we have case 1a equilibrium before the GDPR and case 2a equilibrium after the GDPR.*
- *If  $x \leq 2 - \sqrt{2}$  and  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , then we have case 1b equilibrium before the GDPR and case 2a equilibrium after the GDPR.*
- *If  $x \leq 1$  and  $\alpha \geq \alpha_2(x)$ , then we have case 1b equilibrium before the GDPR and case 2b equilibrium after the GDPR.*

**Proof:** From the preceding discussions, we can identify three possible pairs of equilibria as follows. First, if  $x \leq 1/2$ , then we have (case 1a equilibrium, case 2a equilibrium) if  $\alpha < \alpha_1(x) \leq \alpha_2(x)$ , (case 1b equilibrium, case 2a equilibrium) if  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , and (case 1b equilibrium, case 2b equilibrium) if  $\alpha_1(x) \leq \alpha_2(x) \leq \alpha$ . Second, if  $1/2 < x \leq 2 - \sqrt{2}$ , then we have (case 1b equilibrium, case 2a equilibrium) if  $\alpha < \alpha_2(x)$ , and (case 1b equilibrium, case 2b equilibrium) if  $\alpha \geq \alpha_2(x)$ . Third, if  $2 - \sqrt{2} < x \leq 1$ , then we have (case 1b equilibrium, case 2b equilibrium) for all  $\alpha$ . The lemma follows from this. ■

Our first result is that the equilibrium price for the service increases after the GDPR. It is because the GDPR enables consumers to better manage their privacy. For example, consumers



with high privacy costs who did not purchase the service before the GDPR can choose opt-out purchase after the GDPR. Specifically, suppose a consumer of type  $(v, c)$  did not purchase the service before the GDPR when the price for the service is  $p$ . This implies  $v + x < p + c$ . After the GDPR, the consumer can choose opt-out purchase if  $v \geq p$ . On the other hand, a consumer who purchased the service before the GDPR will continue to do so at the same price after the GDPR; the only difference is that she will choose opt-out purchase if  $v + x > p + c$  but  $x < c$ . Put together, we can conclude that, at any given price, the firm faces a weakly larger demand for its service after the GDPR. This allows the firm to increase price. A higher price can erode the firm's data-based revenue by reducing the amount of data shared. But the effect is smaller after the GDPR because the firm earns data-based revenue only from the opt-in demand after the GDPR whereas its data-based revenue is from the total demand before the GDPR. Consequently, the firm is less constrained by data-based revenue after the GDPR. Thus, it charges a higher price for its service after the GDPR irrespective of the size of its data-based revenue.

**Proposition 1** *The equilibrium price is higher after the GDPR, i.e.,  $\hat{p}^* > \hat{p}^*$ .*

**Proof:** See the appendix.

Our next result relates to how the GDPR changes the equilibrium demand. Other things equal, the GDPR should expand demand because, as explained above, it provides consumers with an additional choice of opt-out purchase. We call this the demand expansion effect. But the price increase following the GDPR as shown in Proposition 1 acts as a countervailing force. If the price effect more than offsets the demand expansion effect, then the GDPR can reduce the equilibrium demand; otherwise, it will increase the equilibrium demand. When  $\alpha$  is small, say  $\alpha < \alpha_1(x)$ , we have case 1a equilibrium before the GDPR and case 2a equilibrium after the GDPR. This is the case where the market is not sufficiently covered before the GDPR in the sense that  $\hat{v}^* > 0$ . Thus, there is significant room for demand expansion. Consequently, the demand expansion effect dominates the price effect, implying that the equilibrium demand increases after the GDPR. When  $\alpha$  is large in that  $\alpha \geq \alpha_2(x)$ , the equilibrium shifts from case 1b equilibrium before the GDPR to case 2b equilibrium after the GDPR. This is the case where the market is sufficiently covered before the GDPR in the sense that  $\hat{v}^* < 0$ , hence there is not much room for demand expansion. In this case, we expect the price effect to dominate and, accordingly, the equilibrium demand decreases after the GDPR.

Although the effect of the GDPR on the total equilibrium demand hinges on the size of the firm's data-based revenue, the effect of the GDPR on the amount of data shared is unambiguously negative. That is, the equilibrium opt-in demand after the GDPR is always smaller than the total equilibrium demand before the GDPR. The intuition is clear. The additional choice of opt-out purchase available after the GDPR allows consumers to better manage their privacy. Accordingly, some consumers with high privacy costs who purchased the service before the GDPR can switch to opt-out purchase after the GDPR. This implies that, even when the GDPR increases the total equilibrium demand, we expect the demand from consumers who choose opt-in purchase to be smaller than that before the GDPR. As discussed previously, the negative effect of the GDPR on the amount of data collected is well documented (Schmitt et al.,

2021; Congiu et al., 2022; Aridor et al., forthcoming). We also note that the marginal consumer with the lowest privacy cost who shares data has a higher valuation for the service after the GDPR, i.e.,  $\tilde{v}^* > \hat{v}^*$ . This follows directly from Proposition 1 since  $\tilde{v}^* = \tilde{p}^* - x > \hat{p}^* - x = \hat{v}^*$ . This is more or less consistent with the empirical findings in Aridor et al. (forthcoming) that, although the GDPR has resulted in a reduction in total cookies, the consumers who continue to be observed after the GDPR tend to be more valuable to advertisers.<sup>23</sup>

**Proposition 2** *Comparing the equilibria before and after the GDPR, we find the following.*

- When  $\alpha < \alpha_1(x)$ , the total demand is larger after the GDPR, i.e.,  $\tilde{D}^* > \hat{D}^*$ .
- When  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , we have  $\tilde{D}^* > \hat{D}^*$  if and only if  $\alpha < \min\{\alpha_2(x), \alpha_4(x)\}$  where  $\alpha_4(x)$  is the value of  $\alpha$  that solves  $4\beta(1-x-\alpha) + 4\alpha^2 + 26x\alpha + 13x^2 - 2 - 8(x+\alpha) = 0$ .
- When  $\alpha \geq \alpha_2(x)$ , we have  $\tilde{D}^* > \hat{D}^*$  if and only if  $\alpha < 1-x$ .
- For all values of  $\alpha$ , the opt-in demand after the GDPR is smaller than the total demand before the GDPR, i.e.,  $\hat{D}^* > \tilde{D}_S^*$ .
- The marginal consumer with lowest privacy cost who shares data has a higher valuation for the service after the GDPR:  $\tilde{v}^* > \hat{v}^*$ .

**Proof:** See the appendix.

Proposition 2 can be better understood with help of Figure 3, where  $\hat{v}^* = \hat{p}^* - x$  and  $\hat{c}^* = 1 + x - \hat{p}^*$  correspond to the equilibrium before the GDPR, and  $\tilde{v}^* = \tilde{p}^* - x$  and  $\tilde{v}_N^* = \tilde{p}^*$  correspond to the equilibrium after the GDPR. Since  $\tilde{p}^* > \hat{p}^*$  as shown in Proposition 1, we have  $\hat{v}^* < \tilde{v}^* < \tilde{v}_N^*$ .

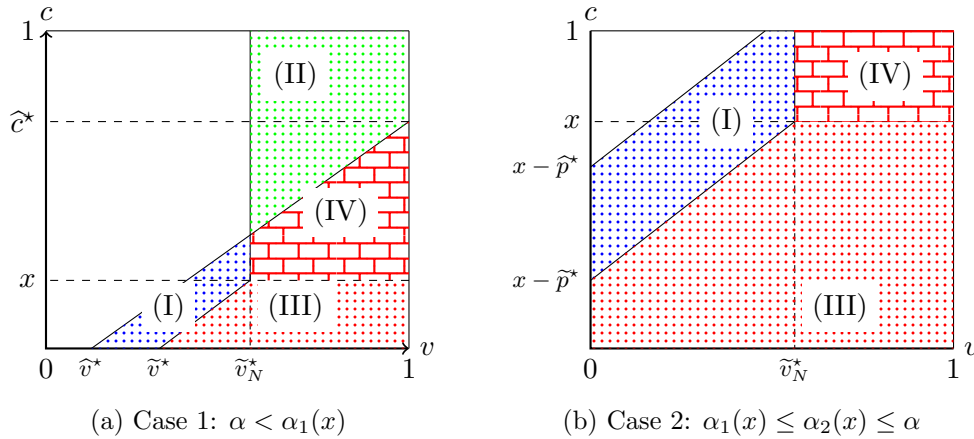


Figure 3: Comparison of equilibrium demands

In Figure 3a, we illustrate the total equilibrium demand when  $\alpha < \alpha_1(x)$  so that the comparison is between case 1a equilibrium before the GDPR and case 2a equilibrium after the

<sup>23</sup>Godinho de Matos and Adjerid (2022) report experimental evidence that suggests an alternative channel that the GDPR can generate benefits to advertisers. That is, the GDPR can lead to varying amounts of data collected from different types of data, which can help improve targeting capabilities by advertisers.

GDPR. Before the GDPR, the total equilibrium demand is represented by the triangle defined by the coordinates  $(\widehat{v}^*, 0)$ ,  $(1, 0)$  and  $(1, \widehat{c}^*)$ , that is, the sum of areas labeled (I), (III) and (IV). In the equilibrium after the GDPR, the opt-out demand is represented by the rectangle defined by the coordinates  $(\widetilde{v}_N^*, x)$ ,  $(\widetilde{v}_N^*, 1)$ ,  $(1, x)$ , and  $(1, 1)$ , i.e., the sum of areas labeled (II) and (IV), and the opt-in demand is given by the trapezoid defined by the coordinates  $(\widetilde{v}^*, 0)$ ,  $(\widetilde{v}_N^*, x)$ ,  $(1, 0)$  and  $(1, x)$ . Thus, the total demand after the GDPR is the sum of areas labeled (II), (III) and (IV). To summarize this case, the sum of areas labeled (III) and (IV) captures consumers who buy the service under both privacy regimes, the area labeled (I) represents consumers who drop out of the market after the GDPR, and the area labeled (II) represents consumers who buy the service only after the GDPR. Proposition 2 shows that the area (II) in case 2a equilibrium is larger than the area (I) in case 1a equilibrium.

These changes in demand can be explained by considering the behavior of four groups of consumers. The first group is consumers in area (I), who drop out of the market after the GDPR because the firm raises its price after the GDPR. They have relatively low valuations for the service and low privacy costs. Thus, the additional opt-out choice made available by the GDPR does not have much value for them. The second group is consumers in area (II), who buy the service only after the GDPR. They have relatively high valuations and high privacy costs. Due to high privacy costs, they do not buy the service before the GDPR but, after the GDPR, they choose opt-out purchase. The third group is consumers in area (III), who buy the service under both privacy regimes and share data as they have relatively low privacy costs ( $c \leq x$ ). The fourth group is consumers in area (IV), who buy the service under both privacy regimes but choose opt-out purchase after the GDPR due to relatively high privacy costs ( $c > x$ ). Thus the GDPR creates winners and losers. The first and third groups of consumers are worse off after the GDPR. The second group of consumers is better off while the welfare change for the fourth group depends on the trade-off between the higher price and better privacy management.

In Figure 3b, we plot the total equilibrium demand when  $\alpha \geq \alpha_2(x)$  so that the comparison is between case 1b equilibrium before the GDPR and case 2b equilibrium after the GDPR. Specifically, we illustrate the case with  $\alpha > 1 - x$ , where the equilibrium demand before the GDPR is the sum of areas (I), (III) and (IV), and the equilibrium demand after the GDPR is the sum of areas (III) and (IV). This is the case where the market is sufficiently covered before the GDPR so that there is no new demand created after the GDPR. Moreover, as shown in the figure, the total equilibrium demand decreases after the GDPR because the price increase after the GDPR is sufficiently high so that consumers with high privacy costs and low valuations who purchased the service before the GDPR drop out of the market after the GDPR. These consumers are represented by area (I) in Figure 3b.<sup>24</sup>

Next, we discuss how the GDPR affects the firm's profit. When the firm's revenue does not depend significantly on data in that  $\alpha < \alpha_1(x)$ , the GDPR benefits the firm. This is the direct consequence of Propositions 1 and 2 since, when  $\alpha < \alpha_1(x)$ , both the equilibrium price and demand increase after the GDPR. As  $\alpha$  becomes large, the increase in price can be countered by the decrease in total demand, making the effect on profit less obvious. But when  $\alpha$  is sufficiently

<sup>24</sup>The remaining case is the comparison of case 1b equilibrium before the GDPR with case 2a equilibrium after the GDPR when  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ . We omit the discussion for brevity.

large, the decrease in demand can dominate the price increase. It is because the equilibrium price decreases in  $\alpha$  as shown in Lemmas 1 and 2. The decrease in demand combined with a small increase in price make the firm worse off when  $\alpha$  is large enough.

**Proposition 3** *Comparing the equilibria before and after the GDPR, we find the following.*

- When  $\alpha < \alpha_1(x)$ , the firm's profit is larger after the GDPR, i.e.,  $\tilde{\Pi}_{2a}^* > \hat{\Pi}_{1a}^*$ .
- When  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , we have  $\tilde{\Pi}_{2a}^* > \hat{\Pi}_{1b}^*$  if and only if  $\alpha \leq \min\{\alpha_2(x), \alpha_5(x)\}$  where  $\alpha_5(x)$  is the value of  $\alpha$  that solves  $\tilde{\Pi}_{2a}^* = \hat{\Pi}_{1b}^*$ .
- When  $\alpha \geq \alpha_2(x)$ , we have  $\tilde{\Pi}_{2b}^* > \hat{\Pi}_{1b}^*$  if and only if  $\alpha_2(x) < \alpha < \alpha_6(x)$  where  $\alpha_6(x)$  is the value of  $\alpha$  that solves  $\tilde{\Pi}_{2b}^* = \hat{\Pi}_{1b}^*$ .

**Proof:** See the appendix.

Figure 4 illustrates Proposition 3. In the figure, the blue shaded region labeled  $\Delta\Pi > 0$  is where the firm's profit increases after the GDPR and the yellow shaded region labeled  $\Delta\Pi < 0$  is where the firm's profit decreases after the GDPR. The curved line dividing these two regions is defined by  $\alpha_5(x)$  and  $\alpha_6(x)$  stated in Proposition 3. The region labeled 'Reg 1' corresponds to the case where  $\alpha < \alpha_1(x)$ , 'Reg 2' corresponds to the case where  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , and 'Reg 3' corresponds to the case where  $\alpha \geq \alpha_2(x)$ . As shown in the figure, for given values of  $x$ , the GDPR benefits the firm for smaller values of  $\alpha$  but hurts the firm for larger values of  $\alpha$ , with smaller values of  $x$  admitting a larger range of  $\alpha$  that benefits the firm. It is because, the smaller  $x$  is, the larger is the mass of consumers who participate in the market through opt-out purchase after the GDPR, as can be checked in Figure 2. This leads to a larger demand expansion effect.

An implication is that the GDPR can benefit digital businesses that earn revenue primarily from selling services. Examples of such usage-driven businesses include ride-hailing platforms such as Uber, Lyft, streaming services such as Netflix, Spotify, and, to some extent, online market places such as Amazon and eBay. But the GDPR can hurt digital businesses whose revenue is in large part based on data. Such data-driven businesses include various navigation and weather apps and, to some extent, Google and Facebook whose main revenue source is advertising that leverages user data.

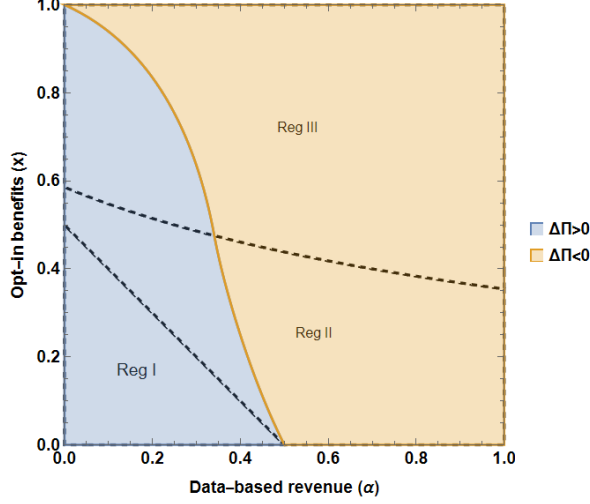


Figure 4: Comparison of equilibrium profits,  $\Delta\Pi = \tilde{\Pi}^* - \hat{\Pi}^*$

Then, how does consumer surplus change after the GDPR? As discussed previously, consumers with low privacy costs are likely to be worse off because of the higher price after the GDPR, but those with high privacy costs are likely to be better off thanks to better privacy management. In particular, all consumers with  $c < x$  are worse off either because they drop out of the market or because they pay a higher price after the GDPR. On balance, the GDPR is more likely to increase consumer surplus when  $\alpha$  is smaller for the following reasons. First, when  $\alpha$  is small, the equilibrium price before the GDPR is already close to the monopoly level and, therefore, the price increase following the GDPR is limited. Second, as explained earlier, there is significant room for demand expansion when  $\alpha$  is small. Thus the GDPR's demand expansion effect is more likely to dominate the price effect for smaller values of  $\alpha$ . Put together, the GDPR's effect on consumer surplus hinges on  $(x, \alpha)$  such that, for smaller value of  $x$ , the GDPR can admit a wider range of  $\alpha$  for which consumer surplus increases after the GDPR. As  $x$  increases, the benefits from privacy management decrease, implying that the range of  $\alpha$  for which consumer surplus increases after the GDPR shrinks.

**Proposition 4** *Comparing the equilibria before and after the GDPR, we find the following.*

- When  $\alpha < \alpha_1(x)$ , the consumer surplus is larger after the GDPR, i.e.,  $\widetilde{CS}_{2a}^* > \widehat{CS}_{1a}^*$ , if and only if  $\alpha < \alpha_7(x)$  where  $\alpha_7(x)$  is the value of  $\alpha$  that solves  $\widetilde{CS}_{2a}^* = \widehat{CS}_{1a}^*$ .
- When  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , we have  $\widetilde{CS}_{2a}^* > \widehat{CS}_{1b}^*$  if and only if  $\alpha < \alpha_8(x)$  where  $\alpha_8(x)$  is the value of  $\alpha$  that solves  $\widetilde{CS}_{2a}^* = \widehat{CS}_{1b}^*$ .
- When  $\alpha \geq \alpha_2(x)$ , we have  $\widetilde{CS}_{2b}^* > \widehat{CS}_{1b}^*$  if and only if  $\alpha < \alpha_9(x)$  where  $\alpha_9(x)$  is the value of  $\alpha$  that solves  $\widetilde{CS}_{2b}^* = \widehat{CS}_{1b}^*$ .

**Proof:** See the appendix.

Figure 5 illustrates Proposition 4. In the figure, the blue shaded region labeled  $\Delta CS > 0$  is where the equilibrium consumer surplus is larger after the GDPR and the yellow shaded region

labeled  $\Delta CS < 0$  is where it is smaller after the GDPR. The line dividing these two regions is made up of three pieces: the lowermost part corresponds to the case where  $\alpha < \alpha_1(x)$  (Region I) and is defined by  $\alpha_7(x)$  stated in Proposition 4; the curve in the middle corresponds to the case where  $\alpha_1(x) \leq \alpha < \alpha_2(x)$  (Region II) and is defined by  $\alpha_8(x)$ ; and the uppermost part corresponds to the case where  $\alpha \geq \alpha_2(x)$  (Region III) and is defined by  $\alpha_9(x)$ .

As shown in the figure, for given values of  $x$ , the GDPR increases consumer surplus for smaller values of  $\alpha$  but decreases it for larger values of  $\alpha$ . The intuition is straightforward. When  $\alpha$  is small, there is much room for demand expansion, which is realized after the GDPR, as shown in Proposition 2. The increased consumer participation following the GDPR leads to larger consumer surplus. When  $\alpha$  is large, however, the GDPR does not increase consumer participation sufficiently enough and, in some cases, even dampens it due to higher price. In addition, the set of  $(x, \alpha)$  that leads to larger consumer surplus is smaller than the set of  $(x, \alpha)$  that leads to larger profit, as shown in Figures 4 and 5. This is because the price increase following the GDPR partially offsets the benefits to consumers. In sum, the implication is that the GDPR can benefit consumers in markets that are largely service-based, but it can hurt consumers when data-based revenue is a significant part of digital business.

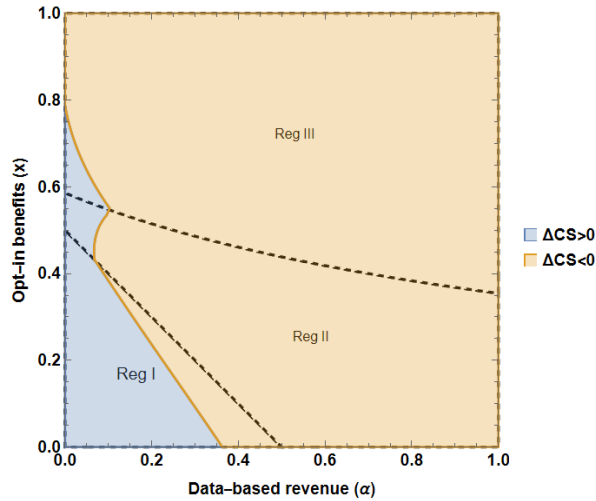


Figure 5: Comparison of consumer surpluses,  $\Delta CS = \widetilde{CS}^* - \widehat{CS}^*$

## 6 Discussions and Extensions

### 6.1 Correlation between $v$ and $c$

Our baseline model assumed that  $v$  and  $c$  are independent. Depending on digital services, it is conceivable that there could be some correlations between the two. Take social media services, for example. A consumer who actively uses social media services may do so because she does not perceive the privacy cost to be high, which suggests a negative correlation between  $v$  and  $c$ .<sup>25</sup> The correlation can be positive in medical services and health apps. These services often require users to share sensitive health data and, therefore, privacy-conscious individuals would

<sup>25</sup>A study by [Pew Research Center \(2019\)](#) suggests that individuals who are less concerned about privacy are more likely to share more personal information on social media.

choose to share such information as they value these services highly. In this section, we allow correlations between  $v$  and  $c$ . In particular, we examine two polar cases of perfect positive and perfect negative correlations. The insight from these two cases may be used for extrapolation in intermediate cases.

□ Positive correlation. Suppose  $c = \phi v$  where  $\phi \in (0, 1)$ , and assume  $x < \phi$  so that there is a positive mass of consumers who opt out after the GDPR, i.e., consumers with  $x < c = \phi v$ , or  $v > \tilde{v} := x/\phi$ . In this case, the GDPR's demand expansion effect does not exist. The reason is that, given the perfect positive correlation, consumers who do not buy the service before the GDPR do not buy the service either after the GDPR. To see this, let  $\hat{v}$  be the marginal consumer who buys the service before the GDPR, i.e.,  $\hat{v} + (x - \phi\hat{v}) - p = 0$  and  $[\hat{v}, 1]$  is the total demand before the GDPR. After the GDPR, the positive correlation between  $v$  and  $c$  implies that consumers on  $[\hat{v}, \tilde{v}]$  continue to choose opt-in purchase while those on  $[\tilde{v}, 1]$  choose opt-out purchase. In addition, by the definition of  $\tilde{v}$  and the fact that  $\tilde{v} > \hat{v}$ , no consumers outside  $[\hat{v}, 1]$  participate in the market through opt-out purchase.

Consequently, the GDPR does not boost demand; it only divides the pre-GDPR total demand into opt-in and opt-out demands. As a result, the equilibrium price does not change after the GDPR. Given that the equilibrium price and total demand do not change, the firm is worse off after the GDPR. It is because the firm earns data-based revenue only from the opt-in demand after the GDPR, which is smaller than the total demand because consumers on  $[\tilde{v}, 1]$  choose opt-out purchase after the GDPR. Finally, consumer surplus increases after the GDPR because the price stays the same while those choosing opt-out purchase benefit from privacy management. To summarize, the GDPR can hurt the firm and benefit consumers when  $v$  and  $c$  are positively correlated. It is worth noting that these results hold regardless of  $\alpha$ .

□ Negative correlation. Suppose now  $c = 1 - \phi v$  where we set  $\phi = 1$  to simplify analysis. Then, there is a positive mass of consumers who opt out after the GDPR because  $x > 0$ . Let  $\tilde{v} = 1 - x$  be the marginal consumer, who is indifferent between opt-in purchase and opt-out purchase if  $\tilde{v} \geq p$ . Then, for all  $v < \tilde{v}$ , we have  $c = 1 - v > 1 - \tilde{v} = x$ . It follows that all consumers on  $[p, \tilde{v}]$  choose opt-out purchase. That is, the negative correlation between  $v$  and  $c$  implies that consumers who choose opt-out purchase are those with lower valuations, hence higher privacy costs, than  $\tilde{v}$ , which is in contrast to the case with positive correlation. This implies that the GDPR can expand demand by inducing participation by consumers with low valuations who would not have opted in before the GDPR. More specifically, since consumers with low valuations choose opt-out purchase, the total demand after the GDPR is given by  $D(p) = 1 - p$ , which is independent of  $x$ . On the other hand, the demand before the GDPR increases in  $x$ . Thus, the GDPR's demand expansion effect depends on  $x$ : the GDPR is more likely to expand demand when  $x$  is smaller and decrease it when  $x$  is larger. To summarize, when  $v$  and  $c$  are negatively correlated, the GDPR's welfare effects depend on  $x$  and  $\alpha$ , in contrast to the case with positive correlation.

**Proposition 5** *Suppose  $v$  and  $c$  have a perfect positive correlation given by  $c = \phi v$ ,  $\phi > x$ ,  $\phi \in (0, 1)$  and assume  $\alpha < 1 + x - \phi$  which ensures that the equilibrium price is positive. Then, comparing the equilibria before and after the GDPR,*

- the price and demand are the same;
- the firm's profit is smaller but the consumer surplus is larger after the GDPR.

Suppose  $v$  and  $c$  have a perfect negative correlation given by  $c = 1 - v$  and assume  $\alpha < 1 + x$  which ensures that the equilibrium price is positive. Then, comparing the equilibria before and after the GDPR,

- the price is higher after the GDPR if and only if  $\alpha > x$ ;
- the total demand is larger after the GDPR if and only if  $\alpha < 1 - x$ ;
- there are threshold values  $\alpha_L$  and  $\alpha_H$  with  $\alpha_L < \alpha_H$  such that the profit is larger after the GDPR if and only if  $\max\{0, \alpha_L\} < \alpha < \min\{\alpha_H, 1 + x\}$ ;
- there is a threshold value  $\alpha_{CS}$  such that the consumer surplus is larger after the GDPR if and only if  $\alpha < \min\{\alpha_{CS}, 1 + x\}$ .

**Proof:** See the appendix.

## 6.2 Data externalities

A consumer's data-sharing decision can impose several types of externalities on others. First, consumers using the firm's service may enjoy network benefits that increase as more consumers share data with the firm (Hagi and Wright, forthcoming). For example, online retailers such as Amazon and Etsy rely on various user-generated data including product reviews and ratings to provide valuable information to consumers. Network benefits may also include improved products and services that sellers on the marketplaces can provide based on customer feedback. Second, data sharing by some consumers can impose negative externalities in the form of larger privacy costs on others who do not share data (Choi et al., 2019; Ichihashi, 2021; Bergemann et al., 2022; Miklós-Thal et al., 2023). For example, some people's shared data can be used to infer information about others who do not share data. As more data is shared, the inference can be improved through machine learning, which leads to larger privacy costs. As another example, hackers can use personal data obtained from one group of users to craft convincing phishing emails or other social engineering attacks targeted at another group of users. Third, there can be negative pecuniary externalities whereby the firm can use the data on opt-in consumers to infer information on and charge higher price to opt-out consumers (Belleflamme and Vergote, 2016; Braghieri, 2019). In this section, we consider the first two, positive network benefits and negative data externalities, and discuss how our main results may change.

□ Positive network benefits. Denote network benefits by  $\theta D_S$  where  $\theta \geq 0$  is an exogenously given value accruing from an additional data-sharing consumer. We assume that network benefits are enjoyed by all consumers that buy the service whether or not they share data with the firm. For example, even if a consumer makes a purchase at an online shop without creating an account, she can still read all the reviews and ratings. Thus, the main difference between opt-in



and opt-out purchase is that a consumer choosing opt-in purchase expects additional opt-in benefits as well as privacy costs, which is the same as in the baseline model.

When there are network benefits, a consumer's opt-in decision depends on her expectation on the mass of opt-in consumers, which we denote by  $D_S^e$ . Then, a consumer of type  $(v, c)$  derives expected utility

$$U(v, c, D_S^e) = \begin{cases} U_S := v + (x - c) + \theta D_S^e - p, & \text{with opt-in purchase,} \\ U_N := v + \theta D_S^e - p, & \text{with opt-out purchase,} \\ U_O := 0, & \text{without purchase.} \end{cases}$$

As is clear from the above, consumers are more likely to purchase the firm's service because both  $U_S - U_O$  and  $U_N - U_O$  increase by  $\theta D_S^e$  compared to the case without network benefits. But network benefits do not affect their opt-in decisions since  $U_S - U_N$  is the same with or without network benefits. Given that the GDPR's main effect in our model is to allow consumers to make opt-in decisions proactively, we expect our key insight from the baseline model to remain unchanged.

□ Negative data externalities. Suppose negative data externalities are proportionate to the consumer's privacy cost and the expected mass of opt-in consumers,  $D_S^e$ . Let  $\gamma \geq 0$  represent the magnitude of negative externalities, so that consumer  $(v, c)$  suffers total negative externalities equal to  $\gamma c D_S^e$ . We can consider two possible ways negative externalities can be introduced to the baseline model. First, suppose negative externalities affect all consumers including those who share data, as in [Choi et al. \(2019\)](#). In this case, all consumers' utility decreases by  $\gamma c D_S^e$ , whether they choose opt-in purchase, opt-out purchase, or no purchase. Consequently, the analysis is exactly the same as before.

Suppose now that negative externalities affect consumers who do not share data, whether or not they buy the service. Then, consumer  $(v, c)$  derives utility

$$U(v, c, D_S^e) = \begin{cases} U_S := v + (x - c) - p, & \text{with opt-in purchase,} \\ U_N := v - \gamma c D_S^e - p, & \text{with opt-out purchase,} \\ U_O := -\gamma c D_S^e, & \text{without purchase.} \end{cases}$$

As is clear from the above, negative externalities make opt-in a more attractive choice than opt-out: compared to the baseline model,  $U_S - U_N$  increases by  $\gamma c D_S^e$ . In addition,  $U_S - U_O$  also increases by  $\gamma c D_S^e$  but  $U_N - U_O$  stays the same. Put together, we can conclude that negative externalities increase the proportion of consumers who choose opt-in purchase. This diminishes the value of an option to choose opt-out purchase, which dampens the effect of the GDPR. But our key insight does not change because the main mechanism that the GDPR can create more demand remains intact even in the presence of negative externalities. That is, the GDPR allows consumers with large privacy costs who stay out of the market before the GDPR to participate in the market through opt-out purchase; negative externalities reduce, but not eliminate, the proportion of such consumers. In addition, by making opt-in a more attractive choice than opt-out, negative externalities also reduce the proportion of consumers who exit the

market after the GDPR because of the high price.

### 6.3 When consumers differ only in valuations of the service

One of our key results is how the GDPR leads to a substitution between data quantity and the value of data as summarized in Proposition 2, i.e.,  $\tilde{D}_S^* < \hat{D}^*$  but  $\tilde{v}^* > \hat{v}^*$ . This result depends crucially on consumer heterogeneity in privacy costs. To verify this claim, we briefly discuss the case without consumer heterogeneity in privacy costs. Suppose  $c$  is constant and denote  $z := x - c$  so that consumer type is represented by  $v$  only.

Before the GDPR, consumer  $v$  buys the service if  $v \geq p - z$ , hence  $D = 1 - (p - z)$ . Maximizing  $\Pi = (p + \alpha)D$  gives us  $\hat{p}^* = (1 + z - \alpha)/2$  and  $\hat{D}^* = (1 + z + \alpha)/2$ . After the GDPR, we have two cases. First, if  $z \geq 0$ , then there is only opt-in purchase, implying that the outcome is the same as that before the GDPR. In this case, the GDPR has no effect on the amount of data shared. Second, if  $z < 0$ , then there is only opt-out purchase. Then  $D = D_N = 1 - p$  and  $D_S = 0$ . Maximizing  $\Pi = pD + \alpha D_S$  leads to  $\tilde{p}^* = \tilde{D}^* = 1/2$  and  $\tilde{D}_S^* = 0$ .

Comparing the above and noting that  $z < 0$  in this case, we find  $\tilde{p}^* > \hat{p}^*$  always holds and  $\tilde{D}^* > \hat{D}^*$  if and only if  $\alpha + z < 0$ . That is, the equilibrium price increases after the GDPR, and the equilibrium demand increases if and only if  $\alpha < c - x$ . It is also straightforward to check that the equilibrium profit increases after the GDPR if and only if  $\alpha < c - x$ . But the quantity of shared data either does not change or decreases to zero. Thus, there is no possibility that the GDPR reduces the quantity of data while increasing the marginal consumer's valuation for the firm's service. This is intuitively clear since, in the absence consumer heterogeneity in privacy costs, there is no sense in which the GDPR can invite consumers with high privacy costs to choose opt-out purchase.

### 6.4 Testable hypotheses

In this section, we summarize testable hypotheses that can be derived from our analysis and discuss the issues relevant to testing these hypotheses.

Our main hypotheses can be summarized as follows. First, for digital businesses whose main revenue source is selling service, empowering consumers to make informed and proactive decisions for data sharing can increase price for the service, consumer activity, business profitability, and consumer surplus. Second, for digital businesses whose main revenue source is data which is collected through selling service, more effective privacy management by consumers can increase price for the service, but decrease consumer activity, business profitability, and consumer surplus. As discussed earlier, the examples of usage-driven businesses include various ride-hailing platforms and streaming services and, to some extent, online market places, while data-driven businesses include various navigation and weather apps and, to some extent, digital platforms whose main revenue source is advertising that leverages user data.

There is an important caveat in testing the above hypotheses. Our model assumes fully-informed consumers who can rationally choose between opt-in and opt-out purchase after the introduction of new privacy regulations, the GDPR being the focus of our paper. However, one may argue that, even after the GDPR, consumers' ability to manage privacy may be hampered for various reasons. First, the GDPR's cookie rules are not refined enough, which has led

to a proliferation of diverse consent management platforms and failed to enable consumers to effectively manage their privacy (Chen, 2022; Chaudhury and Choe, 2023). Second, many websites continue to rely on dark patterns, implied consent, and various forms of nudging to influence consumers' data sharing decisions (Utz et al., 2019; Machuletz and Böhme, 2020; Matte et al., 2020; Nouwens et al., 2020). Third, consumers generally show apathy toward complex and long privacy notices and select the 'quick join' clickwrap to gain access to a website (Obar and Oeldorf-Hirsch, 2020).

In view of the above issues, the GDPR's welfare implications analyzed in this study need to be understood to apply to an environment where the GDPR achieves its stated purpose of enabling consumers to make an informed and unambiguous consent to data sharing. However, identifying GDPR-compliant consent management platforms that are quarantined from the above compounding factors can be challenging. It can also lead to a selection bias. Therefore, our hypotheses may be best tested in carefully designed experiments.

## 6.5 Implications for management

Our analysis generates at least two clear implications for management relevant to digital businesses with some market power. First, empowering customers in privacy management can benefit the business that does not rely heavily on data-based revenue. As we have shown, active privacy management can expand demand as privacy-conscious consumers can choose from multiple options to buy the service. This allows the business to raise price if it has some market power. The implication is that, even in jurisdictions without the GDPR-type opt-in requirements, businesses may consider proactively empowering customers in privacy management. After all, today's savvy consumers can, with a little research, find the information they need to make informed decisions. Providing them upfront with clear information and available options can also minimize potential backlash and reduce customer churn.

Second, we have identified potential conflicts between the usage-based revenue and data-based revenue. As the size of data-based revenue grows, the business has an incentive to lower price to attract more consumers and hence more data. This can erode the usage-based revenue. One possible way to manage such conflicts is to separate the two revenue sources into independent businesses. For example, a firm that provides advertising services can have a separate entity responsible for collecting and monetizing user data. Of course, this is subject to usual considerations relevant to corporate restructuring.<sup>26</sup> The establishment of Alphabet in 2015 may be understood in this context. Alphabet was formed as a holding company that oversees several separate businesses including Google. This separation allowed Google to focus on its core search and advertising business while giving other Alphabet subsidiaries such as Life Sciences and Calico greater autonomy to pursue their own initiatives.<sup>27</sup> Another example involves Twitter and Gnip, an independent social media API aggregation company. Twitter had a data licensing agreement with Gnip in 2010 for selling access to Twitter's data to other companies. This allowed Twitter to focus on its core business of providing a social media

---

<sup>26</sup>Management accounting can also play a crucial role by developing performance measures based on which to provide incentives to each revenue center in a way that can mitigate such conflicts.

<sup>27</sup>See a blog post by Larry Page, Google's co-founder, <https://blog.google/alphabet/google-alphabet/>.

platform while giving Gnip the freedom to focus on monetizing Twitter’s data.<sup>28</sup>

## 7 Conclusion

This paper has studied the GDPR’s opt-in requirement in a simple model with a monopolistic firm and consumers who differ in both valuations of the firm’s service and privacy costs. The firm’s revenue comes from two sources: sales of its service which we call usage-based revenue, and monetization of data which we call data-based revenue. Our basic assumption is that consumers can choose to buy the service without sharing data, called opt-out purchase, only after the GDPR. Relaxing this assumption will not alter our main insight, as long as more consumers proactively manage their privacy after the GDPR, the latter supported by various empirical and experimental studies.

We have shown that the GDPR expands demand for the firm’s service by inviting consumers with high privacy costs to participate in the market through opt-out purchase. This allows the firm to raise price for its service. If the firm’s revenue is largely usage-based, then the equilibrium demand also increases despite the higher price. In this case, the firm benefits from the GDPR thanks to the higher price and larger demand. Consumer surplus also increases mainly due to the increased consumer participation in the market through privacy management. On the other hand, some consumers with high privacy costs and high valuations who bought the service before the GDPR can switch to opt-out purchase after the GDPR. This leads to a smaller quantity of shared data, which erodes the firm’s data-based revenue. Thus, the firm’s profit can decrease after the GDPR if its revenue is largely data-based. Moreover, the higher price after the GDPR can lead some consumers with low valuations to drop out of the market after the GDPR, decreasing consumer surplus as a result. Put together, we conclude that the GDPR has differing effects on the firm and consumers, depending on the firm’s revenue structure. For the firm without significant data-based revenue, the GDPR can increase profits and consumer surplus, and hence can be welfare-improving. But it can decrease welfare when data-based revenue is significant for the firm.

## Appendix

### Proof of Lemma 1.

First, note that  $\partial\beta/\partial x = \partial\beta/\partial\alpha = |x + \alpha - 1|/\beta \in [0, 1)$ . Differentiating the equilibrium price with respect to  $x$ , we have

$$\frac{\partial\hat{p}^*}{\partial x} = \begin{cases} \frac{1}{3} > 0 & \text{if } \alpha < \alpha_1(x), \\ \frac{2}{3} + \frac{1}{3} \left( \frac{\partial\beta}{\partial x} \right) > 0 & \text{if } \alpha \geq \alpha_1(x). \end{cases}$$

---

<sup>28</sup>Twitter eventually acquired by Gnip in 2014 to expand its own data analytics capabilities.

Differentiating the equilibrium price with respect to  $\alpha$ , we obtain

$$\frac{\partial \widehat{p}^*}{\partial \alpha} = \begin{cases} -\frac{2}{3} < 0 & \text{if } \alpha < \alpha_1(x), \\ -\frac{1}{3} + \frac{1}{3} \left( \frac{\partial \beta}{\partial \alpha} \right) < 0 & \text{if } \alpha \geq \alpha_1(x). \end{cases}$$

Differentiating the equilibrium demand with respect to  $k \in \{\alpha, x\}$ , we obtain

$$\frac{\partial \widehat{D}^*}{\partial k} = \begin{cases} \frac{4(1+x+\alpha)}{9} > 0 & \text{if } 0 < \alpha < \alpha_1(x), \\ \frac{(1-x-\alpha+\beta)^2}{9\beta} > 0 & \text{if } \alpha \geq \alpha_1(x). \end{cases}$$

Given that both the equilibrium price and demand increase in  $x$ , it follows that the firm's equilibrium profit also increases in  $x$ . Next, differentiating the equilibrium profit with respect to  $\alpha$ , we obtain

$$\frac{\partial \widehat{\Pi}^*}{\partial \alpha} = \begin{cases} \frac{2(1+x+\alpha)^2}{9} > 0 & \text{if } \alpha < \alpha_1(x), \\ \frac{5+(x+\alpha)(2-x+\beta-\alpha)-\beta}{9} > 0 & \text{if } \alpha \geq \alpha_1(x). \end{cases}$$

To show the second inequality above, let  $\gamma := x + \alpha$  and write the numerator as  $F(\gamma) := 5 + \gamma(2 + \beta - \gamma) - \beta$ . Noting that  $\partial \beta / \partial \gamma = (\gamma - 1) / \beta$ , we have  $F'(\gamma) = (1 - \gamma + \beta)^2 / \beta > 0$ . Next, solving  $F(\gamma) = 0$ , we find a unique solution  $\gamma = 1 - \sqrt{2} < 0$ . Thus, we must have  $F(\gamma) > 0$  for all  $\gamma > 0$ , which proves the second inequality above.

Differentiating the consumer surplus with respect to  $k \in \{\alpha, x\}$ , we have

$$\frac{\partial \widehat{CS}^*}{\partial k} = \begin{cases} \frac{4(1+x+\alpha)^2}{27} > 0 & \text{if } \alpha < \alpha_1(x), \\ \frac{F(\gamma)}{54} \left( 1 + \frac{1-\gamma}{\beta} \right) > 0 & \text{if } \alpha \geq \alpha_1(x). \end{cases}$$

The second inequality above follows from the fact that  $F(\gamma) > 0$  for all  $\gamma > 0$  as shown previously, and  $\beta = \sqrt{6 + (1 - \gamma)^2} > |1 - \gamma|$ . ■

### Proof of Lemma 2.

Differentiating the equilibrium price with respect to  $x$  gives us

$$\frac{\partial \widetilde{p}^*}{\partial x} = \begin{cases} \frac{x-\alpha}{2} & \text{if } \alpha < \alpha_2(x), \\ \frac{2}{3} \left( 1 - \frac{2-2x+\alpha}{\sqrt{6+(2-2x+\alpha)^2}} \right) & \text{if } \alpha \geq \alpha_2(x). \end{cases}$$

The first derivative is negative if and only if  $x < \alpha$ . The second derivative is always positive because  $0 \leq 2 - 2x + \alpha < \sqrt{6 + (2 - 2x + \alpha)^2}$ .

Differentiating the equilibrium price with respect to  $\alpha$ , we obtain

$$\frac{\partial \tilde{p}^*}{\partial \alpha} = \begin{cases} -\frac{x}{2} < 0 & \text{if } \alpha < \alpha_2(x), \\ -\frac{1}{3} \left( 1 - \frac{2 - 2x + \alpha}{\sqrt{6 + (2 - 2x + \alpha)^2}} \right) < 0 & \text{if } \alpha \geq \alpha_2(x). \end{cases}$$

Differentiating the equilibrium opt-in demand with respect to  $x$  leads to

$$\frac{\partial \tilde{D}_S^*}{\partial x} = \begin{cases} \frac{2 + 4x - 3x^2 + 4x\alpha}{4} > 0 & \text{if } \alpha < \alpha_2(x), \\ \frac{1}{9} \left( 9 + 4(2(1 - x) + \alpha) + \frac{12}{\Gamma} - 4\Gamma \right) > 0 & \text{if } \alpha \geq \alpha_2(x) \end{cases}$$

where we recall  $\Gamma := \sqrt{6 + (2 - 2x + \alpha)^2}$ . The first inequality above holds because  $x \in [0, 1]$  and  $\alpha \geq 0$ . To prove the second inequality, define  $z := 2 - 2x + \alpha \geq 0$ , hence  $\Gamma = \sqrt{6 + z^2}$ . Let  $G(z) := 9 + 4z + \frac{12}{\sqrt{6+z^2}} - 4\sqrt{6+z^2}$ . Solving  $G(z) = 0$ , we obtain a unique solution  $\tilde{z} \approx -0.826$ . In addition, we can check  $G'(\tilde{z}) \approx 0.65 > 0$ . These two imply  $G(z) > 0$  for all  $z \geq 0$ .

Differentiating the equilibrium opt-in demand with respect to  $\alpha$ , we have

$$\frac{\partial \tilde{D}_S^*}{\partial \alpha} = \begin{cases} \frac{x^2}{2} > 0 & \text{if } \alpha < \alpha_2(x), \\ \frac{(2 - 2x + \alpha - \Gamma)^2}{9\Gamma} > 0 & \text{if } \alpha \geq \alpha_2(x). \end{cases}$$

Differentiating the equilibrium opt-out demand with respect to  $x$ , we obtain

$$\frac{\partial \tilde{D}_N^*}{\partial x} = \begin{cases} \frac{3x^2 - 2x - 4x\alpha + 2\alpha - 2}{4} & \text{if } \alpha < \alpha_2(x), \\ \frac{1}{3} \left( -5 + 2x - \alpha + \Gamma - 2(1 - x) \left( 1 - \frac{(2 - 2x + \alpha)}{\Gamma} \right) \right) & \text{if } \alpha \geq \alpha_2(x). \end{cases}$$

First, consider the case where  $\alpha < \alpha_2(x)$ . Recall that this case is possible when  $\alpha_2(x) > 0$ , or  $x < 2 - \sqrt{2} \approx 0.58$ , hence the relevant parameter range is  $A := \{(x, \alpha) \in \mathbb{R}_2^+ | x < 2 - \sqrt{2}, \alpha < \alpha_2(x)\}$ . Solving  $\partial \tilde{D}_N^* / \partial x = 0$  for  $\alpha$  yields  $\alpha_3(x) := (2 + 2x - 3x^2) / (2 - 4x)$  and we have  $\alpha_3(x) < \alpha_2(x)$  when  $x < -4 + 3\sqrt{2} \approx 0.24$ . In addition,  $\lim_{x \rightarrow 1/2} \alpha_3(x) = \infty$  and  $\alpha_3(x) < 0$  if  $1/2 < x < 2 - \sqrt{2}$ . Let  $G(\alpha) := 3x^2 - 2x - 4x\alpha + 2\alpha - 2$ . Then  $G'(\alpha) \geq 0$  iff  $x \leq 1/2$ . Based on the above, we can consider by dividing  $A$  into three regions: (i) If  $x < -4 + 3\sqrt{2} < 1/2$ , then we have  $\alpha_3(x) < \alpha_2(x)$  and  $G'(\alpha) > 0$ , hence  $G(\alpha) > 0$  for all  $\alpha_3(x) < \alpha < \alpha_2(x)$ ; (ii) If  $-4 + 3\sqrt{2} \leq x < 1/2$ , then we have  $\alpha_3(x) \geq \alpha_2(x)$  and  $G'(\alpha) > 0$ , hence  $G(\alpha) \leq 0$  for all  $\alpha < \alpha_3(x) < \alpha_2(x)$ ; (iii) If  $1/2 < x < 2 - \sqrt{2}$ , then  $\alpha_3(x) < 0$  and  $G'(\alpha) < 0$ , hence we have  $G(\alpha) < 0$  for all  $\alpha_3(x) < \alpha < \alpha_2(x)$ . From these, we can conclude that, if  $\alpha < \alpha_2(x)$ , then  $\partial \tilde{D}_N^* / \partial x > 0$  if and only if  $\alpha_3(x) < \alpha < \alpha_2(x)$ . Consider next the case where  $\alpha \geq \alpha_2(x)$ . Differentiating  $\partial \tilde{D}_N^* / \partial x$  with respect to  $x$  leads to

$$\frac{\partial^2 \tilde{D}_N^*}{\partial x^2} = \frac{4(6(1 - x) + \Gamma^2)(\Gamma - (2 - 2x + \alpha))}{3\Gamma^3} > 0$$

where the inequality follows because  $\Gamma := \sqrt{6 + (2 - 2x + \alpha)^2} > 2 - 2x + \alpha$  and  $x \leq 1$ . At

$x = 1$ , we have  $\partial \tilde{D}_N^*/\partial x = (-3 - \alpha + \sqrt{6 + \alpha^2})/3 < 0$ . Thus we have  $\partial \tilde{D}_N^*/\partial x < 0$  for all  $x \leq 1$ . Combining the above two cases, we have shown that  $\partial \tilde{D}_N^*/\partial x > 0$  if and only if  $\alpha_3(x) < \alpha < \alpha_2(x)$ .

Differentiating the equilibrium opt-out demand with respect to  $\alpha$ , we have

$$\frac{\partial \tilde{D}_N^*}{\partial \alpha} = \begin{cases} \frac{x(1-x)}{2} > 0 & \text{if } \alpha < \alpha_2(x), \\ \frac{(1-x)}{3} \left( 1 - \frac{2-2x+\alpha}{\sqrt{6+(2-2x+\alpha)^2}} \right) > 0 & \text{if } \alpha \geq \alpha_2(x). \end{cases}$$

Differentiating the equilibrium total demand with respect to  $x$  leads to

$$\frac{\partial \tilde{D}^*}{\partial x} = \begin{cases} \frac{x+\alpha}{2} > 0 & \text{if } \alpha < \alpha_2(x), \\ \frac{14+8x^2+\alpha(2-\alpha)-2x(8+\alpha)-\Gamma(4-4x-\alpha)}{9\Gamma} & \text{if } \alpha \geq \alpha_2(x). \end{cases}$$

Analytically evaluating the sign of the second derivative is demanding. Instead, we numerically calculate the value of  $\partial \tilde{D}^*/\partial x$  and plot the range of  $(x, \alpha)$  for which  $\partial \tilde{D}^*/\partial x > 0$ . We find that  $\partial \tilde{D}^*/\partial x > 0$  for all  $(x, \alpha) \in B := \{(x, \alpha) \in \mathbb{R}_2^+ | x \leq 1, \alpha \geq \alpha_2(x)\}$ , that is, the entire range of  $\alpha$  in this case.

Differentiating the equilibrium total demand with respect to  $\alpha$ , we have

$$\frac{\partial \tilde{D}^*}{\partial \alpha} = \begin{cases} \frac{x}{2} > 0 & \text{if } \alpha < \alpha_2(x), \\ \frac{8+2x^2+\alpha(5+2\alpha)-\Gamma(1+2\alpha-x)-x(4+5\alpha)}{9\Gamma} & \text{if } \alpha \geq \alpha_2(x). \end{cases}$$

For the sign of the second derivative, we again rely on the numerical method and verify that  $\partial \tilde{D}^*/\partial \alpha > 0$  for all  $(x, \alpha) \in B$ .

Comparative statics of the equilibrium profit and consumer surplus can be done in the same way as that for the equilibrium total demand. So we omit the proof. For the case  $\alpha < \alpha_2(x)$ , it is easy to check  $\partial \tilde{\Pi}^*/\partial x > 0$ ,  $\partial \tilde{\Pi}^*/\partial \alpha > 0$ ,  $\partial \tilde{CS}^*/\partial x > 0$ , and  $\partial \tilde{CS}^*/\partial \alpha > 0$ . For the case  $\alpha \geq \alpha_2(x)$ , we can use the numerical method to verify that all the derivatives are positive for all  $(x, \alpha) \in B$ . ■

### Proof of Proposition 1.

Based on Lemma 3, we can calculate the difference in the equilibrium prices before and after the GDPR as follows:

$$\tilde{p}^* - \hat{p}^* = \begin{cases} \frac{2+8\alpha-x(4+6\alpha-3x)}{12} & \text{if } \alpha < \alpha_1(x), \\ \frac{14+4\alpha+3x^2-8x-6x\alpha-4\beta}{12} & \text{if } \alpha_1(x) \leq \alpha < \alpha_2(x), \\ \frac{\Gamma-\beta}{3} & \text{if } \alpha_2(x) \leq \alpha. \end{cases}$$

Consider the case where  $\alpha < \alpha_1(x)$ . This case is possible when  $x \leq 1/2$ , which implies  $2+8\alpha-x(4+6\alpha-3x) > 0$ . Thus we have  $\tilde{p}^* > \hat{p}^*$ . Moreover,  $x \leq 1/2$  implies  $\partial \Delta p^*/\partial \alpha > 0$ .

Consider next the case where  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , which is possible when  $x < 2-\sqrt{2}$ . We will

show  $\tilde{p}^* > \hat{p}^*$ . Denote the numerator of  $\tilde{p}^* - \hat{p}^*$  by  $F(x, \alpha)$ . First, note that  $\partial F(x, \alpha)/\partial x < 0$  because  $\partial F(x, \alpha)/\partial x = 6(x - \alpha) - 8 - 4(\partial\beta/\partial\alpha)$  and  $x < 2 - \sqrt{2}$ ,  $\alpha \geq 0$ , and  $\partial\beta/\partial\alpha > 0$ . Having established  $\partial F(x, \alpha)/\partial x < 0$ , it suffices to show that, for all  $\alpha \in [\alpha_1(x), \alpha_2(x)]$ , we have  $F(x, \alpha) \geq 0$  for the maximum possible value of  $x$  for given  $\alpha$ , which we denote by  $x(\alpha)$ . We can derive  $x(\alpha)$  by solving  $\alpha = \alpha_2(x)$ . Since  $x \in [0, 1]$ , we have  $x(\alpha) = (2 + \alpha) - \sqrt{(2 + \alpha)^2 - 2}$ . Calculating  $F(x(\alpha), \alpha)$ , we have  $F(x(\alpha), \alpha) = 4(2(2 + \alpha) - (\sqrt{(2 + \alpha)^2 - 2} + \sqrt{9 - 2(1 + 2\alpha)\sqrt{(2 + \alpha)^2 - 2} + \alpha(8 + 5\alpha)}))$ . In the following, we show that  $F(x(\alpha), \alpha) > 0$ . First, solving  $F(x(\alpha), \alpha) = 0$ , we obtain two solutions  $\alpha = -1/2$  and  $\alpha = -1/6$ . Next, we can calculate  $(\partial F(x(\alpha), \alpha)/\partial\alpha)|_{\alpha=-1/6} = 72/35 > 0$ . This inequality is sufficient to show  $F(x(\alpha), \alpha) > 0$  for all  $\alpha > 0$  since  $F(x(\alpha), \alpha) = 0$  has only two solutions  $\alpha = -1/2$  and  $\alpha = -1/6 > -1/2$ .

Finally, for the case  $\alpha \geq \alpha_2(x)$ , recall  $\Gamma = \sqrt{6 + (2(1 - x) + \alpha)^2}$  and  $\beta = \sqrt{6 + (1 - x - \alpha)^2}$ . Thus, we have  $\Gamma > \beta$ , hence  $\tilde{p}^* - \hat{p}^* > 0$ . ■

**Proof of Proposition 2.** Comparing the equilibrium total demands before and after the GDPR, we have

$$\tilde{D}^* - \hat{D}^* = \begin{cases} \frac{10 + x^2 - 16x - 2\alpha(8 - x) - 8\alpha^2}{36} & \text{if } \alpha < \alpha_1(x), \\ \frac{4\beta(1 - x - \alpha) + 4\alpha^2 + 26x\alpha + 13x^2 - 2 - 8(x + \alpha)}{(1 - x - \alpha)(3(1 - x) + \beta - \Gamma)} & \text{if } \alpha_1(x) \leq \alpha < \alpha_2(x), \\ \frac{36}{9} & \text{if } \alpha_2(x) \leq \alpha. \end{cases}$$

Consider the case where  $\alpha < \alpha_1(x) = (1 - 2x)/2$ , when we have  $\alpha \leq 1/2$  and  $x \leq 1/2$ . Denote the numerator of  $\tilde{D}^* - \hat{D}^*$  by  $H(x, \alpha)$ . Given  $\alpha \leq 1/2$  and  $x \leq 1/2$ , we have  $\partial H(x, \alpha)/\partial x = 2(x + \alpha - 8) < 0$  for all  $(x, \alpha)$ . We will show that  $H(x, \alpha) \geq 0$  for all  $(x, \alpha)$ . Since  $\partial H(x, \alpha)/\partial x < 0$ , it is sufficient to show  $H(x, \alpha) \geq 0$  for the maximum possible value of  $x$  given  $\alpha$  for all  $\alpha \leq 1/2$ . Such  $x$  solves  $\alpha = \alpha_1(x)$ , hence  $x = (1 - 2\alpha)/2$ . By direct calculation, we have  $H((1 - 2\alpha)/2, \alpha) = (9 - 36\alpha^2)/4 \geq 0$  since  $\alpha \leq 1/2$ . Thus we have shown  $\tilde{D}^* \geq \hat{D}^*$  if  $\alpha < \alpha_1(x)$ .

Consider next the case where  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ . Analytically evaluating the sign of  $\tilde{D}^* - \hat{D}^*$  is quite demanding. Instead, we numerically calculate its value for all  $(x, \alpha)$  such that  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ . In Figure 6, the straight line in the lower border of the shaded area corresponds to  $\alpha = \alpha_1(x)$  and the curved line in the upper border is  $\alpha = \alpha_2(x)$ . Thus, Figure 6 shows how the set  $\{(x, \alpha) | x \in [0, 1], \alpha_1(x) \leq \alpha < \alpha_2(x)\}$  is divided into two regions. The region labeled  $\Delta D > 0$  is where the equilibrium demand is larger after the GDPR and the area labeled  $\Delta D < 0$  is where the equilibrium demand is smaller after the GDPR. The curved line that divides the two areas is where  $\tilde{D}^* = \hat{D}^*$ , which defines  $\alpha_4(x)$ .



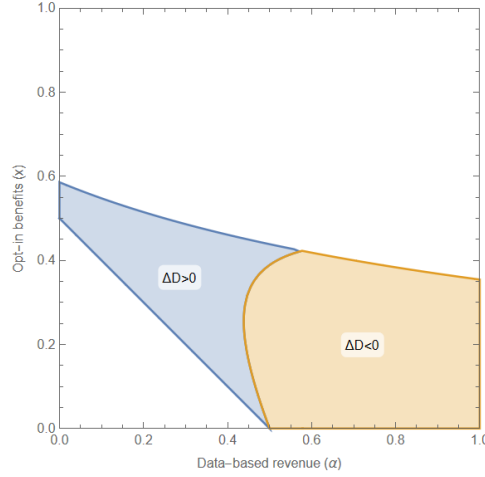


Figure 6: Comparison of total equilibrium demands when  $\alpha_1(x) \leq \alpha < \alpha_2(x)$

Finally, consider the case where  $\alpha \geq \alpha_2(x)$ , which is possible for all  $x \leq 1$ . First, we establish that  $G(x, \alpha) := 3(1 - x) + \beta - \Gamma$  is always positive. Differentiating  $G(x, \alpha)$  with respect to  $x$  yields

$$\frac{\partial G(x, \alpha)}{\partial x} = -3 + \frac{|x + \alpha - 1|}{\beta} + \frac{2(2 - 2x + \alpha)}{\Gamma}.$$

Recall  $\beta = \sqrt{6 + (x + \alpha - 1)^2}$  and  $\Gamma = \sqrt{6 + (2 - 2x + \alpha)^2}$ . Thus we have  $|x + \alpha - 1|/\beta < 1$  and  $(2 - 2x + \alpha)/\Gamma < 1$ , hence  $\partial G(x, \alpha)/\partial x < 0$ . Next, we have  $G(1, \alpha) = 0$ , which implies  $G(x, \alpha) > 0$  for all  $x < 1$  since  $x = 1$  is the maximum possible value of  $x$  and  $\partial G(x, \alpha)/\partial x < 0$ . Thus, the sign of  $\tilde{D}^* - \hat{D}^*$  is the same as the sign of  $(1 - x - \alpha)$ , from which follows  $\tilde{D}^* \geq \hat{D}^*$  iff  $1 - x \geq \alpha$ .

Next, we show  $\tilde{D}_S^* < \hat{D}^*$  for all  $\alpha$ . By direct calculation, we have

$$\tilde{D}_S^* - \hat{D}^* = \begin{cases} -\frac{9x^3 + 8(1 + \alpha)^2 - x(2 - 16\alpha) - 2x^2(5 + 9\alpha)}{36} & \text{if } 0 < \alpha < \alpha_1(x), \\ -\frac{9x^3 + 4(5 - \alpha^2 - \beta + \alpha(2 + \beta)) - 2x(x(11 + 9\alpha) + 5 + 4\alpha - 2\beta)}{36} & \text{if } \alpha_1(x) \leq \alpha < \alpha_2(x), \\ -\frac{12 + 3x^2 - \beta - x(15 + 6\alpha - \beta - 2\Gamma) - 2\Gamma + \alpha(6 + \beta - \Gamma)}{9} & \text{if } \alpha_2(x) \leq \alpha. \end{cases}$$

Consider the case where  $\alpha < \alpha_1(x)$ . It is easy to see  $\partial(\tilde{D}_S^* - \hat{D}^*)/\partial\alpha < 0$  since  $x < 1/2$  in this case. Moreover,  $\tilde{D}_S^* - \hat{D}^* < 0$  when  $\alpha = 0$ . Thus we have  $\tilde{D}_S^* < \hat{D}^*$ . For the other two cases, we numerically calculate the value of  $\tilde{D}_S^* - \hat{D}^*$  and plot the range of  $(x, \alpha)$  for which  $\tilde{D}_S^* < \hat{D}^*$ . When  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , we find that  $\tilde{D}_S^* < \hat{D}^*$  for all  $(x, \alpha) \in \{(x, \alpha) \in \mathbb{R}_2^+ | x < 2 - \sqrt{2}, \alpha_1(x) \leq \alpha < \alpha_2(x)\}$ , i.e., the entire range of  $\alpha$  in this case. When  $\alpha \geq \alpha_2(x)$ , we find that  $\tilde{D}_S^* < \hat{D}^*$  for all  $(x, \alpha) \in \{(x, \alpha) \in \mathbb{R}_2^+ | x \leq 1, \alpha_2(x) \leq \alpha\}$ , i.e., the entire range of  $\alpha$  in this case. ■

**Proof of Proposition 3.** First, when  $\alpha < \alpha_1(x)$ , the difference in equilibrium profits can be

calculated as

$$\tilde{\Pi}^* - \hat{\Pi}^* = \frac{76 + 27x^4 - 96\alpha - 32\alpha^2(3 + \alpha) + 12x^2(1 + \alpha)(1 + 9\alpha)}{432} - \frac{4x^3(8 + 27\alpha) + 24x(4 + 4\alpha^2 - \alpha)}{432}.$$

Denote the numerator of  $\tilde{\Pi}^* - \hat{\Pi}^*$  by  $J(x, \alpha)$ . We will show  $\tilde{\Pi}^* > \hat{\Pi}^*$  by proving  $J(x, \alpha) > 0$  for all  $(x, \alpha)$  such that  $\alpha < \alpha_1(x) = (1 - 2x)/2$  and  $x < 1/2$ . First, note that  $J(x, \alpha_1(x)) = (1 - x)^2 x(4 + 9x)/16 > 0$  for all  $x < 1/2$ . Then, to prove  $J(x, \alpha) > 0$  for all  $(x, \alpha)$ , it is sufficient to show  $\partial J(x, \alpha)/\partial \alpha < 0$  for all  $(x, \alpha)$ . For this, it is sufficient to show  $\partial J(x, 0)/\partial \alpha < 0$  and  $\partial^2 J(x, \alpha)/\partial \alpha^2 < 0$ . Notice that  $\partial J(x, 0)/\partial \alpha = -(8 - x(2 + x(10 - 9x)))/36 < 0$  for all  $x \in (0, 1/2)$ . Thus, it suffices to show  $\partial^2 J(x, \alpha)/\partial \alpha^2 < 0$ . To prove this, it is sufficient to show  $\partial^2 J(x, 0)/\partial \alpha^2 < 0$  and  $\partial^3 J(x, \alpha)/\partial \alpha^3 < 0$ . But  $\partial^2 J(x, 0)/\partial \alpha^2 = -(8 + x(8 - 9x))/18 < 0$  and  $\partial^3 J(x, \alpha)/\partial \alpha^3 = -4/9 < 0$ . Thus, we have shown  $J(x, \alpha) > 0$  for all  $(x, \alpha)$  such that  $\alpha < \alpha_1(x) = (1 - 2x)/2$  and  $x < 1/2$ , hence  $\tilde{\Pi}^* > \hat{\Pi}^*$ .

Next, when  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , the difference in equilibrium profits is given by

$$\begin{aligned} \tilde{\Pi}^* - \hat{\Pi}^* &= \frac{27x^4 + 4x^3(8 - 27\alpha) + 4x^2(3 + 3\alpha(26 + 9\alpha) - 4\beta)}{432} \\ &+ \frac{8x(12\alpha^2 + \alpha(3 - 4\beta) - 2(6 - \beta(2 - \beta)))}{432} \\ &+ \frac{4(67 + 8\alpha^3 - 4\beta(7 - \beta) - 4\alpha^2(6 + \beta) - 4\alpha(6 - \beta(2 - \beta)))}{432}. \end{aligned}$$

Analytically evaluating the sign of  $\tilde{\Pi}^* - \hat{\Pi}^*$  is quite demanding. Instead, we numerically calculate its value for all  $(x, \alpha)$  such that  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , which is plotted in Figure 7, where the straight line in the lower border of the shaded area corresponds to  $\alpha = \alpha_1(x)$  and the curved line in the upper border is  $\alpha = \alpha_2(x)$ . The region labeled  $\Delta\Pi > 0$  is where the equilibrium profit is larger after the GDPR and the area labeled  $\Delta\Pi < 0$  is where the equilibrium profit is smaller after the GDPR. The line that divides the two areas is where  $\Delta\Pi = 0$ , which defines  $\alpha_5(x)$ .

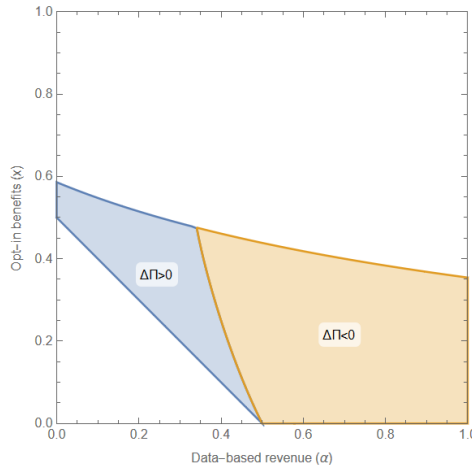


Figure 7: Comparison of equilibrium profits when  $\alpha_1(x) \leq \alpha < \alpha_2(x)$

Finally, when  $\alpha_2(x) \leq \alpha$ , the difference in equilibrium profits is given by

$$\begin{aligned} \tilde{\Pi}^* - \hat{\Pi}^* &= \frac{10\Gamma + 10x^3 + \alpha^3 + \beta^2 - x^2(30 + 6\alpha + \beta - 4\Gamma)}{27} \\ &\quad - \frac{\alpha(27 - 4\Gamma - \beta(2 - \beta)) + \alpha^2(12 + \beta - \Gamma) + \alpha(12 + \beta - \Gamma)}{27} \\ &\quad - \frac{x(8\Gamma + \alpha(4\Gamma + 2\beta - 39) - \beta(2 - \beta) - 36 - 12\alpha^2) + 16 + 7\beta}{27}. \end{aligned}$$

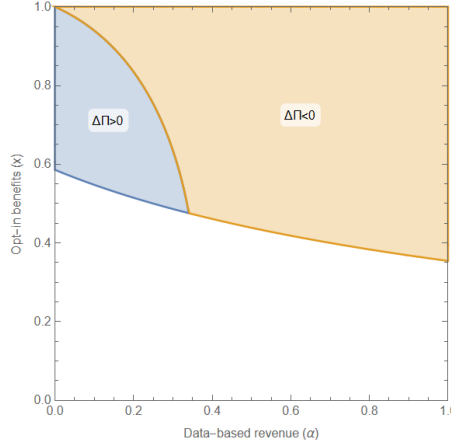


Figure 8: Comparison of equilibrium profits when  $\alpha > \alpha_2(x)$

Once again, we numerically calculate  $\tilde{\Pi}^* - \hat{\Pi}^*$  for all  $(x, \alpha)$  such that  $\alpha_2(x) \leq \alpha$ , which is plotted in Figure 8, where the curved line in the lower border corresponds to  $\alpha = \alpha_2(x)$ . The region labeled  $\Delta\Pi > 0$  is where the equilibrium profit is larger after the GDPR and the area labeled  $\Delta\Pi < 0$  is where the equilibrium profit is smaller after the GDPR. The curved line that divides the two areas is where  $\Delta\Pi = 0$ , which defines  $\alpha_6(x)$ . ■

**Proof of Proposition 4.** The proof is by numerical calculations. For the case  $\alpha < \alpha_1(x)$ , the difference in equilibrium consumer surpluses is given by

$$\begin{aligned} \Delta CS_1 := \widetilde{CS}_{2a}^* - \widehat{CS}_{1a}^* &= \frac{196 - 128\alpha(\alpha + 3) + 3 - 243x^4 + 4(81\alpha + 76)x^3}{2592} \\ &\quad - \frac{24x(16 + \alpha(5 + 16\alpha)) + 12x^2(5 + \alpha(32 - 27\alpha))}{2592}. \end{aligned}$$

In Figure 9, the entire shaded area is the set  $\{(x, \alpha) | x \in [0, 1], \alpha < \alpha_1(x)\}$ , and the region labeled  $\Delta CS > (<) 0$  is the set of  $(x, \alpha)$  for which  $\widetilde{CS}_{2a}^* > (<) \widehat{CS}_{1a}^*$ . The line that divides these two regions is where  $\Delta CS_1 = 0$ , which defines  $\alpha_7(x)$ .

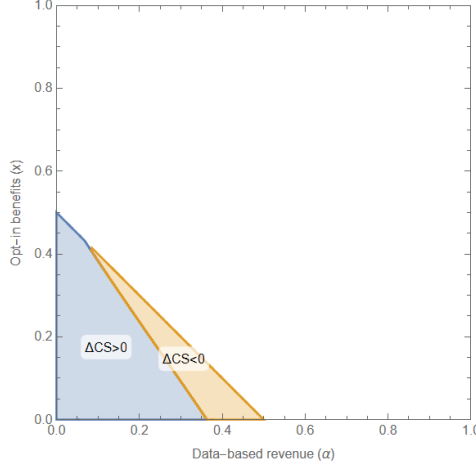


Figure 9: Consumer surplus comparison when  $\alpha < \alpha_1(x)$

When  $\alpha_1(x) \leq \alpha < \alpha_2(x)$ , the difference in equilibrium consumer surpluses is given by

$$\begin{aligned} \Delta CS_2 := \widetilde{CS}_{2a}^* - \widehat{CS}_{1b}^* &= \frac{12x^2(23 + 4\alpha + 27\alpha^2 - 4\beta) + 4x^3(112 + 81\alpha) - 243x^4}{2592} \\ &\quad - \frac{24x(34 - 23\alpha - 2\alpha^2 - \beta(4 - 4\alpha + 2\beta)) + 4(355 - 4\alpha^3 - 204\beta + 12\beta^2)}{2592} \\ &\quad - \frac{4(4\beta^3 + 12\alpha^2(1 + \beta) + 12\alpha(17 - 2\beta + \beta^2))}{2592}. \end{aligned}$$

In Figure 10, the straight line in the lower border of the shaded area corresponds to  $\alpha = \alpha_1(x)$  and the curved line in the upper border is  $\alpha = \alpha_2(x)$ . Thus, Figure 10 shows how the set  $\{(x, \alpha) | x \in [0, 1], \alpha_1(x) \leq \alpha < \alpha_2(x)\}$  is divided into two regions. The region labeled  $\Delta CS > (<) 0$  is the set of  $(x, \alpha)$  for which  $\widetilde{CS}_{2a}^* > (<) \widehat{CS}_{1b}^*$ . The line that divides the two regions is where  $\Delta CS_2 = 0$ , which defines  $\alpha_8(x)$ .

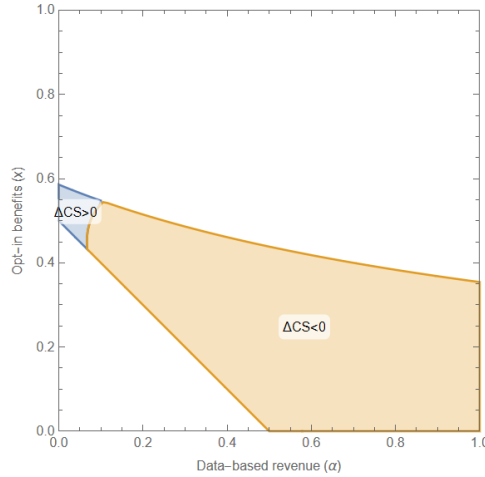


Figure 10: Consumer surplus comparison when  $\alpha_1(x) \leq \alpha < \alpha_2(x)$

When  $\alpha \geq \alpha_2(x)$ , the difference in equilibrium consumer surpluses is given by

$$\Delta CS_3 := \widetilde{CS}_{2b}^* - \widehat{CS}_{1b}^* = \frac{81(1-x)^2 - 54(\Gamma - \beta) + (\Gamma + 7(1-x) - \alpha)(2 - 2x + \alpha - \Gamma)^2}{162} - \frac{(1 + \beta - x - \alpha)^3}{162}.$$

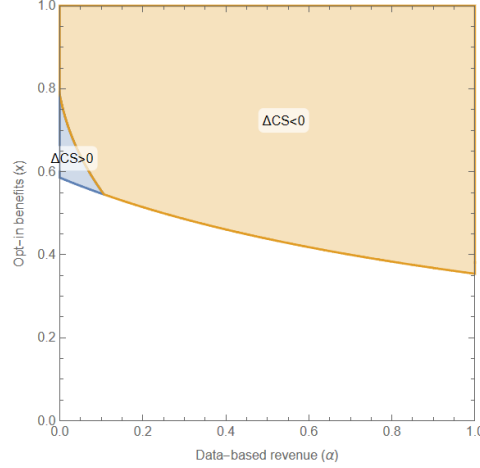


Figure 11: Consumer surplus comparison when  $\alpha > \alpha_2(x)$

In Figure 11, the curved line in the lower border of the shaded area corresponds to  $\alpha = \alpha_2(x)$ . Thus, Figure 11 shows how the set  $\{(x, \alpha) | x \in [0, 1], \alpha > \alpha_2(x)\}$  is divided into two regions. The region labeled  $\Delta CS > (<) 0$  is the set of  $(x, \alpha)$  for which  $\widetilde{CS}_{2b}^* > (<) \widehat{CS}_{1b}^*$ . The line that divides the two regions is where  $\Delta CS_3 = 0$ , which defines  $\alpha_9(x)$ . ■

**Proof of Proposition 5.** Consider the case of positive correlation where  $c = \phi v$  with  $\phi \in (0, 1)$  and  $x < \phi$ . First, before the GDPR, consumer  $v$  is indifferent between purchase and no purchase if and only if  $v + (x - c) - p = 0$ . Substituting  $c = \phi v$  and solving for  $v$ , we obtain  $\widehat{v} = (p - x)/(1 - \phi)$ . Then, the demand is  $\widehat{D}(p) = 1 - \widehat{v} = (1 + x - \phi - p)/(1 - \phi)$ . The firm maximizes  $\widehat{\Pi}(p) = (p + \alpha)\widehat{D}(p)$ , which gives us the equilibrium price and demand as follows:

$$\widehat{p}^* = \frac{1 + x - \alpha - \phi}{2}, \quad \widehat{D}^* = \frac{1 + x + \alpha - \phi}{2(1 - \phi)}.$$

After the GDPR, consumer  $v$  chooses opt-out purchase if  $c = \phi v \geq x$  and  $v \geq p$ . Denote  $\widetilde{v} = x/\phi$ . Then  $\widetilde{v} < 1$  by the assumption  $x < \phi$  and all consumers on  $[\widetilde{v}, 1]$  choose opt-out purchase if  $\widetilde{v} \geq p$ . Next, we show that if consumers on  $[\widetilde{v}, 1]$  choose opt-out purchase, then we must have  $\widetilde{v} \geq \widehat{v}$ . Suppose to the contrary that  $\widetilde{v} - \widehat{v} = (x - \phi p)/(\phi(1 - \phi)) < 0$ , hence  $p > x/\phi = \widetilde{v}$ . This contradicts the fact that consumer  $\widetilde{v}$  is the marginal consumer who chooses opt-out purchase. As shown previously, the marginal consumer who chooses opt-in purchase is  $\widehat{v} = (p - x)/(1 - \phi)$ . Put together, the total demand is  $\widetilde{D}(p) = 1 - \widehat{v} = (1 + x - \phi - p)/(1 - \phi)$ , and the total opt-in demand is  $\widetilde{D}_S(p) = \widetilde{v} - \widehat{v} = (x - \phi p)/(\phi(1 - \phi))$ . The firm maximizes  $\widetilde{\Pi}(p) = p\widetilde{D}(p) + \alpha\widetilde{D}_S(p)$ . Comparing the profits before and after the GDPR, we have  $\widehat{\Pi}(p) - \widetilde{\Pi}(p) = \alpha(\widehat{D}(p) - \widetilde{D}_S(p))$  since  $\widehat{D}(p) = \widetilde{D}(p)$ . It is easy to see  $\widehat{D}'(p) = \widetilde{D}'_S(p)$ , which implies that the equilibrium price

and demand after the GDPR the same as those before the GDPR, i.e.,  $\hat{p}^* = \tilde{p}^*$  and  $\hat{D}^* = \tilde{D}^*$ . Given that the equilibrium price and total demand are the same under both regimes but the firm earns data-based revenue from  $\tilde{D}_S^* < \tilde{D}^* = \hat{D}^*$  after the GDPR, it follows that the firm's equilibrium profit is smaller after the GDPR:  $\hat{\Pi}^* > \tilde{\Pi}^*$ . After the GDPR, consumer surplus does not change for all consumers on  $[0, \tilde{v}^*]$ , but increases for consumers on  $[\tilde{v}^*, 1]$  because they choose opt-out purchase after the GDPR given  $c = \phi v > x$ . Consequently, the consumer surplus is larger in the equilibrium after the GDPR.

Consider now the case of negative correlation where  $c = 1 - v$ . First, before the GDPR, consumer  $v$  is indifferent between purchase and no purchase if and only if  $v + (x - c) - p = 0$ . Substituting  $c = 1 - v$  and solving for  $v$ , we obtain  $\hat{v} = (1 - x + p)/2$ . Then, the demand is  $\hat{D}(p) = 1 - \hat{v} = (1 + x + p)/2$ . The firm maximizes  $\hat{\Pi}(p) = (p + \alpha)\hat{D}(p)$ . From this, we can derive the equilibrium price, demand, profit, and consumer surplus as follows:

$$\hat{p}^* = \frac{1 + x - \alpha}{2}, \quad \hat{D}^* = \frac{1 + x + \alpha}{4}, \quad \hat{\Pi}^* = \frac{(1 + x + \alpha)^2}{8}, \quad \widehat{CS}^* = \frac{(1 + x + \alpha)^2}{16}.$$

In the above, the equilibrium price is positive because of the restriction  $\alpha < 1 + x$ .

After the GDPR, consumer  $v$  chooses opt-out purchase if  $1 - v = c \geq x$  and  $v \geq p$ . Denote  $\tilde{v} = 1 - x$ . Then  $\tilde{v} < 1$  since  $x > 0$ . For all  $v < \tilde{v}$ , we have  $c = 1 - v > 1 - \tilde{v} = x$ . Thus, all consumers on  $[p, \tilde{v}]$  choose opt-out purchase and all consumers on  $[\tilde{v}, 1]$  choose opt-in purchase. This gives us the total demand  $\tilde{D}(p) = 1 - p$  and the opt-in demand  $\tilde{D}_S(p) = 1 - \tilde{v} = x$ . The firm maximizes  $\tilde{\Pi}(p) = p\tilde{D}(p) + \alpha\tilde{D}_S(p)$ , which leads to

$$\tilde{p}^* = \tilde{D}^* = \frac{1}{2}, \quad \tilde{D}_S^* = x, \quad \tilde{\Pi}^* = \frac{1}{4} + \alpha x, \quad \widetilde{CS}^* = \frac{4x^2 - 8x + 9}{8}.$$

Comparing the above, we have  $\tilde{p}^* \geq \hat{p}^*$  if and only if  $\alpha \geq x$  and  $\tilde{D}^* \geq \hat{D}^*$  if and only if  $\alpha \leq 1 - x$ . For the comparison of profits, note that

$$\tilde{\Pi}^* - \hat{\Pi}^* = \frac{-\alpha^2 + 2(3x - 1)\alpha + (1 - 2x - x^2)}{8} := \frac{H(\alpha)}{8}.$$

The equation  $H(\alpha) = 0$  has two solutions denoted by  $\alpha_L(x) := (3x - 1) - \sqrt{8x^2 - 8x + 2}$  and  $\alpha_H(x) = (3x - 1) + \sqrt{8x^2 - 8x + 2}$ . One can check that  $\alpha_L(x)$  increases in  $x$ ,  $\alpha_L(x_L) = 0$  at  $x_L \approx 0.414$ , and  $\alpha_L(x) < 1 + x$  for all  $x \in (0, 1]$ . Also,  $\alpha_H(x)$  increases in  $x$ ,  $\alpha_H(x) > 0$  for all  $x \in (0, 1]$ , and  $\alpha_H(x) \leq 1 + x$  if and only if  $x \leq x_H \approx 0.707$ . From the above, we can conclude the following. First, if  $x < x_L$ , then  $\tilde{\Pi}^* \geq \hat{\Pi}^*$  if and only if  $\alpha \leq \alpha_H(x)$ . Second, if  $x \in [x_L, x_H)$ , then  $\tilde{\Pi}^* \geq \hat{\Pi}^*$  if and only if  $\alpha \in [\alpha_L(x), \alpha_H(x)]$ . Third, if  $x \geq x_H$ , then  $\tilde{\Pi}^* \geq \hat{\Pi}^*$  if and only if  $\alpha \geq \alpha_L$ .

For the comparison of consumer surpluses, note that

$$\widetilde{CS}^* - \widehat{CS}^* = \frac{-\alpha^2 - 2(1 + x)\alpha + 7x^2 - 18x + 17}{16} := \frac{G(\alpha)}{16}.$$

The equation  $G(\alpha) = 0$  has two solutions, one negative and the other positive. The positive solution is given by  $\alpha_{CS}(x) := -(1 + x) + \sqrt{8x^2 - 16x + 18}$ . One can check that  $\alpha_{CS}(x)$  is decreasing in  $x$ ,  $\alpha_{CS}(x) > 0$  for all  $x \in (0, 1]$ , and  $\alpha_{CS}(x) \geq 1 + x$  for all  $x \leq x_{CS} \approx 0.65$ .

where  $1 + x$  is an upper bound on  $\alpha$ . From the above, we can conclude that  $\widehat{CS}^* \geq \widehat{CS}^*$  for all  $\alpha < 1 + x$  if  $x < x_{CS}$ , and for all  $\alpha \leq \alpha_{CS}$  if  $x \geq x_{CS}$ . ■

## References

- Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar (2022) “Too much data: Prices and inefficiencies in data markets,” *American Economic Journal: Microeconomics*, 14 (4), 218–256.
- Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein (2015) “Privacy and human behavior in the age of information,” *Science*, 347 (6221), 509–514.
- Acquisti, Alessandro, Curtis Taylor, and Liad Wagman (2016) “The economics of privacy,” *Journal of Economic Literature*, 54 (2), 442–492.
- Anderson, Simon P. and Joshua S. Gans (2011) “Platform siphoning: Ad-avoidance and media content,” *American Economic Journal: Microeconomics*, 3 (4), 1–34.
- Aridor, Guy, Yeon-Koo Che, and Tobias Salz (forthcoming) “The effect of privacy regulation on the data industry: Empirical evidence from GDPR,” *RAND Journal of Economics*.
- Batikas, Michail, Yi Liu, Hakan Ozlap, and Milan Miric (2023) “Impact of privacy regulation on experimentation and innovation,” *Academy of Management Proceedings*, 2023 (1), 17060.
- Belleflamme, Paul and Wouter Vergote (2016) “Monopoly price discrimination and privacy: The hidden cost of hiding,” *Economics Letters*, 149, 141–144.
- Bergemann, Dirk and Alessandro Bonatti (2015) “Selling cookies,” *American Economic Journal: Microeconomics*, 7 (3), 259–294.
- Bergemann, Dirk, Alessandro Bonatti, and Tan Gan (2022) “The economics of social data,” *RAND Journal of Economics*, 53 (2), 263–296.
- Bessen, James, Stephen Michael Impink, Lydia Reichensperger, and Robert Seamans (2020) “GDPR and the importance of data to AI startups,” Discussion Paper, <http://dx.doi.org/10.2139/ssrn.3576714>.
- Braghieri, Luca (2019) “Targeted advertising and price discrimination in intermediated online markets,” Discussion Paper, <http://dx.doi.org/10.2139/ssrn.3072692>.
- Campbell, James, Avi Goldfarb, and Catherine Tucker (2015) “Privacy regulation and market structure,” *Journal of Economics & Management Strategy*, 24 (1), 47–73.
- Chaudhury, Ratul Das and Chongwoo Choe (2023) “Digital privacy: GDPR and its lessons for Australia,” *Australian Economic Review*, 56 (2), 204–220.
- Chen, Zhijun (2022) “Privacy costs and consumer data acquisition: An economic analysis of data privacy regulation,” Discussion Paper, <http://dx.doi.org/10.2139/ssrn.4085923>.

- Chen, Zhijun, Chongwoo Choe, and Noriaki Matsushima (2020) “Competitive personalized pricing,” *Management Science*, 66 (9), 4003–4023.
- Choe, Chongwoo, Stephen King, and Noriaki Matsushima (2018) “Pricing with cookies: behaviorbased price discrimination and spatial competition,” *Management Science*, 64 (12), 5669–5687.
- Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim (2019) “Privacy and personal data collection with information externalities,” *Journal of Public Economics*, 173, 113–124.
- Cong, Jiajia and Noriaki Matsushima (2023) “The effects of personal data management on competition and welfare,” ISER Discussion Paper 1201.
- Congiu, Raffaele, Lorien Sabatino, and Geza Sapi (2022) “The impact of privacy regulation on web traffic: Evidence from the GDPR,” *Information Economics and Policy*, 61, 101003.
- Fainmesser, Itay P., Andrea Galeotti, and Ruslan Momot (2023) “Digital privacy,” *Management Science*, 69 (6), 3157–3173.
- Ghose, Anindya (2017) *Tap: Unlocking the mobile economy*: The MIT Press.
- Godinho de Matos, Miguel and Idris Adjerid (2022) “Consumer consent and firm targeting after GDPR: The case of a large telecom provider,” *Management Science*, 68 (5), 3330–3378.
- Goldberg, Samuel G., Garrett A. Johnson, and Scott K. Shriver (2022) “Regulating privacy online: An economic evaluation of the GDPR,” Discussion Paper, <http://dx.doi.org/10.2139/ssrn.3421731>.
- Goldfarb, Avi and Catherine Tucker (2012) “Shifts in privacy concerns,” *American Economic Review*, 102 (3), 349–353.
- Gopal, Ram. D., Hooman Hidaji, Sule Nur Kutlu, Raymond A. Patterson, and Niam Yaraghi (forthcoming) “Law, economics, and privacy: Implications of government policies on website and third-party information sharing,” *Information Systems Research*.
- Hagiu, Andrei and Julian Wright (forthcoming) “Data-enabled learning, network effects and competitive advantage,” *RAND Journal of Economics*.
- Ichihashi, Shota (2021) “The economics of data externalities,” *Journal of Economic Theory*, 196, 105316.
- (2023) “Dynamic privacy choices,” *American Economic Journal: Microeconomics*, 15 (2), 1–40.
- Janßen, Rebecca, Reinhold Kesler, Michael E. Kummer, and Joel Waldfogel (2022) “GDPR and the lost generation of innovation apps,” NBER Working Paper 30028.
- Jentzsch, Nicola, Geza Sapi, and Irina Suleymanova (2013) “Targeted pricing and customer data sharing among rivals,” *International Journal of Industrial Organization*, 31, 131–144.



- Jia, Jian, Ginger Zhe Jin, and Liad Wagman (2021) “The short-run effects of the General Data Protection Regulation on technology venture investment,” *Marketing Science*, 40 (4), 661–684.
- Johnson, Eric J., Steven Bellman, and Gerald L. Lohse (2002) “Defaults, framing and privacy: Why opting in-opting out,” *Marketing Letters*, 13 (1), 5–15.
- Johnson, Garrett A., Scott K. Shriver, and Shaoyin Du (2020) “Consumer privacy choice in online advertising: Who opts out and at what cost to industry?” *Marketing Science*, 39 (1), 33–51.
- Johnson, Garrett A., Scott K. Shriver, and Samuel G. Goldberg (forthcoming) “Privacy and market concentration: Intended and unintended consequences of the GDPR,” *Management Science*.
- Ke, T. Tony and K. Sudhir (2023) “Privacy rights and data security: GDPR and personal data markets,” *Management Science*, 69 (8), 4389–4412.
- Lin, Tesary (2022) “Valuing intrinsic and instrumental preferences for privacy,” *Marketing Science*, 41 (4), 235–253.
- Macha, Meghanath, Natasha Zhang Foutz, Beibei Li, and Anindya Ghose (forthcoming) “Personalized privacy preservation in consumer mobile trajectories,” *Information Systems Research*.
- Machuletz, Dominique and Rainer Böhme (2020) “Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR,” *Proceedings on Privacy Enhancing Technologies 2020*, 2020 (2), 481–498.
- Matte, Célestin, Nataliia Bievola, and Cristina Santos (2020) “Do Cookie Banners Respect my Choice?: Measuring Legal Compliance of Banners from IAB Europe’s Transparency and Consent Framework,” *2020 IEEE Symposium on Security and Privacy*, 791–809.
- Miklós-Thal, Jeanine, Avi Goldfarb, Avery M. Haviv, and Catherine Tucker (2023) “Digital hermits,” NBER Working Paper 30920.
- Nouwens, Midas, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal (2020) “Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence,” *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13.
- Obar, Jonathan A. and Anne Oeldorf-Hirsch (2020) “The biggest lie on the Internet: Ignoring the privacy policies and terms of service policies of social networking services,” *Information, Communication & Society*, 23 (1), 128–147.
- Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer (2022) “Regulatory spillovers and data governance: Evidence from the GDPR,” *Marketing Science*, 41 (4), 318–340.

Pew Research Center (2019) “Americans and privacy: Concerned, confused and feeling lack of control over their personal information.”

Schmitt, Julia, Klaus M. Miller, and Bernd Skiera (2021) “The impact of privacy laws on online user behavior,” HEC Paris Research Paper No MKG-2021-1437, <http://dx.doi.org/10.2139/ssrn.3774110>.

Utz, Christine, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz (2019) “(Un)informed Consent: Studying GDPR Consent Notices in the Field,” *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 973–990.