Data Privacy and Temptation*

John Zhuang Liu[†] Michael Sockin[‡] Wei Xiong[§]

July 2020

Abstract

This paper derives a preference for data privacy from consumers' temptation utility. This approach facilitates a welfare analysis of different data privacy regulations, such as the GDPR enacted by the European Union and the CCPA enacted by the state of California, when a fraction of the consumers may succumb to targeted advertising of temptation goods. While sharing consumer data with firms improves firms' matching efficiency of normal consumption goods, it also exposes weak-willed consumers to temptation goods. Despite that the GDPR and the CCPA give each consumer the choice to opt in or out of data sharing, these regulations may not provide sufficient protection for severely tempted consumers because of a negative externality in which the opt-in decision of some consumers reduces the anonymity of those who opt out. Our analysis also shows that the default choices instituted by the GDPR and the CCPA can lead to sharply different outcomes.

*We are grateful to Roland Benabou, Long Chen, Shaowei Ke, LiyanYang, Leeat Yariv, Lixin Ye, and seminar participants at Cambridge, CUHK Shenzhen, Luohan Academy, Northwestern, Princeton, Stanford, UT-Dallas, and the Virtual Finance Seminar for helpful comments and suggestions.

[†]The Chinese University of Hong Kong, Shenzhen. Email: zhuangliu@cuhk.edu.cn.

[‡]University of Texas, Austin. Email: Michael.Sockin@mccombs.utexas.edu.

[§]Princeton University and NBER. Email: wxiong@princeton.edu.

The age of big data brings with it not only substantial benefits, such as dramatically improved access for consumers to products and services, but also undesirable drawbacks. A key concern is that the collection of personal data by digital platforms such as Google, Amazon, and Facebook represents an unprecedented challenge to consumer privacy. Motivated by this concern, the European Union enacted the General Data Privacy Regulation (GDPR) in 2018, and the State of California in the United States enacted the California Consumer Privacy Act (CCPA) in 2020. Both regulations aim to protect consumer privacy, albeit with several important differences that are summarized in Appendix A, and motivate a normative analysis of the welfare consequences of different data collection and sharing schemes. Such a normative analysis, however, requires a systematic framework of why consumers have a preference for privacy. The existing economics literature on consumer privacy tends to focus on the trade-off between matching efficiency and price discrimination: on the one hand, consumer data can increase the social surplus by allowing firms to better match their products with consumer preferences; on the other, such data empower firms to price discriminate against consumers, which also tilts the distribution of the social surplus toward firms.[1] While aversion to price discrimination is an important motivation for a preference for privacy among consumers, this mechanism is indirect and depends intricately on the assumed market structure.

In this paper, we pursue an alternative approach to modeling a preference for privacy among consumers by directly deriving it from their preferences over menus of consumption goods, in the spirit of Kreps (1979), by building on the temptation utility of Gul and Pesendorfer (2001).[2] This utility specification allows consumers to suffer a mental cost from resisting temptation goods on their menus. This can lead to a preference for a smaller menu without temptation goods, in sharp contrast to the standard preference for larger menus in

[1] See Acquisti, Taylor and Wagman (2016), Bergemann and Morris (2019), and Goldfar and Tucker (2019), for recent reviews. For example, both Taylor (2004) and Acquisti and Varian (2005) show that it is optimal for sellers to use consumers' past purchase information to price discriminate only if consumers are naive of how their data is used, but not optimal if consumers are sophisticated and can adapt their purchasing strategies. Ali, Lewis and Vasserman (2019) analyze how consumers can use their data disclosure choices to amplify competition between firms in a competitive setting and to induce price concessions from a seller in a monopolistic setting. Furthermore, Ichibashi (2019) shows that a multi-product seller prefers to commit to not use consumer information for pricing so that consumers truthfully report their information and the seller can recommend to them the best product matches.

[2] Stovall (2010) has expanded this temptation utility representation to include random menus. Such preferences over temptation goods also admit an interpretation as an internal conflict among multiple selves, e.g., Bénabou and Pycia (2002), and are a special case of the random Strotz (1955) utility characterized by Dekel and Lipman (2012).

the absence of self-control issues. As data sharing affects how firms advertise their goods to consumers, a consumer's menu preference eventually determines her preference for privacy when sharing data with digital platforms. This systematic privacy preference allows us to compare how different data sharing schemes affect social welfare.

Self-control problems have become more severe in the age of big data. Consider the story of a compulsive gambler who tries to recover from gambling—he deleted all casino apps from his smart phone; he removed his profile from all of the major gambling sites; he set up a rule in Gmail to automatically delete any emails that are related to gambling. One day, however, he logged on to YouTube, and all his efforts seemed to be in vain: "99% of the ads I see on YouTube are for gambling."[3] This frustration is just one example of how data analytics are increasingly used by firms to target consumers with self-control problems.

Such examples of consumer targeting is widespread. Online casinos and the global video game industry, for instance, use third-party companies to harvest personal data and target advertisements to those who are most likely to be tempted, such as those who previously gambled or played but have stopped.[4] E-cigarette companies use social media platforms and youth-focused advertising strategies to target teenagers and expand the market for their fledgling product (Kim et al. (2019)). Marketers are now creating online alcohol advertisements that are more specifically tailored to their intended audience based on social data, including age, location, gender, interests, and much more (Morris (2019)). The adult film industry uses similar data-driven approaches to cater to consumers' diverse tastes (Raustiala and Sprigman (2019)). This phenomenon also extends well beyond these so-called "sin" industries. Payday lenders use algorithmic scoring to zero-in on consumers when they are likely to be vulnerable to short-term credit products with usurious interest rates and unfavorable terms (Hurley and Adebayo (2017)). The size and influence of these industries are staggering. In 2018, for example, the gross revenue of the gambling industry in the United States was $161 billion;[5] the video game industry generated about $135 billion in revenue (Newzoo (2018));[6] visits to Pornhub, the largest adult film website in the world, totaled 33.5 billion;[7] and 12 million Americans took out payday loans, both online and through about

[3]https://www.reddit.com/r/unitedkingdom/comments/6xdi4d/how_gambling_industry_targets_poor_people_and.

[4]http://theguardian.com/society/2019/oct/09/casumo-ad-banned-for-targeting-people-trying-to-stop-gambling.

[5]See https://www.casino.org/gambling-statistics.

[6]www.newzoo.com/globalgamesreport.

[7]See https://www.pornhub.com/insights/2018-year-in-review.

16,000 storefront offices, borrowing almost $90 billion.[8]

Motivated by these observations, we develop a model to analyze an ecosystem associated with a digital platform, such as Google or Facebook, that may collect the data of consumers on the platform and share the data with sellers. For simplicity, there are two sellers. Seller A sells a normal consumption good, such as music, while seller B sells a temptation good, such as gambling or video games.[9] Each of the sellers can target advertisements to potential buyers of its good at a convex cost. Each consumer may receive advertisements from none, one, or both of the sellers, and then chooses from the menu none, one, or both of the goods.

There are three types of consumers: The first is strong-willed and will always resist the temptation good, while the second is weak-willed and may indulge in the temptation good, as captured by the temptation utility of Gul and Pesendorfer (2001). Both strong-willed and weak-willed consumers benefit from consuming the normal good, while only the weak-willed may succumb to the temptation good. The third type of consumer would never buy either the normal good or the temptation good and serves as noise in the sellers' targeted advertising.

For simplicity, we assume that both strong-willed and weak-willed consumers have a random utility over the normal good. The random utility prevents seller A, even if it has perfect information about consumer types, from using third-degree price discrimination against its potential buyers, which is a potential cost of revealing consumer data that is analyzed in the existing privacy literature. In the absence of such price discrimination (which may also be ensured by perfect competition among multiple normal good sellers), both strong-willed and weak-willed consumers prefer receiving advertisements from seller A. Furthermore, since strong-willed consumers can always resist the temptation good, they do not mind receiving advertisements from seller B. As such, strong-willed consumers prefer a larger menu of goods, which, in turn, leads to a preference for data sharing so that they can be precisely targeted by seller A. Data sharing presents a more intricate trade-off, however, for weak-willed consumers, who benefit from more precise targeting by seller A but suffer from receiving advertisements from seller B. This is a key tension in our model that drives the differences

[8] https://www.occ.gov/news-issuances/bulletins/2018/bulletin-2018-14.html.

[9] While the vast majority of game players would not experience anything close to addiction, a fraction do struggle with gaming addiction—a legitimate medical condition. In 2018, the World Health Organization (WHO) included "gaming disorder" within the 11th Revision of the International Classification of Diseases. Interestingly, Aguiar et al. (2018) estimate that video gaming and other recreational computer activities have reduced labor supply of young men (ages 21–30) in the United States by 1.5 to 3.1 percent since 2004.

in social welfare under different data sharing schemes.

For comparison, we first consider two benchmark data sharing schemes: one without any data sharing and consumers remaining fully anonymous to the sellers, and the other with full data sharing so that the sellers can perfectly identify the type of each consumer. In the former scheme, each seller faces a dark pool of consumers, and consequently the convex cost of advertising determines that each seller only sends advertisements to a subset of potential consumers. This dark pool prevents both the strong-willed and weak-willed from being sufficiently covered by seller A, and at the same time protects the weak-willed from the temptation good of seller B. In the latter scheme with full data sharing, both sellers A and B can precisely target their advertisements to their intended consumers. As such, both the strong-willed and weak-willed benefit from the improved access to the normal good, but the weak-willed suffer from not being able to hide from the temptation good. As a result of this trade-off, our analysis shows that when the temptation of the weak-willed is sufficiently severe, the full data sharing scheme reduces their utilitarian welfare relative to the no data sharing scheme, and the harm to weak-willed consumers may even be greater than the gain of strong-willed consumers and cause lower overall social welfare.

Both the GDPR and the CCPA allow consumers to opt in or out of data collection by any platform, and its subsequent data sharing with sellers, but with an important difference in the default choice. The GDPR requires explicit consumer authorization before the platform can collect consumer data, while the CCPA allows the platform to collect consumer data unless a consumer explicitly opts out. That is, the default choice by the GDPR is opt-out unless a consumer opts in, while the default by the CCPA is opt-in unless a consumer opts out. In our model, since the third type of consumers are indifferent to either opt-in or opt-out, their choices are determined by the default setting of the data privacy regulation. Through this channel, the default data sharing choice impacts the composition of the opt-in and opt-out pools of consumers faced by the sellers.

Despite this difference, both regulations offer each consumer an appealing opt-in or opt-out choice in that strong-willed consumers can opt in to benefit from the improved matching with normal goods sellers while weak-willed consumers can opt out to protect themselves from temptation goods sellers. By the optimality of each consumer's choice, one may naturally expect these regulations to improve social welfare relative to the no data sharing and full data sharing schemes. Indeed, our analysis confirms that the CCPA strictly dominates full data sharing, as it allows seller A to fully identify its intended consumers and at the same time, provides some, albeit imperfect, protection to weak-willed consumers. The comparison of these schemes to no data sharing, however, is more subtle.

Under both the GDPR and the CCPA, all strong-willed consumers opt in for data sharing. As weak-willed consumers face a trade-off when they opt in between improved access to the normal good and intensified exposure to the temptation good, they follow a cut-off strategy in which those severely tempted opt out while those modestly tempted opt in. The equilibrium cutoff depends on the default choice instituted by the GDPR and the CCPA for the third type of consumer. Interestingly, our model highlights a key externality of an individual consumer's data sharing with sellers on other consumers. When one consumer chooses to opt in with data sharing either by an active choice or by default, that consumer's data allows the sellers to infer the preferences of other consumers. This echoes the notion of social data put forth by Acemoglu et al. (2019), Bergemann, Bonatti and Gan (2019), and Easley et al. (2019) that consumer data have an important social dimension.

The data sharing externality we highlight can be negative: when one consumer opts in, he drops out of the opt-out pool and reduces the camouflage available for those severely tempted consumers to hide from seller B. This negative externality increases with the temptation problem of weak-willed consumers. As a result, no data sharing may offer higher social welfare than both the GDPR and the CCPA when the temptation problem of weak-willed consumers is sufficiently severe. This data sharing externality, however, can also be positive: when the third type of consumers share their data under the default choice of the CCPA, their data sharing allows seller A to fully identify and therefore cover its intended consumers, including those weak-willed consumers in the opt-out pool. This positive externality allows the CCPA to dominate both the GDPR and no data sharing when the temptation problem of weak-willed consumers is sufficiently modest. The result that the CCPA may dominate the GDPR is surprising as the GDPR is often regarded as providing stronger protection for consumers. Interestingly, since the GDPR provides more of a balance between the matching efficiency of consumers with seller A and the protection of weak-willed consumers from seller B, there may exist an intermediate range of the temptation problem of weak-willed consumers in which the GDPR is the most desirable scheme.

Our paper adds a new dimension to the privacy literature by highlighting the cost of data sharing imposed on consumers who are vulnerable because of weak self-control. In this respect, it broadens the cost-benefit analysis of privacy protection relevant for policy analysis. Further, since we adopt the self-control utility framework, the weak-will consumers in our model make fully rational information-sharing choices, despite their lack of self-control in their consumption choices. This approach puts our normative analysis of data sharing schemes and privacy regulations on a solid foundation, albeit at the cost of overlooking consumers with even more severe behavioral weakness. For example, the use of hyperbolic discounting may lead consumers not to fully internalize their lack of self-control, as for example in Laibson (1997) and DellaVigna (2009). Our analysis also provides a rationale for the so-called privacy paradox, which states that, although consumers express concerns about data privacy in surveys, they often appear to freely share their data with firms and digital platforms, as, for example, in Athey et al. (2017) and Tang (2019). In our model, even weak-willed consumers may choose to opt in for data sharing to enjoy the benefits from improved matching with normal goods sellers, despite their concerns about intensified exposure to temptation goods. As such, the privacy paradox may well reflect the outcome of a nuanced cost-benefit analysis.

Our paper contributes to a growing body of literature discussing a wide range of economic issues related to data and data privacy. Ali and Benabou (2019) argue that while publicity helps induce pro-social behavior, it crowds out information aggregation, thereby providing an informational rationale for privacy. Tirole (2019) is concerned that political authorities might enlist a social rating that bundles each individual's political attitude and social graph to control society without engaging in severe repression or misinformation. Campbell, Goldfarb and Tucker (2015) show that privacy protection policies can act as de facto barriers to entry that entrench monopolies, while Calzolari and Pavan (2006) show that data sharing across firms can enhance welfare by eroding distortions arising from asymmetric information. Our model motivates the need for privacy protection in order to protect consumers with behavioral weakness and highlights nuanced effects of privacy protection regulations because consumers do not internalize the impact of their data sharing decisions on other consumers. In doing so, our model also provides a microfoundation for future analysis of how data sharing by consumers may serve as a relevant factor for affecting the macroeconomy at the cost of consumer privacy as studied, for instance, in Jones and Tonetti (2020) and Farboodi and Veldkamp (2020).

1 The Model

We consider the ecosystem associated with a digital platform, such as Google or Facebook. As a large number of consumers visit the platform, the platform can collect their digital footprints, which, in turn, reveal useful information about their consumption preferences. There are two types of consumption goods, A and B, each sold by a different goods seller. We consider good A to be a normal good, such as music, and good B to be a temptation good, such as a video game or gambling. There are three types of consumer $\{S, W, O\}$, which represent strong-willed, weak-willed, and others, respectively. Strong-willed consumers can always resist the temptation good, weak-willed consumers may not be able to resist, while the type-O will purchase neither the normal nor the temptation good.

1.1 Consumers

There is a continuum of consumers of each type. The ex ante probability of a consumer being strong-willed is $\pi_S > 0$, of being weak-willed is $\pi_W > 0$, and of being type O is $1 - \pi_S - \pi_W$. In what follows, we assume that

$$\pi_W < 1 - \pi_S,$$

which implies that the fraction of type-O consumers is positive. Both strong- and weak-willed consumers may choose one or both of goods A and B for consumption, depending on their individual preferences and the advertisements they receive from sellers.

We adopt the self-control framework of Gul and Pesendorfer (2001), who provide an axiomatic foundation for temptation. Following Kreps (1979), this framework specifies a consumer's preferences in two steps. Moving backwardly, in the second step, a consumer makes a choice from a given menu N, and, in the first step, the consumer chooses from a set of menus. Specifically, the consumer's preference for a menu N in the first step is given by the following:[10]

$$\max_{x \in N} \; [u(x) + v(x) - p(x)] - \max_{x' \in N} \; v(x'), \tag{1}$$

[10] In this specification, we follow Gul and Pesendorfer (2004) to exclude goods prices from temptation utilities. One may, however, argue that more expensive temptation goods are less tempting, all else being equal. Such a consideration can be incorporated into our framework by, for instance, specifying the consumer's preference instead as:

$$\max_{x \in N} [u(x) + v(x) - 2p(x)] - \max_{x' \in N} [v(x') - p(x)],$$

without qualitatively impacting our key insights. We choose the simpler specification for expositional brevity. We thank Shaowei Ke for pointing out this construction to us.

where x is a possible choice from the menu N, and $u(x)$, $v(x)$, and $p(x)$ are the commitment utility, temptation utility, and price, respectively, of the choice x. The consumer's actual choice from the menu in the second step is determined by the first maximization in Equation (1):

$$x_* = \arg\max_{x \in N} \ [u(x) + v(x) - p(x)],$$

which is a compromise of the commitment utility and the temptation utility. As a result of the compromise, the consumer may not choose the most tempting choice from the menu. If so, that is, $x_* \neq \arg\max_{x' \in N} v(x')$, the consumer exercises self-control. As self-control is costly to the consumer, having the most tempting choice on the menu is undesirable even if it is not eventually chosen. The last term in Equation (1), while it does not directly affect the consumer's actual choice from the menu, affects the consumer's preference for the menu. More precisely, the difference between the temptation utility of the actual choice x_* and the maximal temptation from the menu, $\max_{x' \in N} v(x') - v(x_*)$, represents the cost of self-control incurred by the consumer when it resists the temptation good.[11]

As we will discuss, the menu N faced by a consumer is random and depends on the two sellers' advertising strategies, which, in turn, depend on the platform's data sharing scheme. Our analysis consequently builds directly on the random Gul-Pesendorfer temptation utility of Stovall (2010), which can also be viewed as a special case of the random Strotz (1955) utility characterized by Dekel and Lipman (2012). As such, the ex ante utility of a consumer is the expected utility from all potential menus given the platform's data sharing scheme.

Temptation utility A consumer, with type $\tau_i \in \{S, W, O\}$, has the following commitment and temptation utilities from consuming good A and good B:

	strong-willed		weak-willed		type O	
x	$u_S(x)$	$v_S(x)$	$u_W(x)$	$v_W(x)$	$u_O(x)$	$v_O(x)$
A	$\tilde{u}_A > 0$	0	$\tilde{u}_A > 0$	0	0	0
B	$u_B < 0$	0	$u_B < 0$	$\gamma_i \bar{v} - u_B > 0$	0	0

(2)

with $u_{\tau_i}(\cdot)$ and $v_{\tau_i}(\cdot)$ denoting the commitment and temptation utility of the consumer, respectively. Both strong- and weak-willed consumers have a random utility for good A, $\tilde{u}_A$, which has a uniform distribution $H(\tilde{u}_A) \sim U[0, \bar{u}]$, with $\bar{u} > 0$ as the maximal commitment utility of consumers. One can interpret this random utility for the normal good as a transient taste for the good, such as desiring coffee instead of tea on a given day.

[11] Note that this framework subsumes the standard Von Neumann-Morgenstern utility framework. That is, if $v(x) = 0$, the consumer's choice is fully determined by his commitment utility.

Good B gives a negative commitment utility $u_B < 0$ to both strong- and weak-willed consumers, reflecting that the temptation good is ultimately harmful to consumers. As good B does not give any temptation utility to strong-willed consumers (i.e., $v_S(B) = 0$), they will never buy the temptation good. Good B gives a temptation utility of $\gamma_i \bar{v} - u_B$ to weak-willed consumers, where $\bar{v} > 0$ is a constant measuring the overall temptation of weak-willed consumers to good B, and $\gamma_i \in [0,1]$ measures a consumer's degree of temptation and has a uniform distribution $G(\gamma_i) \sim U[0,1]$ across the population of weak-willed consumers. We specify this particular form of temptation utility coefficient so that a weak-willed consumer's choice of whether to buy good B, when it is on the menu, is determined by a simple expression:

$$\begin{aligned} & \max_{x \in \{B, \emptyset\}} \left[u_W(x) + v_W(x) - p(x) \right] \\ = & \max \left\{ u_W(B) + v_W(B) - p(B), 0 \right\} = \max \left\{ \gamma_i \bar{v} - p(B), 0 \right\}, \end{aligned}$$

which implies that the consumer will choose to buy good B if his temptation coefficient γ_i is sufficiently high, that is, $\gamma_i \geq p(B)/\bar{v}$.

Note that the temptation delivered by good B to a weak-willed consumer is persistent and characterized by a personalized parameter γ_i, while the commitment utility delivered by good A to a consumer (either strong-willed or weak-willed) is random. The random utility delivered by good A prevents price discrimination by seller A even if seller A has full information about consumers.[12] In contrast, information about a weak-willed consumer allows seller B not only to precisely target its advertisements but also to price discriminate against weak-willed consumers. This asymmetric setting allows us to focus on how access to consumer data affects weak-willed consumers through their temptation utility, rather than how price discrimination affects consumers' consumption of normal goods, which has been explored extensively in the literature.

Type-O consumers (with $\tau_i = O$) prefer an outside good, and their commitment utility and temptation utility from either good A or B are both zero. The presence of these consumers makes it costly for sellers A and B to advertise their goods to their intended consumers.

[12] Alternatively, one can view this specification as reflecting a preference by consumers for a specific one of many possible normal goods with the utility benefit indexed on $[0, \bar{u}]$. Under this alternative setting, Ichihashi (2019) shows that a seller who can precommit would choose not to price discriminate when consumers can choose whether to disclose their information.

Menu preference The menu N that a consumer faces is determined by the advertisements the consumer receives from the two sellers. The menu may contain both, one, or none of goods A and B. Note that each consumer has separate and additive utilities for consumption of goods A and B. Furthermore, for simplicity, we assume that each consumer faces no budget constraints and could choose to consume both or one of A and B.[13] That is, each consumer can separately choose each good, even if both goods are on her menu. As a result, we can separately denote the menu faced by consumer i for each of the two goods: $\mathcal{M}_i^A \in \{\{A, \emptyset\}, \emptyset\}$ is the menu for good A, with $\emptyset$ representing the menu when good A is not advertised to the consumer and $\{A, \emptyset\}$ representing the menu when it is advertised, and $\mathcal{M}_i^B \in \{\{B, \emptyset\}, \emptyset\}$ is the menu for good B.

Then, building on the utility framework specified in Equation (1), we derive the choices of a consumer with type $\tau_i \in \{S, W, O\}$ from the menus $\mathcal{M}_i^A$ and $\mathcal{M}_i^B$:

$$\begin{aligned} x_{\tau_i}\left(\mathcal{M}_i^A\right) &= \arg\max_{x \in \mathcal{M}_i^A} \left[\tilde{u}_{\tau_i}(x) - p_{A,\tau_i}(x)\right], \\ y_{\tau_i}\left(\mathcal{M}_i^B\right) &= \arg\max_{y \in \mathcal{M}_i^B} \left[u_{\tau_i}(y) + v_{\tau_i}(y) - p_{B,\tau_i}(y)\right], \end{aligned}$$

where the prices of the two goods $p_{A,\tau_i}(x)$ and $p_{B,\tau_i}(y)$ may be discriminative, depending on the consumer's type and whether the consumer's type is known to the sellers. Each consumer is competitive and takes as given the sellers' advertisement policies and pricing policies.

The consumer's ex ante preference for the full menu is then

$$\begin{aligned} U_{\tau_i}\left(\mathcal{M}_i^A, \mathcal{M}_i^B\right) &= \tilde{u}_{\tau_i}\left(x_{\tau_i}\left(\mathcal{M}_i^A\right)\right) - p_{A,\tau_i}\left(x_{\tau_i}\left(\mathcal{M}_i^A\right)\right) \\ &\quad + u_{\tau_i}\left(y_{\tau_i}\left(\mathcal{M}_i^B\right)\right) + v_{\tau_i}\left(y_{\tau_i}\left(\mathcal{M}_i^B\right)\right) - p_{B,\tau_i}\left(y_{\tau_i}\left(\mathcal{M}_i^B\right)\right) - \max_{y \in \mathcal{M}_i^B} v_{\tau_i}(y). \end{aligned}$$

This menu preference allows us to analyze the welfare implications of the platform's data sharing scheme, which determines the sellers' information about each consumer and consequently their advertising strategies. In our analysis, we will separately examine different schemes regarding whether the platform shares consumer data with the sellers.

1.2 Sellers

There is one seller of good A and one seller of good B in the ecosystem. For simplicity, we assume that both sellers face zero marginal cost of production, but a convex cost of

[13] This assumption simplifies our analysis from potential complications related to the consumer's budget constraint and allows us to focus on how different data sharing schemes with sellers affect the consumer's choice and welfare.

advertising the goods to consumers. Specifically, in order for seller $k \in \{A, B\}$ to reach z_k measure of the consumers, it incurs a cost of $F\frac{z_k}{1-z_k}$ where $F > 0$ is a constant. One may interpret this cost as an advertising fee, with the convexity reflecting that it is increasingly costly to advertise to a broader audience.[14] In what follows, we impose a technical condition $F < \frac{\bar{u}}{4}$ to ensure a nontrivial equilibrium for good A.

In choosing its advertising and pricing policies, seller k maximizes its expected profit:

$$\Pi_k = \sup_{\{p_k, z_k\}} E\left[\int_{i \in Z_k} p_k(i)\, di - F\frac{z_k}{1-z_k} \;\middle|\; \mathcal{I}^k\right], \ k \in \{A, B\},$$

where Z_k is the set of consumers to which seller k advertises its good, $p_k(i)$ is the price that the seller charges consumer i, and z_k is the measure of the set Z_k. We assume that if the seller does not advertise to a consumer, then its good is not on that consumer's menu. Each seller is strategic and can only condition its advertisement and pricing policies on its information set $\mathcal{I}^k$, which may allow the seller to charge different consumers different prices.

Since consumers can always choose to buy nothing, sellers face the following implicit participation constraints:

$$p_A \leq \bar{u}, \ p_B \leq \bar{v}.$$

Violating these price constraints would lead to no sales.

1.3 Rational Expectations Equilibrium

We analyze how different data sharing schemes may affect consumers and sellers by leaving out the incentives of the platform. We implicitly assume that the platform will share all consumer data with the sellers as long as such sharing satisfies each consumer's sharing choice, if applicable. In Section 2, we first analyze two simple data sharing schemes, one without any sharing and the other with full sharing. In both of these schemes, consumers do not have any individual choice over data sharing. In Section 3, we analyze two schemes instituted by the GDPR and the CCPA, both of which allow each consumer to choose whether to share data with the platform, which then shares that data with the sellers.

Under each of these data sharing schemes, an equilibrium in the ecosystem is a set of optimal advertising and pricing policies $\{Z_k, p_k\}$ for each seller $k \in \{A, B\}$, and an optimal

[14] To the extent that consumers have limited attention and online advertisers do not want to flood them with unlimited advertisements, the fees have to rise progressively with the quantity. This is also consistent with the fact that, in practice, sellers need to pay substantial advertising fees to online advertisers.

purchase policy correspondence $\{x_{\tau_i}(\mathcal{M}_i^A), y_{\tau_i}(\mathcal{M}_i^B)\}$ and a data sharing choice s_i for each consumer i such that the following are satisfied:

- Consumer optimization: Given each seller's advertising and pricing policies, each consumer i finds it optimal to first adopt the data sharing choice s_i and then follow the purchase policy $\{x_{\tau_i}(\mathcal{M}_i^A), y_{\tau_i}(\mathcal{M}_i^B)\}$ for a menu set $\{\mathcal{M}_i^A, \mathcal{M}_i^B\}$.
- Seller optimization: Given each consumer's optimal policy, each seller k finds it optimal to choose an optimal advertising policy Z_k and a pricing policy p_k for its good.

To facilitate our welfare analysis, we assume that sellers pay the platform for its advertising services, and consequently the costs of advertising are zero-sum transfers between sellers and the platform, which is also assumed to be owned by consumers. Since consumer preferences are quasi-linear in the cost of their purchases, we can aggregate across consumer utility and seller and platform profits to arrive at the following utilitarian social welfare:

$$\begin{aligned} W &= \int \tilde{u}_A \left(\pi_S \mathbf{1}_{\left\{A\in\mathcal{M}_S^A \,\cap\, x_S=A\right\}} + \pi_W \mathbf{1}_{\left\{A\in\mathcal{M}_W^A \,\cap\, x_W=A\right\}}\right) dH(\tilde{u}_A) \\ &\quad +\pi_W \int \left(u_B \mathbf{1}_{\left\{B\in\mathcal{M}_W^B \,\cap\, x_W=B\right\}} + (u_B - \gamma_i \bar{v})\, \mathbf{1}_{\left\{B\in\mathcal{M}_W^B \,\cap\, x_W=\emptyset\right\}}\right) dG(\gamma_i)\,. \end{aligned} \tag{3}$$

The first term captures the commitment utility of both strong-willed and weak-willed consumers from consuming good A. The second term for weak-willed consumers represents the social deadweight loss from consumption of the temptation good, u_B, and the cost of resisting temptation, $u_B - \gamma_i \bar{v}$, by those who have the temptation good on their menus but choose not to consume it. Note from Equation (1) that for a weak-willed consumer who purchases good B, the realized temptation utility from consuming the good offsets the maximal temptation from the menu, thereby giving the consumer zero temptation utility. The price she pays for the good is a transfer to seller B and does not affect social welfare. As a result, the welfare loss incurred is from the negative commitment utility u_B for those weak-willed consumers who buy the good and from the mental cost of resisting temptation for those who have good B on their menus but resist it.

The social welfare given in Equation (3) reveals a trade-off associated with sharing consumer data with sellers—it increases the matching efficiency of seller A, which improves social welfare through the first term, at the expense of exposing weak-willed consumers to seller B, which reduces social welfare through the second term. This trade-off distinguishes our model from typical models of data privacy that focus on how the availability of consumer data increases the total social surplus through improved matching but also shifts the split of the surplus between consumers and sellers.

To anchor our welfare analysis of different data sharing schemes, it is straightforward to characterize the first-best outcome from the perspective of a planner who maximizes the social welfare in Equation (3). Since advertising is costless from a social perspective, the planner prefers seller A to sell its good to all strong- and weak-willed consumers. In contrast, as the advertisement from seller B brings a cost to each weak-willed consumer, regardless of whether he resists or succumbs to the temptation, the planner prefers seller B not to advertise to any consumer. We summarize this first-best outcome below.

Proposition 1 *In the first-best equilibrium, seller A sells its good to all strong-willed and weak-willed consumers, and seller B advertises to no consumers.*

2 Equilibrium Under Benchmark Schemes

In this section, we characterize the equilibrium of the ecosystem in two benchmark data sharing schemes, one without any sharing and the other with full sharing. Under both schemes, consumers do not have any individual choice to opt in or out of data sharing.

2.1 Consumer Choice

We first analyze the choice of each consumer from a given menu of consumption goods. The policy is simple. A strong-willed consumer may buy good A if its price is below the consumer's reservation value, and always refuses good B. A weak-willed consumer may buy good A if its price is lower than his reservation value, just like a strong-willed consumer, and may buy good B if his temptation coefficient γ_i is sufficiently high relative to the price of the good. The following proposition summarizes these choices in detail, with the proof provided in the appendix.

Proposition 2 *A strong-willed consumer with commitment utility $\tilde{u}_A$ will purchase good A if it is offered at a price below his reservation value $p_A \leq \tilde{u}_A$, and always reject good B. A weak-willed consumer with commitment utility $\tilde{u}_A$ and temptation coefficient γ_i will purchase good A if it is offered at a price below his reservation value $p_A \leq \tilde{u}_A$, and purchase good B if it is on the menu and if his temptation coefficient γ_i is sufficiently high: $\gamma_i \geq \frac{p_B}{\bar{v}}$.*

This proposition reveals that both strong-willed and weak-willed consumers may reject good A if their random utility for the good happens to be lower than its price. This possibility prevents seller A from imposing price discrimination on any consumer. Ex ante, all strong-willed and weak-willed consumers still prefer to receive the advertisement of good A so that they can benefit from a high realization of their random utility for the good. This benefit motivates both strong-willed and weak-willed consumers to share their data with seller A.

Proposition 2 also shows that even when good B is on their menu, only those weak-willed consumers with a sufficiently high temptation coefficient γ_i will buy it. Those with a modest temptation ($\gamma_i < p_B/\bar{v}$) resist it but still suffer a mental cost of $\gamma_i\bar{v} - u_B$ from exercising self-control. Those with strong temptation buy good B and suffer from not only paying the price of p_B to purchase the good, but also from enduring the negative commitment utility of u_B that this purchase entails.

2.2 Equilibrium With No Data Sharing

We first analyze a benchmark scheme in which the platform does not collect or share any consumer data with sellers. As a result, sellers have no information about any consumer's type. The following proposition characterizes the equilibrium.

Proposition 3 *With no data sharing (NS), there exists a unique equilibrium with the following properties:*

1. *Seller A randomly advertises good A to z_A^{NS} measure of consumers:*

$$z_A^{NS} = \min\left\{\max\left\{1 - 2\sqrt{\frac{1}{\pi_S + \pi_W}\frac{F}{\bar{u}}}, 0\right\}, 1\right\}, \tag{4}$$

   at a uniform price: $p_A^{NS} = \frac{1}{2}\bar{u}$.

2. *Seller B randomly advertises good B to z_B^{NS} measure of consumers:*

$$z_B^{NS} = \min\left\{\max\left\{1 - 2\sqrt{\frac{1}{\pi_W}\frac{F}{\bar{v}}}, 0\right\}, 1\right\}, \tag{5}$$

   at a uniform price: $p_B^{NS} = \frac{1}{2}\bar{v}$.

Under this benchmark scheme of no data sharing, the sellers' undirected advertising leads to a source of inefficiency. As a result, seller A limits its advertising to a small pool of potential consumers. Equation (4) shows that seller A's advertising intensity z_A^{NS} decreases with its cost parameter F, and increases with $\pi_S + \pi_W$ (the fraction of intended consumers in the population) and $\bar{u}$ (which determines the price of good A). As shown by Equation (5), anonymity protects weak-willed consumers from being targeted by seller B. Without any knowledge about the reservation value of their consumers, both sellers charge all consumers the same prices for the goods, $p_A^{NS} = \frac{1}{2}\bar{u}$ and $p_B^{NS} = \frac{1}{2}\bar{v}$, which implies that the sellers' advertisements are accepted by their intended consumers half of the time.

2.3 Equilibrium With Full Data Sharing

We now consider a very different scheme under which the platform is able to collect consumers' data and therefore to determine not only the mental state of each consumer $\tau(i)$ but also the severity of each weak-willed customer's temptation coefficient γ_i. While this assumption exaggerates the current power of big data analytics, the rapid development of innovative data analytics over the years is moving us closer to this instructive limiting case. By sharing the data with goods sellers, the platform allows sellers to use different advertising and pricing strategies for different types of consumers.

As strong-willed and weak-willed consumers have the same preference for good A and their purchase decision regarding good A is not affected by good B, there is no need for seller A to differentiate strong-willed and weak-willed consumers. We denote z_A^{FS} as the measure of strong-willed and weak-willed consumers, to whom seller A advertises its good at a price of p_A^{FS}. Proposition 4 derives the seller's optimal z_A^{FS} and p_A^{FS}. Data sharing allows seller A to achieve a higher level of efficiency by avoiding advertising to the type-O consumers who would never buy good A. As a result of the improved efficiency, seller A advertises more with full data sharing than with no sharing, that is, $z_A^{FS} \geq z_A^{NS}$, which in turn implies that both strong-willed and weak-willed consumers have a strictly higher probability of being covered by seller A. As the seller does not know the reservation value of the targeted consumers, it again charges the same price $p_A^{FS} = \frac{1}{2}\bar{u}$.

Access to consumer data also allows seller B to focus its advertising on weak-willed consumers. Furthermore, since seller B also observes the severity of each weak-willed customer's temptation, it will price discriminate against each targeted weak-willed consumer by charging his full reservation value, $p_B(\gamma_i) = \gamma_i \bar{v}$, which is the net utility cost of resisting temptation. Such price discrimination in turn motivates the seller to concentrate its advertising only on the most tempted consumers, that is, those with γ_i higher than a threshold $\hat{\gamma}^{FS}$. As a result, full data sharing allows seller B to precisely target weak-willed consumers at greater intensity than under the no data sharing scheme and to perfectly price discriminate against them.

We summarize the equilibrium in the following proposition.

Proposition 4 *With full data sharing (FS), there exists a unique equilibrium with the following properties:*

1. *Seller A advertises its good to z_A^{FS} measure of strong-willed and weak-willed consumers:*

$$z_A^{FS} = \min\left\{\max\left\{1-2\sqrt{\frac{F}{\bar{u}}},0\right\},\pi_S+\pi_W\right\}$$

   at the same price $p_A^{FS}=\frac{1}{2}\bar{u}$.

2. *Seller B advertises its good to all weak-willed consumers with $\gamma_i \geq \hat{\gamma}^{FS} = 1-\frac{z_B^{FS}}{\pi_W}$, where z_B^{FS} is the total advertising by seller B:*

$$\begin{aligned} z_B^{FS} &= \min\left\{\frac{2+\pi_W}{3}-\sqrt[3]{\left(\frac{1-\pi_W}{3}\right)^3+\frac{\pi_W F}{2\bar{v}}+\sqrt{\left(\left(\frac{1-\pi_W}{3}\right)^3+\frac{\pi_W F}{2\bar{v}}\right)^2-\left(\frac{1-\pi_W}{3}\right)^6}}\right. \\ &\left. -\sqrt[3]{\left(\frac{1-\pi_W}{3}\right)^3+\frac{\pi_W F}{2\bar{v}}-\sqrt{\left(\left(\frac{1-\pi_W}{3}\right)^3+\frac{\pi_W F}{2\bar{v}}\right)^2-\left(\frac{1-\pi_W}{3}\right)^6}},\pi_W\right\}, \end{aligned}$$

   and z_B^{FS} is weakly increasing in $\bar{v}$ and decreasing in F. Furthermore, seller B charges each consumer a price equal to his reservation utility $p_B(\gamma_i)=\gamma_i\bar{v}$.

Data sharing strictly benefits strong-willed consumers by improving their access to the normal good but presents a trade-off to weak-willed consumers. On the one hand, they have better access to the normal good, which improves their welfare; on the other, they are also more exposed to the temptation good, which hurts their welfare. As a consequence, the net effect is ambiguous. As each weak-willed consumer suffers from the negative commitment utility u_B of the temptation good, the utilitarian welfare of weak-willed consumers is increasing in u_B. Proposition 5 shows that when the temptation problem of weak-willed consumers is sufficiently severe, that is, u_B is lower than a critical level, full data sharing reduces the welfare of weak-willed consumers by so much that it even reduces social welfare relative to no data sharing.

Proposition 5 *There exists a critical level of u_B, below which full data sharing lowers social welfare relative to no data sharing.*

The comparison of the schemes with no data sharing and with full data sharing highlights a trade-off brought by data sharing—it improves the efficiency of seller A in covering its intended consumers at the expense of exposing weak-willed consumers to the temptation good. This trade-off motivates the enactment of privacy regulations that allow each consumer to opt in or out of data sharing. We explore two examples of such regulations in the next section.

3 Opt-In and Opt-Out Regulations

The General Data Privacy Regulation (GDPR) enacted by the European Union and the California Consumer Privacy Act (CCPA) aim to protect consumer privacy by giving consumers the right to opt in or out of data sharing with digital platforms. These regulations offer the promise of a Pareto efficient outcome since strong-willed consumers can choose to opt in, and consequently benefit from data sharing, while severely tempted consumers can choose to opt out, and consequently protect themselves from the temptation good. The GDPR and CCPA differ, however, in a key default feature—the GDPR empowers consumers with the initial allocation of rights to their personal data and allows digital platforms to collect consumer data only after explicit consumer authorization, while the CCPA gives businesses the initial rights to collect consumer data and allows consumers to opt out of data collection through an explicit request.

In this section, we analyze whether these regulations can provide sufficient protection to consumers. Specifically, we allow each consumer to choose whether to share his data with the platform, which then shares the data, if allowed by the consumer, with both sellers.[15] Under the GDPR, the default choice is opt-out if a consumer is indifferent, while under the CCPA, the default is opt-in. This default choice directly determines the choice of the type-O consumers in our model.

[15] Note that we do not allow consumers to separately choose whether to share data with each seller. As no one would choose to share data with sellers of temptation goods without other benefits, temptation good sellers, in practice, would bundle their data-share requests with certain conveniences or benefits. In our model, the improved access to the normal good serves as such a benefit.

3.1 The GDPR

We first analyze the equilibrium under the GDPR. Since strong-willed consumers strictly benefit from having their data shared with seller A and are not concerned with seller B, it is straightforward to see that all strong-willed consumers opt in to data sharing. As the type-O consumers are not interested in either good in the platform, there is neither gain nor loss for them from data sharing; as such, they are indifferent between opting in and opting out. As opting out is the default choice of the GDPR, the type-O consumers opt out of data sharing, given that it would take extra effort for an indifferent consumer to opt in. There is now, however, a nontrivial choice for each weak-willed consumer. By opting in for data sharing, weak-willed consumers benefit from the improved access to good A but are also more exposed to the temptation good. It is intuitive to conjecture that weak-willed consumers with sufficiently high temptation coefficient γ_i, that is, higher than a critical level γ_{**}^{GDPR}, will choose to opt out of data sharing, while those with γ_i lower than γ_{**}^{GDPR} will opt in.

The utility of a weak-willed customer that opts in with data sharing is

$$\begin{aligned} U_{W,in}^{GDPR}(\gamma_i) \quad = \quad & \frac{z_{A,in}^{GDPR}}{\pi_S + \pi_W \gamma_{**}^{GDPR}} \int_0^{\bar{u}} \max\left\{\tilde{u}_A - p_{A,in}^{GDPR}, 0\right\} dH(\tilde{u}_A) \\ & + \frac{\hat{z}_{B,in}^{GDPR}(d\gamma_i)}{\pi_W d\gamma_i} \left(u_B - p_{B,in}^{GDPR}(\gamma_i) \mathbf{1}_{\left\{\gamma_i \geq \frac{p_{B,in}^{GDPR}(\gamma_i)}{\bar{v}}\right\}} - \gamma_i \bar{v} \mathbf{1}_{\left\{\gamma_i < \frac{p_{B,in}^{GDPR}(\gamma_i)}{\bar{v}}\right\}} \right), \end{aligned} \tag{6}$$

where $z_{A,in}^{GDPR}$ is the total advertising by seller A to the opt-in pool at price $p_{A,in}^{GDPR}$, $\hat{z}_{B,in}^{GDPR}(d\gamma_i) \in [0, \pi_W] \, d\gamma_i$ is the advertising intensity of seller B to opt-in consumers with temptation coefficient γ_i, and $p_{B,in}^{GDPR}(\gamma_i)$ is the price that seller B charges them. Note that the consumer's utility is determined by his conditional probability of being targeted by both sellers, $\frac{z_{A,in}^{GDPR}}{\pi_S+\pi_W\gamma_{**}^{GDPR}}$ and $\frac{\hat{z}_{B,in}^{GDPR}(d\gamma_i)}{\pi_W d\gamma_i}$. His utility from opt-out is

$$\begin{aligned} U_{W,out}^{GDPR}(\gamma_i) \quad = \quad & \frac{z_{A,out}^{GDPR}}{1 - \pi_S - \pi_W \gamma_{**}^{GDPR}} \int_0^{\bar{u}} \max\left\{\tilde{u}_A - p_{A,out}^{GDPR}, 0\right\} dH(\tilde{u}_A) \\ & + \frac{z_{B,out}^{GDPR}}{1 - \pi_S - \pi_W \gamma_{**}^{GDPR}} \left(u_B - p_{B,out}^{GDPR} \mathbf{1}_{\left\{\gamma_i \geq \frac{p_{B,out}^{GDPR}}{\bar{v}}\right\}} - \gamma_i \bar{v} \mathbf{1}_{\left\{\gamma_i < \frac{p_{B,out}^{GDPR}}{\bar{v}}\right\}} \right), \end{aligned} \tag{7}$$

where $z_{A,out}^{GDPR}$ is the total advertising by seller A to the opt-out pool at price $p_{A,out}^{GDPR}$, and $z_{B,out}^{GDPR}$ is the total advertising by seller B to the opt-out pool at price $p_{B,out}^{GDPR}$. For a consumer to opt in for data sharing, it must be the case that

$$U_{W,in}^{GDPR}(\gamma_i) \geq U_{W,out}^{GDPR}(\gamma_i).$$

The following proposition characterizes the equilibrium under the GDPR.

Proposition 6 *Suppose $\bar{u} < 8(\bar{v} - u_B)$; then there exists an equilibrium under the GDPR with the following properties:*

1. *All strong-willed consumers opt in, and a weak-willed consumer chooses to opt in if $\gamma_i \leq \gamma_{**}^{GDPR}$ and opt out if $\gamma_i > \gamma_{**}^{GDPR}$, where γ_{**}^{GDPR} is the unique root of Equation (22) in $(0,1)$.*

2. *Seller A charges the opt-in and opt-out pools the same price: $p_{A,in}^{GDPR} = p_{A,out}^{GDPR} = \frac{1}{2}\bar{u}$, and adopts a water-filling advertising strategy with priority to the opt-in pool:*

$$\begin{aligned}
z_{A,in}^{GDPR} &= \min\left\{1 - 2\sqrt{\frac{F}{\bar{u}}}, \pi_S + \gamma_{**}^{GDPR}\pi_W\right\}, \\
z_{A,out}^{GDPR} &= \min\left\{\max\left\{1 - 2\sqrt{\frac{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}{\left(1-\gamma_{**}^{GDPR}\right)\pi_W}\frac{F}{\bar{u}}} - z_{A,in}^{GDPR}, 0\right\}, 1 - \pi_S - \gamma_{**}^{GDPR}\pi_W\right\}.
\end{aligned}$$

3. *Seller B also adopts a water-filling advertising strategy with priority to the opt-in pool by targeting tempted consumers with $\gamma_i \in \left[\hat{\gamma}^{GDPR}, \gamma_{**}^{GDPR}\right]$ and charging their reservation utility: $p_{B,in}^{GDPR}(\gamma_i) = \gamma_i\bar{v}$. After it exhausts the most-tempted in the opt-in pool, seller B may also target a measure $z_{B,out}^{GDPR}$ of consumers in the opt-out pool by charging a fixed price of $p_{B,out}^{GDPR} = \max\left\{\frac{1}{2}, \gamma_{**}^{GDPR}\right\}\bar{v}$. Its total advertising in the opt-in pool $z_{B,in}^{GDPR}$ and in the opt-out pool $z_{B,out}^{GDPR}$ is given by Equations (19) and (20).*

4. *It is sufficient, although not necessary, for $\bar{u} < 4F\left(1 - \pi_S - \frac{1}{2}\pi_W\right)^{-2}$ to ensure that the equilibrium is unique.*[16]

Proposition 6 confirms that weak-willed consumers follow a cutoff strategy to opt in and out of data sharing—those with mild temptation (low γ_i) opt in, while those with severe temptation (high γ_i) opt out. This is intuitive since weak-willed customers with mild temptation benefit more from the better coverage from seller A than the temptation they

[16] Although we provide a sufficient condition for a cutoff equilibrium to be unique, we find numerically that such an equilibrium appears to be unique for a much wider range of $\bar{u}$.

suffer from the intensified exposure to seller B. In contrast, severely tempted weak-willed consumers are willing to forego the benefit of better coverage from seller A to mitigate the temptation cost of being targeted by seller B. Since seller B can efficiently target weak-willed consumers in the opt-in pool, it gives higher priority to target the most tempted in the opt-in pool and charges them the full reservation value of their temptation. Only after seller B exhausts the most-tempted in the opt-in pool, and equates the marginal revenues from advertising to the opt-in and opt-out pools, may it also target some in the opt-out pool by charging a fixed price, equal to the maximum between $\frac{1}{2}\bar{v}$ and $\gamma_{**}^{GDPR}\bar{v}$.

Seller B's priority to cover the opt-in pool ensures that the most-tempted consumers will choose opt-out and therefore hide in the opt-out pool from seller B. This protection is weakened by the opt-in decisions of other consumers, that is, strong-willed and modestly weak-willed consumers. Their departure from the opt-out pool reduces the camouflage of those severely weak-willed and increases $\frac{z_{B,out}^{GDPR}}{1-\pi_S-\pi_W\gamma_{**}}$, the probability of weak-willed in the opt-out pool being targeted by seller B. In this sense, there is a *negative* externality in the opt-in decisions of strong-willed and modestly weak-willed consumers, as their decisions do not account for the potential effect on other consumers with severe temptation. This externality echoes the notion of social data put forth by Acemoglu et al. (2019), Bergemann, Bonatti and Gan (2019), and Easley et al. (2019) that data have an important social dimension, as each individual's data also reveals information about others. The presence of this externality suggests that simply allowing consumers to opt in or out of data sharing may not be sufficient for consumers with severe temptation to protect themselves.

3.2 The CCPA

The CCPA makes opt-in the default choice for each consumer and thus makes all type-O consumers opt in for data sharing. This key difference leads to a very different equilibrium under the CCPA than under the GDPR. First, when type-O consumers, who are the consumers that seller A needs to avoid, are all identified, seller A can perfectly target all strong-willed and weak-willed customers, including those that opt out. The default opt-in choice of type-O consumers therefore induces a *positive* externality by allowing seller A to perfectly cover weak-willed consumers in the opt-out pool. This positive externality highlights a subtle benefit of data sharing and provides a justification for the CCPA approach over the GDPR approach, albeit at the expense of removing type-O consumers from the opt-out pool to hide severely weak-willed consumers from seller B.

Second, as seller A symmetrically covers both the opt-in and opt-out pools, strong-willed consumers are also indifferent between opt-in and opt-out and would opt in by default. The opt-in of strong-willed and type-O consumers exposes all weak-willed consumers to seller B. Nevertheless, seller B still faces a challenge in sorting out their degrees of temptation, since seller B cannot directly observe the temptation coefficients of weak-willed consumers in the opt-out pool. As advertising to consumers with prices higher than their reservation values would lead to rejection, it is optimal for seller B to strategically commit not to target weak-willed consumers in the opt-in pool with γ_i lower than a cutoff of γ_{**}^{CCPA}. This strategy bifurcates the pool of weak-willed consumers and segregates the severely tempted consumers above the cutoff in the opt-out pool. Thus, seller B can improve its advertising efficiency to this group of most-tempted consumers and maximize its net profit.

The following proposition characterizes the equilibrium under the CCPA.

Proposition 7 *The equilibrium under the CCPA has the following properties:*

1. *All strong-willed consumers opt in, and all weak-willed consumers follow a cutoff strategy of choosing opt-in if $\gamma_i \leq \gamma_{**}^{CCPA} = \frac{1}{2}$ and opt-out if $\gamma_i > \gamma_{**}^{CCPA} = \frac{1}{2}$.*

2. *Seller A charges consumers in the opt-in and opt-out pools the same prices $p_A^{CCPA} = \frac{1}{2}\bar{u}$, and advertises to the strong and weak-willed consumers in the opt-in and opt-out pools in the same way as under full data sharing:*

$$z_A^{CCPA} = \min\left\{\max\left\{1 - 2\sqrt{\frac{F}{\bar{u}}}, 0\right\}, \pi_S + \pi_W\right\}.$$

3. *Seller B chooses the following advertising intensity for weak-willed consumers in the opt-out pool:*

$$z_{B,out}^{CCPA} = \min\left\{\max\left\{1 - 2\sqrt{\frac{F}{\bar{v}}}, 0\right\}, \frac{1}{2}\right\}.$$

and charges them a fixed price of $p_{B,out}^{CCPA} = \frac{1}{2}\bar{v}$, and commits not to advertise to consumers in the opt-in pool.

Proposition 7 shows the sharply different equilibrium outcomes under the CCPA than under the GDPR. Under the CCPA, seller B strategically commits not to target the half of mildly tempted consumers in the opt-in pool and instead focuses on the other half of severely tempted consumers in the opt-out pool by charging them a high price of $p_{B,out}^{CCPA} = \frac{1}{2}\bar{v}$. In contrast, under the GDPR, seller B gives higher priority to the relatively more tempted consumers in the opt-in pool because its targeting efficiency of those even more tempted consumers in the opt-out pool is relatively poor given the presence of type-O consumers in the pool. Only after it exhausts the more tempted ones in the opt-in pool does it start to target those in the opt-out pool.

More generally, by making opt-in the default choice, the CCPA makes consumer data more accessible to sellers, which brings both a benefit and a cost relative to the GDPR. From a benefit perspective, seller A is able to fully identify its intended consumers, including those weak-willed consumers in the opt-out pool. From a cost perspective, seller B is also able to fully identify all weak-willed consumers in both the opt-in and opt-out pools, albeit without the full information of the degree of temptation of those in the opt-out pool.

3.3 Welfare Comparison

In this subsection, we compare the welfare consequences of the four data sharing schemes that we have analyzed: no data sharing, full data sharing, the GDPR, and the CCPA.

Under a given data sharing scheme, recall from Equation (3) that social welfare is determined by the aggregate utility of all strong-willed and weak-willed consumers over the two consumption goods, under the assumptions that the marginal cost of good production is zero and that the prices of goods and the cost of advertising are all zero-sum transfers within the population. As seller A cannot price discriminate against its customers (with the consumers' random utility for good A), it always charges a price of $\bar{u}/2$ for its good, resulting in only half of the intended consumers having their random utility above $\bar{u}/2$ to consume the good. As a result, the consumers' net utility gain from good A is $\frac{3}{8}\bar{u}\rho_A$, where ρ_A is the measure of strong-willed and weak-willed consumers that receive seller A's advertising. For good B, the weak-willed consumers who purchase the good (with a measure of ρ_B) suffer a negative utility of $u_B < 0$, while those who receive the advertising from seller B but resist the temptation (which we mark in a set S_B) suffer a mental cost of $u_B - \gamma_i \bar{v}$. Taken together, the social welfare is

$$W = \frac{3}{8}\bar{u}\rho_A + u_B\rho_B + \int_{i\in S_B} (u_B - \gamma_i\bar{v})\, dG(\gamma_i).$$

Note that ρ_A, ρ_B, and S_B are determined by the sellers' advertising and pricing strategies under each of the data sharing schemes. Except under no data sharing, seller B chooses an optimal advertising strategy in which its advertising is always accepted by its targeted weak-willed consumers, that is, S_B is empty in equilibrium. As such, across these data sharing schemes, the key trade-off is between the first term (the benefit from good A) and the second and third terms (the cost from good B).

Proposition 8 *The social ranking of full data sharing, no data sharing, the GDPR and the CCPA, has the following properties:*

- *The full data sharing scheme is strictly dominated by the CCPA.*
- *The CCPA gives the highest social welfare if the temptation problem of weak-willed consumers is sufficiently modest, that is, u_B is close to zero.*
- *The no data sharing scheme gives the highest welfare if the temptation problem is sufficiently severe, that is, u_B is sufficiently negative.*
- *There may exist an intermediate range of u_B such that the GDPR gives the highest social welfare.*

Proposition 8 first shows that the full data sharing scheme is always dominated by the CCPA. This is because the CCPA allows seller A to fully identify its intended consumers and thus provides the same coverage to both strong-willed and weal-willed consumers as under full data sharing. At the same time, the CCPA provides some, albeit imperfect, protection to weak-willed consumers against seller B. This ranking consequently supports the common wisdom that giving each consumer the choice to opt out of data sharing helps to improve social welfare. The CCPA, however, may or may not be the most desirable scheme relative to the other two data sharing schemes.

Which scheme among no data sharing, the GDPR, and the CCPA gives the highest social welfare depends on the trade-off between improving the matching efficiency of seller A with both strong-willed and weak-willed consumers and protecting weak-willed consumers from seller B. Generally speaking, by making consumer data the most accessible, the CCPA offers the best matching efficiency with seller A but the worst protection of weak-willed consumers from seller B. The no data sharing scheme lies at the other end of the spectrum—it offers the worst matching efficiency with seller A but the best protection of weak-willed consumers from seller B. The GDPR lies in the middle along both dimensions. As a result of this trade-off, Proposition 8 shows that the CCPA is the most desirable scheme if the temptation problem of weak-willed consumers is sufficiently modest, that is, with u_B close to zero, while no data sharing is the most desirable scheme if the temptation problem is sufficiently severe, that is, u_B sufficiently negative. The GDPR may be the most desirable in an intermediate range of u_B that balances both the benefits and costs of data sharing.

4 Conclusion

In this paper, we develop a model in which data sharing between consumers and firms may be harmful to consumers in the presence of temptation. While data sharing improves the matching between consumers and sellers of normal goods, it also allows sellers of temptation goods to target weak-willed consumers. Our model allows us to analyze the advertising and pricing strategies of normal goods and temptation goods sellers and the data sharing strategies of consumers under different data sharing schemes. These schemes include not only the primitive extremes of no data sharing and full data sharing, but also the more elaborate arrangements adopted by the GDPR and the CCPA that allow each consumer to opt in or opt out of data sharing. Our analysis highlights several general insights about the welfare consequences of data sharing.

First, it is beneficial to give each consumer the choice of opting in or opting out of data sharing, as illustrated by the dominance of the CCPA over the full data sharing scheme.

Second, giving each consumer the data sharing choice does not necessarily lead to the most desirable social efficiency as a result of the presence of data sharing externalities—that the sharing of data with sellers by one consumer may also affect the welfare of other consumers because it allows sellers to infer their preferences and behaviors. This externality can be either positive or negative. By letting normal goods sellers better cover their intended consumers, data sharing leads to a positive externality on other consumers, which provides a justification for making consumer data more widely accessible to firms. By allowing temptation goods sellers to more easily target weak-willed consumers, however, data sharing may also generate a negative externality, which motivates the regulation of data collection and sharing. As a result of this negative externality, no data sharing may deliver the highest social welfare when the temptation problem is sufficiently severe.

Third, the default choice in the data sharing scheme can have substantial effects on the equilibrium outcomes, as reflected by the differences between the equilibria under the GDPR and the CCPA. As part of the consumer population may be indifferent to sharing or not sharing their data, the default choice makes the data of these indifferent consumers automatically available or unavailable to sellers, which in turn affects other consumers because of the presence of data sharing externalities. The CCPA makes opt-in the default choice to maximize the positive externality of data sharing, while the GDPR makes opt-out the default choice, which gives a more balanced trade-off between the positive and negative externalities of data sharing. Interestingly, our analysis shows that the GDPR may deliver the highest social welfare when the temptation problem is in an intermediate range, while the CCPA is most desirable when the temptation problem is sufficiently modest. The result that the CCPA may dominate the GDPR is surprising as the GDPR is commonly regarded to be more protective of consumers.

Fourth, our analysis also highlights that each consumer's data sharing choice represents a subtle trade-off between cost and benefit, which are determined by not only the consumer's own choice but also the choices of others and the overall data sharing scheme of the platform. This subtle trade-off offers a potential explanation for the so-called privacy paradox that, despite the tendency for consumers to state their concerns about their data privacy in surveys, they often appear to freely share their data with firms and digital platforms, see, for example, Athey et al. (2017) and Tang (2019). Our analysis ultimately suggests that consumers may be willing to share their data, despite their concerns about data privacy, if the benefit outweighs the cost.

Appendix A Privacy Law in the EU and the US

The General Data Privacy Regulation (GDPR) came into force across the European Union on May 25, 2018. It applies to the processing of personal data by businesses established within the European Union and, importantly, to businesses outside the European Union if their data collection activities are related to individuals in the European Union. The GDPR gives European Union citizens more control over their personal data. For example, it empowers users with the right to access and get a copy of their data from internet service providers, erase their data from businesses ("the right to be forgotten"), and freely move their data on one internet platform to another (data portability). The GDPR imposes serious fines for infringement of rights and noncompliance, which are as high as $20 million or 4 percent of annual revenue of a firm.

There are two basic models of legal arrangements for privacy and data protection: opt-in and opt-out. Under the opt-in regime, data collectors must obtain consumers' explicit consent before collecting, using, and sharing their personal information. The GDPR adopts the opt-in system, and it also requires consent to be freely given, specific, informed and unambiguous. In contrast, in the opt-out regime, data collectors can collect and share non-public consumer information with third parties, but need to give consumers an opportunity to deny them permission to do so (i.e., opt out). The fundamental difference between opt-in and opt-out regimes is the initial allocation of the property rights over personal information. In the opt-in regime, the rights are assigned to consumers by default, whereas in the opt-out system, the entitlements are allocated to firms, though consumers have the right to withdraw.

The California Consumer Privacy Act (CCPA) became effective on January 1, 2020. Absent a comprehensive federal privacy law, the CCPA is considered to be one of the most significant legislative privacy developments in the United States. Its impact is global given the scale of California's economy. The CCPA adopts the opt-out regime. By default, firms can collect customer data and share them with third parties. The CCPA protects consumers by requiring firms to allow California consumers to make the following requests: (1) to provide information about what personal information firms have collected and whom firms have shared consumer information with, (2) to delete consumer information, or (3) not to sell consumer information. After a consumer opts out, a business cannot sell the consumer's information without the consumer's written consent, and the business cannot ask for that consent for 12 months after the consumer opts out.

Appendix B Proofs of Propositions

B.1 Proof of Proposition 2

We first consider a strong-willed consumer, that is, $\tau(i) = S$, who has the following preferences over different menus:

$$\begin{aligned} U_S(\{A, \emptyset\}) &= \max\{\tilde{u}_A - p_A, 0\}, \\ U_S(\{B, \emptyset\}) &= 0. \end{aligned}$$

Consequently, seller A will buy good A if $\tilde{u}_A \geq p_A$.

Consider now a weak-willed consumer, $\tau(i) = W$, with the following preferences:

$$\begin{aligned} U_W(\{A, \emptyset\}) &= \max\{\tilde{u}_A - p_A, 0\}, \\ U_W(\{B, \emptyset\}) &= u_B + \max\{-p_B, -\gamma_i \bar{v}\}. \end{aligned}$$

Choosing B from the menu $\{B, \emptyset\}$ is optimal if buying B delivers higher utility: $-p_B > -\gamma_i \bar{v}$, which is equivalent to $\gamma_i > \frac{p_B}{\bar{v}}$.

B.2 Proof of Proposition 3

Given the advertising and pricing strategies of seller A, Proposition 2 implies that the quantity of goods sold by seller A is

$$Q_A^{NS} = (\pi_S + \pi_W) z_A^{NS} \left(1 - H\left(p_A^{NS}/\bar{u}\right)\right), \tag{8}$$

and consequently the seller's profit net of the advertisement cost is

$$\Pi_A^{NS} = p_A^{NS} (\pi_S + \pi_W) z_A^{NS} \left(1 - H\left(p_A^{NS}/\bar{u}\right)\right) - F\frac{z_A^{NS}}{1 - z_A^{NS}}. \tag{9}$$

Similarly, the quantity of goods sold by seller B is

$$Q_B^{NS} = \pi_W z_B^{NS} \left(1 - G\left(p_B^{NS}/\bar{v}\right)\right), \tag{10}$$

and the net profit of seller B is

$$\Pi_B^{NS} = p_B^{NS} \pi_W z_B^{NS} \left(1 - G\left(p_B^{NS}/\bar{v}\right)\right) - F\frac{z_B^{NS}}{1 - z_B^{NS}}.$$

Technological feasibility requires that $z_A^{NS} \geq 0$ and $z_B^{NS} \geq 0$.

The first-order condition of Equation (9) with respect to z_A^{NS} is

$$p_A^{NS} Q_A^{NS} = F\frac{z_A^{NS}}{(1 - z_A^{NS})^2}. \tag{11}$$

Then, we have that

$$\Pi_A^{NS} = p_A^{NS} Q_A^{NS} - F \frac{z_A^{NS}}{1 - z_A^{NS}} = F \left(\frac{z_A^{NS}}{\left(1 - z_A^{NS}\right)^2} - \frac{z_A^{NS}}{1 - z_A^{NS}} \right) = F \left(\frac{z_A^{NS}}{1 - z_A^{NS}} \right)^2 .$$

Similarly, the first-order condition with respect to z_B^{NS} is

$$p_B^{NS} Q_B^{NS} = F \frac{z_B^{NS}}{\left(1 - z_B^{NS}\right)^2},$$

which further implies that

$$\Pi_B^{NS} = F \left(\frac{z_B^{NS}}{1 - z_B^{NS}} \right)^2 .$$

The first-order conditions for the goods prices set by the two sellers are

$$Q_A^{NS} = \frac{p_A^{NS}}{\bar{u}} \left(\pi_S + \pi_W\right) z_A^{NS} 1_{\left\{0 \le p_A^{NS} \le \bar{u}\right\}}, \tag{12}$$

$$Q_B^{NS} = \frac{p_B^{NS} \pi_W z_B^{NS}}{\bar{v}} 1_{\left\{0 \le p_B^{NS} \le \bar{v}\right\}}. \tag{13}$$

Note that the expected quantities sold by both sellers, Q_A^{NS} and Q_B^{NS}, are nonnegative, and the net profits with respect to prices are concave, since

$$\frac{d^2 \Pi_A^{NS}}{d\left(p_A^{NS}\right)^2} = -\frac{2}{\bar{u}} \left(\pi_S + \pi_W\right) z_A^{NS} h\left(p_A^{NS}/\bar{u}\right) 1_{\left\{0 \le p_A^{NS} \le \bar{u}\right\}} \le 0,$$

$$\frac{d^2 \Pi_B^{NS}}{d\left(p_B^{NS}\right)^2} = -2\pi_W z_B^{NS} g\left(\gamma_*^{NS}\right) \frac{1}{\bar{v}} 1_{\left\{0 \le p_B^{NS} \le \bar{v}\right\}} \le 0.$$

It follows that optimal prices will always be nonnegative. Since

$$\frac{d^2 \Pi_A^{NS}}{d\left(z_A^{NS}\right)^2} = -2 \frac{F}{\left(1 - z_A^{NS}\right)^3} < 0,$$

and $\frac{d^2 \Pi_A^{NS}}{dp_A^{NS} dz_A^{NS}} = 0$, it follows that the Hessian for seller A's optimization with respect to $\left(p_A^{NS}, z_A^{NS}\right)$ is negative definite and that the FOCs are sufficient.

For strong-willed consumers, there are two possibilities: $p_A^{NS} \in [0, \bar{u}]$ or $p_A^{NS} \notin [0, \bar{u}]$. If $p_A^{NS} \notin [0, \bar{u}]$, then either $p_A^{NS} = 0$ or $p_A^{NS} > \bar{u}$, neither of which generates revenue for seller A, and advertising is costly. Consequently, it must be the case that $p_A^{NS} \in [0, \bar{u}]$. Then, Equations (8) and (12) imply that $p_A^{NS} = \frac{1}{2}\bar{u}$.

Similarly, for seller B, if $p_B^{NS} \notin [0, \bar{v}]$, then either $p_B^{NS} = 0$ or $p_B^{NS} > \bar{v}$. Neither case generates any revenue, but advertising is costly. If $p_B^{NS} \in [0, \bar{v}]$, then Equations (10) and (13) imply $p_B^{NS} = \frac{1}{2}\bar{v}$.

From the FOCs for z_A^{NS} and z_B^{NS}, it then follows that z_A^{NS} and z_B^{NS} satisfy

$$\begin{aligned}\frac{\pi_S+\pi_W}{4F}\bar{u} &= \frac{1}{(1-z_A^{NS})^2},\\ \frac{\pi_W}{4F}\bar{v} &= \frac{1}{(1-z_B^{NS})^2}.\end{aligned}$$

Then, we have

$$z_A^{NS}=1-\sqrt{\frac{1}{\pi_S+\pi_W}\frac{4F}{\bar{u}}},\text{ and } z_B^{NS}=1-\sqrt{\frac{1}{\pi_W}\frac{4F}{\bar{v}}}.$$

Thus, the equilibrium for the two sellers is unique. Note that if $z_A^{NS}\leq 0$, then seller A advertises to zero consumers. Similarly, if $z_B^{NS}\leq 0$, then seller B advertises to zero consumers.

B.3 Proof of Proposition 4

With full data sharing, sellers can now separately advertise to strong-willed and weak-willed consumers. We first consider the optimal advertisement and pricing policies of seller A. It shall be clear that seller A would always avoid advertising to the third type of consumer, and that seller A does not need to differentiate strong-willed and weak-willed consumers. We denote z_A^{FS} as the measure of strong-willed and weak-willed consumers, to which seller A advertises, and p_A^{FS} as the price the seller sets.

Proposition 2 implies that strong-willed and weak-willed consumers use the same threshold $p_A^{FS}/\bar{u}$ in their random utility $\tilde{u}_A$ for purchasing good A. Thus, the sales of seller A is

$$Q_A^{FS}=z_A^{FS}\left[1-H\left(p_A^{FS}/\bar{u}\right)\right],$$

and the net profit of seller A is

$$\Pi_A^{FS}=p_A^{FS}z_A^{FS}\left[1-H\left(p_A^{FS}/\bar{u}\right)\right]-F\frac{z_A^{FS}}{1-z_A^{FS}}.$$

Following the same proof for Proposition 3, it is optimal for seller A to set a price $p_A^{FS}=\frac{1}{2}\bar{u}$. The first-order condition with respect to z_A^{FS} implies that

$$z_A^{FS}=1-2\sqrt{\frac{F}{\bar{u}}}.$$

Like before, if $1-2\sqrt{\frac{F}{\bar{u}}}\leq 0$, it is optimal for the seller to advertise to no consumers. That is, $z_A^{FS}=0$. Furthermore, if $1-2\sqrt{\frac{F}{\bar{u}}}>\pi_S+\pi_W$, then $z_A^{FS}=\pi_S+\pi_W$.

We now consider the policies of seller B. Seller B will advertise only to weak-willed consumers. Since seller B can discriminate by temptation types, it will exercise first-degree price discrimination by charging a weak-willed consumer his full reservation value: $p_B^{FS}(\gamma_i) = \gamma_i \bar{v}$. It can also make its advertising strategy $\hat{z}_B^{FS}$ dependent on γ_i. Since consumers with stronger temptation are willing to pay higher prices, seller B optimally prioritizes strong temptation consumers:

$$d\hat{z}_B^{FS}(\gamma_i) = \begin{cases} 0, & \text{if } \gamma_i < \hat{\gamma}^{FS} \\ \pi_W d\gamma_i, & \text{if } \gamma_i \in \left(\hat{\gamma}^{FS}, 1\right] \end{cases}.$$

Thus, seller B's profit is

$$\Pi_B^{FS} = \bar{v} \int_0^1 \gamma_i \hat{z}_B^{FS}(d\gamma_i) - F \frac{z_B^{FS}}{1 - z_B^{FS}} \quad \text{with} \quad z_B^{FS} = \int_0^1 \hat{z}_B^{FS}(d\gamma_i) \in [0, \pi_W],$$

where $\int_0^1 \gamma_i z_B^{FS}(d\gamma_i)$ is understood as a Riemann-Stieljes integral.

Note that the expected revenue of seller B reduces to $\bar{v} \int_{\hat{\gamma}^{FS}}^1 \pi_W \gamma_i d\gamma_i = \bar{v} \pi_W \frac{1 - \left(\hat{\gamma}^{FS}\right)^2}{2}$, where $\hat{\gamma}^{FS} = 1 - \frac{z_B^{FS}}{\pi_W}$, since $z_B^{FS} \in [0, \pi_W]$. Consequently, the expected revenue of seller B is $\bar{v} z_B^{FS} \left(1 - \frac{1}{2} \frac{z_B^{FS}}{\pi_W}\right)$, which is determined by the seller's total advertising z_B^{FS}. Consequently, we can rewrite seller B's maximization problem as choosing z_B^{FS}:

$$\Pi_B^{FS} = \bar{v} z_B^{FS} \left(1 - \frac{1}{2} \frac{z_B^{FS}}{\pi_W}\right) - F \frac{z_B^{FS}}{1 - z_B^{FS}} \quad \text{with} \quad z_B^{FS} \in [0, \pi_W].$$

The first-order condition for z_B^{FS} is

$$\left(1 - \frac{z_B^{FS}}{\pi_W}\right) \bar{v} = \frac{F}{\left(1 - z_B^{FS}\right)^2},$$

which is a cubic equation with one real, positive root. It then follows that

$$\begin{aligned} z_B^{FS} &= \frac{2 + \pi_W}{3} - \sqrt[3]{\left(\frac{1 - \pi_W}{3}\right)^3 + \frac{\pi_W F}{2\bar{v}} + \sqrt{\left(\left(\frac{1 - \pi_W}{3}\right)^3 + \frac{\pi_W F}{2\bar{v}}\right)^2 - \left(\frac{1 - \pi_W}{3}\right)^6}} \\ &\quad - \sqrt[3]{\left(\frac{1 - \pi_W}{3}\right)^3 + \frac{\pi_W F}{2\bar{v}} - \sqrt{\left(\left(\frac{1 - \pi_W}{3}\right)^3 + \frac{\pi_W F}{2\bar{v}}\right)^2 - \left(\frac{1 - \pi_W}{3}\right)^6}}. \end{aligned}$$

Again, if this solution to the first-order condition moves outside the feasible range $[0, \pi_W]$, it is optimal for the seller to advertise at the corner value. Consequently, the equilibrium is again unique.

Finally, rewriting the cubic equation for z_B^{FS} as

$$\left(\pi_W - z_B^{FS}\right)\left(1 - z_B^{FS}\right)^2 = \frac{\pi_W F}{\bar{v}}, \tag{14}$$

it follows that

$$\left(\pi_W - z_B^{FS}\right)^3 \leq \frac{\pi_W F}{\bar{v}},$$

and consequently

$$\frac{z_B^{FS}}{\pi_W} \geq 1 - \sqrt[3]{\frac{F}{\pi_W^2 \bar{v}}}.$$

Finally, applying the Implicit Function Theorem to Equation (14), one also has that

$$\begin{aligned}
\frac{dz_B^{FS}}{d\bar{v}} &= \frac{\frac{\pi_W F}{\bar{v}^2}}{\left(1 - z_B^{FS}\right)^2 + 2\left(\pi_W - z_B^{FS}\right)\left(1 - z_B^{FS}\right)} \geq 0, \\
\frac{dz_B^{FS}}{dF} &= -\frac{\frac{\pi_W}{\bar{v}}}{\left(1 - z_B^{FS}\right)^2 + 2\left(\pi_W - z_B^{FS}\right)\left(1 - z_B^{FS}\right)} \leq 0.
\end{aligned}$$

B.4 Proof of Proposition 5

It is easy to verify that $z_A^{FS} \geq z_A^{NS}$. Without data sharing, the probability of a strong-willed or weak-willed consumer being covered by seller A is z_A^{NS}; with data sharing, the probability is $\frac{z_A^{FS}}{\pi_S + \pi_W}$. As $z_A^{FS} \geq z_A^{NS}$ and $\pi_S + \pi_W \leq 1$, it follows that $\frac{z_A^{FS}}{\pi_S + \pi_W} \geq z_A^{NS}$, and the inequality is strict if $z_A^{FS} > 0$. Taken together, the conditional probability of a strong-willed or weak-willed consumer being covered by seller A is higher with full data sharing.

Across these two schemes with and without data sharing, seller A charges the same price $p_A^{NS} = p_A^{FS} = \bar{u}/2$ for its good. From Equation (5), the social welfare under no data sharing is given by

$$\begin{aligned}
W^{NS} &= \left(\pi_S + \pi_W\right) z_A^{NS} \int_{p_A}^{\bar{u}} u_A \frac{du_A}{\bar{u}} + \pi_W z_B^{NS} u_B \int_{p_B^{NS}/\bar{v}}^{1} d\gamma_i + \pi_W z_B^{NS} \int_0^{p_B^{NS}/\bar{v}} \left(u_B - \gamma_i \bar{v}\right) d\gamma_i \\
&= \frac{3}{8}\bar{u}\left(\pi_S + \pi_W\right) z_A^{NS} + \pi_W z_B^{NS}\left(u_B - \frac{1}{8}\bar{v}\right).
\end{aligned}$$

With full data sharing, seller B can perfectly price discriminate against each targeted weak-willed consumers. As a result, each targeted weak-willed consumer will purchase good B, and the social welfare is

$$W^{FS} = \frac{3}{8}\bar{u} z_A^{FS} + \pi_W u_B \int_{\hat{\gamma}^{FS}}^{1} d\gamma_i = \frac{3}{8}\bar{u} z_A^{FS} + z_B^{FS} u_B.$$

It then follows that

$$W^{FS} - W^{NS} = \pi_W\left(\frac{z_B^{FS}}{\pi_W} - z_B^{NS}\right) u_B + \frac{1}{8}\pi_W z_B^{NS}\bar{v} + \frac{3}{8}\bar{u}\left(\pi_S + \pi_W\right)\left(\frac{z_A^{FS}}{\pi_S + \pi_W} - z_A^{NS}\right) < 0,$$

if $u_B < u_{B**}$ where

$$u_{B**} = -\frac{3\bar{u}\frac{\pi_S + \pi_W}{\pi_W}\left(\frac{z_A^{FS}}{\pi_S + \pi_W} - z_A^{NS}\right) + \bar{v} z_B^{NS}}{8\left(\frac{z_B^{FS}}{\pi_W} - z_B^{NS}\right)}.$$

That is, social welfare is lower with full data sharing than with no data sharing.

B.5 Proof of Proposition 6

Sellers: We first characterize the optimal strategies of both sellers taking the opt-in cutoff of weak-will consumers γ_{**}^{GDPR} as given. We start with the optimal strategy of seller A. Suppose that seller A advertises to $z_{A,in}^{GDPR}$ measure of strong-willed and weak-willed consumers in the opt-in pool at price $p_{A,in}^{GDPR}$ and $z_{A,out}^{GDPR}$ measure of consumers in the opt-out pool at price $p_{A,out}^{GDPR}$. Then, the seller's expected profit, by the law of large numbers, is given by

$$\begin{aligned}\Pi_A &= \frac{\left(1-\gamma_{**}^{GDPR}\right)\pi_W}{\left(1-\gamma_{**}^{GDPR}\right)\pi_W + 1 - \pi_S - \pi_W} p_{A,out}^{GDPR} z_{A,out}^{GDPR} \left(1 - \frac{p_{A,out}^{GDPR}}{\bar{u}}\right) \\ &\quad + p_{A,in}^{GDPR} z_{A,in}^{GDPR} \left(1 - \frac{p_{A,in}^{GDPR}}{\bar{u}}\right) - F \frac{z_{A,out}^{GDPR} + z_{A,in}^{GDPR}}{1 - z_{A,out}^{GDPR} - z_{A,in}^{GDPR}},\end{aligned}$$

where $z_{A,out}^{GDPR} \in [0, 1-\pi_S-\gamma_{**}\pi_W]$ and $z_{A,in}^{GDPR} \in [0, \pi_S+\gamma_{**}\pi_W]$. Note that an advertisement to the opt-in pool reaches a strong or weak-willed consumer with perfect precision, while one to the opt-out pool reaches a weak-willed consumer (who will buy the good) at a probability of $\frac{\left(1-\gamma_{**}^{GDPR}\right)\pi_W}{\left(1-\gamma_{**}^{GDPR}\right)\pi_W+1-\pi_S-\pi_W}$.

If $z_{A,in}^{GDPR} > 0$ and $z_{A,out}^{GDPR} > 0$, the FOCs for $p_{A,in}^{GDPR}$ and $p_{A,out}^{GDPR}$ reveal that

$$p_{A,in}^{GDPR} = p_{A,out}^{GDPR} = \frac{1}{2}\bar{u}.$$

Then, the seller's profit becomes

$$\Pi_A = \frac{\left(1-\gamma_{**}^{GDPR}\right)\pi_W}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\frac{\bar{u}}{4} z_{A,out}^{GDPR} + \frac{\bar{u}}{4} z_{A,in}^{GDPR} - F\frac{z_{A,out}^{GDPR}+z_{A,in}^{GDPR}}{1-z_{A,out}^{GDPR}-z_{A,in}^{GDPR}}.$$

The marginal profit from $z_{A,in}^{GDPR}$ is strictly higher than that from $z_{A,out}^{GDPR}$, as the advertising efficiency to the opt-in pool is higher. Thus, seller A gives higher priority to the opt-in pool.

The first-order condition with respect to $z_{A,in}^{GDPR}$ gives

$$\frac{\bar{u}}{4} - F\frac{1}{\left(1-z_{A,out}^{GDPR}-z_{A,in}^{GDPR}\right)^2}\begin{cases} < 0 & \text{if } z_{A,in}^{GDPR} = 0 \\ = 0 & \text{if } z_{A,in}^{GDPR} \in \left(0, \pi_S + \gamma_{**}^{GDPR}\pi_W\right) \\ > 0 & \text{if } z_{A,in}^{GDPR} = \pi_S + \gamma_{**}^{GDPR}\pi_W \end{cases}$$

The parameter restriction $F < \frac{\bar{u}}{4}$ ensures that $z_{A,in}^{GDPR} > 0$. As $z_{A,in}^{GDPR}$ has higher priority than $z_{A,out}^{GDPR}$, we have

$$z_{A,in}^{GDPR} = \min\left\{1 - 2\sqrt{\frac{F}{\bar{u}}}, \pi_S + \gamma_{**}^{GDPR}\pi_W\right\}. \tag{15}$$

If $z_{A,in}^{GDPR} = \pi_S + \gamma_{**}^{GDPR}\pi_W$, the seller may have capacity to cover the opt-out pool. The first-order condition for $z_{A,out}^{GDPR}$ in this scenario gives

$$z_{A,out}^{GDPR} = \min\{\max\{1 - 2\sqrt{\frac{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}{(1-\gamma_{**}^{GDPR})\,\pi_W}\frac{F}{\bar{u}}} - \pi_S - \gamma_{**}^{GDPR}\pi_W, 0\}, 1-\pi_S-\gamma_{**}^{GDPR}\pi_W\}.$$

Since seller A gives a higher priority in advertising to the opt-in pool, we can directly prove that each strong-willed consumer would prefer opt-in to opt-out. For simplicity, we skip the proof here.

We now analyze the optimal advertising strategy of seller B. Suppose that seller B advertises with intensity $\hat{z}_{B,in}^{GDPR}(\gamma_i)$ to weak-willed consumers in the opt-in pool at price $p_{B,in}^{GDPR}(\gamma_i) = \gamma_i\bar{v}$ and $z_{B,out}^{GDPR}$ measure of consumers in the opt-out pool at price $p_{B,out}^{GDPR}$. Note that an advertisement to the opt-out pool reaches, with probability of $\frac{(1-\gamma_{**}^{GDPR})\pi_W}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}$, a weak-willed consumer, who will buy the good only if his temptation coefficient γ_i is above $p_{B,out}^{GDPR}/\bar{v}$. Thus, the seller's profit is

$$\begin{aligned}
\Pi_B \;=\; & -F\frac{z_{B,out}^{GDPR}+z_{B,in}^{GDPR}}{1-z_{B,out}^{GDPR}-z_{B,in}^{GDPR}} + \bar{v}\int_0^{\gamma_{**}^{GDPR}} \gamma_i\hat{z}_{B,in}^{GDPR}(d\gamma_i) \\
& + \frac{\pi_W}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W} z_{B,out}^{GDPR} p_{B,out}^{GDPR} \\
& \cdot \left[\left(1-p_{B,out}^{GDPR}/\bar{v}\right) 1_{\{p_{B,out}^{GDPR}\geq\gamma_{**}^{GDPR}\bar{v}\}} + \left(1-\gamma_{**}^{GDPR}\right) 1_{\{p_{B,out}^{GDPR}<\gamma_{**}^{GDPR}\bar{v}\}}\right],
\end{aligned}$$

where $z_{B,out}^{GDPR} \in \left[0, 1-\pi_S-\gamma_{**}^{GDPR}\pi_W\right]$ and $z_{B,in}^{GDPR} = \int_0^{\gamma_{**}^{GDPR}} z_{B,in}^{GDPR}(d\gamma_i) \in \left[0, \gamma_{**}^{GDPR}\pi_W\right]$ is the total advertisement to the opt-in pool.

If $z_{B,out}^{GDPR} > 0$, then the first-order condition for $p_{B,out}^{GDPR}$ gives the following:

$$\begin{aligned}
\text{If } \gamma_{**}^{GDPR} &\leq \frac{1}{2},\ \left(1-2p_{B,out}^{GDPR}/\bar{v}\right) 1_{\{p_{B,out}^{GDPR}\geq\gamma_{**}^{GDPR}\bar{v}\}} = 0, \\
\text{If } \gamma_{**}^{GDPR} &> \frac{1}{2},\ p_{B,out}^{GDPR} = \gamma_{**}^{GDPR}\bar{v} \text{ if } z_{B,out}^{GDPR} \geq 0.
\end{aligned}$$

Thus, the optimal price satisfies

$$p_{B,out}^{GDPR} = \begin{cases} \frac{1}{2}\bar{v} & \text{if } \gamma_{**}^{GDPR} \leq \frac{1}{2} \\ \gamma_{**}^{GDPR}\bar{v} & \text{If } \gamma_{**}^{GDPR} > \frac{1}{2} \end{cases} = \max\left\{\frac{1}{2}, \gamma_{**}^{GDPR}\right\}\bar{v}.$$

Since consumers with stronger temptation are willing to pay higher prices with $p_{B,in}^{GDPR}(\gamma_i) = \gamma_i\bar{v}$, it is optimal for seller B to prioritize consumers with higher γ_i:

$$dz_{B,in}^{GDPR}(\gamma_i) = \begin{cases} 0 & \text{if } \gamma_i < \hat{\gamma}^{GDPR} \\ \pi_W d\gamma_i & \text{if } \gamma_i \in \left(\hat{\gamma}^{GDPR}, \gamma_{**}^{GDPR}\right] \end{cases}.$$

Therefore, the expected revenue of seller B from the opt-in pool reduces to $\bar{v}\int_{\hat{\gamma}^{GDPR}}^{\gamma_{**}^{GDPR}} \pi_W\gamma_i d\gamma_i = \bar{v}\pi_W\frac{\left(\gamma_{**}^{GDPR}\right)^2-\left(\hat{\gamma}^{GDPR}\right)^2}{2}$. As $\hat{\gamma}^{GDPR} = \gamma_{**}^{GDPR} - \frac{z_{B,in}^{GDPR}}{\pi_W}$, the expected revenue of seller B from advertising to the opt-in pool is determined by the seller's total advertising to the opt-in pool $z_{B,in}^{GDPR}$: $\bar{v} z_{B,in}^{GDPR} \left(\gamma_{**}^{GDPR} - \frac{1}{2} \frac{z_{B,in}^{GDPR}}{\pi_W} \right)$. Thus, the expected profit of seller B reduces to

$$\begin{aligned} \Pi_B \quad = \quad & -F \frac{z_{B,out}^{GDPR} + z_{B,in}^{GDPR}}{1 - z_{B,out}^{GDPR} - z_{B,in}^{GDPR}} + z_{B,in}^{GDPR} \left(\gamma_{**}^{GDPR} - \frac{1}{2} \frac{z_{B,in}^{GDPR}}{\pi_W} \right) \bar{v} \\ & + \frac{\pi_W}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} \left[\frac{1}{4} - \left(\gamma_{**}^{GDPR} - \frac{1}{2} \right)^2 1_{\{\gamma_{**}^{GDPR} > \frac{1}{2}\}} \right] \bar{v} z_{B,out}^{GDPR}, \end{aligned} \tag{16}$$

and the seller's choice reduces to choosing $z_{B,in}^{GDPR}$ and $z_{B,out}^{GDPR}$.

The revenue from the opt-in pool is concave in the total advertising to the opt-in pool, $z_{B,in}^{GDPR}$, since seller B targets the highest marginal revenue consumers first, while the revenue from the opt-out pool is linear with respect to the advertising to the opt-out pool $z_{B,out}^{GDPR}$. The first-order condition for $z_{B,in}^{GDPR}$ is

$$\bar{v} \left(\gamma_{**}^{GDPR} - \frac{z_{B,in}^{GDPR}}{\pi_W} \right) - F \frac{1}{\left(1 - z_{B,out}^{GDPR} - z_{B,in}^{GDPR}\right)^2} \begin{cases} < 0 & \text{if } z_{B,in}^{GDPR} = 0 \\ = 0 & \text{if } z_{B,in}^{GDPR} \in \left(0, \pi_W \gamma_{**}^{GDPR}\right) \\ > 0 & \text{if } z_{B,in}^{GDPR} = \pi_W \gamma_{**}^{GDPR} \end{cases}, \tag{17}$$

and the first-order condition for $z_{B,out}^{GDPR}$ is

$$\begin{aligned} & \frac{\pi_W}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} \left[\frac{1}{4} - \left(\gamma_{**}^{GDPR} - \frac{1}{2} \right)^2 1_{\{\gamma_{**}^{GDPR} > \frac{1}{2}\}} \right] \bar{v} - \frac{F}{\left(1 - z_{B,out}^{GDPR} - z_{B,in}^{GDPR}\right)^2} \\ & \qquad \begin{cases} < 0 & \text{if } z_{B,out}^{GDPR} = 0 \\ = 0 & \text{if } z_{B,out}^{GDPR} \in \left(0, 1 - \pi_S - \pi_W \gamma_{**}^{GDPR}\right) \\ > 0 & \text{if } z_{B,out}^{GDPR} = 1 - \pi_S - \pi_W \gamma_{**}^{GDPR} \end{cases}. \end{aligned}$$

Which pool has priority depends on which has higher marginal revenue. The marginal revenue from the opt-in pool when $z_{B,in}^{GDPR} = 0$ is $\bar{v} \gamma_{**}^{GDPR}$, while the marginal revenue from the opt-out pool is $\frac{\pi_W}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} \left[\frac{1}{4} - \left(\gamma_{**}^{GDPR} - \frac{1}{2} \right)^2 1_{\{\gamma_{**}^{GDPR} > \frac{1}{2}\}} \right] \bar{v}$. When $\gamma_{**}^{GDPR} < \frac{1}{2}$, then the opt-in pool has priority whenever

$$\gamma_{**}^{GDPR} > \frac{1}{4} \frac{\pi_W}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W},$$

which is equivalent to

$$\gamma_{**}^{GDPR} \in \left[\frac{1 - \pi_S}{2\pi_W} - \sqrt{\left(\frac{1 - \pi_S}{2\pi_W} \right)^2 - \frac{1}{4}}, \frac{1 - \pi_S}{2\pi_W} + \sqrt{\left(\frac{1 - \pi_S}{2\pi_W} \right)^2 - \frac{1}{4}} \right].$$

This range exists since $1-\pi_S > \pi_W$ (i.e., there are O-type consumers). Thus, the upper end of this range is above $\frac{1}{2}$. Then, the opt-in pool has higher priority if

$$\gamma_{**}^{GDPR} \in \left[\frac{1-\pi_S}{2\pi_W} - \sqrt{\left(\frac{1-\pi_S}{2\pi_W}\right)^2 - \frac{1}{4}}, \frac{1}{2}\right],$$

which is nonempty. When $\gamma_{**}^{GDPR} > \frac{1}{2}$, it is direct to verify that the opt-in pool has priority. Taken together, the opt-in pool has priority if and only if

$$\gamma_{**}^{GDPR} \geq \frac{1-\pi_S}{2\pi_W} - \sqrt{\left(\frac{1-\pi_S}{2\pi_W}\right)^2 - \frac{1}{4}}. \tag{18}$$

If the opt-out pool has higher priority, seller B will devote all resources to the opt-out pool before the opt-in pool:

$$\begin{aligned} z_{B,out}^{GDPR} &= \min\left\{\max\left\{1 - \sqrt{\frac{F}{\pi_W \bar{v}} \frac{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}{\frac{1}{4} - \left(\gamma_{**}^{GDPR} - \frac{1}{2}\right)^2 1_{\{\gamma_{**}^{GDPR} > \frac{1}{2}\}}}},\ 0\right\},\ 1-\pi_S-\gamma_{**}^{GDPR}\pi_W\right\}, \\ z_{B,in}^{GDPR} &= \min\left\{\max\left\{1 - \sqrt{\frac{F}{\pi_W \bar{v}} \frac{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}{\frac{1}{4} - \left(\gamma_{**}^{GDPR} - \frac{1}{2}\right)^2 1_{\{\gamma_{**}^{GDPR} > \frac{1}{2}\}}}} - z_{B,out}, 0\right\},\ \pi_W \gamma_{**}^{GDPR}\right\}. \end{aligned}$$

If, instead, the opt-in pool has priority, seller B will devote resources to the opt-in pool until its first-order condition for $z_{B,in}^{GDPR}$ is satisfied or the marginal products of the two pools are equal, whichever occurs first. Since the marginal revenue of the opt-in pool decreases from $\bar{v}\gamma_{**}^{GDPR}$ to 0, the two marginal revenues will intersect at a unique level $z_{B,in}^{GDPR} = z_*$, where

$$z_* = \pi_W \gamma_{**}^{GDPR} - \pi_W^2 \frac{\frac{1}{4} - \left(\gamma_{**}^{GDPR} - \frac{1}{2}\right)^2 1_{\{\gamma_{**}^{GDPR} > \frac{1}{2}\}}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}.$$

The first-order condition for $z_{B,in}^{GDPR}$ in Equation (17) when $z_{B,out}^{GDPR} = 0$ gives a cubic equation to determine a unique, positive level for z_{**}:

$$z_{**} = 1 - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} + \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{(1-\pi_W\gamma_{**}^{GDPR})^3}{27}}} - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} - \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{(1-\pi_W\gamma_{**}^{GDPR})^3}{27}}}.$$

Consequently, it follows that

$$z_{B,in}^{GDPR} = \min\left\{z_*, z_{**}\right\}. \tag{19}$$

If $z_* < z_{**}$, seller B would also target the opt-out pool with

$$z_{B,out}^{GDPR} = \min\left\{\max\left\{1 - \sqrt{\frac{F}{\pi_W \bar{v}} \frac{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}{\frac{1}{4} - \left(\gamma_{**}^{GDPR} - \frac{1}{2}\right)^2 1_{\{\gamma_{**}^{GDPR} > \frac{1}{2}\}}}} - z_*, 0\right\}, 1-\pi_S-\gamma_{**}^{GDPR}\pi_W\right\}. \tag{20}$$

If $z_* \geq z_{**}$, seller B does not target the opt-out pool.

Taken together, the level of γ_{**}^{GDPR} determines whether the opt-in or opt-out pool has higher priority to seller B. In either case, its optimal advertising policy exists and is unique given γ_{**}^{GDPR}.

Weak-willed customers: We first verify that, if other weak-willed customers follow the conjectured cutoff strategy with cutoff γ_{**}^{GDPR}, that it is optimal for a weak-willed consumer with temptation γ_i to follow the same cutoff strategy. We then characterize the equilibrium γ_{**}^{GDPR}.

Consider a weak-willed consumer with temptation index γ_i. Following Equation (6), his expected utility from opt-in is

$$U_{W,in}^{GDPR}(\gamma_i) = \frac{z_{A,in}^{GDPR}}{\pi_S + \gamma_{**}^{GDPR}\pi_W}\frac{\bar{u}}{8} + \frac{\hat{z}_{B,in}^{GDPR}(\gamma_i)}{\pi_W}(u_B - \gamma_i \bar{v}).$$

This expression shows that $U_{W,in}^{GDPR}$ increases with $z_{A,in}^{GDPR}$ but decreases with $\hat{z}_{B,in}^{GDPR}(\gamma_i)$. Following Equation (7), his expected utility from opt-out is

$$\begin{aligned} U_{W,out}^{GDPR}(\gamma_i) \;=\; & \frac{z_{A,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\frac{\bar{u}}{8} + \frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}u_B - \frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W} \\ & \cdot\bar{v}\left[\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\mathbf{1}_{\left\{\gamma_i>\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\right\}} + \gamma_i\mathbf{1}_{\left\{\gamma_i\leq\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\right\}}\right], \end{aligned}$$

which increases with $z_{A,out}^{GDPR}$ and decreases with $z_{B,out}^{GDPR}$. Then,

$$\begin{aligned} & U_{W,in}^{GDPR}(\gamma_i) - U_{W,out}^{GDPR}(\gamma_i) \qquad (21)\\ =\; & \frac{\bar{u}}{8}\left[\frac{z_{A,in}^{GDPR}}{\pi_S+\gamma_{**}^{GDPR}\pi_W} - \frac{z_{A,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\right] + \left(\frac{\hat{z}_{B,in}^{GDPR}(\gamma_i)}{\pi_W} - \frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\right)u_B \\ & +\bar{v}\left[\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\left(\gamma_i\mathbf{1}_{\left\{\gamma_i\leq\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\right\}} + \max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\mathbf{1}_{\left\{\gamma_i>\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\right\}}\right) - \frac{\hat{z}_{B,in}^{GDPR}(\gamma_i)}{\pi_W}\gamma_i\right] \end{aligned}$$

Note that $\frac{z_{A,in}^{GDPR}}{\pi_S+\gamma_{**}^{GDPR}\pi_W} \geq \frac{z_{A,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}$ from our earlier analysis of seller A's strategy. Therefore, whether $U_{W,in}^{GDPR}(\gamma_i) - U_{W,out}^{GDPR}(\gamma_i)$ crosses zero depends on the second and third terms.

Note that if $z_{B,in}^{GDPR} = 0$, then $U_{W,in}(\gamma_i) > U_{W,out}(\gamma_i)$ for all γ_i, and $\gamma_{**}^{GDPR} = 1$. Consequently, it must be the case that the opt-in pool has priority, or γ_{**}^{GDPR} satisfies Equation (18). It then follows that, unless the equilibrium is trivial for seller B (i.e., advertising costs are forbiddingly high and the seller does not advertise at all), then $z_{B,in}^{GDPR} > 0$.

Since the marginal consumer must be indifferent to opt-in and opt-out, $U_{W,in}^{GDPR}\left(\gamma_{**}^{GDPR}\right) - U_{W,out}^{GDPR}\left(\gamma_{**}^{GDPR}\right) = 0$, which imposes the following condition on γ_{**}^{GDPR}:

$$C\left(\gamma_{**}^{GDPR}\right)\begin{cases} <0 & \text{if } \gamma_{**}^{GDPR} = 0 \\ =0 & \text{if } \gamma_{**}^{GDPR} \in (0,1) \\ >0 & \text{if } \gamma_{**}^{GDPR} = 1 \end{cases}. \qquad (22)$$

where, with some manipulation of Equation (21),

$$
\begin{aligned}
C\left(\gamma_{**}^{GDPR}\right) \quad = \quad & \left(\pi_S+\gamma_{**}^{GDPR}\pi_W\right)^3 \frac{\hat{z}_{B,in}^{GDPR}\left(\gamma_{**}^{GDPR}\right)}{\pi_W} \\
& -\left(\left(1+\pi_S+\frac{\pi_W u_B}{\bar{v}}\right)\frac{\hat{z}_{B,in}^{GDPR}\left(\gamma_{**}^{GDPR}\right)}{\pi_W}-z_{B,out}^{GDPR}\right)\left(\pi_S+\gamma_{**}^{GDPR}\pi_W\right)^2 \\
& +\left(\left(\pi_S+\frac{\pi_W u_B}{\bar{v}}\right)\left(\frac{\hat{z}_{B,in}^{GDPR}\left(\gamma_{**}^{GDPR}\right)}{\pi_W}-z_{B,out}^{GDPR}\right)-\frac{\pi_W\bar{u}}{8\bar{v}}\left(z_{A,in}^{GDPR}+z_{A,out}^{GDPR}\right)\right) \\
& \cdot\left(\pi_S+\gamma_{**}^{GDPR}\pi_W\right)+\frac{\pi_W\bar{u}}{8\bar{v}}z_{A,in}^{GDPR}.
\end{aligned} \tag{23}
$$

We now verify the optimality of the cutoff strategy for weak-willed customers to opt in. We substitute the indifference condition in Equation (22), assuming an interior $\gamma_{**}^{GDPR}\in(0,1)$, into Equation (21) to obtain

$$
\begin{aligned}
& U_{W,in}^{GDPR}(\gamma_i)-U_{W,out}^{GDPR}(\gamma_i) \\
= \quad & \left(\frac{\hat{z}_{B,in}^{GDPR}(\gamma_i)}{\pi_W}-1\right)u_B+\bar{v}\left(\gamma_{**}^{GDPR}-\frac{\hat{z}_{B,in}^{GDPR}(\gamma_i)}{\pi_W}\gamma_i\right) \\
& +\bar{v}\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\left(\gamma_i\mathbf{1}_{\left\{\gamma_i\le\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\right\}}+\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\mathbf{1}_{\left\{\gamma_i>\max\left\{\frac{1}{2},\gamma_{**}^{GDPR}\right\}\right\}}-\gamma_{**}^{GDPR}\right).
\end{aligned}
$$

We begin with $\gamma_i>\gamma_{**}^{GDPR}$. Note that $\hat{z}_{B,in}^{GDPR}(\gamma_i)=\pi_W$, since this more tempted weak-willed consumer will be targeted by seller B if he opts in. If $\gamma_{**}^{GDPR}\ge\frac{1}{2}$, then

$$
U_{W,in}^{GDPR}(\gamma_i)-U_{W,out}^{GDPR}(\gamma_i)=\bar{v}\left(\gamma_{**}^{GDPR}-\gamma_i\right)<0.
$$

If instead $\gamma_{**}^{GDPR}<\frac{1}{2}$ and $\gamma_i\le\frac{1}{2}$, then

$$
U_{W,in}^{GDPR}(\gamma_i)-U_{W,out}^{GDPR}(\gamma_i)=\left(1-\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\right)\bar{v}\left(\gamma_{**}^{GDPR}-\gamma_i\right)\le 0,
$$

since $\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\le 1$. Finally, if $\gamma_{**}^{GDPR}<\frac{1}{2}$ and $\gamma_i>\frac{1}{2}$, then

$$
\begin{aligned}
& U_{W,in}^{GDPR}(\gamma_i)-U_{W,out}^{GDPR}(\gamma_i) \\
= \quad & \left(1-\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\right)\bar{v}\left(\gamma_{**}^{GDPR}-\gamma_i\right)+\bar{v}\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\left(\frac{1}{2}-\gamma_i\right)<0,
\end{aligned}
$$

since $\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W}\le 1$ and $\gamma_i>\frac{1}{2}$. Therefore, all weak-willed consumers with $\gamma_i>\gamma_{**}^{GDPR}$ opt out, regardless of the level of γ_{**}^{GDPR}.

We now consider $\gamma_i<\gamma_{**}^{GDPR}$. Note that for $\gamma_i<\hat{\gamma}^{GDPR}$, the threshold γ_i below which seller B does not advertise ($\hat{z}_{B,in}^{GDPR}(\gamma_i)=0$ for $\gamma_i<\hat{\gamma}^{GDPR}$), it is trivial that $U_{W,in}^{GDPR}(\gamma_i)-$

$U_{W,out}^{GDPR}(\gamma_i) > 0$, since the consumer benefits from both higher advertising by seller A and lower (zero) advertising by seller B. Consequently, all weak-willed consumers with $\gamma_i < \hat{\gamma}^{GDPR}$ opt in. For $\gamma_i \in \left[\hat{\gamma}^{GDPR}, \gamma_{**}^{GDPR}\right]$, $z_{B,in}^{GDPR}(\gamma_i) = \pi_W$, and their opt-in / opt-out condition reduces to

$$U_{W,in}^{GDPR}(\gamma_i) - U_{W,out}^{GDPR}(\gamma_i) = \left(1 - \frac{z_{B,out}^{GDPR}}{1 - \pi_S - \gamma_{**}^{GDPR}\pi_W}\right)\bar{v}\left(\gamma_{**}^{GDPR} - \gamma_i\right) > 0,$$

since $\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_{**}^{GDPR}\pi_W} \leq 1$. Therefore, all $\gamma_i \in \left[\hat{\gamma}^{GDPR}, \gamma_{**}^{GDPR}\right]$ opt in. Consequently, all $\gamma_i < \gamma_{**}^{GDPR}$ opt in, which verifies the optimality of the cutoff strategy. Importantly, the optimality of the cutoff strategy holds regardless of the equilibrium value of γ_{**}^{GDPR}.

Existence of γ_{**}^{GDPR}: An interior solution for Equation (22) is such that $C\left(\gamma_{**}^{GDPR}\right) = 0$. If $C(0) < 0$, then $\gamma_{**}^{GDPR} = 0$, while if $C(1) > 0$, then $\gamma_{**}^{GDPR} = 1$. We next recognize that

$$\begin{aligned}\frac{C(1)}{\pi_S + \pi_W} &= -\pi_W\left(1 - \frac{u_B}{\bar{v}}\right)\left(\frac{\hat{z}_{B,in}^{GDPR}(1)}{\pi_W}(1 - \pi_S - \pi_W) - z_{B,out}^{GDPR}\right) \\ &\quad + \frac{1 - \pi_S - \pi_W}{\pi_S + \pi_W}\frac{\pi_W \bar{u}}{8\bar{v}} z_{A,in}^{GDPR} - \frac{\pi_W \bar{u}}{8\bar{v}} z_{A,out}^{GDPR}.\end{aligned}$$

Suppose $\frac{\bar{u}}{8\bar{v}} \leq 1 - u_B$. Note that when $\gamma_{**}^{GDPR} = 1$, then $\frac{\hat{z}_{B,in}^{GDPR}(1)}{\pi_W} = 1$ (the most tempted customer that opts-in is targeted) since the opt-in pool has priority, and

$$C(1) = \left(\frac{\bar{u}}{8\bar{v}}\frac{z_{A,in}^{GDPR}}{\pi_S + \pi_W} - 1 + \frac{u_B}{\bar{v}}\right)(1 - \pi_S - \pi_W)\pi_W - \frac{\pi_W \bar{u}}{8\bar{v}} z_{A,out}^{GDPR} < 0,$$

since $z_{B,out}^{GDPR} = 0$ when $\gamma_{**}^{GDPR} = 1$, $u_B < 0$, and $\frac{z_{A,in}^{GDPR}}{\pi_S + \pi_W} \leq 1$. Consequently, it follows that $\gamma_{**}^{GDPR} < 1$.

At the other end, note that

$$C(0) = \frac{\pi_W u_B}{\bar{v}}\left(\frac{\hat{z}_{B,in}^{GDPR}(0)}{\pi_W}(1 - \pi_S) - z_{B,out}^{GDPR}\right)\pi_S + \frac{\pi_W \bar{u}}{8\bar{v}}\pi_S(1 - \pi_S)\left(\frac{z_{A,in}^{GDPR}}{\pi_S} - \frac{z_{A,out}^{GDPR}}{1 - \pi_S}\right) > 0,$$

since $\frac{z_{A,in}^{GDPR}}{\pi_S} - \frac{z_{A,out}^{GDPR}}{1-\pi_S} > 0$ because the opt-in pool has higher advertising efficiency for seller A, $u_B < 0$, and $\frac{\hat{z}_{B,in}^{GDPR}(0)}{\pi_W} = 0$ because only the most mildly tempted opt in (condition (18) fails). Consequently, $C(\pi_S) > 0$ and therefore $\gamma_{**}^{GDPR} > 0$.

Notice now that all advertising policies, $z_{A,in}^{GDPR}$, $z_{A,out}^{GDPR}$, $\hat{z}_{B,in}^{GDPR}$, and $z_{B,out}^{GDPR}$ are (piece-wise) continuous in γ_{**}^{GDPR} when γ_{**}^{GDPR} has an interior solution. As such, $C\left(\gamma_{**}^{GDPR}\right)$ is continuous. Since $C(0) > 0$ while $C(1) < 0$, by the Intermediate Value Theorem, there exists a $\gamma_{**}^{GDPR} \in (0, 1)$. Given the nonlinearity of $C\left(\gamma_{**}^{GDPR}\right)$, however, there may be multiple values in $(0, 1)$ with $C\left(\gamma_{**}^{GDPR}\right) = 0$, and consequently multiple equilibria.

We further establish that there exists an equilibrium in which $\gamma_{**}^{GDPR} \geq \gamma_+ = \frac{1-\pi_S}{\pi_W} - \sqrt{\left(\frac{1-\pi_S}{\pi_W}\right)^2 - \frac{1-\pi_S}{\pi_W}} > \frac{1}{2}$. Direct manipulation of the definition of $C(\cdot)$ gives that

$$\begin{aligned} C\left(\gamma_+\right) &= \left((1-\pi_S)\left(1-\sqrt{1-\frac{\pi_W}{1-\pi_S}}\right) - \frac{\pi_W u_B}{\bar{v}}\right)\tilde{\gamma}_+\left(1-\tilde{\gamma}_+\right)\left(1-\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_+\pi_W}\right) \\ &\quad + \frac{\pi_W \bar{u}}{8\bar{v}}\tilde{\gamma}_+\left(1-\tilde{\gamma}_+\right)\left(\frac{z_{A,in}^{GDPR}}{\pi_S+\gamma_+\pi_W} - \frac{z_{A,out}^{GDPR}}{1-\pi_S-\gamma_+\pi_W}\right) > 0, \end{aligned}$$

where $\tilde{\gamma}_+ = \pi_S + \gamma_+\pi_W$. Note that $u_B < 0$, and therefore $(1-\pi_S)\left(1-\sqrt{1-\frac{\pi_W}{1-\pi_S}}\right) - \frac{\pi_W u_B}{\bar{v}} > 0$, $\frac{z_{B,out}^{GDPR}}{1-\pi_S-\gamma_+\pi_W} \leq 1$ by the definition of advertising efficiency, and the last term is positive since advertising efficiency is higher in the opt-in pool than in the opt-out pool for seller A. Since $C\left(\gamma_+\right) > 0$ and $C(1) < 0$, it follows that there is an equilibrium for which $\gamma_{**}^{GDPR} \geq \gamma_+ > \frac{1}{2}$. Thus, the optimal advertising policy of seller B for the opt-in and opt-out pools is given by Equations (19) and (20).

Uniqueness: We finally provide a sufficient condition under which an equilibrium with $\gamma_{**}^{GDPR} \geq \gamma_+$ is the unique equilibrium. Since $\gamma_{**}^{GDPR} \geq \frac{1}{2}$, it is sufficient for

$$1 - 2\sqrt{\frac{F}{\bar{u}}} < \pi_S + \frac{1}{2}\pi_W,$$

or equivalently

$$\bar{u} < 4F\left(1-\pi_S-\frac{1}{2}\pi_W\right)^{-2},$$

to ensure $z_{A,out}^{GDPR} = 0$. Intuitively, it is too costly for seller A to advertise to more than $\pi_S + \frac{1}{2}\pi_W \leq \pi_S + \gamma_{**}^{GDPR}\pi_W$ consumers. Given this sufficient condition, notice that $z_{A,in}^{GDPR}$ from Equation (15) is also insensitive to γ_{**}^{GDPR}.

Next, we view γ_{**}^{GDPR} as a fixed point determined by Equation (22) through $z_{A,in}^{GDPR}$, $z_{A,out}^{GDPR}$, $z_{B,in}^{GDPR}$, $z_{B,out}^{GDPR}$, which are functions of γ_{**}^{GDPR}. With $z_{A,in}^{GDPR}$ being independent of γ_{**}^{GDPR} and $z_{A,out}^{GDPR} = 0$, $C\left(\gamma_{**}^{GDPR}\right)$ in Equation (23) is now only a function of $z_{B,out}^{GDPR}$. Conditional on $z_{B,out}^{GDPR}$, $C\left(\gamma_{**}^{GDPR}\right)$ is a cubic equation in γ_{**}^{GDPR}. By taking $z_{B,out}^{GDPR}$ as given, we apply the Implicit Function Theorem to Equation (22) to obtain

$$\frac{d\gamma_{**}^{GDPR}}{dz_{B,out}^{GDPR}} = -\frac{\pi_W\left(\gamma_{**}^{GDPR} - \frac{u_B}{\bar{v}}\right)\left(\pi_S + \gamma_{**}^{GDPR}\pi_W\right)}{C'\left(\gamma_{**}^{GDPR}\right)}.$$

Since $C\left(\gamma_{**}^{GDPR}\right) = 0$, and $C\left(\gamma_{**}^{GDPR}\right)$ is a cubic equation with one negative and two positive real roots, it follows that $C'\left(\gamma_{**}^{GDPR}\right) < 0$. Furthermore, since $u_B < 0$, it follows that

$$\frac{d\gamma_{**}^{GDPR}}{dz_{B,out}^{GDPR}} > 0.$$

Consequently, we have established that the solution for $\gamma_{**}^{GDPR} \in \left(\gamma_+, 1\right)$ is continuous and monotonically increasing in $z_{B,out}^{GDPR}$.

Note next that either $z_{B,out}^{GDPR} = 0$ or, from Equations (19) and (20),

$$z_{B,out}^{GDPR} = \min\left\{1 - \sqrt{\frac{F}{\pi_W \bar{v}} \frac{1 - \pi_S - \gamma_{**}^{GDPR}\pi_W}{\gamma_{**}^{GDPR}\left(1 - \gamma_{**}^{GDPR}\right)}} - \frac{\pi_W \gamma_{**}^{GDPR}\left(1 - \pi_S - \pi_W\right)}{1 - \pi_S - \gamma_{**}^{GDPR}\pi_W}, 1 - \pi_S - \gamma_{**}^{GDPR}\pi_W\right\}.$$

It then follows that $z_{B,out}^{GDPR}$ as a function of γ_{**}^{GDPR} that is decreasing in γ_{**}^{GDPR} for $\gamma_{**}^{GDPR} \geq \gamma_+$, since at an interior solution

$$z_{B,out}^{GDPR\prime}\left(\gamma_{**}^{GDPR}\right) = \frac{1}{2}\frac{F}{\bar{v}} \frac{\frac{1-\pi_S}{\pi_W}\left(1 - 2\gamma_{**}^{GDPR}\right) + \left(\gamma_{**}^{GDPR}\right)^2}{\left(1 - z_{B,out}^{GDPR}\right)\left(\gamma_{**}^{GDPR}\left(1 - \gamma_{**}^{GDPR}\right)\right)^2} - \frac{\pi_W\left(1 - \pi_S\right)\left(1 - \pi_S - \pi_W\right)}{\left(1 - \pi_S - \gamma_{**}^{GDPR}\pi_W\right)^2}.$$

The first term has a root between $(0,1)$ at γ_+. Consequently, since $\gamma_{**}^{GDPR} \geq \gamma_+$, and the other root of $\frac{dz_{B,out}^{GDPR}}{d\gamma_{**}^{GDPR}}$ is above 1, it follows that $z_{B,out}^{GDPR}$ is monotonically decreasing in γ_{**}^{GDPR}. If, instead, $z_{B,out}^{GDPR}$ is at a corner solution, $1 - \pi_S - \gamma_{**}^{GDPR}\pi_W$, it is again decreasing in γ_{**}^{GDPR}. Consequently, by continuity, $z_{B,out}^{GDPR}$ is decreasing and then flat at 0 in γ_{**}^{GDPR}.

Since the map from $z_{B,out}^{GDPR}$ to γ_{**}^{GDPR}, Equation (22), is monotonically increasing, while that from γ_{**}^{GDPR} to $z_{B,out}^{GDPR}$ from Equation (20) is (weakly) monotonically decreasing, it follows these two curves on the $\left(z_{B,out}^{GDPR}, \gamma_{**}^{GDPR}\right)$ plane intersect at most at one point, and therefore the equilibrium is unique.

B.6 Proof of Proposition 7

Since all type-O consumers opt in by default, seller A no longer faces an inference problem when targeting consumers and has no incentive to distinguish between the opt-in and opt-out pools. As such, seller A's advertising, z_A^{CCPA}, and pricing policy, p_A^{CCPA}, correspond to those under the full data sharing scheme from Proposition 4, with $p_A^{CCPA} = \frac{1}{2}\bar{u}$, and $z_A^{CCPA} = z_A^{FS}$. Strong-willed consumers are also indifferent to opting in or out, because seller A does not distinguish between the two pools. Thus, they also opt in by the default policy. The probability of a consumer of type S or W receiving an advertisement from seller A is consequently $\frac{z_A^{CCPA}}{\pi_S + \pi_W}$.

As seller A does not distinguish between the opt-in and opt-out pools, weak-willed consumers do not face the cost of losing improved advertisement targeting by seller A if they opt out. As such, they need only consider how it impacts their interaction with seller B.

Seller B must now set a price schedule for consumers $p_{B,in}^{CCPA}\left(\gamma_i\right)$ and an advertising policy function $\hat{z}_{B,in}^{CCPA}\left(\gamma_i\right)$ for consumers that opt in. By observing the level of temptation of weak-willed consumer i in the opt-in pool, seller B will charge his full reservation value for good B as in Proposition 4: $p_{B,in}^{CCPA}\left(\gamma_i\right) = \gamma_i \bar{v}$. For the opt-out pool, seller B must set a uniform price $p_{B,out}^{CCPA}$ and advertising intensity $z_{B,out}^{CCPA}$.

Let us conjecture an equilibrium cutoff γ_{**}^{CCPA} such that weak-willed consumers opt in if $\gamma_i \leq \gamma_{**}^{CCPA}$, and opt out if $\gamma_i > \gamma_{**}^{CCPA}$. By similar arguments to those in Proposition 6, seller B charges a price:

$$p_{B,out}^{CCPA} = \max\left\{\frac{1}{2}, \gamma_{**}^{CCPA}\right\}\bar{v},$$

to consumers in the opt-out pool.

Note that any weak-willed consumer that would be targeted by seller B if seller B knew his γ_i from the opt-in pool is better off by opting out. To see this, we recognize that, if he opts in, he receives utility $-\gamma_i\bar{v}$ from buying the temptation good; if instead he opts out, then his expected utility is $-\frac{z_{B,out}}{(1-\gamma_{**}^{CCPA})\pi_W}\min\left\{p_{B,out}^{CCPA}, \gamma_i\bar{v}\right\}$. Since $\frac{z_{B,out}^{CCPA}}{(1-\gamma_{**}^{CCPA})\pi_W} \leq 1$, the consumer prefers opt-out, as he may not receive the advertising from seller B.

In contrast, suppose that seller B commits to leaving this consumer alone if he opts in. In this case, the consumer prefers opt-in, and this preference is strict when $z_{B,out}^{CCPA} > 0$, under which case he may be targeted by seller B in the opt-out pool. Since the equilibrium is trivial if $z_{B,out}^{CCPA} = 0$, we hereafter only consider the case when $z_{B,out}^{CCPA} > 0$. Therefore, consumers that would be left alone by seller B strictly prefer opt-in, and those that would be targeted by seller B if they opt in prefer opt-out.

Consider now the optimal advertising policy of seller B. By committing to leaving weak-willed consumers in the opt-in pool alone, that is, $\hat{z}_{B,in}^{CCPA}(\gamma_i) = 0$, seller B can bifurcate the pool of weak-willed consumers to improve its efficiency in targeting the more-tempted consumers in the opt-out pool. To find the optimal cutoff γ_{**}^{CCPA}, consider the profit for seller B, Π_B. As the profit from the opt-in pool under this strategy is zero, we have

$$\begin{aligned}
\Pi_B &= p_{B,out}^{CCPA}\frac{z_{B,out}^{CCPA}}{\left(1-\gamma_{**}^{CCPA}\right)\pi_W}\pi_W\int_{\gamma_{**}^{CCPA}}^{1}\mathbf{1}_{\left\{\gamma_i \geq p_{B,out}^{CCPA}/\bar{v}\right\}}d\gamma_i - F\frac{z_{B,out}^{CCPA}}{1-z_{B,out}^{CCPA}} \\
&= \gamma_{**}^{CCPA}\bar{v}z_{B,out}^{CCPA}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}>\frac{1}{2}\right\}} + \frac{1}{1-\gamma_{**}^{CCPA}}\frac{\bar{v}}{4}z_{B,out}^{CCPA}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}\leq\frac{1}{2}\right\}} - F\frac{z_{B,out}^{CCPA}}{1-z_{B,out}^{CCPA}},
\end{aligned}$$

where $z_{B,out}^{CCPA} \in \left[0, \left(1-\gamma_{**}^{CCPA}\right)\pi_W\right]$ .The first-order condition for $z_{B,out}^{CCPA}$ is

$$\gamma_{**}^{CCPA}\bar{v}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}>\frac{1}{2}\right\}} + \frac{1}{1-\gamma_{**}^{CCPA}}\frac{\bar{v}}{4}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}\leq\frac{1}{2}\right\}} - \frac{F}{\left(1-z_{B,out}^{CCPA}\right)^2}\begin{cases} <0 & \text{if } z_{B,out}^{CCPA} = 0 \\ =0 & \text{if } z_{B,out}^{CCPA} \in \left(0, \left(1-\gamma_{**}^{CCPA}\right)\pi_W\right) \\ >0 & \text{if } z_{B,out}^{CCPA} = \left(1-\gamma_{**}^{CCPA}\right)\pi_W \end{cases},$$

from which follows that

$$z_{B,out}^{CCPA} = \min\left\{\max\left\{1-\sqrt{\frac{F}{\gamma_{**}^{CCPA}\bar{v}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}>\frac{1}{2}\right\}} + \frac{1}{1-\gamma_{**}^{CCPA}}\frac{\bar{v}}{4}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}\leq\frac{1}{2}\right\}}}}, 0\right\}, \left(1-\gamma_{**}^{CCPA}\right)\pi_W\right\}.$$

Assume that seller B's problem is nontrivial, that is, $z_{B,out}^{CCPA} \neq 0$. At an interior solution, seller B's profit is given by

$$\Pi_B = \left(\sqrt{\gamma_{**}^{CCPA}\bar{v}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}>\frac{1}{2}\right\}} + \frac{1}{1-\gamma_{**}^{CCPA}}\frac{\bar{v}}{4}\mathbf{1}_{\left\{\gamma_{**}^{CCPA}\leq\frac{1}{2}\right\}}} - F\right)^2.$$

Note that Π_B when $\gamma_{**}^{CCPA} > \frac{1}{2}$ is maximized at the corner $\gamma_{**}^{CCPA} = 1$, in which case profit is arbitrarily close to zero, since $z_{B,out}^{CCPA} = \left(1 - \gamma_{**}^{CCPA}\right)\pi_W \to 0$ since revenue per consumer is bounded by $\bar{v}$. Consequently, this cannot be the optimal choice of γ_{**}^{CCPA}. When $\gamma_{**}^{CCPA} \leq \frac{1}{2}$, this is maximized at $\gamma_{**}^{CCPA} = \frac{1}{2}$, since $\frac{1}{1-\gamma_{**}^{CCPA}}$ is increasing in γ_{**}^{CCPA}. Consequently,

$$z_{B,out}^{CCPA} = \min\left\{\max\left\{1 - \sqrt{\frac{2F}{\bar{v}}}, 0\right\}, \frac{1}{2}\pi_W\right\}.$$

One may be concerned that seller B faces a time-consistency problem and has incentive to search the opt-in pool after announcing not to advertise to those consumers. Note that the marginal revenue to advertising to the most tempted weak-willed customer in the opt-in pool is weakly less than $\frac{\bar{v}}{2}$, while the marginal revenue to targeting the opt-out pool is $\frac{\bar{v}}{2}$. Thus, seller B prefers targeting the opt-out pool, even though it has incentives to target the opt-in pool after exhausting the opt-out pool. Thus, the commitment is not an issue if seller B faces a high cost to cover the consumers. Only when it is optimal for the seller to cover more than half of the weak-willed consumers, the commitment is an issue. In this case, if seller B cannot commit to leaving consumers in the opt-in pool alone, those weak-willed consumers with γ_i right below $\frac{1}{2}$, anticipating being targeted by the seller, would choose opt-out. This in turn reduces the effective γ_{**}^{CCPA} to a level below $\frac{1}{2}$, the optimal level, thus hurting the seller. We assume that the seller has the power to precommit in this case.

Since seller B finds it optimal to separate the two pools according to the cutoff γ_{**}^{CCPA}, this confirms the conjectured cutoff equilibrium. Furthermore, since the price is $p_{B,out}^{CCPA} = \frac{\bar{v}}{2}$ and $\gamma_{**}^{CCPA} = \frac{1}{2}$, it follows that all weak-willed consumers that opt out buy good B when it is advertised to them.

B.7 Proof of Proposition 8

We compare the social welfare under four data sharing schemes: no data sharing, full data sharing, the CCPA and the GDPR.

No data sharing: From the proof of Proposition 5, the social welfare is

$$W^{NS} = \frac{3}{8}\bar{u}\left(\pi_S + \pi_W\right) z_A^{NS} + \pi_W z_B^{NS}\left(u_B - \frac{\bar{v}}{8}\right).$$

Full data sharing: From the proof of Proposition 5, the social welfare is

$$W^{FS} = \frac{3}{8}\bar{u} z_A^{FS} + z_B^{FS} u_B.$$

CCPA: From (1), the social welfare is given by

$$W^{CCPA} = \frac{3}{8}\bar{u} z_A^{CCPA} + z_{B,out}^{CCPA} u_B.$$

We first compare to full data sharing. It is straightforward to see that the benefit from sharing data with seller A is the same under both schemes ($z_A^{FS} = z_A^{CCPA}$), while the cost of sharing with seller B is worse under full data sharing $z_B^{FS} \geq z_{B,out}^{CCPA}$. It therefore follows that

$$W^{CCPA} \geq W^{FS}.$$

This inequality is sharp whenever seller B's advertising policy is nontrivial. Thus, the CCPA strictly dominates the scheme of full data sharing.

Comparing to no data sharing, it is apparent that the utility benefit of good A is higher with the CCPA, since seller A can fully cover weak-willed consumers ($z_A^{CCPA} \geq z_A^{NS}$), while the utility cost from good B is more severe $\left(u_B - \frac{\bar{v}}{2}\right) z_{B,out}^{CCPA} < \left(u_B - \frac{3}{8}\bar{v}\right) \pi_W z_B^{NS}$ and $\frac{z_{B,out}^{CCPA}}{\pi_W} \geq z_B^{NS}$. Consequently, for

$$u_B < u_{B*} = -\frac{3\bar{u}\left(z_A^{CCPA} - (\pi_S + \pi_W) z_A^{NS}\right) + \pi_W z_B^{NS} \bar{v}}{8\left(z_{B,out}^{CCPA} - \pi_W z_B^{NS}\right)},$$

we have $W^{CCPA} < W^{NS}$, and $W^{CCPA} \geq W^{NS}$ otherwise. This bound u_{B*} is well-defined since z_A^{CCPA}, z_A^{NS}, z_B^{NS}, and $z_{B,out}^{CCPA}$ are all independent of u_B.

GDPR: From a social welfare perspective, we have

$$\begin{aligned} W^{GDPR} &= \left(z_{A,in}^{GDPR} + \left(1 - \gamma_{**}^{GDPR}\right) \pi_W \frac{z_{A,out}^{GDPR}}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W}\right) \frac{3}{8}\bar{u} \\ &\quad + \left(z_{B,in}^{GDPR} + \pi_W \left(1 - \gamma_{**}^{GDPR}\right) \frac{z_{B,out}^{GDPR}}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W}\right) u_B. \end{aligned}$$

As the scheme of full data sharing is dominated by the CCPA, we do not compare the GDPR with full data sharing, but rather the CCPA.

We now compare the GDPR to no data sharing:

$$\begin{aligned} W^{GDPR} - W^{NS} &= \left(z_{A,in}^{GDPR} + \left(1 - \gamma_{**}^{GDPR}\right) \pi_W \frac{z_{A,out}^{GDPR}}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} - z_A^{NS}\right) \frac{3}{8}\bar{u} \\ &\quad + \left(z_{B,in}^{GDPR} + \left(1 - \gamma_{**}^{GDPR}\right) \pi_W \frac{z_{B,out}^{GDPR}}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} - z_B^{NS}\right) u_B, \end{aligned}$$

where z_A^{NS} and z_B^{NS} are independent of u_B. The first term is positive, representing the improved matching with seller A under the GDPR, while the second is negative, due to the increased exposure of weak-willed consumers to seller B.

Notice that, when $u_B = 0$, it must be the case that $W^{GDPR} > W^{NS}$ due to the improved matching with seller A. For sufficiently negative u_B, in contrast, the most-tempted weak-willed consumers lose the camouflage of not only all strong-willed consumers, but also the more-mildly tempted weak-willed consumers. The social benefit of the GDPR for increased matching with seller A accrues as $z_{A,in}^{GDPR}\frac{\bar{u}}{4}$, which is bounded from above. In contrast, with γ_{**}^{GDPR} bounded from below, the cost to the weak-willed consumers in the opt-out pool becomes arbitrarily large. Since they have less camouflage than in the no data sharing scheme because the strong-willed and mildly weak-willed consumers all opt in, then $W^{GDPR} < W^{NS}$. Since the objectives are continuous, it follows that there exist critical values of u_B, u_{B**}, such that $W^{GDPR} < W^{NS}$ when $u_B \leq u_{B**}$.

We now compare the GDPR with the CCPA. The difference in the social welfare is given by

$$\begin{aligned} W^{GDPR} - W^{CCPA} &= \left(z_{A,in}^{GDPR} + \left(1 - \gamma_{**}^{GDPR}\right) \pi_W \frac{z_{A,out}^{GDPR}}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} - z_A^{CCPA} \right) \frac{3}{8}\bar{u} \\ &\quad + \left(z_{B,in}^{GDPR} + z_{B,out}^{GDPR} - \frac{1 - \pi_S - \pi_W}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} z_{B,out}^{GDPR} - z_{B,out}^{CCPA} \right) u_B. \end{aligned}$$

Note that under the CCPA, seller A has higher advertising efficiency and therefore is able to better cover its intended consumers, that is, the first term is negative. Since total advertising by seller B under the GDPR is less than under the CCPA, $z_{B,in}^{GDPR} + z_{B,out}^{GDPR} - z_{B,out}^{CCPA} < 0$, because of seller B's less efficient targeting of the most-tempted customers, it follows that the coefficient of u_B in the last term is negative.

Consequently, there exists a critical u_{B*} such that $W^{GDPR} > W^{CCPA}$ if $u_B \leq u_{B***}$ (and $W^{GDPR} < W^{CCPA}$ otherwise).

Ranking the four schemes: Suppose u_B is sufficiently mild ($u_B > \max\{u_{B*}, u_{B***}\}$), then $W^{CCPA} > W^{GDPR},\ W^{NS}$. Since $W^{CCPA} \geq W^{FS}$, it follows that the CCPA delivers the highest social welfare.

In contrast, suppose u_B is sufficiently severe ($u_B < \min\{u_{B*}, u_{B**}\}$), then $W^{NS} > W^{GDPR},\ W^{CCPA}$. Since $W^{CCPA} \geq W^{FS}$, it follows that no data sharing delivers the highest social welfare.

Finally, it is sufficient, although not necessary, for u_B to be in an intermediate range ($u_B < u_{***}$ and $u_B > u_{B**}$), for $W^{GDPR} > W^{CCPA},\ W^{NS}$. Since $W^{CCPA} \geq W^{FS}$, it follows that GDPR delivers the highest social welfare.

References

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar (2019), Too Much Data: Prices and Inefficiencies in Data Markets, National Bureau of Economic Research Working Paper 26296.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman (2016), The Economics of Privacy, Journal of Economic Literature 54, 442–492.

Acquisti, Alessandro, and Hal R. Varian. (2005), Conditioning Prices on Purchase History, Marketing Science 24, 367–381.

Aguiar, Mark, Mark Bils, Kerwin Kofi Charles, and Erik Hurst (2018), Leisure Luxuries and the Labor Supply of Young Men, mimeo Princeton University, University of Chicago Booth School of Business.

Ali, S. Nageeb, and Roland Bénabou (2019), Image versus Information: Changing Societal Norms and Optimal Privacy, American Economic Journal, Microeconomics, forthcoming.

Ali, S. Nageeb, Greg Lewis, and Shoshana Vasserman (2019), Voluntary Disclosure and Personalized Pricing, mimeo Penn State University.

Athey, Susan, Christian Catalini, and Catherine Tucker (2017), The Digital Privacy Paradox: Small Money, Small Costs, Small Talk, NBER Working Paper No. 23488.

Bénabou, Roland, and Marek Pycia (2002), Economic Letters 77, 419–424.

Bergemann, Dirk, and Stephen Morris (2019), Information Design: A Unified Perspective, Journal of Economic Literature 57, 44–95.

Bergemann, Dirk, Alessandro Bonatti, and Tan Gan (2019), The Economics of Social Data, mimeo Yale University.

Calzolari, Giacomo, and Alessandro Pavan (2006), On the Optimality of Privacy in Sequential Contracting, Journal of Economic Theory 130, 168–204.

Campbell, James, Avi Goldfarb, and Catherine Tucker (2015), Privacy Regulation and Market Structure, Journal of Economics and Management Strategy 24, 47–73.

Deckel, Eddie, and Barrton L. Lipman (2012), Costly Self-Control and Random Self-Indulgence, Econometrica 80, 1271–1302.

DellaVigna, Stefano (2009), Psychology and Economics: Evidence from the Field, Journal of Economic Literature 47, 315–72.

Easley, David, Shiyang Huang, Liyan Yang, and Zhuo Zhong (2019), The Economics of Data, mimeo Cornell University.

Farboodi, Maryam, and Laura Veldkamp (2020), A Growth Model of the Data Economy, mimeo MIT Sloan.

Goldfarb, Avi, and Catherine Tucker (2019), Digital Economics, Journal of Economic Literature 57, 3–43.

Gul, Faruk, and Wolfgang Pesendorfer (2001), Temptation and Self-Control, Econometrica 69, 365-396.

– (2004), Self-Control and the Theory of Consumption, Econometrica 72, 119–158.

Hurley, Mikella, and Julius Adebayo (2016), Credit Scoring in the Era of Big Data. Yale Journal of Law and Technology. 18:148–216.

Ichihashi, Shota (2019), Online Privacy and Information Disclosure by Consumers, American Economic Review, forthcoming.

Jones, Charles, and Chris Tonetti (2020), Nonrivalry and the Economics of Data, American Economic Review, forthcoming.

Kim, A. E., Chew, R., Wenger, M., Cress, M., Bukowski, T., Farrelly, M., and Hair, E. (2019). Estimated Ages of JUUL Twitter Followers. JAMA pediatrics.

Kreps, David M (1979), A Representation Theorem for "Preference for Flexibility", Econometrica, 565–577.

Morris, Chris (2019), Big Data Meets the Beer Industry, Fortune. Jan.

Raustiala, Kal, and Jon Christopher Sprigman (2019), The Second Digital Disruption: Streaming & the Dawn of Data-Driven Creativity. New York University Law Review. 94:1555–1622.

Stovall, J. (2010), Multiple Temptations, Econometrica 78, 349–376.

Strotz, R. (1955), Myopia and Inconsistency in Dynamic Utility Maximization, Review of Economic Studies 23, 165–180.

Tang, Huan (2019), The Value of Privacy: Evidence from Online Borrowers, mimeo HEC.

Taylor, Curtis R. (2004), Consumer Privacy and the Market for Customer Information, RAND Journal of Economics 35, 631–650.

Tirole, John (2019), Digital Dystopia, mimeo Toulouse School of Economics.