Data Privacy and Consumer Vulnerability*

John Zhuang Liu†  Michael Sockin‡  Wei Xiong§

December 2021

Abstract

This paper microfound a consumer’s preference for data privacy as a mechanism for concealing behavioral vulnerabilities. This approach facilitates a welfare analysis of different data privacy regulations, such as the GDPR and CCPA. Sharing data with a digital platform benefits a consumer through improved matching efficiency with normal consumption goods at the expense of exposing those with self-control issues to temptation goods. Although the GDPR and CCPA empower consumers to opt in or out of data sharing, our analysis also highlights the limitations of these regulations because of nuanced data sharing externalities induced by consumers’ active and default choices.

*We are grateful to Roland Benabou, Dirk Bergmann, Long Chen, Will Cong, Shaowei Ke, Liyan Yang, Leeat Yariv, Lixin Ye, and seminar participants at American Finance Association Meetings, Cambridge, CUHK Shenzhen, Luohan Academy, Northwestern, Princeton, Stanford, Toulouse, UCL, 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 has introduced not only substantial benefits for consumers, such as improved access to products and services, but also imposed undesirable costs through the collection and exploitation of personal data. This amassing of consumer data by digital platforms such as Google, Amazon, and Facebook represents an unprecedented challenge to consumer privacy. Motivated by this concern, many countries have enacted privacy regulations. For example, 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. Despite the increased focus on consumer privacy, and the wave of new regulations that it has inspired, academics and policymakers still lack a systematic framework for analyzing why consumers have a preference for privacy and how different privacy regulations affect their welfare.

The existing economics literature on consumer privacy tends to emphasize 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. Although aversion to price discrimination is an important motivation for a privacy preference among consumers, this mechanism tilts the distribution of the social surplus towards firms without necessarily inducing a net social loss. As such, it does not provide a compelling argument for pervasive data privacy regulations.

This conventional approach to consumer privacy ignores the more immediate issue of protecting vulnerable consumers who are subject to behavioral biases. The recent report

---

1 More recently, Virginia and Colorado have passed the Virginia Consumer Data Protection Act (VCDPA) and the Colorado Privacy Act (CPA), respectively, that are similar in spirit to the CCPA.

2 See Acquisti, Taylor and Wagnman (2016) and Goldfar and Tucker (2019), for recent reviews of this literature, which shows that the effects of price discrimination on consumer surplus and social welfare depend on the competitive landscape. 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 naïve regarding 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 (2020) 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.

3 The 2015 report "Big Data and Differential Pricing" by the Obama White House Economic Advisors concludes that price discrimination associated with the use of big data has not risen to a first-order concern yet. The recent survey by Chen et al. (2021) also shows that price discrimination is not the most relevant concern for Alipay users in sharing personal data with third-party mini-programs.
of the Stigler Committee (2019), for instance, argues that such consumers are particularly susceptible to exploitative behavior by digital platforms. Because the digitalization of commerce has substantially increased the capacity of digital platforms to influence consumers at a personal level, they are able to not only discover but also cater to each consumer’s specific biases and sources of vulnerability. The OECD (2019) report on consumer protection policies emphasizes the important role of data privacy in safeguarding vulnerable consumers in the digital age. Legal scholars, e.g., Zarsky (2019) and Spencer (2020), have suggested that the core objection to such online manipulation of consumers is not its manipulative nature but its implementation—intense data collection, personalization, and real-time implementation made possible by the internet. Rather than pursuing direct regulation alongside traditional consumer protection laws, they propose that issues of consumer vulnerability are better addressed through a comprehensive data privacy legislation. This notion is also reflected in the GDPR and the CCPA that were, in part, designed to protect consumers through informed consent requirements. Despite the importance of data privacy to protecting vulnerable consumers, there is no systematic analysis of how the two interact.

In this paper, we develop a model to fill this gap. We adopt a specific form of consumer vulnerability—limited self-control—to illustrate how consumer biases interact with data privacy. Because of this behavioral weakness, some consumers may be unable to resist the temptation of purchasing a certain good when a seller advertises it to them, even though purchasing it does not improve (and may even harm) their utility. The desire of such weak-willed consumers to avoid being targeted by temptation goods sellers provides a rationale for a preference for data privacy. We model this limited self-control using the temptation utility of Gul and Pesendorfer (2001).[^4] This utility specification allows consumers to suffer a mental cost from resisting temptation goods on their menus, which, in turn, may lead to a preference for a smaller menu without temptation goods, in sharp contrast to the standard preference for larger menus in 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 analyze how different data sharing schemes affect social welfare.

Our model features an ecosystem associated with a digital platform, such as Google or

[^4]: Stovall (2010) has expanded this 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).
Facebook, that may collect the data of consumers on the platform and share the data with sellers. There are two sellers. Seller $A$ sells a normal consumption good, such as music, while seller $B$ sells a temptation good. This temptation good can be viewed narrowly as addictive content, such as gambling or video games, or more broadly as a good or service that the consumer may be persuaded to buy but ultimately does not want. Each seller can target advertisements to potential buyers of its good at a convex cost. A consumer may receive advertisements from none, one, or both of the sellers, and then chooses from the menu none, one, or both 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 good and serves as noise in sellers’ targeted advertising.

For simplicity, we assume that seller $A$ cannot price discriminate its intended consumers as both strong-willed and weak-willed consumers have a random utility over the normal good. As a result, both strong-willed and weak-willed consumers prefer receiving advertisements from seller $A$. Furthermore, because 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 also suffer from receiving advertisements from seller $B$. This is a key tension in our model that drives the differences in social welfare under different data sharing schemes. As weak-willed consumers are aware of their vulnerability, they may choose to protect themselves when given the option to opt out of data sharing.$^5$

To illustrate the role of data privacy, we characterize two benchmark data sharing schemes: one without any data sharing, in which consumers remain fully anonymous to sellers, and one with full data sharing, in which sellers can perfectly infer each consumer’s

$^5$Because we adopt the self-control utility framework, the weak-willed consumers in our model make fully rational data sharing choices, despite their lack of self-control in 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, in particular those who are naive about their weakness, not to fully internalize their lack of self-control, e.g., O’Donoghue and Rabin (1999), and DellaVigna and Malmendier (2004).
type. In the former scheme, which is analogous to conventional advertising before the era of big data, each seller faces a dark pool of consumers. As a result, the convex cost of advertising restricts each seller to sending advertisements only to a random 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 being targeted by 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, when the temptation of the weak-willed is sufficiently severe, the full data sharing scheme reduces their welfare relative to the no data sharing scheme, and the harm to weak-willed consumers may even exceed the gain of strong-willed consumers and lower the overall social welfare.

Given these two benchmarks, we next investigate whether data privacy regulations, such as the GDPR and the CCPA, can protect vulnerable consumers. Both privacy regulations 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 any data collection by the platform, while the CCPA allows the platform to collect consumer data unless a consumer explicitly opts out. That is, the default choice of the GDPR is opt-out unless a consumer opts in, while that of the CCPA is opt-in unless a consumer opts out. In our model, because the third type of consumers are indifferent to either opt-in or opt-out, their choices are determined by the default setting of the implemented 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 both sellers.

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 exceptionally tempted opt out while those more 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.

Because both privacy regulations offer consumers an appealing opt-in or opt-out option,
one might expect that they raise social welfare through a revealed preference argument. That is, by the optimality of each consumer’s choice, these regulations should improve upon the no data sharing and full data sharing schemes. Indeed, our analysis confirms that the CCPA strictly dominates full data sharing because it allows seller $A$ to fully identify its intended consumers while simultaneously providing some, albeit imperfect, protection to weak-willed consumers. The comparison of these schemes to no data sharing, however, is more nuanced because of the presence of data sharing externalities—when one consumer chooses to opt in with data sharing either by an active choice or by default, the consumer’s data allow sellers to infer the preferences of other consumers.

The data sharing externalities we highlight can be either negative or positive. When one consumer opts in, she drops out of the opt-out pool and reduces the camouflage available for vulnerable consumers in the opt-out pool. 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. On the other hand, 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 target 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. Because the GDPR provides more of a balance between the matching efficiency of consumers with the normal good and the protection of vulnerable consumers from the temptation good, there may exist an intermediate range of the temptation problem faced by weak-willed consumers in which the GDPR is the most desirable scheme. Taken together, our analysis highlights the limitations of these privacy regulations because of the data sharing externalities.

**Related Literature** Our paper adds to the literature on the economics of privacy and data sharing. Beside the aforementioned price discrimination mechanism, Ali and Benabou (2020) and Jann and Schottmuller (2020) have also highlighted social discrimination as another rationale for individuals to prefer privacy. That is, when an individual’s actions are publicly observable, the public might use the observed actions to infer the individual’s unobservable type, which in turn induces pro-social behavior and discourages the individual
from acting on her private information and personal preferences. Tirole (2021) goes further to argue that in the absence of privacy protections, 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. Our model motivates the need for privacy protection to protect vulnerable consumers with behavioral weakness and highlights nuanced effects of data privacy regulations because a consumer does not internalize the impact of their data sharing decisions on other consumers.

The data sharing externalities highlighted by our model echo the notion of social data put forth by Acemoglu et al. (2019), Bergemann, Bonatti and Gan (2019), and Easley et al. (2019). These papers commonly formulate a hypothetical data market through which a digital platform purchases personal data from consumers. As the data sold by one consumer also reveals information about other consumers, such data externality may drive down the equilibrium data price, leading to excessive data sharing. Although these models provide sharp insights, they omit important features of sharing personal data with digital platforms in practice—consumers typically face the choice of whether to authorize data sharing in exchange for the convenience of using the free services offered by the platform. In this exchange, there is not an explicit data price. Instead, both the cost and benefit of data sharing are determined not only by a consumer’s preferences and vulnerabilities but also data sharing by other consumers. Our model explicitly formulates such a setting and further highlights that each consumer’s data sharing decision may intricately depend on the platform’s data-driven advertising policies and the default choice set by data privacy regulations. 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), Farboodi and Veldkamp (2020), and Cong, Xie and Zhang (2020), which emphasize the nonrivalry and increasing returns-to-scale nature of data sharing but treat the cost of data sharing as exogenous.

Our paper also adds privacy protection as a new dimension to the extensive literature on the protection of vulnerable consumers. For example, there is ample evidence of firms over-pricing leisure goods, such as credit card-finance consumption, and under-pricing investment goods, such as health club membership, to target time-inconsistent consumers (e.g., DellaVigna and Malmendier (2004, 2006)), lenders targeting impulsive borrowers for payday loans with high interest rates (e.g., Bertrand and Morse (2011), Melzer (2011)), firms us-
ing add-on pricing to target inattentive consumers (e.g., Gabaix and Laibson (2006)), and inattention causing consumers to pay banks tens of billions each year in overdraft fees (e.g., Stango and Zinman (2014)). The internet and digital technologies not only make it easier for firms to target consumers on these existing vulnerabilities (see, for example, Di Maggio and Yao (2020) for empirical evidence of FinTech lenders using payment information to target borrowers with self-control issues and He, Huang and Zhou (2021) for a related model), but have also created new forms of vulnerabilities such as digital addiction. 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. Allcott et al. (2020) find evidence from a randomized experiment of Facebook users to show that social media are addictive. Allcott, Gentzkow and Song (2021) estimate a structural model through a randomized experiment to show that self-control problems contribute to 31 percent of social media use. As we discuss at the end of the paper, despite the limits of data privacy regulations, protecting data privacy ex ante may be the most effective way to protect vulnerable consumers on digital platforms.

1 The Model

We consider the ecosystem associated with a digital platform. By collecting the digital history of consumers in the ecosystem, the platform can infer each consumer’s biases and weakness. The ecosystem contains a mix of consumers, some with self-control and some without. To capture the realistic data sharing problem faced by each consumer, we suppose that the platform owns two goods sellers. One sells a normal good (good A), while the other sells a good (good B) that we refer to as a temptation good. If a consumer shares personal data with the platform, the platform will then relay the data to both sellers and allow them to target their intended consumers. One may literally interpret these two goods as merchandise. For example, the normal good could be music or any other product that is desirable to consumers, while the temptation good can be video gaming, gambling or any product that is potentially addictive and harmful to consumers. One may also broadly interpret the normal good as convenience provided by the platform to attract users, such as free search provided by Google, while the temptation good as a certain kind of potential harm that data sharing might induce, such as impulse consumption or addictive content.

It is a key feature of our model that each consumer’s data sharing decision is bundled
with these two goods. That is, by sharing personal data with the platform, a consumer makes herself more accessible to the normal good seller, even though this convenience comes at the expense of also exposing herself to the temptation good seller.\(^6\)

### 1.1 Consumers

There are three types of consumers \( \{S, W, O\} \), 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. 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 and their presence may serve as potential camouflage for weak-willed consumers in certain data-sharing schemes that we will analyze later. 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\(^7\)

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

\[\text{(1)}\]

\(^6\)We implicitly assume that the platform has perfect data security so that consumers do not face the risk that their data may be hacked by an adversarial third-party. See Fainmesser, Galeotti and Momot (2021) for an analysis of how a platform’s data security and data collection strategies may jointly affect a user’s activity on the platform, subject to the endogenous risk of third-party hacks.

\(^7\)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:

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

without qualitatively impacting our key insights. We choose the simpler specification for expository brevity. We thank Shaowei Ke for pointing out this construction to us.
where $x$ is a choice from menu $N$, and $u(x)$, $v(x)$, and $p(x)$ are the commitment utility, temptation utility, and price, respectively, of this choice. The consumer’s actual choice from the menu in the second step is determined by the first maximization in Equation (1):

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

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.

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.

Alternatively, one may also model the consumer’s self-control problem through the formulation of hyperbolic discounting popularized by Laibson (1997), O’Donoghue and Rabin (1999), and DellaVigna and Malmendier (2004). In a dynamic setting with multiple periods, the use of hyperbolic discounting causes a consumer to over-weight the enjoyment from consuming the temptation good in the current period and under-weight the cost in the future period, leading to a self-control problem. If the consumer is sufficiently sophisticated to anticipate that when the temptation good appears on the menu in a future period, the future self will not be able to correctly account for the trade-off between the cost and benefit, the sophisticated consumer with the self-control problem may also prefer to keep the temptation good off her menu. See DellaVigna (2009) for a review of this approach. Interestingly, Benabou and Pycia (2002) have formally shown that the temptation utility representation

---

8 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.
is equivalent to a formulation of an internal conflict among multiple selves. Thus, in terms of welfare analysis, we view the consumer with temptation utility as consistent to a sophisticated consumer with hyperbolic discounting. We opt for the temptation utility approach as it is simpler and does not require a dynamic setting.

**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 $B$:

$$
\begin{array}{|c|c|c|c|c|c|c|}
\hline
x & \text{strong-willed} & \text{weak-willed} & \text{type } O \\
\hline
A & u_A > 0 & v_A(x) & u_A > 0 & v_A(x) & u_O(x) & v_O(x) \\
B & u_B < 0 & v_B(x) & u_B < 0 & v_B(x) & 0 & 0 \\
\hline
\end{array}
$$

(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 $H(\tilde{u}_A) \sim U[0, \tilde{u}]$, with $\tilde{u} > 0$ 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.

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 $\gamma_i \tilde{v} - u_B$ to weak-willed consumers, where $\tilde{v} > 0$ 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:

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

---

$^9$Our approach does not capture naive consumers with hyperbolic discounting. As such consumers do not recognize their own weakness, they may not value the need to protect their data privacy. Thus, the opt-in and opt-out choice instituted by either the GDPR or the CCPA does not concern them. Nevertheless, our approach captures sufficiently sophisticated consumers who recognize their own weakness and thus the need to protect their data privacy.
which implies that the consumer will choose to buy good $B$ if his temptation coefficient $\gamma_i$ 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 $\gamma_i$, 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.\textsuperscript{[10]} 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.

**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 $B$.\textsuperscript{[11]} 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: $M_i^A \in \{\{A, \emptyset\}, \emptyset\}$ 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 $M_i^B \in \{\{B, \emptyset\}, \emptyset\}$ is the menu for good $B$.

\textsuperscript{[10]} 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 (2020) shows that a seller who can precommit would choose not to price discriminate when consumers can choose whether to disclose their information.

\textsuperscript{[11]} This assumption simplifies our analysis by abstracting from a consumer’s budget constraint, and allows us to focus on how different data sharing schemes with sellers affect a consumer’s choice and welfare. Imposing a budget constraint would introduce an additional distortion from temptation goods crowding out normal good consumption for weak-willed consumers.
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 \( \mathcal{M}^A_i \) and \( \mathcal{M}^B_i \):

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

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

\[
U_{\tau_i} (\mathcal{M}^A_i, \mathcal{M}^B_i) = \tilde{u}_{\tau_i}(x_{\tau_i}(\mathcal{M}^A_i)) - p_{A,\tau_i}(x_{\tau_i}(\mathcal{M}^A_i)) + u_{\tau_i}(y_{\tau_i}(\mathcal{M}^B_i)) + v_{\tau_i}(y_{\tau_i}(\mathcal{M}^B_i)) - p_{B,\tau_i}(y_{\tau_i}(\mathcal{M}^B_i)) - \max_{y \in \mathcal{M}^B_i} v_{\tau_i}(y).
\]

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 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 its convexity reflecting, as in Grossman and Shapiro (1984), that it is increasingly costly to reach a broader audience.\(^{12}\) In what follows, we impose a technical condition \( F < \frac{\theta}{4} \) to ensure a nontrivial equilibrium for good \( A \).

Figure 1 shows the composition of consumers faced by the two sellers. Seller \( A \) would like to target both strong-willed and weak-willed consumers, but only type-O consumers, while seller \( B \) aim to target only weak-willed consumers. How much they can differentiate the consumers depend on the data sharing schemes that we will discuss. Thus, each seller

\(^{12}\text{To the extent that consumers have limited attention and online advertisers do not want to flood them with unlimited advertisements, the fees have to rise progressively with the quantity. This is also consistent with the fact that, in practice, sellers need to pay substantial advertising fees to online advertisers.}\)
Figure 1: The composition of consumers faced by the sellers.

\( k \in \{A, B\} \) chooses its advertising and pricing policies based on its information set about the consumers to maximize its expected profit:

\[
\Pi_k = \sup_{\{p_k, z_k\}} E \left[ \int_{i \in Z_k} p_k(i) di - F \frac{z_k}{1 - z_k} \right | T^k],
\]

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 \( T^k \), which may allow the seller to charge different consumers different prices.

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

\[
p_A \leq \bar{u}, \quad p_B \leq \bar{v}.
\]

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 purchase policy correspondence \( \{ x_{r_i}(M^A_i), y_{r_i}(M^B_i) \} \) 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 \( \{ x_{r_i}(M^A_i), y_{r_i}(M^B_i) \} \) for a menu set \( \{ M^A_i, M^B_i \} \).

- **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 ultimately owned by consumers. Because 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:

\[
W = \int \bar{u}_A \left( \pi_S 1_{\{ A \in M^A_{W} \cap x_S = A \}} + \pi_W 1_{\{ A \in M^A_{W} \cap x_W = A \}} \right) dH(\bar{u}_A) + \pi_W \int \left( u_B 1_{\{ B \in M^B_{W} \cap x_W = B \}} + (u_B - \gamma_i \bar{v}) 1_{\{ B \in M^B_{W} \cap x_W = \emptyset \}} \right) dG(\gamma_i).
\]

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). Because 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 $\gamma_i$ 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 \( \gamma_i \) 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 \( \gamma_i \) is sufficiently high: \( \gamma_i \geq \frac{p_B}{\bar{u}} \).

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 \( \gamma_i \) will buy it. Those with a modest temptation \( (\gamma_i < p_B/\bar{u}) \) resist it but still suffer a mental cost of \( \gamma_i \bar{u} - u_B \) 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 and thus faces a dark pool for its advertising. That is, the probability of each seller’s advertising reaching its intended consumers is the same as the unconditional probability, as shown in Figure 1. One may interpret this setting as the practice before the era of big data. 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 \( z^{NS}_A \) measure of consumers:

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

at a uniform price: \( p^{NS}_A = \frac{1}{2} \bar{u} \).
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 $A$). To the extent that seller $A$ does not advertise good $A$ to all strong-willed and weak-willed consumers, as required by the first-best equilibrium, the no data sharing scheme leads to inefficient matching of goods $A$. This inefficiency motivates more data sharing. At the same time, Equation (5) shows that anonymity discourages seller $B$ from targeting all weak-willed consumers. This is a source of welfare gain, which increases with the fraction of weak-willed consumers. 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.

\subsection*{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 infer each consumer’s type. One may simply interpret this scheme as the practice used by most digital platforms before any data privacy regulations were enacted. For simplicity, we suppose that the data collected by the platform allow it to determine not only the mental state of each consumer $i$ but also the severity of each weak-willed customer’s temptation coefficient $\gamma_i$. 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 and, in particular, to target vulnerable consumers, as summarized by Nadler and McGuigan (2018) and the Stigler Committee Report (2019).\footnote{Our model simplifies many nuances of how platforms target vulnerable consumers in practice. By using big data in real time to identify each consumer’s current mental state, a digital platform is able not only to}
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 \( z_{FS}^A \) as the measure of strong-willed and weak-willed consumers, to whom seller \( A \) advertises its good at a price of \( p_{FS}^A \). Proposition 4 derives the seller’s optimal \( z_{FS}^A \) and \( p_{FS}^A \). 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, \( z_{FS}^A \geq z_{NS}^A \), 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 \( p_{FS}^A = \frac{1}{2} \bar{u} \).

Access to consumer data also allows seller \( B \) to perfectly target weak-willed consumers. Furthermore, because seller \( B \) also observes the severity of each weak-willed customer’s temptation, it will price discriminate against each 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 \( \gamma_i \) higher than a threshold \( \gamma_{FS} \). 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 \( z_{FS}^A \) measure of strong-willed and weak-willed consumers:

   \[
   z_{FS}^A = \min \left\{ \max \left\{ 1 - 2\sqrt{\frac{F}{\bar{u}}}, 0 \right\}, \pi_S + \pi_W \right\}
   \]

   at the same price \( p_{FS}^A = \frac{1}{2} \bar{u} \).

identify a consumer’s cognitive and affective vulnerabilities, but also to determine the most effective strategy to manipulate her decision-making when she is most vulnerable. A platform may also adopt strategies that proactively trigger consumer vulnerability, for example, by exhausting a consumer’s willpower through aggressive advertising and the use of "dark patterns". Social media platforms, for instance, tailor nudges and content to maximize consumer engagement, at times bordering on inducing addictive behavior (e.g., Allcott et al. (2020)).
2. Seller B advertises its good to all weak-willed consumers with \( \gamma_i \geq z^{FS} = 1 - \frac{z_B^{FS}}{\pi_W} \), where \( z_B^{FS} \) is the total advertising by seller B:

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

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 for 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 harms it. 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 vulnerability 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 enough that it 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 matching efficiency of the normal good with 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, instead of forcing all consumers to take the same arrangement. 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 each consumer can make the best choice for herself. That is, strong-willed consumers can choose to opt in and consequently benefit from data sharing, while exceptionally 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 default allocation of rights to their personal data and allows digital platforms to collect consumer data only after explicit authorization. The CCPA, in contrast, gives businesses the default rights to collect consumer data and allows consumers to opt out of data collection through an explicit request.

Before presenting our analysis of these specific data sharing schemes, it is useful to note several features of our analysis. First, even though a vulnerable consumer may protect herself from the temptation good by opting out of data sharing with the platform, this protection comes at the cost of losing the convenient access to the normal good. This is because the platform has bundled each consumer’s data sharing choice for the two goods sellers together. As no one would choose to share data with sellers that exploit consumer vulnerability, the platform will always bundle its data sharing request with certain conveniences or benefits. In our model, this is the improved access to the normal good.

Second, opting out of data sharing does not provide perfect protection for a vulnerable consumer because the opt-out choice itself already provides useful information to the platform. How much the platform can use this signal to infer the type of the opt-out consumer depends on the composition of the opt-out pool. The opt-out pool provides effective protection only if it includes a large fraction of type-Ω consumers who are not interested in either the normal or temptation good. Interestingly, whether this is the case depends on the default choices of the specific data privacy regulations. The GDPR sets the default to be opt-out, while the CCPA sets it to be opt-in. As a type-Ω consumer is indifferent to information sharing on the platform, his data sharing choice is set by the default. Such an assumption is consistent with the role of nudges and framing from the behavioral choice architecture literature (e.g., Thaler et al. (2013)). The default choice therefore plays a vital role in determining the information structure on the platform and the protection that the

\[14\text{Such regulations do have an impact in practice. Goldberg et al. (2019), for instance, find that spending and visits fell by as much as 7\% for EU visitors in 2018 relative to 2017 because of the GDPR, and this decline was significantly more pronounced for smaller firms.}

\[15\text{Like the CCPA, the newly enacted VCPDA by Virginia and the CPA by Colorado also feature opt-in as the default choice.}\]
opt-out pool affords vulnerable consumers.

Third, despite the seeming appeal of the GDPR and the CCPA to deliver Pareto efficient outcomes compared to the no and full data sharing schemes, the welfare ranking of these data sharing schemes is far from clear. This is because of the nuanced externalities induced by the default or active data sharing choice of each consumer on other consumers.

3.1 The GDPR

We first analyze the equilibrium under the GDPR. Figure 2 illustrates the equilibrium. Because strong-willed consumers strictly benefit from having their data shared with seller $A$ and are unconcerned about seller $B$, all strong-willed consumers opt in to data sharing. Because type-O consumers are not interested in either good on 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, type-O consumers will opt out of data sharing, given that it would take 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 $B$. We conjecture that weak-willed consumers with sufficiently high temptation coefficient $\gamma_i$, that is, higher than a critical level $\gamma^{GDPR}_{**}$,
will choose to opt out, while those with \( \gamma_i \) lower than \( \gamma_{\text{GDPR}}^{**} \) will opt in.

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

\[
U_{W,\text{in}}^{\text{GDPR}} (\gamma_i) = \frac{z^{\text{GDPR}}_{A,\text{in}}}{\pi_S + \pi_W \gamma^{\text{GDPR}}_{**}} \int_{0}^{\bar{u}} \max \left\{ \tilde{u}_A - p^{\text{GDPR}}_{A,\text{in}}, 0 \right\} dH (\tilde{u}_A) + \frac{z^{\text{GDPR}}_{B,\text{in}} (d\gamma_i)}{\pi_W d\gamma_i} \left( u_B - p^{\text{GDPR}}_{B,\text{in}} (\gamma_i) \mathbf{1}_{\left\{ \gamma_i \geq \frac{p^{\text{GDPR}}_{B,\text{in}} (\gamma_i)}{\gamma_{**}} \right\}} - \gamma_i \mathbf{1}_{\left\{ \gamma_i < \frac{p^{\text{GDPR}}_{B,\text{in}} (\gamma_i)}{\gamma_{**}} \right\}} \right),
\]

where \( z^{\text{GDPR}}_{A,\text{in}} \) is the total advertising by seller A to the opt-in pool at price \( p^{\text{GDPR}}_{A,\text{in}} \), \( z^{\text{GDPR}}_{B,\text{in}} (d\gamma_i) \in [0, \pi_W] d\gamma_i \) is the advertising intensity of seller B to opt-in consumers with temptation coefficient \( \gamma_i \), and \( p^{\text{GDPR}}_{B,\text{in}} (\gamma_i) \) 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, \( z_{A,\text{in}}^{\text{GDPR}} \) and \( z_{B,\text{in}}^{\text{GDPR}} (d\gamma_i) \).

His utility from opt-out is

\[
U_{W,\text{out}}^{\text{GDPR}} (\gamma_i) = \frac{z^{\text{GDPR}}_{A,\text{out}}}{1 - \pi_S - \pi_W \gamma^{\text{GDPR}}_{**}} \int_{0}^{\bar{u}} \max \left\{ \tilde{u}_A - p^{\text{GDPR}}_{A,\text{out}}, 0 \right\} dH (\tilde{u}_A) + \frac{z^{\text{GDPR}}_{B,\text{out}}}{1 - \pi_S - \pi_W \gamma^{\text{GDPR}}_{**}} \left( u_B - p^{\text{GDPR}}_{B,\text{out}} \mathbf{1}_{\left\{ \gamma_i \geq \frac{p^{\text{GDPR}}_{B,\text{out}} (\gamma_i)}{\gamma_{**}} \right\}} - \gamma_i \mathbf{1}_{\left\{ \gamma_i < \frac{p^{\text{GDPR}}_{B,\text{out}} (\gamma_i)}{\gamma_{**}} \right\}} \right),
\]

where \( z^{\text{GDPR}}_{A,\text{out}} \) is the total advertising by seller A to the opt-out pool at price \( p^{\text{GDPR}}_{A,\text{out}} \), and \( z^{\text{GDPR}}_{B,\text{out}} \) is the total advertising by seller B to the opt-out pool at price \( p^{\text{GDPR}}_{B,\text{out}} \). For a weak-willed consumer to opt in for data sharing, it must be the case that

\[
U_{W,\text{in}}^{\text{GDPR}} (\gamma_i) \geq U_{W,\text{out}}^{\text{GDPR}} (\gamma_i).
\]

The following proposition characterizes the equilibrium under the GDPR.

**Proposition 6** Suppose \( \bar{u} < 8 (\bar{v} - u_B) \). 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 \( \gamma_i \leq \gamma_{\text{GDPR}}^{**} \) and opt out if \( \gamma_i > \gamma_{\text{GDPR}}^{**} \), where \( \gamma_{\text{GDPR}}^{**} \) is the unique root of Equation 22 in \((0, 1)\).

2. Seller A charges the opt-in and opt-out pools the same price: \( p_{A,\text{in}}^{\text{GDPR}} = p_{A,\text{out}}^{\text{GDPR}} = \frac{1}{2} \bar{u} \),

22
and adopts a water-filling advertising strategy with priority to the opt-in pool:

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

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

3. Seller $B$ also adopts a water-filling advertising strategy with priority to the opt-in pool by targeting the most-tempted consumers in the pool with $\gamma_i \in \left[ z_{A,in}^{GDPR}, \gamma^{GDPR}_{**} \right]$ 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), respectively.

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 $\gamma_i$) opt in, while those with severe temptation (high $\gamma_i$) opt out. This is intuitive because weak-willed customers with mild temptation benefit more from the better coverage from seller $A$ than the temptation they suffer from intensified exposure to seller $B$. In contrast, very 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 $B$.

Because seller $B$ can efficiently target weak-willed consumers in the opt-in pool, it gives higher priority to 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, will it also advertise to 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 to hide in the opt-out pool from seller $B$. This protection is weakened,\(^{16}\)

\(^{16}\) 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 $\bar{u}$. 

23
however, by the opt-in decisions of other consumers, that is, those who are strong-willed and modestly weak-willed. Their departure from the opt-out pool reduces the camouflage of those very weak-willed and increases $e_{GDPR}^{B_{out}} \cdot \pi_{W\,\pi_{W\,\pi_{W}}}$, 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 impact on the most vulnerable consumers. 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 vulnerable consumers to protect themselves.

There is evidence that consumers select into the opt-in and opt-out pools. De Matos and Adjerid (2021), for instance, find in a field experiment conducted by a large telecommunications provider that sales, marketing communication efficacy, and contractual lock-ins increased after new data authorizations because of better targeting of interested consumers. Aridor et al. (2020) show that although the volume of observable consumers dropped by 12.5% for firms in the online travel industry because of GDPR’s opt-in requirement, those that chose opt-in could be more persistently identified and efficiently targeted.

The default choice of the GDPR scheme plays an important role in protecting vulnerable consumers, not by affecting their own choices but through the choices of type-$O$ consumers. By guiding type-$O$ consumers to the opt-out pool, the default choice strengthens the camouflage of the vulnerable consumers in the opt-out pool. Thaler and Sunstein (2008) elegantly highlighted the importance of default options in influencing consumers’ choices, and thus their welfare, and the recent Stigler Committee Report (2019) has also argued the importance of the default data sharing option in protecting inattentive or biased consumers who may fail to make the optimal choice. Different from these arguments, our analysis highlights an important externality of the default choice on other consumers, specifically on the protection provided by the opt-out pool to vulnerable consumers.

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, as illustrated by Figure 3. 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 very weak-willed consumers from seller B.

Second, the opt-in of strong-willed and type-O consumers exposes all weak-willed consumers to seller B\textsuperscript{17}. Nevertheless, seller B still faces a challenge in sorting out their degrees of temptation because 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 $\gamma_i$ lower than a cutoff of $\gamma_{**}^{CCPA}$. This strategy bifurcates the pool of weak-willed consumers and segregates those exceptionally tempted above the cutoff in the opt-out pool. As a result, seller B can still

\textsuperscript{17}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.
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:

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

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 very 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 the more vulnerable 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 positive and negative externalities relative to the GDPR. From a positive perspective, seller $A$ is able to fully identify its intended consumers, including those weak-willed consumers in the opt-out pool. From a negative perspective,
seller B is also able to fully identify all weak-willed consumers in the opt-in and opt-out pools, albeit without knowing the degree of a consumer’s temptation 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 goods 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 \( \rho_A \) 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 \( \rho_B \)) 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 \( \rho_A, \rho_B, \) 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 weak-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 conventional 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 data sharing scheme relative to the other two 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 (i.e., \( u_B \) close to zero). In contrast, no data sharing is the most desirable scheme if the temptation problem is sufficiently severe (i.e., \( u_B \) is 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 and Implications

This paper adopts a new approach to analyze consumers’ preference for privacy based on a desire to protect themselves from their own vulnerabilities. This approach facilitates a welfare analysis of different data privacy regulations, such as the GDPR and the CCPA. Although sharing consumer data with a digital platform improves the matching efficiency of consumers
with normal consumption goods, it also exposes weak-willed consumers to the predatory advertising of temptation goods. Although 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 to exceptionally vulnerable consumers because of the data sharing externalities induced by the default and active data-sharing choice of each consumer on other consumers.

Our analysis also offers useful insights about several important issues:

**The Data Privacy Paradox**  
Our analysis provides a rationale for the data privacy paradox that although consumers express concerns about data privacy in surveys, they often appear to freely share their data with firms and digital platforms, as highlighted, for example, by Gross and Acquisti (2005), Goldfarb and Tucker (2012), Athey et al. (2017), Tang (2019), and Chen et al. (2021). The literature, as recently reviewed by Acquisti, Brandimarte, and Loewenstein (2020), tends to explain the data privacy paradox by certain consumer biases in making data sharing decisions, such as ignorance about the consequences of data sharing, present bias that causes consumers to overweight the immediate convenience from using digital services and underweight the future cost of sharing personal data, and an illusion of control in which consumers feel more in control than they actually are when making data-sharing choices. Acemoglu et al. (2019) posits that data-sharing externalities among consumers drives their reservation price for privacy to zero.

Different from these explanations, our analysis attributes the data privacy paradox to an intricate trade-off between cost and benefit that is strongly influenced by the data-sharing choices of other consumers. Despite intensified exposure to potential predatory advertising, a vulnerable consumer in our model may nonetheless opt in if the benefit of sharing his data with the platform (i.e., improved access to normal goods) outweighs this cost. A subtle feature of our model is that opting-out provides only limited protection because of the externalities induced by the data sharing decisions of other consumers. In particular, the protection provided by opting-out crucially depends on the composition of consumers in the opt-out pool, which, in turn, depends on the default choice and the structure of the data-sharing environment.

Chen et al. (2021) provide evidence consistent with this trade-off in data-sharing. Combining both survey and administrative data of a sample of Alipay users, they confirm the data privacy paradox by showing that Alipay users display no difference in the number of third-party mini-programs with which they authorize data-sharing, regardless of whether they are
concerned about their privacy. More important, they also show that privacy-concerned users tend to use their authorized mini-programs more frequently and extensively, revealing greater demand for digital services provided by the authorized mini-programs. Such findings suggest that the greater demand of privacy-concerned users for digital services may sufficiently offset their privacy concerns to explain the data privacy paradox.

Limits of Data Privacy Regulations  Our analysis reveals the limited capacity for data privacy regulations, such as the GDPR and the CCPA, to protect vulnerable consumers on digital platforms. These data privacy regulations are constrained for two key reasons. First, the bundling by the digital platform of data sharing with both normal and temptation goods sellers makes it costly for vulnerable consumers to opt out of data sharing because the benefits of sharing data with seller A can outweigh the costs of sharing data with seller B. Although we model the benefits of data sharing to be improved matching with seller A; in practice digital platforms also offer a variety of free services, such as email, messaging, and entertainment content that they bundle with data sharing to induce users to share their data. Furthermore, large platforms, such as Amazon and Facebook, also take advantage of network effects to make it particularly costly to opt-out or switch platforms. The ability to bundle data sharing authorizations also introduces increasing returns-to-scale to data for digital platforms as one consumer authorization is effectively shared with many sellers.

Second, the social nature of data sharing limits the protection afforded to vulnerable consumers who opt out. That is, data sharing by one consumer may also affect the welfare of other consumers because the shared data allow sellers to infer other consumers’ preferences and vulnerabilities. Such externalities 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 digital platforms. 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.

Even enforcement of data privacy regulation has its issues. Matte et al. (2020) and Nouwens et al. (2020), for instance, find significant violations of GDPR data authorization policy after its implementation, including fraudulent authorizations, implied consent, and dark patterns.
Given the limits of data privacy regulations in protecting vulnerable consumers, one may argue that there are more direct remedies for mitigating consumer harm, such as banning the sale of certain goods, giving consumers recourse when they are exploited, or promoting platform competition. In contrast to issues of price discrimination or fraud, it is difficult to measure the harm of consumer exploitation because one would have to prove that a consumer did or purchased something that they would not have otherwise. Calo (2013), for instance, emphasizes that current consumer protection legal system, which is mainly based on fraud and misrepresentation, is ill-equipped to address issues of exploitation of consumer vulnerability; as a result, such manipulation is rarely policed (Sunstein (2015)). In addition, increased competition may not help protect consumers because it can push online platforms to make their content even more addictive (e.g., Stigler Committee (2019), Ichihashi and Kim (2021)), and to engage in more exploitative practices to compete for revenue. As such, despite its limits, protecting data privacy ex ante may be the most effective way to protect consumers on online platforms ex post.  

**The Default Choice**  

Our analysis also shows that the default choice in a data privacy regulation can have substantial effects on equilibrium outcomes. Such a distinction is relevant as European data privacy laws (e.g., GDPR) set opt-out as the default choice while U.S. laws (e.g., CCPA, VCDPA, CPA) set opt-in as the default. As part of the consumer population may be indifferent toward sharing their data, the default choice makes the data of these indifferent consumers automatically available or unavailable to digital platforms. This, in turn, affects other consumers because of data sharing externalities. It is important to note that this is a different channel from that advocated by Thaler and Sunstein (2008) to improve the welfare of consumers by using default choices to correct their biases.

Instead, the default choice in our setting may induce both positive and negative externalities on other consumers. 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 in our model is in an intermediate range, while the CCPA

---

19 Another emerging trend for digital platforms to commit to consumer protection is to decentralize through tokenization. Sockin and Xiong (2021) argue that a platform may use blockchain technology to give control back to users through the sale of utility tokens. This commitment not to exploit users, however, comes at the cost of not having owners to subsidize user participation to maximize the platform’s network effect.
is most desirable when the temptation problem is sufficiently modest. The result that the CCPA may dominate the GDPR is surprising given that the GDPR is commonly regarded to be more protective of consumers. Our analysis consequently highlights that how privacy regulation is framed can have unintended consequences.

References


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


Chen, Long, Yadong Huang, Shumiao Ouyang, and Wei Xiong (2021), Data Privacy Paradox and Digital Demand, mimeo, Princeton University.

Cong Will, Danxia Xie, and Longtian Zhang (2020), Knowledge Accumulation, Privacy, and Growth in a Data Economy, Management Science, forthcoming.


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

Fainmesser, Itay, Andrea Galeotti, and Ruslan Momot (2021), Digital Privacy, mimeo.


He, Zhiguo, Jing Huang, and Jidong Zhou (2021), Open Banking: Credit Market Competition When Borrowers Own the Data, NBER Working Paper #28118.

Ichihashi, Shota, and Byung-Cheol Kim (2021), Addictive Platforms, SSRN 3853961.


Matte, Celestin, Nataliia Bielova and Cristiana Santos (2020), Do Cookie Banners Respect my Choice?


Sockin, Michael and Wei Xiong (2021), Decentralization through Tokenization, mimeo, UT Austin and Princeton University.


Stigler Committee (2019), Stigler Committee Report on Digital Platforms, University of Chicago.


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


A Proofs of Propositions

A.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:

\[
U_S (\{A,\emptyset\}) = \max \{\tilde{u}_A - p_A, 0\},
U_S (\{B,\emptyset\}) = 0.
\]

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:

\[
U_W (\{A,\emptyset\}) = \max \{\tilde{u}_A - p_A, 0\},
U_W (\{B,\emptyset\}) = u_B + \max \{-p_B, -\gamma_i \bar{v}\}.
\]

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 \(\gamma_i > \frac{p_B}{\bar{v}}\).

A.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

\[
Q_A^{NS} = (\pi_S + \pi_W) z_A^{NS} (1 - H (p_A^{NS} / \bar{u})) ,
\]
and consequently the seller’s profit net of the advertisement cost is

\[
\Pi_A^{NS} = p_A^{NS} (\pi_S + \pi_W) z_A^{NS} (1 - H (p_A^{NS} / \bar{u})) - F \frac{z_A^{NS}}{1 - z_A^{NS}} .
\]

Similarly, the quantity of goods sold by seller B is

\[
Q_B^{NS} = \pi_W z_B^{NS} (1 - G (p_B^{NS} / \bar{v})) ,
\]
and the net profit of seller B is

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

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

\[
p_A^{NS} Q_A^{NS} = F \frac{z_A^{NS}}{(1 - z_A^{NS})^2}.
\]
Then, we have that
\[ \Pi_A^{NS} = p_A^{NS} Q_A^{NS} - F \frac{z_A^{NS}}{1 - z_A^{NS}} = F \left( \frac{z_A^{NS}}{(1 - z_A^{NS})^2} - \frac{z_A^{NS}}{1 - z_A^{NS}} \right) = F \left( \frac{z_A^{NS}}{1 - z_A^{NS}} \right)^2. \]

Similarly, the first-order condition with respect to \( z_B^{NS} \) is
\[ p_B^{NS} Q_B^{NS} = F \frac{z_B^{NS}}{(1 - z_B^{NS})^2}, \]
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
\[
\begin{align*}
Q_A^{NS} &= \frac{p_A^{NS}}{\bar{u}} (\pi_s + \pi_W) z_A^{NS} 1\{0 \leq p_A^{NS} \leq \bar{u}\}, \quad (12) \\
Q_B^{NS} &= \frac{p_B^{NS}}{\bar{v}} \pi_W z_B^{NS} 1\{0 \leq p_B^{NS} \leq \bar{v}\}. \quad (13)
\end{align*}
\]

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
\[
\begin{align*}
\frac{d^2 \Pi_A^{NS}}{d \left( p_A^{NS} \right)^2} &= -\frac{2}{\bar{u}} \left( \pi_s + \pi_W \right) z_A^{NS} h \left( p_A^{NS} / \bar{u} \right) 1\{0 \leq p_A^{NS} \leq \bar{u}\} \leq 0, \\
\frac{d^2 \Pi_B^{NS}}{d \left( p_B^{NS} \right)^2} &= -2 \pi_W z_B^{NS} g \left( \gamma_s^{NS} \right) \frac{1}{\bar{v}} 1\{0 \leq p_B^{NS} \leq \bar{v}\} \leq 0.
\end{align*}
\]

It follows that optimal prices will always be nonnegative. Since
\[
\frac{d^2 \Pi_A^{NS}}{d \left( z_A^{NS} \right)^2} = -2 \frac{F}{(1 - z_A^{NS})^3} < 0,
\]
and \( \frac{d^2 \Pi_A^{NS}}{d p_A^{NS} d z_A^{NS}} = 0 \), it follows that the Hessian for seller A's optimization with respect to \( (p_A^{NS}, z_A^{NS}) \) 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 \( z_{NS}^A \) and \( z_{NS}^B \), it then follows that \( z_{NS}^A \) and \( z_{NS}^B \) satisfy

\[
\frac{\pi_S + \pi_W}{4F} \bar{u} = \frac{1}{(1 - z_{NS}^A)^2}, \quad \frac{\pi_W}{4F} \bar{v} = \frac{1}{(1 - z_{NS}^B)^2}.
\]

Then, we have

\[z_{NS}^A = 1 - \sqrt{\frac{1}{\pi_S + \pi_W} \frac{4F}{\bar{u}}}, \quad \text{and} \quad z_{NS}^B = 1 - \sqrt{\frac{1}{\pi_W} \frac{4F}{\bar{v}}}.\]

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

### A.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 \( z_{FS}^A \) as the measure of strong-willed and weak-willed consumers, to which seller A advertises, and \( p_{FS}^A \) as the price the seller sets.

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

\[Q_{FS}^A = z_{FS}^A \left[ 1 - H \left( p_{FS}^A/\bar{u} \right) \right],\]

and the net profit of seller A is

\[\Pi_{FS}^A = p_{FS}^A z_{FS}^A \left[ 1 - H \left( p_{FS}^A/\bar{u} \right) \right] - F \frac{z_{FS}^A}{1 - z_{FS}^A}.\]

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

\[z_{FS}^A = 1 - 2 \sqrt{\frac{F}{\bar{u}}}.\]

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, \( z_{FS}^A = 0 \). Furthermore, if \( 1 - 2 \sqrt{\frac{F}{\bar{u}}} > \pi_S + \pi_W \), then \( z_{FS}^A = \pi_S + \pi_W \).

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 \( \gamma_i \). Since consumers with stronger temptation are willing to pay higher prices, seller B optimally prioritizes strong temptation consumers:

\[
d_{z_B^{FS}} (\gamma_i) = \begin{cases} 
0, & \text{if } \gamma_i < \hat{\gamma}_i^{FS} \\
\pi_W d\gamma_i, & \text{if } \gamma_i \in (\hat{\gamma}_i^{FS}, 1] 
\end{cases}
\]

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 \( \int_0^{1 \gamma_i \hat{z}_B^{FS}} (d\gamma_i) \) is understood as a Riemann-Stieltjes integral.

Note that the expected revenue of seller B reduces to

\[
\bar{v} \int_{z_B^{FS}}^{1 \gamma_i \hat{z}_B^{FS}} \pi_W \gamma_i \, d\gamma_i = \bar{v} \pi_W \frac{1 - (z_B^{FS})^2}{2},
\]

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

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

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

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

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

\[
z_B^{FS} = \frac{2 + \pi_W}{3} - 3 \sqrt{\left( \frac{1 - \pi_W}{3} \right)^3 + \frac{\pi_W F}{2 \bar{v}}} + \sqrt{\left( \frac{1 - \pi_W}{3} \right)^3 + \frac{\pi_W F}{2 \bar{v}}} - \frac{1 - \pi_W}{3}^6
\]

\[
- 3 \sqrt{\left( \frac{1 - \pi_W}{3} \right)^3 + \frac{\pi_W F}{2 \bar{v}}} - \sqrt{\left( \frac{1 - \pi_W}{3} \right)^3 + \frac{\pi_W F}{2 \bar{v}}} - \frac{1 - \pi_W}{3}^6.
\]

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

\[
(\pi_W - z_B^{FS}) (1 - z_B^{FS})^2 = \frac{\pi_W F}{\bar{v}},
\]

(14)
it follows that
\[ (\pi_W - z_{FS}^B)^3 \leq \frac{\pi_W F}{v}, \]
and consequently
\[ \frac{z_{FS}^B}{\pi_W} \geq 1 - \frac{3}{\sqrt{2}} \frac{F}{\pi_W^2 v}. \]

Finally, applying the Implicit Function Theorem to Equation (14), one also has that
\[
\begin{align*}
\frac{dz_{FS}^B}{d\bar{v}} &= \frac{\pi_W F}{2 \pi_W} \frac{(1 - z_{FS}^B)^2 + 2 (\pi_W - z_{FS}^B)}{(1 - z_{FS}^B)^2 + 2 (\pi_W - z_{FS}^B)} \geq 0, \\
\frac{dz_{FS}^B}{dF} &= -\frac{\pi_W}{2 \pi_W} \frac{(1 - z_{FS}^B)^2 + 2 (\pi_W - z_{FS}^B)}{(1 - z_{FS}^B)^2 + 2 (\pi_W - z_{FS}^B)} \leq 0.
\end{align*}
\]

### A.4 Proof of Proposition 5

It is easy to verify that \( z_{FS}^A \geq z_{NS}^A \). Without data sharing, the probability of a strong-willed or weak-willed consumer being covered by seller A is \( z_{NS}^A \); with data sharing, the probability is \( \frac{z_{FS}^A}{\pi_S + \pi_W} \). As \( z_{FS}^A \geq z_{NS}^A \) and \( \pi_S + \pi_W \leq 1 \), it follows that \( \frac{z_{FS}^A}{\pi_S + \pi_W} \geq z_{NS}^A \), and the inequality is strict if \( z_{FS}^A > 0 \). 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 \( p_{NS}^A = p_{FS}^A = \bar{u}/2 \) for its good. From Equation (15), the social welfare under no data sharing is given by
\[
W_{NS} = (\pi_S + \pi_W) z_{NS}^A \int_{p_A}^{\bar{u}} u_A \frac{du_A}{\bar{u}} + \pi_W z_{NS}^B u_B \int_{p_B^{FS}}^{p_B^{NS}} d\gamma_i + \pi_W z_{NS}^B u_B \int_0^1 (u_B - \gamma_i \bar{v}) \ d\gamma_i.
\]

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
\[
W_{FS} = \frac{3}{8} \bar{u} z_{FS}^A + \pi_W u_B \int_{\gamma_i}^{1} d\gamma_i = \frac{3}{8} \bar{u} z_{FS}^A + z_{FS}^B u_B.
\]

It then follows that
\[
W_{FS} - W_{NS} = \pi_W \left( \frac{z_{FS}^B}{\pi_W} - z_{NS}^B \right) u_B + \frac{1}{8} \pi_W z_{NS}^B \bar{v} + \frac{3}{8} \bar{u} (\pi_S + \pi_W) \left( \frac{z_{FS}^A}{\pi_S + \pi_W} - z_{NS}^A \right) < 0,
\]
if \( u_B < u_{B**} \) where
\[
u_{B**} = -\frac{3\bar{u} \pi_S + \pi_W}{8 \left( \frac{z_{FS}^A}{\pi_S + \pi_W} - z_{NS}^A \right)} \frac{(\pi_S + \pi_W) \left( \frac{z_{FS}^A}{\pi_S + \pi_W} - z_{NS}^A \right) + \bar{u} z_{NS}^B}{\pi_W (z_{FS}^B - z_{NS}^B)}.
\]

That is, social welfare is lower with full data sharing than with no data sharing.
A.5 Proof of Proposition 6

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

$$\Pi_A = \frac{(1 - \gamma^{GDPR}_{**}) \pi_W}{\gamma^{GDPR}_{**}} \pi_W + 1 - \pi_S - \pi_W \left( \frac{z^{GDPR}_{A,out}}{\pi_W} \right) \left( 1 - \frac{p^{GDPR}_{A,in}}{\pi_W} \right) + p^{GDPR}_{A,in} \frac{z^{GDPR}_{A,in}}{\pi_W} \left( 1 - \frac{z^{GDPR}_{A,in}}{\pi_W} \right) - F \frac{z^{GDPR}_{A,out} + z^{GDPR}_{A,in}}{1 - z^{GDPR}_{A,out} - z^{GDPR}_{A,in}},$$

where $z^{GDPR}_{A,out} \in [0, 1 - \pi_S - \gamma^{GDPR}_{**} \pi_W]$ and $z^{GDPR}_{A,in} \in [0, \pi_S + \gamma^{GDPR}_{**} \pi_W]$. 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 $\frac{(1 - \gamma^{GDPR}_{**}) \pi_W}{\pi_W + 1 - \pi_S - \pi_W}$.

If $z^{GDPR}_{A,in} > 0$ and $z^{GDPR}_{A,out} > 0$, the FOCs for $p^{GDPR}_{A,in}$ and $p^{GDPR}_{A,out}$ reveal that

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

Then, the seller’s profit becomes

$$\Pi_A = \frac{(1 - \gamma^{GDPR}_{**}) \pi_W}{1 - \pi_S - \gamma^{GDPR}_{**} \pi_W} \bar{u} z^{GDPR}_{A,out} + \bar{u} z^{GDPR}_{A,in} - F \frac{z^{GDPR}_{A,out} + z^{GDPR}_{A,in}}{1 - z^{GDPR}_{A,out} - z^{GDPR}_{A,in}}.$$

The marginal profit from $z^{GDPR}_{A,in}$ is strictly higher than that from $z^{GDPR}_{A,out}$, 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 $z^{GDPR}_{A,in}$ gives

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

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

$$z^{GDPR}_{A,in} = \min \left\{ 1 - 2 \sqrt{\frac{F}{\bar{u}}}, \pi_S + \gamma^{GDPR}_{**} \pi_W \right\}.$$  \hspace{1cm} (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

\[
\begin{align*}
    z_{A,out}^{GDPR} &= \min\{\max\{1 - 2 \left(1 - \frac{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W}{1 - \gamma_{**}^{GDPR}}\right) \frac{F}{\pi_W} - \pi_S - \gamma_{**}^{GDPR} \pi_W, 0\}, 1 - \pi_S - \gamma_{**}^{GDPR} \pi_W\}.
\end{align*}
\]

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 \( z_{B,in}^{GDPR} (\gamma_i) \) 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 \( \gamma_i \) is above \( p_{B,out}^{GDPR} / \bar{v} \). Thus, the seller’s profit is

\[
    \Pi_B = -F \frac{z_{B,out}^{GDPR} + z_{B,in}^{GDPR}}{1 - z_{B,out}^{GDPR} - z_{B,in}^{GDPR}} + \bar{v} \int_0^{\gamma_{**}^{GDPR}} \gamma_i z_{B,in}^{GDPR} (d\gamma_i)
    \]

\[
    + \frac{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W}{z_{B,out}^{GDPR} p_{B,out}^{GDPR}} \left[(1 - p_{B,out}^{GDPR} / \bar{v}) 1_{\{p_{B,out}^{GDPR} = \gamma_{**}^{GDPR} \}} + (1 - \gamma_{**}^{GDPR}) 1_{\{p_{B,out}^{GDPR} < \gamma_{**}^{GDPR}\}}\right],
\]

where \( z_{B,out}^{GDPR} \in [0, 1 - \pi_S - \gamma_{**}^{GDPR} \pi_W] \) and \( z_{B,in}^{GDPR} = \int_0^{\gamma_{**}^{GDPR}} z_{B,in}^{GDPR} (d\gamma_i) \in [0, \gamma_{**}^{GDPR} \pi_W] \) 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:

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

Thus, the optimal price satisfies

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

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 \( \gamma_i \):

\[
    d\gamma_{B,in}^{GDPR} (\gamma_i) = \\begin{cases} \\
    0 & \text{if } \gamma_i < \gamma_{**}^{GDPR} \\
    \pi_W d\gamma_i & \text{if } \gamma_i \in (\gamma_{**}^{GDPR}, \gamma_{**}^{GDPR})
\end{cases}
\]

Therefore, the expected revenue of seller \( B \) from the opt-in pool reduces to \( \bar{v} \int \gamma_{**}^{GDPR} \pi_W \gamma_i d\gamma_i = \bar{v} \pi_W \frac{(\gamma_{**}^{GDPR})^2 - (\gamma_{**}^{GDPR})^2}{2} \). As \( \gamma_{**}^{GDPR} = \gamma_{**}^{GDPR} - \frac{z_{B,in}^{GDPR} \pi_W}{p_{B,in}^{GDPR}} \), 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 \( z_{GDPR}^{B,in} \): 

\[
\Pi_B = -F \left( \frac{z_{GDPR}^{B,in} + z_{GDPR}^{B,out}}{1 - z_{GDPR}^{B,out} - z_{GDPR}^{B,in}} \right) + z_{GDPR}^{B,in} \left( \gamma_{GDPR}^{**} - \frac{1}{2} \frac{z_{GDPR}^{B,in}}{\pi_W} \right) \bar{v} 
\]

and the seller’s choice reduces to choosing \( z_{GDPR}^{B,in} \) and \( z_{GDPR}^{B,out} \):

The revenue from the opt-in pool is concave in the total advertising to the opt-in pool, \( z_{GDPR}^{B,in} \), 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 \( z_{GDPR}^{B,out} \).

The first-order condition for \( z_{GDPR}^{B,in} \) is

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

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

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

Which pool has priority depends on which has higher marginal revenue. The marginal revenue from the opt-in pool when \( z_{GDPR}^{B,in} = 0 \) 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

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

which is equivalent to

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

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

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:

$$z_{B, out}^{GDPR} = \min \left\{ \max \left\{ 1 - \frac{F}{\pi_W \bar{u}} \frac{1 - \pi_S - \gamma_{GDPR}^{**} \pi_W}{1 - \frac{1}{2} \left\{ \gamma_{GDPR}^{**} \right\}^2 \{ \gamma_{GDPR}^{**} > \frac{1}{2} \}}, 0 \right\}, 1 - \pi_S - \gamma_{GDPR}^{**} \pi_W \right\},$$

$$z_{B, in}^{GDPR} = \min \left\{ \max \left\{ 1 - \frac{F}{\pi_W \bar{u}} \frac{1 - \pi_S - \gamma_{GDPR}^{**} \pi_W}{1 - \frac{1}{2} \left\{ \gamma_{GDPR}^{**} \right\}^2 \{ \gamma_{GDPR}^{**} > \frac{1}{2} \}} - z_{B, out}^{GDPR}, 0 \right\}, \pi_W \gamma_{GDPR}^{**} \right\}. $$

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

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

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 $z_{**}$:

$$z_{**} = 1 - \frac{1}{2} \left( \frac{\pi_W F}{2 \bar{u}} \right)^2 - \frac{1 - \pi_S - \gamma_{GDPR}^{**} \pi_W}{27} - \frac{1}{2} \left( \frac{\pi_W F}{2 \bar{u}} \right)^2 - \frac{1 - \pi_S - \gamma_{GDPR}^{**} \pi_W}{27}. $$

Consequently, it follows that

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

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

$$z_{B, out}^{GDPR} = \min \left\{ \max \left\{ 1 - \frac{F}{\pi_W \bar{u}} \frac{1 - \pi_S - \gamma_{GDPR}^{**} \pi_W}{1 - \frac{1}{2} \left\{ \gamma_{GDPR}^{**} \right\}^2 \{ \gamma_{GDPR}^{**} > \frac{1}{2} \}} - z_*, 0 \right\}, 1 - \pi_S - \gamma_{GDPR}^{**} \pi_W \right\}. \tag{20}$$
If \( z_s \geq z_{**} \), seller \( B \) does not target the opt-out pool.

Taken together, the level of \( \gamma^{GDPR}_{**} \) 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 \( \gamma^{GDPR}_{**} \).

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

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

\[
U^{GDPR}_{W,in} (\gamma_i) = \frac{z^{GDPR}_{A,in}}{\pi_S + \gamma^{GDPR}_{**} \pi_W} \bar{u} + \frac{z^{GDPR}_{B,in} (\gamma_i)}{\pi_W} (u_B - \gamma_i \bar{v}).
\]

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

\[
U^{GDPR}_{W,out} (\gamma_i) = \frac{z^{GDPR}_{A,out}}{1 - \pi_S - \gamma^{GDPR}_{**} \pi_W} \bar{u} + \frac{z^{GDPR}_{B,out} (\gamma_i)}{1 - \pi_S - \gamma^{GDPR}_{**} \pi_W} u_B - \frac{z^{GDPR}_{B,out}}{1 - \pi_S - \gamma^{GDPR}_{**} \pi_W} \bar{v} \max \left\{ \frac{1}{2}, \gamma^{GDPR}_{**} \right\} \left\{ \gamma_i \geq \max \left\{ \frac{1}{2}, \gamma^{GDPR}_{**} \right\} \right\} + \gamma_i \max \left\{ \frac{1}{2}, \gamma^{GDPR}_{**} \right\} \left\{ \gamma_i \leq \max \left\{ \frac{1}{2}, \gamma^{GDPR}_{**} \right\} \right\},
\]

which increases with \( z^{GDPR}_{A,out} \) and decreases with \( z^{GDPR}_{B,out} \). Then,

\[
U^{GDPR}_{W,in} (\gamma_i) - U^{GDPR}_{W,out} (\gamma_i) = \frac{\bar{u}}{8} \left[ \frac{z^{GDPR}_{A,in}}{\pi_S + \gamma^{GDPR}_{**} \pi_W} - \frac{z^{GDPR}_{A,out}}{1 - \pi_S - \gamma^{GDPR}_{**} \pi_W} \right] + \left( \frac{z^{GDPR}_{B,in} (\gamma_i)}{\pi_W} - \frac{z^{GDPR}_{B,out}}{1 - \pi_S - \gamma^{GDPR}_{**} \pi_W} \right) u_B + \bar{v} \left[ \frac{z^{GDPR}_{B,out}}{1 - \pi_S - \gamma^{GDPR}_{**} \pi_W} \left\{ \gamma_i \geq \max \left\{ \frac{1}{2}, \gamma^{GDPR}_{**} \right\} \right\} + \max \left\{ \frac{1}{2}, \gamma^{GDPR}_{**} \right\} \left\{ \gamma_i \leq \max \left\{ \frac{1}{2}, \gamma^{GDPR}_{**} \right\} \right\} - \frac{z^{GDPR}_{B,out} (\gamma_i)}{\pi_W} \gamma_i \right].
\]

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

Note that if \( z^{GDPR}_{B,in} = 0 \), then \( U^{GDPR}_{W,in} (\gamma_i) > U^{GDPR}_{W,out} (\gamma_i) \) for all \( \gamma_i \), and \( \gamma^{GDPR}_{**} = 1 \). Consequently, it must be the case that the opt-in pool has priority, or \( \gamma^{GDPR}_{**} \) 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 \( z^{GDPR}_{B,in} > 0 \).

Since the marginal consumer must be indifferent to opt-in and opt-out, \( U^{GDPR}_{W,in} (\gamma^{GDPR}_{**}) - U^{GDPR}_{W,out} (\gamma^{GDPR}_{**}) = 0 \), which imposes the following condition on \( \gamma^{GDPR}_{**} \):

\[
C (\gamma^{GDPR}_{**}) = \begin{cases} 
0 & \text{if } \gamma^{GDPR}_{**} = 0 \\
0 & \text{if } \gamma^{GDPR}_{**} \in (0, 1) \\
> 0 & \text{if } \gamma^{GDPR}_{**} = 1
\end{cases}.
\]
where, with some manipulation of Equation (21),

$$C(\gamma^{\text{GDPR}}_{**}) = (\pi_S + \gamma^{\text{GDPR}}_{**} \pi_W)^3 \frac{z_{B,\text{in}}^{\text{GDPR}}(\gamma^{\text{GDPR}}_{**})}{\pi_W} \cdot \left(1 + \pi_S + \frac{\pi_W u_B}{\bar{v}}\right) \frac{z_{B,\text{in}}^{\text{GDPR}}(\gamma^{\text{GDPR}}_{**})}{\pi_W} - z_{B,\text{out}}^{\text{GDPR}} - \left(\pi_S + \gamma^{\text{GDPR}}_{**} \pi_W\right)^2$$

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 $\gamma^{\text{GDPR}}_{**} \in (0, 1)$, into Equation (21) to obtain

$$U_{W,\text{in}}^{\text{GDPR}}(\gamma_i) - U_{W,\text{out}}^{\text{GDPR}}(\gamma_i) = \left(\frac{z_{B,\text{in}}^{\text{GDPR}}(\gamma_i)}{\pi_W} - 1\right) u_B + \bar{v} \left(\gamma_{GDPR} - \frac{z_{B,\text{in}}^{\text{GDPR}}(\gamma_i)}{\pi_W}\gamma_i\right) + \bar{v} \frac{z_{B,\text{out}}^{\text{GDPR}}}{1 - \pi_S - \gamma^{\text{GDPR}}_{**} \pi_W} \left(\gamma_i \mathds{1}_{\gamma_i \leq \max\{\frac{1}{2}, \gamma^{\text{GDPR}}_{**}\}} + \max\left\{\frac{1}{2}, \gamma^{\text{GDPR}}_{**}\right\}\mathds{1}_{\gamma_i > \max\{\frac{1}{2}, \gamma^{\text{GDPR}}_{**}\}}\right) - \gamma^{\text{GDPR}}_{**}.$$  

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

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

If instead $\gamma^{\text{GDPR}}_{**} < \frac{1}{2}$ and $\gamma_i \leq \frac{1}{2}$, then

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

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

$$U_{W,\text{in}}^{\text{GDPR}}(\gamma_i) - U_{W,\text{out}}^{\text{GDPR}}(\gamma_i) = \left(1 - \frac{z_{B,\text{out}}^{\text{GDPR}}}{1 - \pi_S - \gamma^{\text{GDPR}}_{**} \pi_W}\right) \bar{v} \left(\gamma^{\text{GDPR}}_{**} - \gamma_i\right) + \bar{v} \frac{z_{B,\text{out}}^{\text{GDPR}}}{1 - \pi_S - \gamma^{\text{GDPR}}_{**} \pi_W} \left(\frac{1}{2} - \gamma_i\right) < 0,$$

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

We now consider $\gamma_i < \gamma^{\text{GDPR}}_{**}$. Note that for $\gamma_i < \gamma^{\text{GDPR}}_{**}$, the threshold $\gamma_i$ below which seller $B$ does not advertise ($z_{B,\text{in}}^{\text{GDPR}}(\gamma_i) = 0$ for $\gamma_i < \gamma^{\text{GDPR}}_{**}$), it is trivial that $U_{W,\text{in}}^{\text{GDPR}}(\gamma_i) -$
$U^\text{GDPR}_{W,\text{in}}(\gamma_i) > 0$, since the consumer benefits from both higher advertising by seller $A$ and lower (zero) advertising by seller $B$. Consequently, all weak-willed consumers with $\gamma_i < \gamma^\text{GDPR}_{**}$ opt in. For $\gamma_i \in [\hat{\gamma}^\text{GDPR}, \gamma^\text{GDPR}_{**}]$, $z^\text{GDPR}_{B,\text{in}}(\gamma_i) = \pi_W$, and their opt-in / opt-out condition reduces to

$$U^\text{GDPR}_{W,\text{in}}(\gamma_i) - U^\text{GDPR}_{W,\text{out}}(\gamma_i) = \left(1 - \frac{z^\text{GDPR}_{B,\text{out}}}{1 - \pi_S - \gamma^\text{GDPR}_{**} \pi_W}\right) \bar{v} \left(\gamma^\text{GDPR}_{**} - \gamma_i\right) > 0,$$

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

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

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

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

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

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

At the other end, note that $C (0) = \frac{\pi_W u_B}{\bar{v}} \left(\frac{z^\text{GDPR}_{B,\text{in}} (0)}{\pi_W} \left(1 - \pi_S\right) - \frac{z^\text{GDPR}_{B,\text{out}}}{\pi_S + \pi_W} \right) \pi_S + \frac{\pi_W \bar{v}}{8 \bar{v}} \pi_S \left(1 - \pi_S\right) \left(\frac{z^\text{GDPR}_{A,\text{in}}}{\pi_S} - \frac{z^\text{GDPR}_{A,\text{out}}}{1 - \pi_S}\right) > 0,$$since $\frac{z^\text{GDPR}_{A,\text{in}}}{\pi_S} - \frac{z^\text{GDPR}_{A,\text{out}}}{1 - \pi_S} > 0$ because the opt-in pool has higher advertising efficiency for seller $A$, $u_B < 0$, and $\frac{z^\text{GDPR}_{A,\text{in}} (0)}{\pi_W} = 0$ because only the most mildly tempted opt in (condition (18) fails). Consequently, $C (\pi_S) > 0$ and therefore $\gamma^\text{GDPR}_{**} > 0$.

Notice now that all advertising policies, $z^\text{GDPR}_{A,\text{in}}$, $z^\text{GDPR}_{A,\text{out}}$, $z^\text{GDPR}_{B,\text{in}}$, and $z^\text{GDPR}_{B,\text{out}}$ are (piecewise) continuous in $\gamma^\text{GDPR}_{**}$ when $\gamma^\text{GDPR}_{**}$ has an interior solution. As such, $C (\gamma^\text{GDPR}_{**})$ is continuous. Since $C (0) > 0$ while $C (1) < 0$, by the Intermediate Value Theorem, there exists a $\gamma^\text{GDPR}_{**} \in (0, 1)$. Given the nonlinearity of $C (\gamma^\text{GDPR}_{**})$, however, there may be multiple values in $(0, 1)$ with $C (\gamma^\text{GDPR}_{**}) = 0$, and consequently multiple equilibria.
We further establish that there exists an equilibrium in which \( \gamma^{\text{GDPR}} \geq \gamma_+ = \frac{1 - \pi_s}{\pi_W} - \sqrt{\left(\frac{1 - \pi_s}{\pi_W}\right)^2 - \frac{1 - \pi_s}{\pi_W}} > \frac{1}{2} \). Direct manipulation of the definition of \( C(\cdot) \) gives that

\[
C(\gamma_+) = \left(1 - \pi_s\right) \left(1 - \sqrt{1 - \frac{\pi_W}{1 - \pi_s}} - \frac{\pi_W u_B}{\bar{u}}\right) \tilde{\gamma}_+ \left(1 - \tilde{\gamma}_+\right) \left(1 - \frac{1}{1 - \pi_s - \gamma_+ \pi_W}\right)
\]

\[
+ \frac{\pi_W \bar{u}}{8\bar{u}} \tilde{\gamma}_+ \left(1 - \tilde{\gamma}_+\right) \left(\frac{z_{A;\text{in}}^{\text{GDPR}}}{\pi_S + \gamma_+ \pi_W} - \frac{z_{A;\text{out}}^{\text{GDPR}}}{1 - \pi_S - \gamma_+ \pi_W}\right) > 0,
\]

where \( \tilde{\gamma}_+ = \pi_S + \gamma_+ \pi_W \). Note that \( u_B < 0 \), and therefore \( \left(1 - \pi_s\right) \left(1 - \sqrt{1 - \frac{\pi_W}{1 - \pi_s}} - \frac{\pi_W u_B}{\bar{u}}\right) \tilde{\gamma}_+ \left(1 - \tilde{\gamma}_+\right) \left(1 - \frac{1}{1 - \pi_s - \gamma_+ \pi_W}\right) \\
+ \frac{\pi_W \bar{u}}{8\bar{u}} \tilde{\gamma}_+ \left(1 - \tilde{\gamma}_+\right) \left(\frac{z_{A;\text{in}}^{\text{GDPR}}}{\pi_S + \gamma_+ \pi_W} - \frac{z_{A;\text{out}}^{\text{GDPR}}}{1 - \pi_S - \gamma_+ \pi_W}\right) > 0 \\
\]

Uniqueness: We finally provide a sufficient condition under which an equilibrium with \( \gamma^{\text{GDPR}} \geq \gamma_+ \) is the unique equilibrium. Since \( \gamma^{\text{GDPR}} \geq \frac{1}{2} \), 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 \( z_{A;\text{out}}^{\text{GDPR}} = 0 \). Intuitively, it is too costly for seller \( A \) to advertise to more than \( \pi_S + \frac{1}{2} \pi_W \leq \pi_S + \gamma^{\text{GDPR}} \pi_W \) consumers. Given this sufficient condition, notice that \( z_{A;\text{in}}^{\text{GDPR}} \) from Equation (15) is also insensitive to \( \gamma^{\text{GDPR}} \).

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

\[
\frac{d\gamma^{\text{GDPR}}}{dz_{B;\text{out}}^{\text{GDPR}}} = -\frac{\pi_W \left(\gamma^{\text{GDPR}} - \frac{u_B}{\bar{u}}\right) \left(\pi_S + \gamma^{\text{GDPR}} \pi_W\right)}{C' \left(\gamma^{\text{GDPR}}\right)}.
\]

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

\[
\frac{d\gamma^{\text{GDPR}}}{dz_{B;\text{out}}^{\text{GDPR}}} > 0.
\]
Consequently, we have established that the solution for $\gamma_{**}^{GDPR} \in (\gamma_+, 1)$ 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),

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

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

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

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

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

A.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 $z_{A}^{CCPA} = \frac{\pi_S}{\pi_S + \pi_W}$.

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 $p_{B, in}^{CCPA} (\gamma_i)$ and an advertising policy function $z_{B, in}^{CCPA} (\gamma_i)$ 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 $p_{B, in}^{CCPA} (\gamma_i) = \gamma_i \bar{u}$. 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 $\gamma_{**}^{CCPA}$ 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:

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

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 $\gamma_i$ 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 $-\frac{z_{B,\text{out}}^{\text{CCPA}}}{(1-\gamma_{**}^{\text{CCPA}}) \pi_W} \min \left\{ p_{B,\text{out}}^{\text{CCPA}}, \gamma_i \bar{v} \right\}$. Since $\frac{z_{B,\text{out}}^{\text{CCPA}}}{(1-\gamma_{**}^{\text{CCPA}}) \pi_W} \leq 1$, 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 $\gamma_i = 0$, as he may be targeted by seller $B$. 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, $z_{B,\text{in}}^{\text{CCPA}}(\gamma_i) = 0$, 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 $\gamma_{**}^{\text{CCPA}}$, consider the profit for seller $B$, $\Pi_B$. As the profit from the opt-in pool under this strategy is zero, we have

$$\Pi_B = p_{B,\text{out}}^{\text{CCPA}} \frac{z_{B,\text{out}}^{\text{CCPA}}}{(1-\gamma_{**}^{\text{CCPA}}) \pi_W} \int_{\gamma_{**}^{\text{CCPA}}}^{1} 1_{\left\{ \gamma_i \geq p_{B,\text{out}}^{\text{CCPA}} / \bar{v} \right\}} d\gamma_i - F \frac{z_{B,\text{out}}^{\text{CCPA}}}{1-\gamma_{**}^{\text{CCPA}}},$$

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

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

from which follows that

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

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

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

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

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 $\gamma_i$ right below $\frac{1}{2}$, anticipating being targeted by the seller, would choose opt-out. This in turn reduces the effective $\gamma_{**}^{\text{CCPA}}$ 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 $\gamma_{**}^{\text{CCPA}}$, this confirms the conjectured cutoff equilibrium. Furthermore, since the price is $p_{B,\text{out}}^{\text{CCPA}} = \frac{\bar{v}}{2}$ and $\gamma_{**}^{\text{CCPA}} = \frac{1}{2}$, it follows that all weak-willed consumers that opt out buy good $B$ when it is advertised to them.

A.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

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

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

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

51
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 ($\frac{2}{u_{B}} z_{B, out}^{CCPA} < \frac{1}{u_{B}} z_{B}^{NS}$ and $\frac{2}{u_{B}} z_{B, out}^{CCPA} \geq z_{B}^{NS}$). Consequently, for

$$u_{B} < u_{B^*} = -\frac{3\bar{u}}{8} \left( \frac{z_{A}^{CCPA} - (\pi_{S} + \pi_{W}) z_{A}^{NS}}{z_{B, out}^{CCPA} - \pi_{W} z_{B}^{NS}} \right),$$

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

$$W^{GDPR} = \left( z_{A, in}^{GDPR} + (1 - \gamma_{**}^{GDPR}) \pi_{W} \frac{z_{A, out}^{GDPR}}{1 - \pi_{S} - \gamma_{**}^{GDPR} \pi_{W}} \right) \frac{3}{8} \bar{u} + \left( z_{B, in}^{GDPR} + \pi_{W} (1 - \gamma_{**}^{GDPR}) \frac{z_{B, out}^{GDPR}}{1 - \pi_{S} - \gamma_{**}^{GDPR} \pi_{W}} \right) u_{B},$$

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:

$$W^{GDPR} - W^{NS} = \left( z_{A, in}^{GDPR} + (1 - \gamma_{**}^{GDPR}) \pi_{W} \frac{z_{A, out}^{GDPR}}{1 - \pi_{S} - \gamma_{**}^{GDPR} \pi_{W}} - z_{A}^{NS} \right) \frac{3}{8} \bar{u} + \left( z_{B, in}^{GDPR} + (1 - \gamma_{**}^{GDPR}) \pi_{W} \frac{z_{B, out}^{GDPR}}{1 - \pi_{S} - \gamma_{**}^{GDPR} \pi_{W}} - z_{B}^{NS} \right) u_{B},$$

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, because of the increased exposure of weak-willed consumers to seller $B$.

Notice that, when $u_{B} = 0$, it must be the case that $W^{GDPR} > W^{NS}$ because of 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 $z_{A,in}^{GDPR} - \frac{z_{A,in}^{GDPR}}{4}$, which is bounded from above. In contrast, with $\gamma_{**}^{GDPR}$ 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

$$W^{GDPR} - W^{CCPA} = \left( z_{A,in}^{GDPR} + (1 - \gamma_{**}^{GDPR}) \pi_W \frac{z_{A,out}^{GDPR}}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} - z_A^{CCPA} \right) \frac{3}{8} \bar{u}$$

$$+ \left( z_{B,in}^{GDPR} + z_{B,out}^{GDPR} - \frac{1 - \pi_S - \pi_W}{1 - \pi_S - \gamma_{**}^{GDPR} \pi_W} z_{B,out}^{GDPR} - z_{B,out}^{CCPA} \right) u_B.$$ 

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.