Data Privacy and Temptation*

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

February 2020

Abstract

This paper analyzes how different data-sharing schemes of a digital platform may affect consumer surplus and social surplus when a fraction of the consumers have weak self-control and suffers from targeted advertising of temptation goods, such as gambling and video games. While sharing consumer data with firms improves the efficiency of matching consumers with normal consumption goods, it also exposes weak-willed consumers to temptation goods. Despite the seeming appeal of the opt-in policy of allowing each consumer to opt in or out of data sharing, our analysis shows that this policy may not be effective in protecting severely tempted consumers. When other consumers, motivated by the improved access to normal goods, choose to share their data, their opt-in reduces the anonymity of the weak-willed consumers who choose to opt out. To alleviate this externality, privacy protection regulation needs to limit the bundling of the consumer authorization to share data with normal good and temptation good sellers.

*PRELIMINARY DRAFT.

[†]The Chinese University of Hong Kong, Shenzhen. Email: zhuangliu@cuhk.edu.cn.
[‡]University of Texas, Austin. Email: Michael.Sockin@mccombs.utexas.edu.
[§]Princeton University and NBER. Email: wxiong@princeton.edu.

The age of big data brings with it not only substantial benefits, such as dramatically improved access for consumers to products and services, but also undesirable drawbacks. Consider the story of a compulsive gambler who tries to recover from gambling—he deleted all the casino apps from his smart phone; he removed his profile from all of the major gambling sites; he set up a rule in Gmail to automatically delete any emails that are related to gambling. One day, however, he logged on to YouTube, and all his efforts seemed to be in vain: "99% of the ads I see on YouTube are for gambling."[1] This frustration is just one example of how data analytics are increasingly used by firms to target consumers with self-control problems.

Online casinos and the global video game industry use third-party companies to harvest personal data and send advertisements to target those who are most likely to be tempted, such as those who previously gambled or played but have stopped (The Guardian (2019)). E-cigarette companies use social media platforms and youth-focused advertising strategies to target teenagers and expand the market for their fledgling product (Kim et al. (2019)). Marketers are now creating online alcohol advertisements that are more specifically tailored to their intended audience using social data, including age, location, gender, interests, and much more (Morris (2019)). The adult film industry uses similar data-driven approaches to cater to consumers' diverse tastes (Raustiala & Sprigman (2019)). This phenomenon extends beyond these so-called "sin" industries. Payday lenders use algorithmic scoring to zero-in on consumers when they are likely to be vulnerable to short-term credit products with usurious interest rates and unfavorable terms (Hurley and Adebayo (2017)). The size and influence of these industries are staggering. In 2018, for example, the gross revenue of the gambling industry in the U.S. was $161 billion;[2] the video game industry generated about $135 billion in revenue (Newzoo (2018));[3] visits to Pornhub, the largest adult film website in the world, totaled 33.5 billion;[4] and 12 million Americans took out payday loans, both online and through about 16,000 storefront offices, borrowing almost $90 billion (OCC (2018)).

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

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

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

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

The fact that consumer data can be used by firms to target consumers with self-control problems lies at the heart of the ongoing public and policy discussions on data privacy and data regulations. In particular, there is a growing concern that the collection of personal data by digital platforms such as Google, Amazon, and Facebook represents an unprecedented challenge to consumer privacy. An extensive economics literature analyzes how the use of vast data collected by these platforms may affect social surplus and consumer surplus.[5] At the risk of over-simplifying the debate, this literature tends to focus on two key effects—on the one hand, consumer data can increase the social surplus by allowing firms to cater more precisely to consumer preferences, which improves the matching between consumers and sellers; on the other hand, consumer data empower firms to price discriminate consumers and thereby tilt the distribution of the social surplus toward firms. These effects create an intricate tradeoff between the sum and the distribution of the economic gains from trade, which makes the impact of the use of consumer data on consumer surplus *ambiguous* and largely dependent on the market setting.

Much of the privacy literature, however, builds on a key assumption that consumers are rational and are therefore able to make the best choice for themselves given a choice set. In doing so, the literature overlooks the aforementioned frustration faced by the compulsive gambler who tries to restrain himself and, more generally, the *unambiguous* cost imposed by the much expanded access of firms to data on consumers with behavioral biases and character flaws. There is already extensive evidence from the behavioral economics literature characterizing a variety of behavioral biases and flaws among consumers.[6] In particular, evidence from both laboratory and field data shows that consumers often fail to exercise sufficient self-control and succumb to temptation. Profit-maximizing firms design contracts and products to exploit such consumers, e.g., DellaVigna and Malmandier (2004), Gabaix and Laibson (2006), and Heidhues and Koszegi (2017). Innovative data analytics powered by increasingly accessible consumer data allow firms to more precisely identify and exploit consumers with self-control problems or other vices. Through this mechanism, easily accessible

[5] See Acquisti, Taylor and Wagman (2016), Bergemann and Morris (2019), and Goldfar and Tucker (2019), for recent reviews. For example, both Taylor (2004) and Acquisti and Varian (2005) show that it is optimal for sellers to use consumers' past purchase information to price discriminate them only if consumers are naive of how their data is used, but not optimal if consumers are sophisticated and can adapt their purchasing strategies. Interestingly, Ichibashi (2019) shows that a multi-product seller prefers to commit to not use consumer information for pricing so that consumers would truthfully report their information and the seller can recommend to them the best matching products.

[6] See DellaVigna (2009) for a review of the evidence.

consumer data by firms unambiguously hurts these weak-willed consumers. This important aspect has been largely ignored by the ongoing discussions of privacy protection and privacy regulations.

This paper fills this gap by developing a model to examine how different sharing rules of consumer data on a digital platform may impact consumer welfare when a subset of the consumer population suffers from temptation or vice to different extents. Specifically, the model features an ecosystem associated with a digital platform, such as Google or Facebook, which collects the data of consumers on the platform and shares the data with sellers. For simplicity, there are two sellers. Seller A sells a normal consumption good, such as music, while seller B sells a temptation good. One may think of the temptation good as alcohol, e-cigarette, gambling, or more modern incarnations of addiction, such as mobile gaming. Each of the sellers can target advertisements to potential buyers of its good at a convex cost. Each consumer may receive advertisements from none, one or both of the sellers, and then chooses from the received advertisements none, one or both of the goods.

There are three types of consumers: The first type is strong-willed and will always resist the temptation good, while the second type is weak-willed and may indulge in the temptation good. Both the 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 consumers would never buy either the normal good or the temptation good, and serves as noise in the sellers' targeted advertising.

For simplicity, we assume that both strong-willed and weak-willed consumers have a random utility over the normal good. The random utility prevents seller A, even if it has perfect information about consumer types, from using third-degree price discrimination against its potential buyers, which is a key cost of revealing consumer data, as widely analyzed in the aforementioned privacy literature. In the absence of such price discrimination, both strong-willed and weak-willed consumers prefer receiving the advertisement from seller A. Furthermore, since strong-willed consumers can always resist the temptation good, they do not mind receiving advertisement from seller B. As such, strong-willed consumers strictly prefer 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. On the one hand, they benefit from more precise targeting by seller A. On the other, as they are vulnerable to temptation, they suffer from receiving the advertisement from seller B. This is a key tension in our model that drives the differences in the weak-willed consumer surplus and utilitarian social surplus under different data-sharing schemes.

Weak-willed consumers suffer from varying degrees of temptation and randomly receive advertisements for a temptation good that they must then decide whether to buy. As such, they face a lottery of potential menus on the platform. To analyze the preferences of these consumers, we builds on the self-control utility of Gul and Pesendorfer (2001), which has been expanded to a specification of random temptation utility by Stovall (2010). 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). Temptation in this utility formulation imposes a mental cost of exerting self-control. When facing the same price, those with more mild temptation resist the temptation good when it is on their menu, and suffer a mental cost for exercising self-control, while those with more severe temptation succumb and buy it. Since this utility formulation has an axiomatic underpinning, it anchors our analysis of consumer surplus on a systematic foundation. In particular, each consumer's menu preference determines his preference for the platform's data sharing scheme, which in turn determines each seller's advertising strategy and, consequently, the menu faced by each consumer.

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

The General Data Protection Regulation (GDPR) enacted by the European Union in 2018 requires digital platforms to explicitly ask for permission of each user for data collection and data use. That is, each consumer can choose whether to opt in or opt out of data collection by a platform and its subsequent data sharing with sellers. This policy is appealing as it allows strong-willed to opt in and benefit from the improved matching with the normal good seller and weak-willed to opt out and protect themselves. Our model permits a welfare analysis of this policy. Intuitively, all strong-willed opt in for data sharing, as they unambiguously benefit from revealing themselves to seller A. By opting in, each weak-willed consumer, however, faces a trade-off between the improved access to the normal good and intensified targeting from seller B. In equilibrium, modestly tempted consumers with temptation coefficient below an equilibrium cutoff opt in, while severely tempted consumers opt out. Surprisingly, while each weak-willed consumer can choose to opt out, this opt-out choice may not sufficiently protect him. Our analysis shows that, when the aggregate temptation of weak-willed consumers is sufficiently severe, this opt-in policy lowers welfare for weak-willed consumers than the benchmark setting without any data sharing, despite its seeming advantages.

This ranking of weak-willed consumer welfare with the opt-in policy arises because the composition of the opt-out pool, which determines the extent that the opt-out pool can act as camouflage for weak-willed consumers, is endogenously determined by the choices of all consumers. When a strong-willed consumer chooses to opt in, he drops out of the opt-out pool, which reduces the camouflage available for those severely tempted consumers. This externality echoes the notion of social data put forth by Acemoglu et al (2019) and Bergemann, Bonatti and Gan (2019) that data have an important social dimension as each individual's data is also indicative of others' behavior.

Our paper adds a new dimension to the privacy literature by highlighting the cost of data sharing imposed on consumers with behavioral weakness. This cost expands the cost-benefit analysis of privacy protection. Our analysis suggests that while allowing consumers to opt in or out of data sharing helps to protect tempted consumers, its effectiveness is undermined by the externality of other consumers' opt-in choices. In particular, our analysis shows that by bundling authorization of data sharing with temptation good sellers with other highly desirable normal goods sellers, a platform may inadvertently isolate severely tempted consumers in the opt-out pool. It is therefore important for privacy protection regulations to further limit such bundling in order to ensure the effectiveness of the opt-in policy.

Our paper contributes to a growing literature discussing a wide range of economic issues related to privacy. Ali and Benabou (2019) argue that while publicity helps induce pro-social behavior, it crowds out information aggregation, thus providing an informational rationale for privacy. Tirole (2019) is concerned that political authorities might enlist a social rating that bundles each individual's political attitude and social graph to control society without engaging in brutal repression or misinformation. Campbell, Goldfarb, and Tucker (2015) show that privacy protection policies can act as "de facto" barriers to entry that entrench monopolies, while Calzolari and Pavan (2006) show that data sharing across firms can enhance welfare by eroding distortions arising from asymmetric information. Our model motivates the need for privacy protection in order to protect consumers with behavioral weakness and highlights a nuanced role for privacy protection policies because consumers do not internalize that their participation in data sharing with normal good sellers indirectly cross-subsidizes the advertisement policies of temptation good sellers.

1 The Model

We consider the ecosystem associated with a digital platform, such as Google or Facebook. As a large number of consumers visit the platform, the platform can collect the consumers' digital footprints, which in turn reveal useful information about the consumers' preferences. To examine the consequences of the platform sharing consumer information with goods sellers, we consider an ecosystem with a continuum of consumers and two goods sellers. There are two types of consumption goods, A and B. We think of good A as a normal good, such as music, while we consider good B to be a temptation good, such as video game and gambling. There is a continuum of consumers from three types $\{S, W, O\}$, which represents strong willed, weak willed, and outside, respectively. The strong-willed consumers can always resist the temptation to purchase the temptation good, while the weak-willed consumers may not be able to resist. The third type of consumers would never purchase either good.

1.1 Consumers

There is a continuum of consumers in each type. The ex ante probability for a consumer to be strong-willed is π_S, to be weak-willed is π_W, and to be type O is $1-\pi_S-\pi_W$. 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.

To facilitate our normative analysis of how privacy regulations may affect consumers with temptation, 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. In the second step, a consumer makes a choice from a given menu N, and, moving backwardly, the consumer chooses from a set of menus. Since advertisements of the temptation good may be random, such a preference is actually over a set of lotteries over potential menus.

Under the Gul-Pesendorfer framework, the consumer's preference for the menu N is given by the following:

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

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

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

As a result of the compromise, the consumer may not choose the most tempting choice from the menu. If so, i.e., $x_* \neq \arg\max_{x'\in N} v(x')$, and the consumer has to exercise 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 (1), while it does not directly affect the consumer's actual choice, determines 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 choice set: $\max_{x'\in N} v(x') - v(x_*)$ represents the cost of self-control incurred by the consumer. The ex ante utility of the consumer is then the expected utility from all potential menus N given the prevailing data sharing policy.[7]

As we will discuss, the menu N faced by a consumer is random and depends on the two sellers' advertising strategies, which in turn depends on the platform's data sharing scheme. Our analysis thus builds directly on the random Gul-Pesendorfer temptation utility

[7] 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, i.e. Von Neumann-Morgenstern 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). The net utility determines the consumer's ex ante preference of the platform's data sharing scheme.

Temptation utility A strong-willed consumer, with type $\tau_i = S$, has the following utilities from consuming good A and good B:

x	$u_S(x)$	$v_S(x)$
A	$\tilde{u}_A > 0$	0
B	$u_B < 0$	0

(2)

with $u_S(\cdot)$ and $v_S(\cdot)$ denoting the commitment utility and temptation utility of a strong-willed consumer, respectively. Good B gives a negative commitment utility $u_B < 0$, reflecting that the temptation good, such as tobacco or alcohol, is ultimately harmful to consumers. The strong-willed consumer will never buy good B, as he does not get any temptation utility from it. Both strong and weak-willed consumers have a random utility for good A, $\tilde{u}_A$, which has a uniform distribution $H(\tilde{u}_A) \sim U[0, \bar{u}]$ with $\bar{u} > 0$ as a constant representing 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.

A weak-willed consumer, with type $\tau_i = W$, has the following utilities:

x	$u_W(x)$	$v_W(x)$
A	$\tilde{u}_A > 0$	0
B	$u_B < 0$	$\gamma_i \bar{v} - u_B > 0$

(3)

with the subscript W denoting weak-willed consumers. Good A also gives a commitment utility of $\tilde{u}_A$ to the weak-willed consumer and no temptation utility, just like to strong-willed consumers. Good B gives a negative commitment utility $u_B < 0$, and a temptation utility of $\gamma_i \bar{v} - u_B$ to the weak-willed consumer, where $\bar{v} > 0$ is a constant measuring the overall temptation of weak-willed consumers to good B, and $\gamma_i \in [0, 1]$ measures the 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 the weak-willed consumer's choice of whether to buy good B, when it is on the menu, is determined by a simple expression:

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

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

Note that the temptation utility delivered by good B to a weak-willed consumer is persistent and characterized by a personalized parameter γ_i, while the commitment utility delivered by good A to a consumer (either strong-willed or weak-willed) is random. The random utility delivered by good A prevents price discrimination by seller A even if seller A has full information about consumers.[8] In contrast, information about a weak-willed consumer allows seller B not only to precisely target its advertisement but also to price discriminate the weak-willed consumer. This asymmetric setting allows us to focus on how access to consumer data affects weak-willed consumers through their temptation utility, rather than how more general price-discrimination affects consumers' consumption of normal goods.

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

Menu preferences 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 goods A and B.[9] That is, each consumer can separately make his choice of each good, even if both goods are on his menu. As a result, we can separately denote the menu of consumer i for each of the two goods: $\mathcal{M}_i^A \in \{\{A, \emptyset\}, \emptyset\}$ is the menu faced by the consumer for good A, with $\emptyset$ representing the menu when good A is not advertised to the consumer and $\{A, \emptyset\}$ the menu when it is advertised, and $\mathcal{M}_i^B \in \{\{B, \emptyset\}, \emptyset\}$ is his menu for good B.

Then, we can build on the utility framework specified in (1) to analyze the actual choice

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

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

of a consumer with type $\tau_i \in \{S, W, O\}$ from the menus $\mathcal{M}_i^A$ and $\mathcal{M}_i^B$,

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

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 advertisement policy and pricing policy of the sellers.

The consumer's ex ante preference of the full menu is

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

This menu preference allows us to analyze the welfare implications of the platform's data sharing rule, 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. In the first case, the platform does not share any data and thus consumers are completely anonymous to the sellers. In the second case, the platform shares all data sharing with the sellers and the consumers' temptation types are therefore known to the sellers. In this case, the sellers may charge consumers different prices based on their temptation types. We will also examine a more complicated third case in which sellers only know the types of some consumers who choose to share their data.

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 a search cost that sellers pay to find consumers through advertising, and the convexity reflecting that it is increasingly difficult to reach a broader audience.[10]

[10] Note that while the physical cost of sending out electronic messages and advertisements is almost zero, in practice sellers need to send advertisements through advertising firms/platforms and pay substantial advertising fees. 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.

For the baseline case with anonymous consumers, sellers cannot price discriminate across customers, and we drop the type subscripts from goods prices here for ease of exposition. As a consequence, seller k maximizes

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

where Q_k is the quantity sold by seller k in the market, p_k is the price of its good, and z_k is the measure of the consumer population to whom the seller advertises. We assume that, if a seller does not advertise to a consumer, then its good is not on that consumer's menu. Each seller is strategic, and can only condition its advertisement and pricing policies on its information set $\mathcal{I}^k$ that it knows about the consumers. In what follows, we will analyze different settings with different information sets for the sellers.

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

$$p_A \leq \bar{u} \text{ and } p_B \leq \bar{v}.$$

Violating these price constraints would lead to no sales by the sellers.

1.3 Rational Expectations Equilibrium

An equilibrium in the ecosystem is an optimal advertising policy z_k and an optimal pricing policy p_k for each seller $k \in \{A, B\}$ and an optimal purchase policy correspondence $\{x_{\tau_i}\left(\mathcal{M}_i^A\right), y_{\tau_i}\left(\mathcal{M}_i^B\right)\}$ for each consumer i such that the following are satisfied.

- Consumer optimization: Taking as given each seller's advertising policy z_k and pricing p_k, a consumer i with temptation type τ_i finds it optimal to follow the purchase policy $\{x_{\tau_i}\left(\mathcal{M}_i^A\right), y_{\tau_i}\left(\mathcal{M}_i^B\right)\}$ for a choice set $\{\mathcal{M}_i^A, \mathcal{M}_i^B\}$.

- Seller optimization: Taking as given each consumer's optimal policy and the other seller's advertising and pricing policies, each seller k finds it optimal to choose an optimal advertising policy z_k and charges price p_k for its good.

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

$$\begin{aligned} W &= \int \tilde{u}_A \left(\pi_S \mathbf{1}_{\left\{A \in \mathcal{M}_S^A \,\cap\, x_S = A\right\}} + \pi_W \mathbf{1}_{\left\{A \in \mathcal{M}_W^A \,\cap\, x_W = A\right\}} \right) dH\left(\tilde{u}_A\right) \\ &\quad + \pi_W \int \left(u_B \mathbf{1}_{\left\{B \in \mathcal{M}_W^B \,\cap\, x_W = B\right\}} + \left(u_B - \gamma_i \bar{v}\right) \mathbf{1}_{\left\{B \in \mathcal{M}_W^B \,\cap\, x_W = \emptyset\right\}} \right) dG\left(\gamma_i\right). \end{aligned} \tag{4}$$

The first term captures the commitment utility of both strong-willed and weak-willed consumers from consuming good A, and 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 by those who have the temptation good on their menus but choose not to consume it, $-\left(\gamma_i \bar{v} - u_B\right)$. Note from (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 utility. The price it pays for the good is a transfer to seller B and does not affect the welfare. As a consequence, the welfare loss incurred is from the negative commitment utility 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 (4) reveals a tradeoff associated with sharing consumer data with sellers—it increases the search 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 second term 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 (4). Since advertising is costless from a social perspective, the planner would want seller A to give all strong and weak-willed consumers access to good A. 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 would prefer seller B not to advertise to any consumer. We summarize this first-best outcome below.

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

2 Equilibrium

In this section, we characterize the equilibrium of the ecosystem in two contrasting data-sharing settings, one without any data sharing with sellers, and thus the consumer types remaining anonymous to them, and the other with full data sharing so that the consumer types are fully known to sellers.

2.1 Consumer Choice

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

Proposition 2 *A strong-willed consumer, $\tau(i) = S$ with commitment utility $\tilde{u}_A$, will choose as follows:*

- *purchase good A if it is offered at a price below his reservation value $p_A \leq \tilde{u}_A$;*
- *always reject good B.*

A weak-willed consumer, $\tau_i = W$, with commitment utility $\tilde{u}_A$ and temptation coefficient γ_i will choose as follows:

- *purchase good A if it is offered at a price below his reservation value $p_A \leq \tilde{u}_A$;*
- *purchase good B if it is on the menu and if his temptation coefficient γ_i is sufficiently high relative to the price p_B scaled by the maximum temptation $\bar{v}$ of good B:*

$$\gamma_i \geq \gamma_* = \frac{p_B}{\bar{v}}.$$

The 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. Ex ante, however, 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 to have better access to the advertisements of 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 type γ_i will buy it. Those with a modest temptation ($\gamma_i < \gamma_*$) avoid buying good B, but still suffer a mental cost, $\gamma_i \bar{v} - u_B$, from resisting the temptation. Those with strong temptation who buy good B suffer even more from not only paying a high price of $p_B = \gamma_* \bar{v}$ to purchase the good, but also from enduring the negative commitment utility of u_B that this purchase entails.

2.2 Equilibrium without Data Sharing

We first analyze a baseline setting in which the platform does not collect or share any consumer data with the two sellers. As a result, consumers are anonymous and sellers have no information about each consumer's type. The following proposition characterizes the equilibrium.

Proposition 3 *In the absence of any data sharing, there exists a unique equilibrium with the following properties:*

1. *Seller A advertises good A to z_A measure of consumers:*

$$z_A = \min \left\{ \max \left\{ 1 - 2\sqrt{\frac{1}{\pi_S + \pi_W} \frac{F}{\bar{u}}}, 0 \right\}, 1 \right\}, \tag{5}$$

   at a uniform price:

$$p_A = \frac{1}{2}\bar{u}.$$

2. *Seller B advertises good B to z_B measure of consumers:*

$$z_B = \min \left\{ \max \left\{ 1 - 2\sqrt{\frac{1}{\pi_W} \frac{F}{\bar{v}}}, 0 \right\}, 1 \right\}. \tag{6}$$

   at a uniform price:

$$p_B = \frac{1}{2}\bar{v}.$$

In this baseline setting, the sellers' undirected search of anonymous consumers leads to a source of inefficiency. As a result of the search frictions, seller A limits its advertising to a small pool of potential consumers. Equation (5) shows that seller A's advertising intensity decreases with its search cost parameter F, and increases with $\pi_S + \pi_W$ (the probability of interested consumers in the population) and $\bar{u}$ (which determines the price of good A). The search frictions also deter seller B from fully covering weak-willed consumers, as shown by Equation (6). As such, anonymity protects weak-willed customers from being targeted by seller B.

2.3 Equilibrium with Full Data Sharing

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

Specifically, seller A can now avoid the costly advertisement to the third type of consumers who would never buy good A. As strong-willed and weak-willed consumers have the same preference for good A and their purchase decision of 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_{AD} as the measure of strong-willed and weak-willed consumers, to whom seller A advertises its good at a price of p_A. Proposition 4 derives the seller's optimal z_{AD} and p_A. While seller A faces the same cost function of advertising to consumers as before, data sharing allows it to achieve a higher level of efficiency by avoiding advertising to the third type of consumers who would never buy good A.

Access to consumer data also allows seller B to focus its advertisement on weak-willed consumers. Furthermore, since seller B also observes the severity of each weak-willed customer's temptation, it concentrates its advertising only on the most tempted consumers, i.e., those with γ_i higher than a threshold γ^*. It will also charge each weak-willed consumer his full reservation value, $p_B(\gamma_i) = \gamma_i \bar{v}$, which is the net utility cost of resisting temptation. Thus, full data sharing allows seller B to precisely target weak-willed consumers and to perfectly price discriminate against them.

We derive the equilibrium in the following proposition.

Proposition 4 *In the setting with full data sharing, there exists a unique equilibrium with the following properties:*

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

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

at the same price:

$$p_A = \frac{1}{2}\bar{u}.$$

2. *Seller B advertises good B to all weak-willed consumers with $\gamma_i \geq \gamma^* = 1-\frac{\bar{z}_{FS}}{\pi_W}$, where*

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

is the total advertising by seller B with $\frac{\bar{z}_{FS}}{\pi_W} \geq 1-\sqrt[3]{\frac{F}{\pi_W^2\bar{v}}}$, and $\bar{z}_{FS}$ being weakly increasing in $\bar{v}$ and decreasing in F. Furthermore, seller B charges each weak-willed consumer a price equal to his reservation utility $p_B(\gamma_i) = \gamma_i\bar{v}$.

The improved search efficiency afforded by full data sharing allows seller A to better cover both strong-willed and weak-willed consumers, which strictly benefits them. Data sharing, however, also allows seller B not only to more precisely target weak-willed consumers, but also to discriminate against them by charging them a price equal to their reservation utility. As a result, while data sharing strictly benefits strong-willed consumers, it presents a tradeoff for weak-willed consumers. On the one hand, they have better access to good A, which improves their welfare; on the other, they are also more exposed to the temptation good, which hurts their welfare. The net effect is ambiguous. As each weak-willed consumer potentially suffers from both good B's negative commitment utility u_B and temptation utility $(\gamma_i\bar{v}-u_B)$, the utilitarian welfare of weak-willed consumers is increasing with u_B and decreasing with $\bar{v}$. The following proposition consequently shows that, when temptation in the population is sufficiently strong ($\bar{v}$ is higher than a critical level) or the commitment utility of good B is sufficiently negative (u_B is lower than a critical level), data sharing reduces the welfare of weak-willed consumers. Furthermore, if the impairment to the welfare of weak-willed consumers is sufficiently severe, then data sharing may even reduce social welfare, as established by the following proposition.

Proposition 5 *The utilitarian welfare of strong-willed consumers is higher with data sharing. There exist critical levels u_B^* and $\bar{v}^*$ such that the utilitarian welfare of weak-willed consumers is lower with data sharing if u_B is less than u_B^* or $\bar{v}$ is greater than $\bar{v}^*$, and higher otherwise. Furthermore, there exist critical levels of $u_B^{**} < u_B^*$ and $\bar{v}^{**} > \bar{v}^*$ such that the utilitarian social welfare of the ecosystem is lower with data sharing if u_B is lower than u_B^{**} or $\bar{v}$ is greater than $\bar{v}^{**}$.*

Proposition 5 shows that data sharing may harm the welfare of weak-willed consumers in the presence of temptation goods, even though it always improves the welfare of strong-willed consumers. In typical models of data sharing and price discrimination in which temptation is absence, data sharing tends to improve social surplus, even though it may or may not improve consumer surplus. In contrast, data sharing in our model may even reduce social surplus if the temptation problem of weak-willed consumers is sufficiently severe (i.e., either u_B is sufficiently negative or $\bar{v}$ is sufficiently high). This outcome motivates better data sharing policies that can protect weak-willed consumers while allowing strong-willed consumers to benefit from data sharing. We explore one such policy in the next section in which we allow each consumer to opt in or out of data collection and sharing on the platform.

3 Consumer Opt-in & Opt-out

A key feature of the GDPR is to protect consumers by empowering consumers with the initial allocation of rights to their personal data and therefore greater capacity to deny data collection and data sharing with others by businesses. Under the GDPR, a consumer must explicitly opt in and give a business consent to collect and use the consumer's data, otherwise the business cannot do so. (See Appendix A for a more extensive summary of the GDPR.) This policy motivates us to explore in this section whether giving consumers' such opt-in and opt-out rights is sufficient to protect consumers in the presence of temptation. Specifically, we now give each consumer the option to opt in or out of data collection by the platform so that the platform can collect and share data with both sellers about a consumer only after obtaining his consent. Our analysis will later permit us to discuss the relevance of data portability.

As strong-willed consumers strictly benefit from having their data shared with seller A, it is straightforward to see that all strong-willed opt in for data collection. As the third type of consumers are not interested in either good in the platform, there is neither gain nor loss for them from the data collection. We assume that they would always opt-out of the data collection by the platform as a result of their preference for privacy. There is now a non-trivial choice for each weak-willed consumer. By opting in for the data collection, the consumer benefits from the improved access to good A but also is more exposed to the temptation good. It is intuitive to conjecture that weak-willed consumers with sufficiently strong temptation problem will choose to protect themselves by opting out of data collection. We therefore conjecture that weak-willed consumers with temptation coefficient γ_i lower than a critical level γ_{**} will opt in, while those above will opt out.

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

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

where $z_{B,in}(d\gamma_i) \in [0, \pi_W]\, d\gamma_i$ is the advertising intensty to opt-in consumers with temptation coefficient γ_i and $p_{B,in}(\gamma_i)$ is the price that seller B charges them. Note that it is the conditional probability of a weak-willed consumer being targeted by both sellers, $\frac{z_{A,in}}{\pi_S+\pi_W\gamma_{**}}$ and $\frac{z_{B,in}(\gamma_i)}{\pi_W}$, that matters for his utility. If the consumer instead opts out of data collection, his utility is

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

For a consumer to voluntarily opt-in to data sharing, it must therefore be the case that

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

The following proposition characterizes the equilibrium when each consumer can choose to opt in or out of data collection on the platform.

Proposition 6 *Under the parameter restriction that* $F < \frac{\bar{u}}{4}$, *the equilibrium with each consumer choosing to opt in or out of data collection 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 $\gamma_i \leq \gamma_{**}$ and opt-out if $\gamma_i > \gamma_{**}$.*

2. *Seller A charges consumers in the opt-in and opt-out pools the same prices*

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

and adopts a water-filling advertising strategy that gives priority to the opt-in pool:

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

3. *Seller B charges consumers in the opt-out pool a fixed price of*

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

and charges weak-willed consumers in the opt-in pool their reservation utility:

$$p_{B,in}(\gamma_i) = \gamma_i \bar{v}.$$

Furthermore, seller B chooses an advertising intensity for weak-willed consumers in the opt-in pool:

$$dz_{B,in}(\gamma_i) = \begin{cases} \pi_W d\gamma_i & \text{if } \gamma_i \in [\gamma_*, \gamma_{**}] \\ 0 & \text{otherwise} \end{cases},$$

where $\gamma_* = \gamma_{**} - \frac{\bar{z}_{B,in}}{\pi_W}$ and $\bar{z}_{B,in} = \int_{\gamma_*}^{\gamma_{**}} dz_{B,in}(\gamma_i)$ is the total advertising of seller B to weak-willed consumers in the opt-in pool. $\bar{z}_{B,in}$ and $z_{B,out}$ (seller B's advertising to the opt-out pool) are given by (17) if $\gamma_{**} \geq \frac{1-\pi_S}{2\pi_W} - \sqrt{\left(\frac{1-\pi_S}{2\pi_W}\right)^2 - \frac{1}{4}}$, and by (16) otherwise.

4. *It is sufficient, although not necessary, for $\pi_S \geq 1 - 2\sqrt{\frac{F}{\bar{u}}}$, $\bar{u} \leq \frac{24\bar{v}}{\pi_W}$, and $u_B \geq \frac{1.8-\pi_S}{\pi_W}\bar{v}$ to ensure that at least one equilibrium cutoff γ_{**} for weak-willed consumers exists.*[11] *When it exists, γ_{**} is given by*

$$\gamma_{**} = \min\left\{\frac{2}{\pi_W}\sqrt{-\frac{p}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2p}\sqrt{-\frac{3}{p}}\right) - \frac{2\pi k}{3}\right) - \frac{1+4\pi_S - z_{B,out}}{3\pi_W}, 1\right\},$$

where $k \in \{0,1\}$, p and q are given by (18), and there can be multiple equilibria.

[11] Although we provide a sufficient condition for a cutoff equilibrium to exist, we find numerically that such an equilibrium appears to exist for a much wider range of π_S.

Proposition 6 confirms that weak-willed customers follow a cutoff strategy to opt in and out of data collection—those with mild temptation (low γ_i) opt in, while those with severe temptation (high γ_i) opt out. This is intuitive since weak-willed customers with mild temptation benefit more from the better coverage from seller A than the temptation they suffer from the potential advertising from seller B. In contrast, severely tempted weak-willed customers are willing to forego the benefit of better coverage from seller A because the cost of being targeted and exploited by seller B is higher. Since seller B can efficiently target weak-willed consumers in the opt-in pool, it targets the most tempted among them and charges them their full reservation value of temptation, while it charges one price, the maximum between $\frac{1}{2}\bar{v}$ and $\gamma_{**}\bar{v}$, to those in the opt-out pool.

While consumers with severe temptation can choose to opt out and therefore to hide within the opt-out pool from seller B, this protection is weakened by the opt-in decisions of other consumers, i.e., strong-willed and modestly weak-willed consumers. Their departure from the pool reduces the camouflage of those that are severely weak-willed, as reflected by the $\frac{z_{B,out}}{1-\pi_S-\pi_W\gamma_{**}}$ probability of a weak-willed consumer to receive the advertising of seller B, and consequently increases their exposures to the temptation good seller. In this sense, there is an externality in the opt-in decisions of strong-willed and modestly weak-willed consumers, as their decisions do not account for the potential effect on consumers with severe temptation. The presence of this externality suggests that simply allowing consumers to opt in and out of data collection is not sufficient to for consumers with temptation to be able to protect themselves.

The following proposition formally compares the profit of seller B and the welfare of weak-willed consumers across the three data sharing settings that we have examined.

Proposition 7 *There is a strict ranking of the profit of seller B across the three data sharing schemes:*

$$\Pi_B^{FS} > \Pi_B^{OP} > \Pi_B^{NS},$$

which decreases from the scheme with full data sharing (FS) to the scheme with partial data sharing of consumers choosing opt-in or opt-out (OP), and then to the scheme with no data sharing (NS). There exist critical values $\bar{v}_*$ and $\bar{v}_{**}$, and u_{B*} and u_{B**}, such that the social surplus (utilitarian social welfare) of the ecosystem under the scheme with consumers choosing opt-in or opt-out (OP) is higher than that under the scheme with full data sharing

(FS):

$$W^{OP} > W^{FS},$$

if the temptation problem of weak-willed consumers is sufficiently severe (i.e., $\bar{v}$ is greater than $\bar{v}_*$ or if u_B is less than u_{B*}), and lower than that under the scheme with no data sharing (NS):

$$W^{OP} < W^{NS}.$$

again if the temptation problem of weak-willed consumers is sufficiently severe (i.e., $\bar{v}$ is greater than $\bar{v}_{**}$ or u_B is less than u_{B**}).

It is perhaps not surprising that the profit of seller B is strictly increasing with the extent of their access to consumer data. Its profit is highest in the setting with full data sharing by the platform, and lowest in the setting with no data sharing. Data sharing brings weak-willed consumers a trade-off between the improved access to seller A and the greater exposure to seller B, and so there is not a strict ranking for the welfare of weak-willed consumers across these data sharing schemes. When the temptation problem of weak-willed consumers is sufficiently severe (i.e., $\bar{v}$ sufficiently high or u_B sufficiently negative), the welfare of weak-willed consumers is decreasing with the extent of data sharing in the platform—being highest with no data sharing and lowest with full data sharing. While allowing consumers to opt in or out of data collection protects weak-willed consumers to some extent, this protection is insufficient in restoring their welfare in the benchmark without any data sharing.

Proposition 7 reveals an even strong result that when the temptation problem of weak-willed consumers is sufficiently severe, social welfare, which accounts for the profits of seller B and the welfare of strong-willed and weak-willed consumers, displays the same ranking, decreasing from the setting with no data sharing to that with partial data sharing and then to that with full data sharing. This ranking is surprising—despite the great advantages promised by the opt-in policy to allow strong-willed consumers to take advantage of the benefit of data sharing and still give weak-willed consumers the opt-out option to protect themselves, it can deliver lower social surplus than the benchmark scheme of no data sharing at all. The key reason for the worse welfare outcome of the opt-in policy is exactly the externality discussed earlier. As this policy enables strong-willed and mildly tempted weak-willed consumers to opt in, their opt-in decisions lead to less anonymity for these highly tempted consumers in the opt-out pool, which can make these consumers sufficiently worse off that the social surplus is lower with the opt-in policy.

Multiple temptations For simplicity, our analysis has focused on a setting with only one temptation good. In practice, however, there are many temptations, ranging from alcohol and cigarettes to gambling, pornography, and online gaming. In principle, a strong-willed customer toward one temptation good may be weak-willed toward another. That is, even though the weak-willed population toward any one vice is small, with so many temptations available online, a majority of population may choose to opt out of data sharing on online platforms to avoid different vices. As such, the opt-in policy may be sufficient to insulate the weak-willed population from temptation goods sellers. While compelling, this argument ignores several issues. First, given the great benefit of opt-in from the normal good seller, which proxies for the many conveniences brought by data sharing, addiction needs to be severe enough to motivate a consumer to opt out of data sharing. Perhaps for this reason, few choose to opt out in practice. As such, even with many temptations, the size of the opt-out pool still may not be considerably large enough to hide weak-willed consumers. Second, severe addictions tend to be correlated, which is why casinos offer complimentary alcohol and exotic dancers as entertainment, and why alcohol abusers have a higher incidence of polysubstance abuse, such as with cocaine or sedatives. For these reasons, the presence of multiple temptations may not necessarily enlarge the pool of severely addicted consumers who would choose to opt out of data collection so that they can provide camouflage for each other. Taken together, our analysis remains relevant even with multiple temptation goods.

Unbundling the authorization The ineffectiveness of the opt-in policy is crucially related to the fact that each consumer's data-sharing authorization bundles the temptation good seller with the normal good seller. This is similar to broad opt-in policies in practice, such as the GDPR and the California Consumer Privacy Act (CCPA), which apply to all of a consumer's data and potential uses of that data. In our setting, the normal good proxies for many advantages that data sharing with a digital platform brings to a consumer. For example, by letting Google tracks a user's prior search history, Google can help the user to quickly find what he might want to search for in the future. This simple convenience would already make it difficult for many users to opt out of data sharing with Google, not to mention that Google also offers many other conveniences such as free use of Gmail, Youtube, and Google map. Similarly, other digital platforms such as Facebook, Alibaba, WeChat, Amazon, and Apple, all offer many conveniences through the services they provide to users.[12] As a consequence, it is a difficult choice for any user to opt out of data sharing with these platforms.[13] This dilemma is exactly the issue with data sharing. Phrased differently, if consumers directly face a request to share data with a temptation good seller, such as a gambling site, most of them are likely to opt out. As a result, sharing consumer data with temptation good sellers necessitates entangling the authorization with that of other firms or platforms that offer consumers benefits that are difficult to refuse. Taken together, in order to make the opt-in policy more effective in protecting weak-willed consumers, privacy protection regulations need to limit the bundling of data-sharing authorization.

4 Conclusion

In this paper, we introduce a setting in which data sharing between consumers and firms can be harmful to consumers in the presence of temptation. While data sharing improves the matching between consumers and sellers of normal goods, it also improves the matching between weak-willed consumers and sellers of temptation goods, such as gambling and mobile gaming. Since those who can resist temptation effectively cross-insure those who cannot by making it more difficult for temptation good sellers to target them, broad opt-in/opt-out data sharing policies that give consumers control of their data, such as the GDPR, can harm susceptible consumers by unraveling anonymity. Our analysis suggests that policies that promote data portability, however, by unbundling how consumer data is used, can be effective in guarding consumers against temptations when such temptations can be clearly identified.

An important issue, however, is that what constitutes a temptation good is subjective and individual-specific. The lack of clear-cut designation may create difficulty for policy making. For instance, while alcohol, drugs, gambling, and pornography are widely recognized as common addictions, some people are addicted to exercise and others to drinking coffee, online

[12] Data sharing is an ubiquitous part of the many services that these firms provide. Google, for instance, scans the text in emails in Gmail to aid in advertisement targeting and records search and web browser histories with its Chrome web browser. Facebook monitors likes and searches on its websites (Facebook, Instagram, etc.) and messages through their messaging app and WhatsApp, while Apple compiles data from its apps and gives access to third party app developers.

[13] Opting-out from such platforms in reality may also be more difficult than individuals may perceive, even with policies like the GDPR. Google, for instance, is partnering with hospitals to gain access to individual medical records through "Project Nightingale" without needing patients' consent, while Facebook has expanded into VR gaming through Oculus Rift.

shopping, or posting on social media. Whether such activities should also be considered temptation goods because some subset of the population is addicted to them complicates the crafting of concrete data sharing policy prescriptions.

Appendix A The General Data Privacy Regulation

The General Data Protection Regulation (GDPR) is the core of the European digital privacy legislation. In January 2012, the European Commission set out ambitious plans for data protection reform across the European Union in order to make Europe fit for the digital age. After four years of preparation and debate, the GDPR was approved by the European Parliament in April 2016. On May 25, 2018, this legislation came into force across the European Union. This new framework is the largest change in data protection laws in Europe since 1995, in replacement of the EU Data Protection Directive 95/46/EC. It automatically becomes part of each member-state's legal framework, and applies to the processing of personal data by businesses established within the EU – and, importantly, businesses outside the EU if their data processing activities relate to individuals in the EU.

The GDPR aims to strengthen the protection of personal data in the EU by giving EU citizens more control over their personal data. For example, it empowers users with the right to access and get a copy of their data from internet service providers, erase their data from the businesses ("right to be forgotten"), and freely move their data on one internet platform to another (data portability). The GDPR imposes serious fines to infringement of rights and non-compliance, which is as high as $20 million or 4 percent of annual turnover of a firm, whichever is greater.

As for privacy and data protection, there are two basic models of legal arrangements: opt-in and opt-out. Under the opt-in regime, data collectors must obtain consumers' explicit consent before collecting, using, and sharing their personal information. The GDPR adopts the opt-in system, and it also requires consent to be freely given, specific, informed and unambiguous. In contrast, in the opt-out regime, data collectors can collect and share non-public customer information with third parties, but need to give consumers an opportunity to deny them to do so, or opt out (Lacker 2002). This opt-out regime is generally adopted by regulators in the United States. As a result, internet service providers in the United States are allowed to collect, share and sell their customers' web browsing data without their consent. For example, Skyscanner, TripAdvisor, and MyFitnessPal, among other popular apps for Android smartphones, transmit personal data to Facebook without the consent of users. Some apps such as travel site Kayak also send detailed information about users' flight searches to Facebook, including travel dates, whether a user has children and which flights and destinations a user searches for (Murgia 2018). These practices are prohibited under the opt-in regime of the GDPR. The fundamental difference between opt-in and opt-out regimes is the initial allocation of the property rights over personal information. In the opt-in regime, the rights are assigned to consumers by default, whereas in the opt-out system, the entitlements are allocated to firms, though consumers have the right to withdraw.

The right to data portability is another fundamental right of data subjects (consumers) in the GDPR. Unlike many other rights, data portability is novel. The right allows data subjects to receive personal data they provided to a controller (firm) in a structured, commonly used and machine-readable format, and to transmit those data to another controller. Implementing this idea is still challenging in practice, as it is still ambiguous and requires further clarification before it can be implemented (de Hert et al. 2018). But the definition of the right is clear. It is designed not only as a right to empower individual consumers, but also as a tool to foster competition between firms by supporting the free flow of personal data (The Article 29 Working Party, 2017, p3).

Appendix B Proofs of Propositions

B.1 Proof of Proposition 2

We first consider a strong-willed consumer, i.e., $\tau(i) = S$, who has the following preferences over different choice sets:

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

Consequently, seller A will buy good A if

$$\tilde{u}_A \geq \tilde{u}_{A*} = p_A.$$

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

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

Choosing B from the menu $\{B, \emptyset\}$ is optimal if buying B delivers higher utility:

$$-p_B > -\gamma_i \bar{v},$$

which is equivalent to

$$\gamma_i > \gamma_* = \frac{p_B}{\bar{v}}.$$

B.2 Proof of Proposition 3

Given the advertising and pricing strategies of seller A, the quantity of goods sold by seller A is

$$Q_A = (\pi_S + \pi_W) z_A (1 - H(\tilde{u}_{A*}/\bar{u})), \tag{7}$$

from which it follows that the seller's profit net of the advertisement cost is

$$\Pi_A = p_A (\pi_S + \pi_W) z_A (1 - H(\tilde{u}_{A*}/\bar{u})) - F\frac{z_A}{1 - z_A}. \tag{8}$$

Similarly, the quantity of goods sold by seller B is

$$Q_B = \pi_W z_B (1 - G(\gamma_*)), \tag{9}$$

and the net profit of seller B is

$$\Pi_B = p_B \pi_W z_B (1 - G(\gamma_*)) - F\frac{z_B}{1 - z_B}.$$

Technological feasibility requires that $z_A \geq 0$ and $z_B \geq 0$.

It then follows that the first order conditions for prices set by the two sellers are

$$\begin{aligned} p_A &: Q_A = \frac{p_A}{\bar{u}} (\pi_S + \pi_W) z_A h(\tilde{u}_{A*}/\bar{u}) 1_{\{0 \leq \tilde{u}_{A*}/\bar{u} \leq 1\}}, \\ p_B &: Q_B = \frac{p_B \pi_W z_B}{\bar{v}} g(\gamma_*) 1_{\{0 \leq \gamma_* \leq 1\}}. \end{aligned}$$

Given that the expected quantities sold by both sellers, Q_A and Q_B, are nonnegative, and the net profits with respect to prices are concave, since

$$\begin{aligned} \frac{d^2\Pi_A}{dp_A^2} &= -\frac{2}{\bar{u}} (\pi_S + \pi_W) z_A h(\tilde{u}_{A*}/\bar{u}) 1_{\{0 \leq \tilde{u}_{A*}/\bar{u} \leq 1\}} \leq 0, \\ \frac{d^2\Pi_B}{dp_B^2} &= -2\pi_W z_B g(\gamma_*) \frac{1}{\bar{v}} 1_{\{0 \leq \gamma_* \leq 1\}} \leq 0, \end{aligned}$$

it follows that prices will always be nonnegative.

The FOC of (8) with respect to z_A is

$$z_A : p_A Q_A = F\frac{z_A}{(1 - z_A)^2}, \tag{10}$$

and with respect to z_B is

$$z_B : p_B Q_B = F\frac{z_B}{(1 - z_B)^2}.$$

Then, we have, from the FOC for z_A, that

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

Similarly,

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

Since

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

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

Since γ_i and $\tilde{u}_A/\bar{u}$ are uniformly distributed, $g(\gamma_*) = 1$ and $G(\gamma_*) = \gamma_*$, and $h(\tilde{u}_{A*}/\bar{u}) = 1$ and $H(\tilde{u}_{A*}/\bar{u}) = \tilde{u}_{A*}/\bar{u}$, it follows that the FONCs for p_A and p_B reduce to

$$p_A \quad : \quad Q_A = \frac{p_A}{\bar{u}}(\pi_S + \pi_W) z_A 1_{\{0 \le \tilde{u}_{A*}/\bar{u} \le 1\}}, \tag{11}$$

$$p_B \quad : \quad Q_B = \frac{p_B \pi_W z_B}{\bar{v}} 1_{\{0 \le \gamma_* \le 1\}}. \tag{12}$$

Note from Proposition 2 that

$$\tilde{u}_{A*} = p_A, \text{ and } \gamma_* = \frac{p_B}{\bar{v}}.$$

For strong-willed consumers, there are two possibilities: $\tilde{u}_{A*}/\bar{u} \in [0,1]$ or $\tilde{u}_{A*}/\bar{u} \notin [0,1]$. If $\tilde{u}_{A*}/\bar{u} \notin [0,1]$, then either $p_A = 0$ or $p_A > \bar{u}$, neither of which generates revenue for seller A, and advertising is costly. Consequently, it must be the case that $\tilde{u}_{A*}/\bar{u} \in [0,1]$. Then, equations (7) and (11) imply that

$$p_A = \frac{1}{2}\bar{u}.$$

Similarly, for seller B, if $\gamma_* \notin [0,1]$, then either $p_B = 0$ or $p_B > \bar{v}$. Neither case generates any revenue, but advertising is costly. If $\gamma_* \in [0,1]$, then equations (9) and (12) imply

$$p_B = \frac{1}{2}\bar{v}.$$

From the FOCs for z_A and z_B, it then follows that z_A and z_B satisfy

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

Then, we have

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

Therefore, the equilibrium for the two sellers is unique. Note that if $z_A \le 0$, then seller A chooses to advertise to 0 consumers. Similarly, if $z_B \le 0$, then seller B chooses to advertise to 0 consumers.

B.3 Proof of Proposition 4

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 consumers, who would buy neither good A nor B. As strong-willed and weak-willed consumers have the same preference for good A and a consumer's purchase of good A is not affected by his purchase of good B even if the consumer is weak willed. As a result, seller A does not need to differentiate strong-willed and weak-willed consumers. We denote z_{AD} as the measure of strong-willed and weak-willed consumers, to which seller A advertises, and p_A as the price the seller sets. The sales to these consumers by seller A is given by

$$Q_{AD} = z_{AD} \left[1 - H\left(\tilde{u}_{A*D}/\bar{u}\right)\right].$$

It follows from Proposition 2 that strong-willed and weak-willed consumers have the same threshold $\tilde{u}_{A*D} = p_A \in [0, 1]$ for purchasing good A. Thus, the net profit of seller A is

$$\Pi_A = p_A z_{AD} \left[1 - H\left(\tilde{u}_{A*D}/\bar{u}\right)\right] - F\frac{z_{AD}}{1 - z_{AD}}.$$

Following the same proof for Proposition 3, it is optimal for seller A to set a price

$$p_A = \frac{1}{2}\bar{u}.$$

The FOC with respect to z_{AD} is the same:

$$\frac{\bar{u}}{4} = F\frac{1}{\left(1 - z_{AD}\right)^2},$$

which implies that

$$z_{AD} = 1 - 2\sqrt{\frac{F}{\bar{u}}}.$$

Like before, if $1 - \sqrt{\frac{4F}{\bar{u}}} \leq 0$, it is optimal for the seller to advertise to no consumers. That is, $z_{AD} = 0$. Furthermore, if $1 - \sqrt{\frac{4F}{\bar{u}}} > \pi_S + \pi_W$, then

$$z_{AD} = \pi_S + \pi_W.$$

We now consider the policies of seller B. As strong-willed consumers always reject good 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 with temptation type γ_i his full reservation value: $p_{FS}(\gamma_i) = \gamma_i \bar{v}$. Substituting for $p_{FS}(\gamma_i)$, we can express seller B's profit as

$$\Pi_B = \bar{v}\int_0^1 \gamma_i z_{FS}\left(d\gamma_i\right) - F\frac{\int_0^1 z_{FS}\left(d\gamma_i\right)}{1 - \int_0^1 z_{FS}\left(d\gamma_i\right)} \quad \text{such that} \quad \int_0^1 z_{FS}\left(d\gamma_i\right) \in \left[0, \pi_W\right],$$

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

It will be helpful to first characterize the optimal advertising, conditional on an allocation of total advertising capacity. With the potential of abusing the notation, we denote

$$\bar{z}_{FS} = \int_0^1 z_{FS}(d\gamma_i)$$

as the total advertising by seller B. Since seller B knows the temptation types of all consumers and consumers with stronger temptation are willing to pay higher prices, it follows that the solution will be a water-filling strategy that prioritizes strong temptation consumers. It follows that:

$$dz_{FS}(\gamma_i) = \begin{cases} 0, & \text{if } \gamma_i < \gamma_* \\ \pi_W d\gamma_i, & \text{if } \gamma_i \in (\gamma_*, 1] \end{cases}.$$

Therefore, the expected revenue of seller B reduces to $\bar{v}\int_{\gamma_*}^1 \pi_W \gamma_i d\gamma_i = \bar{v}\pi_W \frac{1-\gamma_*^2}{2}$, where $\gamma_* = 1 - \frac{\bar{z}_{FS}}{\pi_W}$, since $\bar{z}_{FS} \in [0, \pi_W]$. Consequently, the expected revenue of seller B from advertising is $\bar{v}\bar{z}_{FS}\left(1 - \frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)$, which is determined by the seller's total advertising $\bar{z}_{FS}$.

Consequently, the expected profit of seller B reduces to

$$\Pi_B = \bar{z}_{FS}\left(1 - \frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)\bar{v} - F\frac{\bar{z}_{FS}}{1-\bar{z}_{FS}} \quad \text{such that } \bar{z}_{FS} \in [0, \pi_W],$$

and the choice of advertising reduces to choosing $\bar{z}_{FS}$. The FOC for $\bar{z}_{FS}$ is

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

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

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

and $\gamma^* = 1 - \frac{\bar{z}_{FS}}{\pi_W}$. Again, if this solution to the FOC 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 $\bar{z}_{FS}$ as

$$(\pi_W - \bar{z}_{FS})(1-\bar{z}_{FS})^2 = \frac{\pi_W F}{\bar{v}}, \tag{13}$$

it follows that

$$(\pi_W - \bar{z}_{FS})^3 \le \frac{\pi_W F}{\bar{v}},$$

and consequently

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

Finally, applying the Implicit Function Theorem to equation (13), one also has that

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

B.4 Proof of Proposition 5

It is easy to verify that $z_{AD} \geq z_A$. Seller A chooses to advertise to consumers in the setting with data sharing when $\frac{4F}{\bar{u}} \leq 1$, and in the setting without data sharing when $\frac{1}{\pi_S+\pi_W}\frac{4F}{\bar{u}} \leq 1$. Without data sharing the probability of a strong-willed or weak-willed consumer being covered by seller A is z_A, while with data sharing the probability is $\frac{z_{AD}}{\pi_S+\pi_W}$. As $z_{AD} \geq z_A$ and $\pi_S + \pi_W \leq 1$, it follows that $\frac{z_{AD}}{\pi_S+\pi_W} \geq z_A$, and the inequality is strict if $z_{AD} > 0$. 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 the two settings with and without data sharing, seller A charges the same price p_A for its good to its consumers. As a result, the welfare of the consumers varies across the two settings, directly driven by the sellers' advertising intensity. In the baseline setting with anonymous consumers, the utilitarian welfare of strong-willed consumers is given by

$$\begin{aligned} U_S(z_A) &= \pi_S z_A \left(\int_{\tilde{u}_{A*}/\bar{u}}^{1} \tilde{u}_A dH(\tilde{u}_A/\bar{u}) - p_A(1 - H(\tilde{u}_{A*}/\bar{u})) \right) \\ &= \pi_S z_A \left(\frac{\bar{u}}{2}(1 + \tilde{u}_{A*}/\bar{u}) - p_A \right)(1 - \tilde{u}_{A*}/\bar{u}) = \frac{1}{8}\pi_S z_A \bar{u}, \end{aligned}$$

which increases with z_A, the advertising intensity of seller A. One can interpret z_A as the probability of each strong-willed consumer being covered by seller A's advertisement. In the setting with data sharing, seller A advertises to a measure of z_{AD} strong-willed and weak-willed consumers. Each strong-willed consumer has a probability $\frac{z_{AD}}{\pi_S+\pi_W}$ of being covered by the advertisement of seller A. Thus, following the same procedure as above, one can show that the strong-willed consumers' welfare is given by $U_S\left(\frac{z_{AD}}{\pi_S+\pi_W}\right)$. From Proposition 4, $U_S\left(\frac{z_{AD}}{\pi_S+\pi_W}\right) \geq U_S(z_A)$ and the inequality is strict if $z_{AD} > 0$.

In the setting without data sharing, the utilitarian welfare of weak-willed consumers is given by

$$\begin{aligned} U_W^{NS} &= \pi_W z_A \left(\int_{\tilde{u}_{A*}/\bar{u}}^{1} \tilde{u}_A dH(\tilde{u}_A/\bar{u}) - p_A(1 - H(\tilde{u}_{A*}/\bar{u})) \right) \\ &\quad + u_B \pi_W z_B - p_B \pi_W z_B (1 - G(\gamma_*)) - \pi_W z_B \int_0^{\gamma_*} \gamma_i \bar{v} dG(\gamma_i), \end{aligned}$$

which, after substituting for $H\left(\tilde{u}_{A*}/\bar{u}\right)$ and $G\left(\gamma_i\right)$, reduces to

$$U_W^{NS}=\pi_W z_A\left(\frac{1}{2}\left(1+\frac{\tilde{u}_{A*}}{\bar{u}}\right)\bar{u}-p_A\right)\left(1-\tilde{u}_{A*}/\bar{u}\right)+u_B\pi_W z_B-\pi_W p_B z_B\left(1-\gamma_*\right)-\frac{1}{2}\pi_W z_B\bar{v}\gamma_*^2.$$

By substituting for $\tilde{u}_{A*}/\bar{u}$ and γ_*, we arrive at

$$U_W^{NS}=\frac{1}{8}\pi_W z_A\bar{u}+\pi_W z_B u_B-\frac{1}{4}\pi_W z_B\left(\bar{v}+\frac{1}{2}\bar{v}\right)=\frac{1}{8}\pi_W z_A\bar{u}+\pi_W z_B u_B-\frac{3}{8}\pi_W z_B\bar{v}.$$

The welfare of weak-willed consumers is also increasing with z_A the advertising intensity of seller A, but is decreasing with z_B the advertising intensity of seller B.

In the setting with data sharing, the most tempted weak-willed consumers are targeted by seller B. Their utilitarian welfare is given by

$$\begin{aligned}U_W^{FS} &= z_{AD}\frac{\pi_W}{\pi_S+\pi_W}\frac{\bar{u}}{8}+u_B\int_{\gamma_*}^1\pi_W d\gamma_i-\bar{v}\int_{\gamma_*}^1\pi_W\gamma_i d\gamma_i\\ &= z_{AD}\frac{\pi_W}{\pi_S+\pi_W}\frac{\bar{u}}{8}+\bar{z}_{FS}u_B-\bar{z}_{FS}\left(1-\frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)\bar{v}.\end{aligned}$$

Increased targeting by seller A raises the welfare of weak-willed consumers with data sharing, while the more surgical targeting by seller B lowers it. Thus, the impact of big data on the welfare of weak-willed consumers is nuanced.

The difference in welfare is given by

$$U_W^{FS}-U_W^{NS}=\left(\frac{z_{AD}}{\pi_S+\pi_W}-z_A\right)\frac{\pi_W\bar{u}}{8}+\left(\bar{z}_{FS}-\pi_W z_B\right)u_B+\left(\frac{3}{8}z_B-\frac{\bar{z}_{FS}}{\pi_W}\left(1-\frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)\right)\pi_W\bar{v}.$$

The impact of full data sharing on the welfare of weak-willed consumers is negative if

$$\frac{3\bar{v}}{\bar{u}}\left(\frac{4}{3}\frac{\bar{z}_{FS}}{\pi_W}\left(2-\frac{\bar{z}_{FS}}{\pi_W}\right)-z_B\right)+\frac{8}{\bar{u}}\left(z_B-\frac{\bar{z}_{FS}}{\pi_W}\right)u_B>\frac{z_{AD}}{\pi_S+\pi_W}-z_A, \tag{14}$$

and positive otherwise.

Notice that $\left(z_B-\frac{\bar{z}_{FS}}{\pi_W}\right)u_B\geq 0$ since $u_B<0$ and $\frac{\bar{z}_{FS}}{\pi_W}>z_B$. That $\frac{\bar{z}_{FS}}{\pi_W}>z_B$ follows from revealed preference since $\frac{\bar{z}_{FS}}{\pi_W}$ represents targeting efficiency. Since seller B can target the most tempted weak-willed customers, it maximizes the revenue extraction, while with anonymous customers its expected revenue is lower since the expected revenue per targeted consumer is lower. Since $\frac{8}{\bar{u}}\left(z_B-\frac{\bar{z}_{FS}}{\pi_W}\right)u_B>0$, it immediately follows, since z_{AD}, z_A, $\bar{z}_{FS}$ and z_B are independent of u_B, that for u_B sufficiently negative, the welfare of weak-willed consumers is lower with data sharing.

Notice now from Propositions 3 and 4 that

$$\begin{aligned}&\frac{z_{AD}}{\pi_S+\pi_W}-z_A\\ =\;&\min\left\{\max\left\{\frac{1}{\pi_S+\pi_W}\left(1-\sqrt{\frac{4F}{\bar{u}}}\right),0\right\},1\right\}-\min\left\{\max\left\{1-\sqrt{\frac{1}{\pi_S+\pi_W}\frac{4F}{\bar{u}}},0\right\},1\right\},\end{aligned}$$

which is independent of $\vec{v}$ and u_B.

Consider the case when both $\bar{z}_{FS}$ and z_B have interior solutions. Since $z_B = \max\left\{1 - \sqrt{\frac{1}{\pi_W}\frac{4F}{\bar{v}}}, 0\right\}$, differentiating the LHS of equation (14) with respect to $\bar{v}$, it follows that

$$\begin{aligned}\frac{dLHS}{d\bar{v}} &= \frac{3}{\bar{u}}\left(\frac{4}{3}\frac{\bar{z}_{FS}}{\pi_W}\left(2 - \frac{\bar{z}_{FS}}{\pi_W}\right) + \left(\frac{1}{2} + \frac{4}{3}\frac{u_B}{\bar{v}}\right)\sqrt{\frac{4F}{\pi_W \bar{v}}} - 1\right)\\ &\quad + \frac{8\bar{v}}{\bar{u}}\left(1 - \frac{\bar{z}_{FS}}{\pi_W} - \frac{u_B}{\bar{v}}\right)\frac{1}{\pi_W}\frac{d\bar{z}_{FS}}{d\bar{v}}\\ &\geq \frac{3}{\bar{u}}\left(\frac{4}{3}\frac{\bar{z}_{FS}}{\pi_W}\left(2 - \frac{\bar{z}_{FS}}{\pi_W}\right) + \left(\frac{1}{2} + \frac{4}{3}\frac{u_B}{\bar{v}}\right)\sqrt{\frac{4F}{\pi_W \bar{v}}} - 1\right)\end{aligned}$$

since $\frac{\bar{z}_{FS}}{\pi_W} \leq 1$, $u_B \leq 0$, and $\frac{d\bar{z}_{FS}}{d\bar{v}} \geq 0$.

We now recognize that $\frac{4}{3}\frac{\bar{z}_{FS}}{\pi_W}\left(2 - \frac{\bar{z}_{FS}}{\pi_W}\right)$ is concave and increasing in $\bar{z}_{FS}$ and, by Proposition 4, in $\bar{v}$. Notice that when $\frac{\bar{z}_{FS}}{\pi_W} \geq \frac{1}{2}$ that $\frac{4}{3}\frac{\bar{z}_{FS}}{\pi_W}\left(2 - \frac{\bar{z}_{FS}}{\pi_W}\right) \geq 1$, and $\frac{dLHS}{d\bar{v}} \geq 0$. Since, from Proposition 4, $\frac{\bar{z}_{FS}}{\pi_W} \geq 1 - \sqrt[3]{\frac{F}{\pi_W^2 \bar{v}}}$, it is therefore sufficient that

$$1 - \sqrt[3]{\frac{F}{\pi_W^2 \bar{v}}} \geq \frac{1}{2}, \text{ or } \bar{v} \geq \frac{8F}{\pi_W^2},$$

for $\frac{\bar{z}_{FS}}{\pi_W} \geq \frac{1}{2}$. In addition, if $\bar{v} \geq -\frac{8u_B}{3}$, then $\frac{1}{2} + \frac{4}{3}\frac{u_B}{\bar{v}} \geq 0$. Consequently, it is sufficient that $\bar{v} \geq 8\max\left\{\frac{F}{\pi_W^2}, \frac{-u_B}{3}\right\}$ for $\frac{dLHS}{d\bar{v}} \geq 0$.

Notice now that $z_B > 0$ when $\bar{v} > \frac{4F}{\pi_W}$. Substituting $\bar{v} = \frac{4F}{\pi_W}$ into our expression for $\bar{z}_{FS}$ from Proposition 4, it follows that

$$\begin{aligned}\bar{z}_{FS} &= 1 - \frac{1-\pi_W}{3} - \sqrt[3]{\left(\frac{1-\pi_W}{3}\right)^3 + \frac{\pi_W^2}{8} + \sqrt{\left(\left(\frac{1-\pi_W}{3}\right)^3 + \frac{\pi_W^2}{8}\right)^2 - \left(\frac{1-\pi_W}{3}\right)^6}}\\ &\quad - \sqrt[3]{\left(\frac{1-\pi_W}{3}\right)^3 + \frac{\pi_W^2}{8} - \sqrt{\left(\left(\frac{1-\pi_W}{3}\right)^3 + \frac{\pi_W^2}{8}\right)^2 - \left(\frac{1-\pi_W}{3}\right)^6}} > 0,\end{aligned}$$

for any π_W. Consequently, $\bar{z}_{FS} > 0$ whenever $z_B > 0$, since $\bar{z}_{FS}$ is increasing in $\bar{v}$ and $\bar{v} \geq 8\max\left\{\frac{F}{\pi_W^2}, \frac{-u_B}{3}\right\}$.

When $\bar{z}_{FS} > 0$ and $z_B = 0$, then the LHS of equation (14) is $\frac{3\bar{v}}{\bar{u}}\frac{4}{3}\frac{\bar{z}_{FS}}{\pi_W}\left(2 - \frac{\bar{z}_{FS}}{\pi_W}\right) - \frac{8}{\bar{u}}\frac{\bar{z}_{FS}}{\pi_W}u_B$, which is increasing in $\bar{v}$.

Lastly, when $\frac{\bar{z}_{FS}}{\pi_W} = 1$, then the LHS of equation (14) is $\frac{3\bar{v}}{\bar{u}}\left(\frac{4}{3} - z_B\right) - \frac{8}{\bar{u}}\left(1 - z_B\right)u_B$, which is again increasing in $\bar{v}$ since $z_B \leq 1$. Consequently, it is sufficient that $\bar{v} \geq 8\max\left\{\frac{F}{\pi_W^2}, \frac{-u_B}{3}\right\}$ for the LHS of equation (14) to be increasing in $\bar{v}$.

Since $\frac{z_{AD}}{\pi_S+\pi_W} - z_A$ is independent of $\bar{v}$, while the LHS of equation (14) is increasing in $\bar{v}$ for $\bar{v} \geq \frac{8F}{\pi_W^2}$, and diverges to infinity as $\bar{v}$ goes to infinity, it follows that there exists a critical level $\bar{v}^*$ such that equation (14) is satisfied for $\bar{v} > \bar{v}^*$. That is, provided that $\bar{v}$ is sufficiently high, the welfare of weak-willed consumers is lower with full data sharing. This holds since the u_B term is nonnegative.

Finally, from (1), social welfare with anonymous consumers is given by

$$\begin{aligned} W^{NS} &= \frac{u_A}{4}(\pi_S+\pi_W) z_A + \pi_W z_B u_B \int_{\gamma_*}^{1} d\gamma_i + \pi_W z_B \int_0^{\gamma_*} (u_B - \gamma_i \bar{v})\, d\gamma_i \\ &= \frac{\bar{u}}{4}(\pi_S+\pi_W) z_A + \pi_W z_B u_B (1-\gamma_*) + \pi_W z_B \left(u_B\gamma_* - \frac{\gamma_*^2}{2}\bar{v}\right) \\ &= \frac{\bar{u}}{4}(\pi_S+\pi_W) z_A + \pi_W z_B \left(u_B - \frac{1}{8}\bar{v}\right) \end{aligned}$$

since $p_A = \frac{\bar{u}}{2}$ and $\gamma_* = \frac{p_B}{\bar{v}} = \frac{1}{2}$, while with data sharing, it is instead

$$W^{FS} = \frac{\bar{u}}{4} z_{AD} + \pi_W u_B \int_{\gamma^*}^{1} d\gamma_i = \frac{\bar{u}}{4} z_{AD} + \bar{z}_{FS} u_B.$$

It then follows that

$$W^{FS} - W^{NS} = \bar{z}_{FS} u_B - \pi_W z_B \left(u_B - \frac{1}{8}\bar{v}\right) - \frac{\bar{u}}{4}(\pi_S+\pi_W)\left(z_A - \frac{z_{AD}}{\pi_S+\pi_W}\right).$$

By similar arguments to those above, if u_B is sufficiently negative, then $W^{FS} - W^{NS} < 0$. Interestingly, because $\bar{v}$ impacts the cost of resisting temptation, for $\bar{v}$ sufficiently high, then $W^{FS} - W^{NS} > 0$.

B.5 Proof of Proposition 6

We start with the optimal strategy of seller A. Suppose that seller A advertises to $z_{A,in}$ measure of consumers who opt in at price $p_{A,in}$ and $z_{A,out}$ measure of consumers who opt out at price $p_{A,out}$. Then, the seller's expected profit, by the law of large numbers, is given by

$$\begin{aligned} \Pi_A \;&=\; \frac{(1-\gamma_{**})\pi_W}{(1-\gamma_{**})\pi_W + 1 - \pi_S - \pi_W} p_{A,out} z_{A,out}\left[1 - H\left(\tilde{u}_{A*,out}/\bar{u}\right)\right] + p_{A,in} z_{A,in}\left[1 - H\left(\tilde{u}_{A*,in}/\bar{u}\right)\right] \\ &\quad - F\frac{z_{A,out}+z_{A,in}}{1-z_{A,out}-z_{A,in}}, \\ s.t. \;&:\; z_{A,out} \in [0, 1-\pi_S-\gamma_{**}\pi_W], \text{ and } z_{A,in} \in [0, \pi_S + \gamma_{**}\pi_W], \end{aligned}$$

where, from Proposition 2, $\tilde{u}_{A*,in} = p_{A,in}$ and $\tilde{u}_{A*,out} = p_{A,out}$. We can rewrite the seller's profit as

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

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

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

Then, the seller's profit becomes

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

Note that the marginal profit to $z_{A,in}$ is strictly higher than that to $z_{A,out}$ because consumers in the opt-in pool are precisely identified while those in the opt-out pool may belong to the third-type that won't buy good A. Thus, seller A gives higher priority to the opt-in pool.

The FOC with respect to $z_{A,in}$ gives

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

As $z_{A,in}$ has higher priority than $z_{A,out}$, we have

$$z_{A,in} = \min\left\{\max\left\{1-2\sqrt{\frac{F}{\bar{u}}}, 0\right\}, \pi_S + \gamma_{**}\pi_W\right\}.$$

To ensure that $z_{A,in} > 0$, we impose the parameter restriction:

$$F < \frac{\bar{u}}{4}.$$

Then,

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

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

$$\frac{(1-\gamma_{**})\,\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\frac{\bar{u}}{4} - F\frac{1}{(1-z_{A,out}-z_{A,in})^2} \geq 0 \;\; (= \;\; if \;\; z_{A,out} < 1-\pi_S-\gamma_{**}\pi_W)$$

Thus,

$$z_{A,out} = \begin{cases} 0 & \text{if } z_{A,in} < \pi_S + \gamma_{**}\pi_W \\ \begin{array}{c}\min\{\max\{1-2\sqrt{\frac{1-\pi_S-\gamma_{**}\pi_W}{(1-\gamma_{**})\pi_W}\frac{F}{\bar{u}}} - \pi_S - \gamma_{**}\pi_W, 0\}, \\ 1-\pi_S-\gamma_{**}\pi_W\}\end{array} & \text{if } z_{A,in} = \pi_S + \gamma_{**}\pi_W \end{cases}.$$

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 $z_{B,in}(\gamma_i)$ to consumers who opt in with temptation coefficient γ_i, at price $p_{B,in}(\gamma_i)$, and $z_{B,out}$ measure of consumers who opt out at price $p_{B,out}$. The seller's profit is

$$\begin{aligned}
\Pi_B &= \int_0^{\gamma_{**}} p_{B,in}(\gamma_i) z_{B,in}(d\gamma_i) + \frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W} p_{B,out} z_{B,out}\left(1-G\left(\gamma_{*,out},[\gamma_{**},1]\right)\right) \\
&\quad -F\frac{z_{B,out}+\int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)}{1-z_{B,out}-\int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)}, \\
s.t. &: z_{B,out}\in[0,1-\pi_S-\gamma_{**}\pi_W], \text{ and } \int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)\in[0,\gamma_{**}\pi_W],
\end{aligned}$$

where $G\left(\gamma_{*,out},[\gamma_{**},1]\right) = \frac{\gamma_{*,out}-\gamma_{**}}{1-\gamma_{**}} 1_{\{\gamma_{*,out}\geq\gamma_{**}\}}$ as the truncated distribution of the weak-willed that opt out, and $[0,\gamma_{**}]$ as the truncated distribution of the weak-willed that opt in. In addition, we recognize from Proposition 2, that $\gamma_{*,out} = \frac{p_{B,out}}{\bar{v}}$. Furthermore, since seller B knows an opt-in consumer's temptation type, γ_i, it can perfectly discriminate against its consumers and exercise first-degree price discrimination by charging each consumer's full reservation utility: $p_{B,in}(\gamma_i) = \gamma_i\bar{v}$. Substituting for these distributions and expressions, we can express seller B's profit as

$$\begin{aligned}
\Pi_B &= \bar{v}\int_0^{\gamma_{**}} \gamma_i z_{B,in}(d\gamma_i) \\
&\quad +\frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W} z_{B,out} p_{B,out}\left[\left(1-\frac{p_{B,out}}{\bar{v}}\right)1_{\{p_{B,out}\geq\gamma_{**}\bar{v}\}}+(1-\gamma_{**})1_{\{p_{B,out}<\gamma_{**}\bar{v}\}}\right] \\
&\quad -F\frac{z_{B,out}+\int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)}{1-z_{B,out}-\int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)}, \\
s.t. &: z_{B,out}\in[0,1-\pi_S-\gamma_{**}\pi_W], \text{ and } \int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)\in[0,\gamma_{**}\pi_W],
\end{aligned}$$

where $\int_0^{\gamma_{**}} \gamma_i z_{B,in}(d\gamma_i)$ is understood as a Riemann-Stieljes integral. If $z_{B,out} > 0$, then the FOC for $p_{B,out}$ gives the following:

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

Thus, the optimal price satisfies

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

By substituting the price into the objective, we now determine the seller's optimal advertising policy:

$$\begin{aligned}
\Pi_B &= \bar{v}\int_0^{\gamma_{**}} \gamma_i z_{B,in}(d\gamma_i) + \frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\left[\frac{1}{4}-\left(\gamma_{**}-\frac{1}{2}\right)^2 1_{\left\{\gamma_{**}>\frac{1}{2}\right\}}\right]\bar{v}z_{B,out} \\
&\quad -F\frac{z_{B,out}+\int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)}{1-z_{B,out}-\int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)} \\
s.t. &: \; z_{B,out}\in[0,1-\pi_S-\gamma_{**}\pi_W], \text{ and } \int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)\in[0,\gamma_{**}\pi_W].
\end{aligned}$$

It will be helpful to first characterize the optimal advertising to the opt-in pool, conditional on an allocation of total advertising capacity. With the potential of abusing the notation, we denote

$$\bar{z}_{B,in} = \int_0^{\gamma_{**}} z_{B,in}(d\gamma_i)$$

as the total advertising to the opt-in pool by seller B. Since seller B knows the temptation types of consumers in the opt-in pool, γ_i, and consumers with stronger temptation are willing to pay higher prices with $p_{B,in}(\gamma_i)=\gamma_i\bar{v}$, it follows that the solution will be a water-filling strategy that prioritizes strong temptation consumers. It follows that:

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

Therefore, the expected revenue of seller B from the opt-in pool reduces to $\bar{v}\int_{\gamma_*}^{\gamma_{**}}\pi_W\gamma_i d\gamma_i = \bar{v}\pi_W\frac{\gamma_{**}^2-\gamma_*^2}{2}$, where $\gamma_*=\gamma_{**}-\frac{\bar{z}_{B,in}}{\pi_W}$, since $\bar{z}_{B,in}\in[0,\gamma_{**}\pi_W]$. Consequently, the expected revenue of seller B from advertising to the opt-in pool is $\bar{v}\bar{z}_{B,in}\left(\gamma_{**}-\frac{1}{2}\frac{\bar{z}_{B,in}}{\pi_W}\right)$, which is determined by the seller's total advertising to the opt-in pool $\bar{z}_{B,in}$.

Consequently, the expected profit of seller B reduces to

$$\begin{aligned}
\Pi_B &= \bar{z}_{B,in}\left(\gamma_{**}-\frac{1}{2}\frac{\bar{z}_{B,in}}{\pi_W}\right)\bar{v} + \frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\left[\frac{1}{4}-\left(\gamma_{**}-\frac{1}{2}\right)^2 1_{\left\{\gamma_{**}>\frac{1}{2}\right\}}\right]\bar{v}z_{B,out} \\
&\quad -F\frac{z_{B,out}+\bar{z}_{B,in}}{1-z_{B,out}-\bar{z}_{B,in}} \\
s.t. &: \; z_{B,out}\in[0,1-\pi_S-\gamma_{**}\pi_W], \text{ and } \bar{z}_{B,in}\in[0,\gamma_{**}\pi_W],
\end{aligned}$$

and the choice of advertising reduces to choosing $\bar{z}_{B,in}$ and $z_{B,out}$. The revenue from the opt-in pool is concave in the aggregate advertising to the opt-in pool, $\bar{z}_{B,in}$, 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 $z_{B,out}$. It follows that the

FOCs for the advertising intensities are:

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

$$z_{B,out} \quad : \quad \frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\left[\frac{1}{4} - \left(\gamma_{**}-\frac{1}{2}\right)^2 1_{\{\gamma_{**}>\frac{1}{2}\}}\right]\bar{v} - F\frac{1}{(1-z_{B,out}-\bar{z}_{B,in})^2}$$

$$\begin{cases} = 0 & \text{if } z_{B,out} \in (0, 1-\pi_S-\pi_W\gamma_{**}) \\ < 0 & \text{if } z_{B,out} = 0 \\ > 0 & \text{if } z_{B,out} = 1-\pi_S-\pi_W\gamma_{**} \end{cases}.$$

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

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

which is equivalent to

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

Note that this range exists under the condition $\frac{1-\pi_S}{\pi_W} > 1$, which implies that the upper end of this range is above $\frac{1}{2}$. Thus, the opt-in pool has higher priority if

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

Since $\frac{1-\pi_S}{2\pi_W} - \sqrt{\left(\frac{1-\pi_S}{2\pi_W}\right)^2 - \frac{1}{4}} \leq \frac{1}{2}$, this set is nonempty, although it may be singular at $\frac{1}{2}$.

When $\gamma_{**} > \frac{1}{2}$, the opt-in pool has priority whenever $\pi_W < 1 - \pi_S$. Therefore, under the condition that $\pi_W < 1 - \pi_S$, which is imposed throughout this paper, the opt-in pool has priority whenever

$$\gamma_{**} \geq \frac{1-\pi_S}{2\pi_W} - \sqrt{\left(\frac{1-\pi_S}{2\pi_W}\right)^2 - \frac{1}{4}}. \tag{15}$$

It is clear that if the opt-out pool has higher priority, seller B will devote all resources to the opt-out pool before the opt-in pool:

$$
\begin{aligned}
z_{B,out} &= \min\left\{\max\left\{1-\sqrt{\frac{F}{\pi_W \bar{v}}\frac{1-\pi_S-\gamma_{**}\pi_W}{\frac{1}{4}-\left(\gamma_{**}-\frac{1}{2}\right)^2 1_{\{\gamma_{**}>\frac{1}{2}\}}}},\ 0\right\},\ 1-\pi_S-\gamma_{**}\pi_W\right\}, \\
\bar{z}_{B,in} &= \min\left\{\max\left\{1-\sqrt{\frac{F}{\pi_W \bar{v}}\frac{1-\pi_S-\gamma_{**}\pi_W}{\frac{1}{4}-\left(\gamma_{**}-\frac{1}{2}\right)^2 1_{\{\gamma_{**}>\frac{1}{2}\}}}}-z_{B,out}, 0\right\},\ \pi_W\gamma_{**}\right\}. \qquad (16)
\end{aligned}
$$

If, instead, the opt-in pool has priority, then seller B will devote resources to the opt-in pool until the marginal product of the two pools are equal or its FOC condition is satisfied, depending on which occurs first. Since the marginal revenue of the opt-in pool decreases from $\bar{v}\gamma_{**}$ to 0, the two marginal products will intersect at the unique level $\bar{z}_{B,in} = z_*$, where

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

Now define $x = 1-\bar{z}_{B,in}$, and the FOC for $\bar{z}_{B,in}$ when $z_{B,out}=0$ is

$$
x^3 - \left(1-\pi_W\gamma_{**}\right)x^2 - \frac{\pi_W F}{\bar{v}} = 0,
$$

which gives a unique, positive root

$$
\bar{z}_{B,in} = 1 - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} + \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{\left(1-\pi_W\gamma_{**}\right)^3}{27}}} - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} - \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{\left(1-\pi_W\gamma_{**}\right)^3}{27}}}.
$$

Consequently, it follows that

$$
\begin{aligned}
\bar{z}_{B,in} = \min\Bigg\{ & 1 - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} + \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{\left(1-\pi_W\gamma_{**}\right)^3}{27}}} \\
& - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} - \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{\left(1-\pi_W\gamma_{**}\right)^3}{27}}},\ z_* \Bigg\}.
\end{aligned}
$$

This second case also characterizes the solution when both $\bar{z}_{B,in} > 0$ and $z_{B,out} > 0$, with

$$\begin{aligned}
\bar{z}_{B,in} &= \min\left\{1 - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} + \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{(1-\pi_W\gamma_{**})^3}{27}}}\right. \\
&\qquad \left. - \sqrt[3]{\frac{\pi_W F}{2\bar{v}} - \sqrt{\left(\frac{\pi_W F}{2\bar{v}}\right)^2 - \frac{(1-\pi_W\gamma_{**})^3}{27}}}, z_*\right\} \\
z_{B,out} &= \min\left\{\max\left\{1 - \sqrt{\frac{F}{\pi_W \bar{v}}\frac{1-\pi_S-\gamma_{**}\pi_W}{\frac{1}{4} - \left(\gamma_{**}-\frac{1}{2}\right)^2 1_{\left\{\gamma_{**}>\frac{1}{2}\right\}}}} - z_*, 0\right\}, 1-\pi_S-\gamma_{**}\pi_W\right\} \quad (17)
\end{aligned}$$

Consequently, these are two possible solutions to the advertising intensities, depending on whether the opt-in or opt-out pool has higher priority. Then, given γ_{**}, the advertising policy of seller B exists and is unique.

We now consider a weak-willed consumer with temptation index γ_i. The expected utility from opting in is

$$\begin{aligned}
U_{W,in}(\gamma_i) &= \frac{z_{A,in}}{\pi_S + \gamma_{**}\pi_W}\int_{\frac{p_A}{\bar{u}}}^{1} (\tilde{u}_A - p_A)\, d\left(\frac{\tilde{u}_A}{\bar{u}}\right) + \frac{z_{B,in}(\gamma_i)}{\pi_W}(u_B - \gamma_i \bar{v}) \\
&= \frac{z_{A,in}}{\pi_S + \gamma_{**}\pi_W}\frac{\bar{u}}{8} + \frac{z_{B,in}(\gamma_i)}{\pi_W}(u_B - \gamma_i\bar{v}).
\end{aligned}$$

This expression shows that $U_{W,in}$ increases with $z_{A,in}$ but decreases with $z_{B,in}(\gamma_i)$. His expected utility from opting out is

$$\begin{aligned}
U_{W,out}(\gamma_i) &= \frac{z_{A,out}}{1-\pi_S-\gamma_{**}\pi_W}\int_{\frac{p_A}{\bar{u}}}^{1} (\tilde{u}_A - p_A)\, d\left(\frac{\tilde{u}_A}{\bar{u}}\right) \\
&\quad + \frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\left(u_B - p_{B,out}\mathbf{1}_{\left\{\gamma_i > \frac{p_{B,out}}{\bar{v}}\right\}} - \gamma_i\bar{v}\mathbf{1}_{\left\{\gamma_i \le \frac{p_{B,out}}{\bar{v}}\right\}}\right) \\
&= \frac{z_{A,out}}{1-\pi_S-\gamma_{**}\pi_W}\frac{\bar{u}}{8} + \frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}u_B \\
&\quad - \frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\bar{v}\left[\max\left\{\frac{1}{2},\gamma_{**}\right\}\mathbf{1}_{\left\{\gamma_i>\max\left\{\frac{1}{2},\gamma_{**}\right\}\right\}} + \gamma_i\mathbf{1}_{\left\{\gamma_i\le\max\left\{\frac{1}{2},\gamma_{**}\right\}\right\}}\right],
\end{aligned}$$

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

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

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

Notice if $\bar{z}_{B,in} = 0$, then $U_{W,in}(\gamma_i) > U_{W,out}(\gamma_i)$ for all γ_i, and $\gamma_{**} = 1$. Consequently, from equation (15), the opt-in pool has priority. 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), $\bar{z}_{B,in} > 0$.

When $\bar{z}_{B,in} > 0$, then for the marginal consumer that opts in with temptation index γ_{**}, $z_{B,in}(\gamma_{**}) = \pi_W$ (probability 1 of being targeted in the opt-in pool), because seller B has the strongest incentive to cover the consumer with the strongest temptation and thus the highest willingness to pay. Since the marginal consumer must be indifferent to opting in and opting out, $U_{W,in}(\gamma_{**}) - U_{W,out}(\gamma_{**}) = 0$, which imposes that

$$\frac{\bar{u}}{8}\left[\frac{z_{A,in}}{\pi_S+\gamma_{**}\pi_W} - \frac{z_{A,out}}{1-\pi_S-\gamma_{**}\pi_W}\right] + (\bar{v}\gamma_{**} - u_B)\left[\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W} - 1\right] = 0.$$

Define $\tilde{\gamma} = \pi_S + \pi_W\gamma_{**}$. Then, we can rewrite this condition for γ_{**} when there is an interior solution as the solution to the cubic function:

$$\begin{aligned} 0 &= \tilde{\gamma}^3 - \left(1+\pi_S+\frac{\pi_W u_B}{\bar{v}} - z_{B,out}\right)\tilde{\gamma}^2 \\ &\quad + \left(\left(\pi_S+\frac{\pi_W u_B}{\bar{v}}\right)(1-z_{B,out}) - \frac{\pi_W\bar{u}}{8\bar{v}}(z_{A,in}+z_{A,out})\right)\tilde{\gamma} + \frac{\pi_W\bar{u}}{8\bar{v}}z_{A,in}, \end{aligned}$$

which we can further rewrite as the depressed cubic polynomial:

$$0 = x^3 + tx + q,$$

where $x = \tilde{\gamma} + \frac{1}{3}(1+\tilde{\pi}_S - z_{B,out})$, $\tilde{\pi}_S = \pi_S + \frac{\pi_W u_B}{\bar{v}}$ and

$$\begin{aligned} t &= -\frac{1}{3}\left((1-z_{B,out}-\tilde{\pi}_S)^2 + \tilde{\pi}_S(1-z_{B,out})\right) - \frac{\pi_W\bar{u}}{8\bar{v}}(z_{A,in}+z_{A,out}) < 0, \\ q &= \frac{1+\tilde{\pi}_S - z_{B,out}}{27}\left(\tilde{\pi}_S(1-z_{B,out}) - 2(1-z_{B,out}-\tilde{\pi}_S)^2\right) \\ &\quad + \frac{\pi_W\bar{u}}{8\bar{v}}z_{A,in} - \frac{\pi_W\bar{u}}{8\bar{v}}\frac{1+\tilde{\pi}_S - z_{B,out}}{3}(z_{A,in}+z_{A,out}). \end{aligned} \tag{18}$$

If $q < 0$, then there is one positive real root, while if $q > 0$, then there are 2 positive real roots. When there are two positive solutions, this depressed cubic has roots given by

$$x_k = 2\sqrt{-\frac{t}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2t}\sqrt{-\frac{3}{t}}\right) - \frac{2\pi k}{3}\right),\ k \in \{0,1,2\},$$

and consequently, since $k = 2$ corresponds to the negative root,

$$\gamma_{**k} = \min\left\{\frac{2}{\pi_W}\sqrt{-\frac{t}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2t}\sqrt{-\frac{3}{t}}\right) - \frac{2\pi k}{3}\right) - \frac{1+4\pi_S - z_{B,out}}{3\pi_W}, 1\right\},\ k \in \{0,1\},$$

since γ_{**} is bounded from above by 1. The first two roots are positive with the second root being the smaller positive root. Notice that both positive roots are continuous in $z_{B,in}$ and $z_{B,out}$.

For two positive roots to exist, we require that the discriminant of the depressed cubic polynomial be positive, which is the case if

$$\left(-\frac{t}{3}\right)^3 > \left(\frac{q}{2}\right)^2.$$

Writing

$$\begin{aligned} t &= -3\left(\frac{\tilde{\pi}_S}{3}\right)^2\left((\tilde{z}_{B,out}-1)^2+\tilde{z}_{B,out}\right)-\frac{\pi_W\bar{u}}{8\bar{v}}\left(z_{A,in}+z_{A,out}\right), \\ q &= 2\left(\frac{\tilde{\pi}_S}{3}\right)^3\frac{1+\tilde{z}_{B,out}}{2}\left(\tilde{z}_{B,out}-2\left(\tilde{z}_{B,out}-1\right)^2\right)+\frac{\pi_W\bar{u}}{8\bar{v}}z_{A,in} \\ &\quad -\frac{\pi_W\bar{u}}{8\bar{v}}\frac{1+\tilde{\pi}_S-z_{B,out}}{3}\left(z_{A,in}+z_{A,out}\right), \end{aligned}$$

and defining $\tilde{z}_A=\frac{1}{3}\frac{\pi_W\bar{u}}{8\bar{v}}\left(z_{A,in}+z_{A,out}\right)$, $\tilde{z}_{A,out}=\frac{1}{3}\frac{\pi_W\bar{u}}{8\bar{v}}z_{A,out}$, and $\tilde{z}_{B,out}=\frac{1-z_{B,out}}{\tilde{\pi}_S}\in\left(0,\frac{1}{\tilde{\pi}_S}\right)$, we can express t and q as

$$\begin{aligned} -\frac{t}{3} &= \frac{\tilde{\pi}_S^2}{9}\left((\tilde{z}_{B,out}-1)^2+\tilde{z}_{B,out}\right)+\tilde{z}_A \\ \frac{q}{2} &= \frac{\tilde{\pi}_S^3}{27}\frac{1+\tilde{z}_{B,out}}{2}\left(\tilde{z}_{B,out}-2\left(\tilde{z}_{B,out}-1\right)^2\right)+\frac{1}{2}\left(3-\tilde{\pi}_S\left(1+\tilde{z}_{B,out}\right)\right)\tilde{z}_A-\frac{3}{2}\tilde{z}_{A,out}, \end{aligned}$$

where $\tilde{z}_A\in\left[0,\frac{\pi_W\bar{u}}{24\bar{v}}\right]$ and $\tilde{z}_{A,out}\in\left[0,\tilde{z}_A\right]$. It then follows that

$$\begin{aligned} &\left(-\frac{t}{3}\right)^3-\left(\frac{q}{2}\right) \\ = &\left(\frac{\tilde{\pi}_S}{3}\right)^6\left[\left((\tilde{z}_{B,out}-1)^2+\tilde{z}_{B,out}\right)^3-\left(\frac{1+\tilde{z}_{B,out}}{2}\right)^2\left(\tilde{z}_{B,out}-2\left(\tilde{z}_{B,out}-1\right)^2\right)^2\right] \\ &+\frac{\tilde{\pi}_S^4}{27}\left[\left(1+\tilde{z}_{B,out}^2-\tilde{z}_{B,out}\right)^2+\left(1+\tilde{z}_{B,out}\right)\left(1+\tilde{z}_{B,out}^2-\frac{5}{2}\tilde{z}_{B,out}\right)\left(\frac{3}{\tilde{\pi}_S}-1-\tilde{z}_{B,out}\right)\right]\tilde{z}_A \\ &+\tilde{z}_A^3+\frac{\tilde{\pi}_S^2}{3}\left((\tilde{z}_{B,out}-1)^2+\tilde{z}_{B,out}\right)\tilde{z}_A^2+\frac{\tilde{\pi}_S^3}{9}\frac{1+\tilde{z}_{B,out}}{2}\left(\tilde{z}_{B,out}-2\left(\tilde{z}_{B,out}-1\right)^2\right)\tilde{z}_{A,out} \\ &-\frac{1}{4}\left(\left(3-\tilde{\pi}_S\left(1+\tilde{z}_{B,out}\right)\right)\tilde{z}_A-3\tilde{z}_{A,out}\right)^2. \end{aligned}$$

Comparing the first two terms we recognize that

$$\left((\tilde{z}_{B,out}-1)^2+\tilde{z}_{B,out}\right)^3-\left(\frac{1+\tilde{z}_{B,out}}{2}\right)^2\left(\tilde{z}_{B,out}-2\left(\tilde{z}_{B,out}-1\right)^2\right)^2\geq 0,$$

which is straightforward to verify since it is a function of only $\tilde{z}$. Similarly, it is straightforward to confirm that the second term is nonnegative whenever $\tilde{\pi}_S \geq 1.05$. Consider the special case that $\tilde{z}_{A,out} = 0$. Suppose $\bar{v} \geq \frac{\pi_W \bar{u}}{24}$ and $\tilde{\pi}_S \geq 1.05$, then since $\tilde{z}_A < 1$

$$\begin{aligned} & \left(-\frac{t}{3}\right)^3 - \left(\frac{q}{2}\right) \\ \geq \; & \frac{\tilde{\pi}_S^4}{27}\left[\left(1+\tilde{z}_{B,out}^2 - \tilde{z}_{B,out}\right)^2 + \left(1+\tilde{z}_{B,out}\right)\left(1+\tilde{z}_{B,out}^2 - \frac{5}{2}\tilde{z}_{B,out}\right)\left(\frac{3}{\tilde{\pi}_S} - 1 - \tilde{z}_{B,out}\right)\right]\tilde{z}_A^2 \\ & + \left[\frac{\tilde{\pi}_S^2}{3}\left(1+\tilde{z}_{B,out}^2 - \tilde{z}_{B,out}\right) - \frac{1}{4}\left(3 - \tilde{\pi}_S\left(1+\tilde{z}_{B,out}\right)\right)^2\right]\tilde{z}_A^2 + \tilde{z}_A^3. \end{aligned}$$

Then, one can verify that it is now sufficient for $\tilde{\pi}_S \geq 1.8$ for the first two terms to be nonnegative. It is therefore sufficient in this special case that $\tilde{\pi}_S = \pi_S + \frac{\pi_W u_B}{\bar{v}} \geq 1.8$, or $u_B \geq \frac{1.8-\pi_S}{\pi_W}\bar{v}$, for a positive real solution to the cubic equation to always exist.

Notice that $\tilde{z}_{A,out} = 0$ when $\pi_S \geq 1 - 2\sqrt{\frac{F}{\bar{u}}}$. Consequently, if $\pi_S \geq 1 - 2\sqrt{\frac{F}{\bar{u}}}$, $\bar{u} \leq \frac{24\bar{v}}{\pi_W}$, and $u_B \geq \frac{1.8-\pi_S}{\pi_W}\bar{v}$ then a solution to the cubic equation always exists.

We next establish existence with our additional assumptions. Let $\Omega = [0,1] \times [0, \pi_W \gamma_{**}] \times [0, 1-\pi_S - \pi_W\gamma_{**}]$ be the compact, convex set in which the equilibrium triple $(\gamma_{**}, \bar{z}_{B,in}, z_{B,out})$ lies. Notice the maps from γ_{**} to $(\bar{z}_{B,in}, z_{B,out})$ are upper-semicontinuous if we allow for the correspondences to be set-valued at the single point, $\frac{1-\pi_S}{2\pi_W} - \sqrt{\left(\frac{1-\pi_S}{2\pi_W}\right)^2 - \frac{1}{4}}$, at which the opt-in and opt-out pools change priority. Furthermore, they always map to a $(\bar{z}_{B,in}, z_{B,out})$ in the set $[0, \pi_W\gamma_{**}] \times [0, 1-\pi_S-\pi_W\gamma_{**}]$. Similarly, $(z_{B,in}, z_{B,out}) \rightarrow \gamma_{**}$ is a continuous, compact-valued correspondence on $[0,1]$ in the two positive roots. It follows that we have an upper-semicontinuous correspondence from $\Omega \rightarrow \Omega$. Then, by Kakutani's Fixed Point Theorem, an equilibrium exists.

We now verify the optimality of a cutoff strategy for weak-willed customers to opt-in. In what follows in the comparison of $U_{W,in}(\gamma_i) - U_{W,out}(\gamma_i)$, we will substitute the indifference condition $U_{W,in}(\gamma_{**}) = U_{W,out}(\gamma_{**})$ to simplify the resulting expression for the difference in opt-in and opt-out utilities

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

We begin with $\gamma_i > \gamma_{**}$. We recognize that $z_{B,in}(\gamma_i) = \pi_W$, since off-equilibrium seller B would immediately and fully target consumers with stronger temptation than γ_{**}, since it can extract $(\gamma_i - \gamma_{**})\bar{v}$ more in revenue. Suppose first $\gamma_{**} \geq \frac{1}{2}$, then all $\gamma_i > \gamma_{**}$ are in the opt-out pool, and the opt-in/opt-out condition for γ_i reduces to

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

and all $\gamma_i > \gamma_{**}$ opt out when $\gamma_{**} \geq \frac{1}{2}$. If instead, $\gamma_{**} < \frac{1}{2}$, and further if $\gamma_i \leq \frac{1}{2}$

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

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

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

since $\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W} \leq 1$ and $\gamma_i > \frac{1}{2}$. Therefore, all $\gamma_i > \gamma_{**}$ opt out when $\gamma_{**} < \frac{1}{2}$. Consequently, all $\gamma_i > \gamma_{**}$ opt out.

We now consider $\gamma_i < \gamma_{**}$. Notice for $\gamma_i < \gamma_*$, the threshold γ_i below which seller B does not advertise ($z_{B,in}(\gamma_i) = 0$ for $\gamma_i < \gamma_*$), it is trivial that $U_{W,in}(\gamma_i) - U_{W,out}(\gamma_i) > 0$, since the consumer benefits from both higher search by seller A and lower (zero) search by seller B. Consequently, all $\gamma_i < \gamma_*$ opt in. For $\gamma_i \in [\gamma_*, \gamma_{**}]$, $z_{B,in}(\gamma_i) = \pi_W$, and their opt-in/opt-out condition (since $\gamma_i \leq \gamma_{**}$) reduces to

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

since $\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W} \leq 1$ and $\gamma_i < \gamma_{**}$. Therefore, all $\gamma_i \in [\gamma_*, \gamma_{**}]$ opt in. Consequently, all $\gamma_i < \gamma_{**}$ opt in, verifying the optimality of the cutoff opt-in/opt-out strategy. Importantly, the optimality of the cutoff strategy holds regardless of the equilibrium value of γ_{**}.

B.6 Proof of Proposition 7

Seller B's profit is

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

Substituting with the FOCs with complementary slackness

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

we arrive at

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

That seller B's profit is lower with the opt-in / opt-out policy than with big data follows since seller B could always implement the opt-in / opt-out pricing and advertising policy with big data if it were optimal. Since the big data policy is different, it follows that

$$\Pi_B^{OP} \leq \Pi_B^{FS}.$$

Since seller B under opt-in / opt-out cannot replicate the optimal pricing and advertising policy with big data, it is a constrained optimization, and it follows that generically this inequality is sharp.

That seller B's profit with opt-in / opt-out is higher than with anonymous customers follows by an analogous argument. It therefore follows that

$$\Pi_B^{FS} > \Pi_B^{OP} > \Pi_B^{NS}.$$

For a weak-willed consumer with temptation coefficient γ_i, his utility is determined by the probabilities of being covered by both sellers A and B, which we denote ρ_A and ρ_B:

$$U_{\gamma_i} = \rho_A \left(\int_{p_A/\bar{u}}^{1} \tilde{u}_A - p_A dH\left(\tilde{u}_A\right) \right) + \rho_B \left[u_B - p_B 1_{\{\gamma_i > p_B/\bar{v}\}} - \gamma_i \bar{v} 1_{\{\gamma_i \leq p_B/\bar{v}\}} \right].$$

We compare all three settings: anonymous consumers, full data sharing, and opt-in / opt-out.

In the setting without any data sharing, $p_A = \frac{1}{2}\bar{u}$, $\rho_A = z_A = 1 - 2\sqrt{\frac{1}{\pi_S + \pi_W}\frac{F}{\bar{u}}}$, $p_B = \frac{1}{2}\bar{v}$, and $\rho_B = z_B = 1 - 2\sqrt{\frac{1}{\pi_W}\frac{F}{\bar{v}}}$. Thus,

$$\begin{aligned} U_W^{NS} &= \pi_W \left(1 - 2\sqrt{\frac{1}{\pi_S + \pi_W}\frac{F}{\bar{u}}} \right) \int_{1/2}^{1} \left(\tilde{u}_A - \bar{u}/2\right) dH\left(\tilde{u}_A\right) \\ &\quad + \pi_W \left(1 - 2\sqrt{\frac{1}{\pi_W}\frac{F}{\bar{v}}} \right) \left[u_B - \frac{1}{4}\bar{v} - \int_0^{1/2} \gamma_i \bar{v} d\gamma_i \right] \\ &= \frac{1}{8}\pi_W \left(1 - 2\sqrt{\frac{1}{\pi_S + \pi_W}\frac{F}{\bar{u}}} \right) \bar{u} + \pi_W \left(1 - 2\sqrt{\frac{1}{\pi_W}\frac{F}{\bar{v}}} \right) \left(u_B - \frac{3}{8}\bar{v} \right). \end{aligned}$$

In the setting with full data sharing, $p_A = \frac{1}{2}\bar{u}$, $\rho_A = \frac{1}{\pi_S + \pi_W} z_{AD} = \min\left\{ \frac{1 - 2\sqrt{\frac{F}{\bar{u}}}}{\pi_S + \pi_W}, 1 \right\}$, $p_B(\gamma_i) = \gamma_i \bar{v}$, and $\rho_B = \mathbf{1}_{\left\{\gamma_i \geq \gamma_* = 1 - \frac{\bar{z}_{FS}}{\pi_W}\right\}}$. Thus,

$$U_W^{FS} = \frac{1}{8}\pi_W \min\left\{ \frac{1 - 2\sqrt{\frac{F}{\bar{u}}}}{\pi_S + \pi_W}, 1 \right\} \bar{u} + \bar{z}_{FS} u_B - \bar{z}_{FS} \left(1 - \frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W} \right) \bar{v}$$

In the setting with opt-in/opt-out, in the opt-out pool, $p_A = \frac{1}{2}\bar{u}$, $\rho_A = \frac{z_{A,out}}{1 - \pi_S - \gamma_{**}\pi_W}$, $p_{B_{out}} =$

$\max\left\{\frac{1}{2},\gamma_{**}\right\}\bar{v}$, and $\rho_B = \frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}$. Thus,

$$\begin{aligned}
U_W^{OP} &= \left[z_{A,in}\frac{\gamma_{**}\pi_W}{\pi_S+\gamma_{**}\pi_W} + z_{A,out}\frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\right]\frac{\bar{u}}{8} - \bar{v}\int_{\gamma_*}^{\gamma_{**}}\pi_W\gamma_i d\gamma_i \\
&\quad +\pi_W(\gamma_{**}-\gamma_*)u_B + z_{B,out}\frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W}u_B \\
&\quad -z_{B,out}\frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\left[\int_{\gamma_{**}}^{\max\left\{\frac{1}{2},\gamma_{**}\right\}}\gamma_i d\gamma_i + \left[\frac{1}{4}-\left(\gamma_{**}-\frac{1}{2}\right)^2 1_{\{\gamma_{**}>1/2\}}\right]\right]\bar{v} \\
&= \left[z_{A,in}\frac{\gamma_{**}\pi_W}{\pi_S+\gamma_{**}\pi_W} + z_{A,out}\frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\right]\frac{\bar{u}}{8} - \bar{z}_{B,in}\left(\gamma_{**}-\frac{1}{2}\frac{\bar{z}_{B,in}}{\pi_W}\right)\bar{v} \\
&\quad +\left(\bar{z}_{B,in} + z_{B,out}\frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\right)u_B \\
&\quad -z_{B,out}\frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\left\{\begin{array}{c}\frac{1}{2}\left[\max\left\{\frac{1}{2},\gamma_{**}\right\}^2-\gamma_{**}^2\right] \\ +\left(1-\max\left\{\frac{1}{2},\gamma_{**}\right\}\right)\max\left\{\frac{1}{2},\gamma_{**}\right\}\end{array}\right\}\bar{v},
\end{aligned}$$

since $\gamma_* = \gamma_{**} - \frac{\bar{z}_{B,in}}{\pi_W}$. Substituting the indifference condition for the marginal weak-willed customer γ_{**}, the above can be rewritten as

$$\begin{aligned}
U_W^{OP} &= z_{A,in}\frac{\pi_W}{\pi_S+\gamma_{**}\pi_W}\frac{\bar{u}}{8} - \bar{v}(1-\gamma_{**})\pi_W\gamma_{**} - \bar{z}_{B,in}\left(\gamma_{**}-\frac{1}{2}\frac{\bar{z}_{B,in}}{\pi_W}\right)\bar{v} \\
&\quad +\bar{v}z_{B,out}\frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\left(\left(1-\frac{1}{2}\gamma_{**}\right)\gamma_{**}+\frac{1}{2}\max\left\{\frac{1}{2},\gamma_{**}\right\}^2-\max\left\{\frac{1}{2},\gamma_{**}\right\}\right) \\
&\quad +\left(\pi_W\left(\frac{\bar{z}_{B,in}}{\pi_W}-\gamma_{**}\right)+z_{B,out}\frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\right)u_B \\
&= z_{A,in}\frac{\pi_W}{\pi_S+\gamma_{**}\pi_W}\frac{\bar{u}}{8} - \bar{v}(1-\gamma_{**})\pi_W\gamma_{**} - \bar{z}_{B,in}\left(\gamma_{**}-\frac{1}{2}\frac{\bar{z}_{B,in}}{\pi_W}\right)\bar{v} \\
&\quad -z_{B,out}\frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\left(\frac{3}{8}-\left(1-\frac{1}{2}\gamma_{**}\right)\gamma_{**}\right)\bar{v}\mathbf{1}_{\left\{\gamma_{**}<\frac{1}{2}\right\}},
\end{aligned}$$

which can be rewritten as

$$\begin{aligned}
U_W^{OP} &= \frac{1}{8}\pi_W\min\left\{\frac{1-2\sqrt{\frac{F}{\bar{u}}}}{\pi_S+\gamma_{**}\pi_W},1\right\}\bar{u} - \left[1-\left(1-\frac{\bar{z}_{B,in}}{\pi_W\gamma_{**}}\right)\gamma_{**}-\frac{1}{2}\left(\frac{\bar{z}_{B,in}}{\pi_W\gamma_{**}}\right)^2\gamma_{**}\right]\pi_W\gamma_{**}\bar{v} \\
&\quad -\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\left(\frac{3}{8}-\left(1-\frac{1}{2}\gamma_{**}\right)\gamma_{**}\right)\pi_W\bar{v}\mathbf{1}_{\left\{\gamma_{**}<\frac{1}{2}\right\}} \\
&\quad +\left(\pi_W\left(\frac{\bar{z}_{B,in}}{\pi_W}-\gamma_{**}\right)+z_{B,out}\frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\right)u_B.
\end{aligned}$$

Since, fixing γ_{**}, the second term achieves its minimum when $\frac{\bar{z}_{B,in}}{\pi_W\gamma_{**}}=1$, it follows that we can bound the second term from below by

$$-\left[1-\left(1-\frac{\bar{z}_{B,in}}{\pi_W\gamma_{**}}\right)\gamma_{**}-\frac{1}{2}\left(\frac{\bar{z}_{B,in}}{\pi_W\gamma_{**}}\right)^2\gamma_{**}\right]\pi_W\gamma_{**}\bar{v} < -\pi_W\gamma_{**}\left(1-\frac{1}{2}\gamma_{**}\right)\bar{v},$$

and consequently

$$
\begin{aligned}
U_W^{OP} \geq\ & \frac{1}{8}\pi_W \min\left\{\frac{1-2\sqrt{\frac{F}{\bar{u}}}}{\pi_S+\pi_W},1\right\}\bar{u}-\pi_W\gamma_{**}\left(1-\frac{1}{2}\gamma_{**}\right)\bar{v} \\
& -\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\left(\frac{3}{8}-\left(1-\frac{1}{2}\gamma_{**}\right)\gamma_{**}\right)\pi_W\bar{v}\mathbf{1}_{\{\gamma_{**}<\frac{1}{2}\}} \\
& +\left(\pi_W\left(\frac{\bar{z}_{B,in}}{\pi_W}-\gamma_{**}\right)+z_{B,out}\frac{\pi_W}{1-\pi_S-\gamma_{**}\pi_W}\right)u_B.
\end{aligned}
$$

Since $U_W^{FS}=\frac{1}{8}\pi_W\min\left\{\frac{1-2\sqrt{\frac{F}{\bar{u}}}}{\pi_S+\pi_W},1\right\}\bar{u}+\bar{z}_{FS}u_B-\bar{z}_{FS}\left(1-\frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)\bar{v}$, we arrive at

$$
\begin{aligned}
U_W^{OP} \geq\ & U_W^{FS}+\bar{z}_{FS}\left(1-\frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)\bar{v}-\pi_W\gamma_{**}\left(1-\frac{1}{2}\gamma_{**}\right)\bar{v} \\
& -\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\left(\frac{3}{8}-\left(1-\frac{1}{2}\gamma_{**}\right)\gamma_{**}\right)\pi_W\bar{v}\mathbf{1}_{\{\gamma_{**}<\frac{1}{2}\}} \\
& +\left(\bar{z}_{B,in}+z_{B,out}\frac{\pi_W(1-\gamma_{**})}{1-\pi_S-\gamma_{**}\pi_W}-\bar{z}_{FS}+\left(\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}-1\right)\pi_W\gamma_{**}\right)u_B.
\end{aligned}
$$

Notice that $\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\leq 1$, since at most the opt-out pool has full coverage of unity. Further notice that

$$
\bar{z}_{B,in}+z_{B,out}\frac{\pi_W(1-\gamma_{**})}{1-\pi_S-\gamma_{**}\pi_W}\leq \bar{z}_{FS},
$$

since with full data sharing, seller B can target and price discriminate against the highest temptation consumers and can always choose to pursue the same policy as with opt-in / opt-out. By revealed preference, it would (weakly) choose a higher level of total advertising coverage. Consequently, the last term is nonnegative, since $u_B\leq 0$.

Consider the case when $\gamma_{**}>\frac{1}{2}$. Then, since the last term is nonnegative, the above reduces to

$$
U_W^{OP}\geq U_W^{FS}+\frac{\bar{z}_{FS}}{\pi_W}\left(1-\frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)\pi_W\bar{v}-\gamma_{**}\left(1-\frac{1}{2}\gamma_{**}\right)\pi_W\bar{v}.
$$

Notice if $\frac{\bar{z}_{FS}}{\pi_W}>\gamma_{**}$, then $\frac{\bar{z}_{FS}}{\pi_W}\left(1-\frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)>\gamma_{**}\left(1-\frac{1}{2}\gamma_{**}\right)$, since $\frac{\bar{z}_{FS}}{\pi_W}\left(1-\frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)$ is monotonically increasing in $\frac{\bar{z}_{FS}}{\pi_W}$. From Proposition 4,

$$
\frac{\bar{z}_{FS}}{\pi_W}\geq 1-\sqrt[3]{\frac{F}{\pi_W^2\bar{v}}},
$$

and consequently it is sufficient that

$$
\gamma_{**}\leq 1-\sqrt[3]{\frac{F}{\pi_W^2\bar{v}}}.
$$

The RHS of this condition is increasing in $\bar{v}$, while γ_{**} is (weakly) decreasing in $\bar{v}$, since stronger temptation reduces the benefit of opting-in for all consumers. As such, there exists a critical level $\bar{v}_*$, such that $U_W^{OP} > U_W^{FS}$ for all $\bar{v} > \bar{v}_*$. Since γ_{**} is decreasing in $\bar{v}$, however, the bound may not be satisfied for $\gamma_{**} \geq \frac{1}{2}$. Note that $U_W^{OP} > U_W^{FS}$ may hold for a wider range of $\bar{v}$ than above $\bar{v}_*$ since our bound is not tight on the difference in utility.

Suppose now that $\gamma_{**} < \frac{1}{2}$, then, since $\frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W} \leq 1$ and $-\left(\frac{3}{8} - \left(1 - \frac{1}{2}\gamma_{**}\right)\gamma_{**}\right) \leq 0$, it follows since the last term is nonnegative that

$$U_W^{OP} \geq U_W^{FS} + \bar{z}_{FS}\left(1 - \frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)\bar{v} - \frac{3}{8}\pi_W\bar{v}.$$

Similarly, notice if $\frac{\bar{z}_{FS}}{\pi_W} > \frac{1}{2}$, then $\frac{\bar{z}_{FS}}{\pi_W}\left(1 - \frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right) > \frac{3}{8}$, since $\frac{\bar{z}_{FS}}{\pi_W}\left(1 - \frac{1}{2}\frac{\bar{z}_{FS}}{\pi_W}\right)$ is monotonically increasing in $\frac{\bar{z}_{FS}}{\pi_W}$. From Proposition 4,

$$\frac{\bar{z}_{FS}}{\pi_W} \geq 1 - \sqrt[3]{\frac{F}{\pi_W^2\bar{v}}},$$

and consequently it is sufficient that

$$\bar{v} > \frac{8F}{\pi_W^2}.$$

If $\bar{v} > \frac{8F}{\pi_W^2}$, then

$$U_W^{OP} > U_W^{FS}.$$

Comparing these two conditions, it follows that there exists a $\bar{v}_* \geq \frac{8F}{\pi_W^2}$ (the max of the two critical lower bounds on $\bar{v}$), such that, if $\bar{v} \geq \bar{v}_*$, then

$$U_W^{OP} > U_W^{FS}.$$

Comparing to the anonymous consumer case, we recognize that

$$\begin{aligned}
U_W^{OP} - U_W^{NS} &= \frac{1}{8}\pi_W\left(\min\left\{\frac{1-2\sqrt{\frac{F}{\bar{u}}}}{\pi_S+\gamma_{**}\pi_W}, 1\right\} - z_A\right)\bar{u} + \frac{3}{8}\pi_W z_B\bar{v} \\
&\quad - \left[1 - \left(1 - \frac{\bar{z}_{B,in}}{\pi_W\gamma_{**}}\right)\gamma_{**} - \frac{1}{2}\left(\frac{\bar{z}_{B,in}}{\pi_W\gamma_{**}}\right)^2\gamma_{**}\right]\gamma_{**}\pi_W\bar{v} \\
&\quad - \frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\left(\frac{3}{8} - \left(1 - \frac{1}{2}\gamma_{**}\right)\gamma_{**}\right)\pi_W\bar{v}\mathbf{1}_{\{\gamma_{**}<\frac{1}{2}\}} \\
&\quad + \left(\bar{z}_{B,in} + z_{B,out}\frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W} - \pi_W\left(1 - 2\sqrt{\frac{1}{\pi_W}\frac{F}{\bar{v}}}\right)\right)u_B.
\end{aligned}$$

where $z_B = \max\left\{1 - 2\sqrt{\frac{1}{\pi_W}\frac{F}{\bar{v}}}, 0\right\}$. With opt-in, the first term is nonnegative, since there is more efficient matching with seller A with data sharing. Notice as $\bar{v}$ becomes arbitrarily large, then $\gamma_{**} \to 0$, since the utility cost to being targeted by seller B becomes arbitrarily high, and with some manipulation

$$U_W^{OP} - U_W^{NS} \to \left(\frac{z_{A,out}}{1-\pi_S} - z_A\right)\frac{1}{8}\pi_W \bar{u} + \left(\frac{z_{B,out}}{1-\pi_S} - z_B\right)\pi_W u_B = \left(\frac{z_{A,out}}{1-\pi_S} - z_A\right)\frac{1}{8}\pi_W \bar{u} < 0,$$

since $z_B = \frac{z_{B,out}}{1-\pi_S} = 1$ and, because all weak-willed opt-out while all strong-willed opt-in, $\frac{z_{A,out}}{1-\pi_S} - z_A < 0$. Consequently, for $\bar{v}$ sufficiently large, $U_W^{OP} < U_W^{NS}$.

For any $\bar{v}$, the utility cost with opt-in/opt-out must be higher (the negative $\bar{v}$ terms) than that without data sharing, $\frac{3}{8}\pi_W z_B \bar{v}$, since the utility cost directly corresponds to the rent extraction of seller B. Seller B could always implement then anonymous customer pricing and advertising policies, and do strictly better than without data sharing because all strong-willed customers opt-in. Since seller B pursues different pricing and advertising strategies with opt-in/opt-out, its rent extraction must therefore be (weakly) higher. As such

$$\begin{aligned}
&\frac{3}{8}\pi_W z_B - \left[1 - \left(1 - \frac{\bar{z}_{B,in}}{\pi_W \gamma_{**}}\right)\gamma_{**} - \frac{1}{2}\left(\frac{\bar{z}_{B,in}}{\pi_W \gamma_{**}}\right)^2 \gamma_{**}\right]\gamma_{**} \\
&+ \left(\bar{z}_{B,in} + z_{B,out}\frac{(1-\gamma_{**})\pi_W}{1-\pi_S-\gamma_{**}\pi_W} - \pi_W\left(1 - 2\sqrt{\frac{1}{\pi_W}\frac{F}{\bar{v}}}\right)\right) u_B \\
&\quad - \frac{z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\left(\frac{3}{8} - \left(1 - \frac{1}{2}\gamma_{**}\right)\gamma_{**}\right)\mathbf{1}_{\left\{\gamma_{**} < \frac{1}{2}\right\}} \leq 0.
\end{aligned}$$

It then follows for there exists a critical $\bar{v}_{**}$ such that, for $\bar{v} > \bar{v}_{**}$, then $U_W^{OP} < U_W^{NS}$. From the limiting case as $\bar{v}$ becomes arbitrarily large, we know that $U_W^{OP} < U_W^{NS}$, and consequently the difference between U_W^{OP} and U_W^{NS} does not vanish asymptotically because strong-willed consumers still opt-in to benefit from increased matching with seller A.

Since opt-ing in exposes weak-willed consumers to a higher likelihood of suffering the commitment utility u_B than the anonymous case, and less so than in the full-data sharing case, by analogous arguments to those for $\bar{v}$ and for u_B in Proposition 5, there exists a u_{B*} and a u_{B**} such that, if $u_B > u_{B*}$, then $U_W^{OP} > U_W^{FS}$, and if $u_B > u_{B**}$, then $U_W^{OP} < U_W^{NS}$.[14]

From a social welfare perspective, aggregating the welfare of strong and weak-willed consumers, as well as the revenues of both sellers, we arrive at

$$\begin{aligned}
W^{OP} \;=\; & \left((\pi_S + \pi_W \gamma_{**})\frac{z_{A,in}}{\pi_S + \gamma_{**}\pi_W} + (1-\gamma_{**})\pi_W \frac{z_{A,out}}{1-\pi_S-\gamma_{**}\pi_W}\right)\frac{\bar{u}}{4} + \bar{z}_{B,in} u_B \\
& + \frac{\pi_W z_{B,out}}{1-\pi_S-\gamma_{**}\pi_W}\left((1-\gamma_{**}) u_B - \bar{v}\left(\frac{1}{8} - \frac{\gamma_{**}^2}{2}\right)\mathbf{1}_{\left\{\gamma_{**} \leq \frac{1}{2}\right\}}\right).
\end{aligned}$$

Comparing this to social welfare under no information sharing

$$W^{NS} = \frac{\bar{u}}{4}(\pi_S + \pi_W) z_A + \pi_W z_B \left(u_B - \frac{\bar{v}}{8}\right).$$

[14] For instance, at each $\bar{v}_*$ such $U_W^{OP} = U_W^{FS}$, then for $u_B + \varepsilon$, it must be the case $U_W^{OP} > U_W^{FS}$. In such a manner, one can define $\bar{v}_* (u_B)$ and $u_{B*}(v)$.

Notice that when $u_B = \bar{v} = 0$, then it must be the case that $W^{OP} > W^{NS}$. For sufficiently high $\bar{v}$ or negative u_B, then only the weak-willed population with the mildest γ_i opt-in, and are left alone by seller B because it prioritizes the opt-out pool ($\bar{z}_{B,in} = 0$). The social benefit of opt-in / opt-out for increased matching with seller A then accrues as $z_{A,in}\frac{\bar{u}}{4}$, which is bounded from above. In contrast, with γ_{**} small, the cost to the weak-willed in the opt-opt pool becomes arbitrarily large. Since they have less camouflage than in the no information sharing case because the strong-willed and mildly weak-willed all opt-in, then $W^{OP} < W^{NS}$. Since the objectives are continuous, it follows that there exist critical values of u_B and $\bar{v}$, u_B^{**} and $\bar{v}^{**}$, such that $W^{OP} < W^{NS}$ when $u_B \leq u_B^{**}$ or $v \geq \bar{v}^{**}$.

Comparing W^{OP} to social welfare under full data sharing

$$W^{FS} = \frac{\bar{u}}{4} z_{AD} + \bar{z}_{FS} u_B,$$

we recognize that both z_{AD} and $\bar{z}_{FS}$ are independent of u_B from Proposition 4. Consequently, as u_B decreases, W^{FS} becomes arbitrarily negative. It is instructive to consider two limits. First, when $u_B \to 0$, then $W^{FS} > W^{OP}$ since there is no social cost of temptation, since the cost of the temptation good represents a zero-sum transfer. Second, when u_B becomes arbitrarily negative, then again only the weak-willed population with the mildest γ_i opt-in, and are left alone by seller B because it prioritizes the opt-out pool ($\bar{z}_{B,in} = 0$). In this case, to first-order

$$\begin{aligned} W^{FS} &\approx \bar{z}_{FS} u_B, \\ W^{OP} &\approx \pi_W (1 - \gamma_{**}) \frac{z_{B,out}}{1 - \pi_S - \gamma_{**}\pi_W} u_B, \end{aligned}$$

ignoring $\bar{u}$ and $\bar{v}$ terms, which can be made an order of magnitude smaller. Since there is still the indifferent population in the opt-out pool and, conditional on γ_{**}, seller B does not respond to a more negative u_B by searching more (higher $z_{B,out}$), it follows that $W^{OP} > W^{FS}$. Consequently, there exists a u_B^* such that $W^{OP} > W^{FS}$ if $u_B < u_B^*$.

References

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

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

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

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

Alan, Sule, Mehmet Cemalcilar, Dean Karlan, and Jonathan Zinman (2018), Unshrouding: Evidence from Bank Overdrafts in Turkey, Journal of Finance 73, 481-522.

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

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

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

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

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

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

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

De Hert, Paul, Vagelis Papakonstantinou, Gianclaudio Malgieri, Laurent Beslay, and Ignacio Sanchez (2018), The Right to Data Portability in the GDPR: Towards User-centric Interoperability of Digital Services. Computer Law & Security Review 34, 193-203.

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

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

DellaVigna, Stefano and Ulrike Malmandier (2004), Contract Design and Self-Control: Theory and Evidence, Quarterly Journal of Economics 119, 353-402.

Gabaix, Xavier, and David Laibson (2006), Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets, Quarterly Journal of Economics 121, 505-540.

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

Grubb, Michael D. and Matthew Osborne (2015), Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock, American Economic Review 105, 234-271.

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

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

Heidhues, Paul and Botond Koszegi (2017), Naivete- Based Discrimination, Quarterly Journal of Economics, 1019-1054.

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

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

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

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

Loveman, Gary W. (2003), Diamonds in the Data Mine, Harvard Business Review. May.

Marr, Bernard (2015), Big Data at Caesars Entertainment - A One Billion Dollar Asset?, Forbes, May 18, at https://www.forbes.com/sites/bernardmarr/2015/05/18/when-big-data-becomes-your-most-valuable-asset/#770ee66f1eef.

Matz, S. C., M. Kosinski, G. Nave, and D. J. Stillwell (2017), Psychological Targeting as an Effective Approach to Digital Mass Persuasion, Proceedings of the National Academy of Sciences of the United States of America, 114 (48).

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

Murgia, Madhumita (2018), Popular Apps Share Data with Facebook without User Consent, Financial Times, December 31.

Newzoo, April 2018 Quarterly Update, at www.newzoo.com/globalgamesreport.

OCC (Office of the Comptroller of the Currency). 2018. Installment Lending: Core Lending Principles for Short-Term, Small-Dollar Installment Lending. Available at https://www.occ.gov/news-issuances/bulletins/2018/bulletin-2018-14.html.

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

Ru, Hong and Antoinette Schoar (2017), Do Credit Card Companies Screen for Behavioral Biases? mimeo MIT Sloan School of Management and Nanyang Technological University.

Stango, Victor and Jonathan Zinman, Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees, Review of Financial Studies 27, 990-1030.

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

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

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

The Article 29 Working Party (2017), WP29 Guidelines on the Right to Data Portability in the GDPR, Available at https://iapp.org/media/pdf/resource% _center/WP29-2017-04-data-portability-guidance.pdf.

The Guardian (2019), Online Casino Advert Banned for Targeting Problem Gamblers (Oct. 2019), at http://theguardian.com/society/2019/oct/09/casumo-ad-banned-for-targeting-people-trying-to-stop-gambling.

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

Upturn (2015), Led Astray: Online Lead Generation and Payday Loans (Oct. 2015), https://www.teamupturn.com/reports/2015/led-astray [https://perma.cc/42VF-VEW4].