Revealing Transactions Data to Third Parties: Implications of Privacy Regimes for Welfare in Online Markets

by

Michael R. Baye* and David E. M. Sappington**

Abstract

We examine the effects of privacy policies regarding transactions (e.g., price/quantity) data on online shopping platforms. Disclosure of transactions data induces consumer signaling behavior that affects merchant pricing decisions and the welfare of platform participants. A profit-maximizing platform prefers the disclosure policy that maximizes social welfare. Although this policy benefits sophisticated (fully rational) consumers, it harms unsophisticated consumers. Consequently, the welfare effects of alternative privacy policies, data breaches, willful violations of stated privacy policies, and opt-in/opt-out requirements differ sharply, depending on the level of consumer sophistication and on other factors such as the prevailing status quo.

JEL Nos: D04, D18, D4, D6, D8, L00, L5, Keywords: Information Economics, Privacy, Signaling, Consumer Protection

October 2018

* Kelley School of Business, Indiana University, 1309 East Tenth Street, Bloomington, IN 47401 (mbaye@indiana.edu).

** Department of Economics, P.O. Box 117140, University of Florida, Gainesville, Florida 32611-7140 (sapping@ufl.edu).

We thank Paul Belleflamme, Rick Harbaugh, Jeff Prince, Dan Sacks, Curtis Taylor, and Liad Wagman for helpful discussions. We are indebted to Roger Blair for suggesting a related line of research and assisting with early model formulation.

1 Introduction

Price transparency is considered a virtue in competitive markets. At least since Adam Smith (1776), economists have recognized that prices provide valuable signals about how and where to allocate resources to produce the goods and services that society values most highly. The benefits of a transparent price system are memorialized in the first- and second-fundamental theorems of welfare economics.[1] However, at least since George Stigler (1964), the literature has also recognized that price transparency can be a two-edged sword in settings with more limited competition. In particular, price transparency can facilitate collusion in part by enabling oligopolistic sellers to better detect and punish departures from jointly-profitable super-competitive prices.[2]

More recently, the literature has examined the potential gains and losses from making more transparent price and quantity data from consumer transactions. The literature demonstrates that restricting the use or limiting the sharing of transactions data can either benefit or harm consumers.[3] For example, restricting a platform's ability to track the purchases of individual consumers can harm consumers by limiting the platform's ability to efficiently match consumers with products and/or advertisers (Evans, 2009).[4] However, the same privacy policy can benefit consumers by preventing a monopolist from exploiting in one transaction information it learns about a customer in a different transaction (Acquisti and Varian, 2005). The literature also notes that the effects of such a privacy policy can depend on whether consumers are sophisticated, i.e., whether they fully anticipate how their present purchase decisions may affect the prices they face in subsequent transactions (Taylor, 2004).[5]

We analyze a model that includes many of the key elements of Taylor (2004)'s and Acquisti and Varian (2005)'s models to examine the welfare effects of several platform privacy

[1] See Blaug (2007), for instance.

[2] Asker, Fershtman, Jeon, and Pakes (2016) provide both an excellent overview of this literature and a novel contribution to the literature.

[3] Taylor and Wagman (2014) employ four common models of oligopoly competition to demonstrate that winners and losers from privacy policies can vary with the prevailing form of market competition. Acquisti, Taylor and Wagman (2016) provide an excellent survey of the literature on privacy.

[4] Campbell, Goldfarb, and Tucker (2015) demonstrate that policies that require firms to protect consumer data (e.g., prevent third parties from accessing it) can harm consumers by placing smaller firms at a competitive disadvantage, thereby affecting market structure adversely.

[5] Tucker (2012) provides an excellent survey of the empirical literature on these and other tradeoffs.

policies. In particular, we examine the effects of mandates to adopt the privacy policy that maximizes the welfare of particular types of consumers (either sophisticated or unsophisticated consumers). We also analyze the performance of a *laissez faire* policy that allows a profit-maximizing online shopping platform (such as Amazon or eBay) to adopt its preferred privacy policy. In addition, we examine "opt-in" mandates (that require consumers to give their explicit consent before platforms can share transactions data with third parties) and "opt-out" mandates (that require platforms to allow consumers to request and thereby receive a personal exemption from default sharing of transactions data). We also analyze the effects of data breaches (caused by hackers, for example) and willful violations of announced platform policies.

In our model, consumers purchase two distinct (non-competing) products from different merchants on an online platform. When transactions data are shared on the platform, a consumer's interaction with one merchant may reveal to other merchants the consumer's reservation value for their products. The other merchants may modify the prices they charge the consumer accordingly. A sophisticated consumer who recognizes this effect of data sharing takes it into account when interacting with all merchants. In contrast, when he decides whether to purchase a merchant's product, an unsophisticated consumer only considers whether the price the merchant sets exceeds his reservation value for the product.

We find that sophisticated consumers and the platform generally benefit when the platform shares all transactions data with third parties (i.e., other merchants on the platform). The data sharing provides a channel through which sophisticated consumers can credibly signal when their reservation values for the merchants' products are low. Such signaling induces price concessions from merchants.[6] When the platform does not share transactions data, it effectively closes the signaling channel, thereby harming sophisticated consumers. Closing the channel also can reduce platform profit and total welfare by limiting the consummation

[6] Belleflamme and Vergote (2016) document in a distinct setting the consumer welfare gains that can arise when a merchant is better able to discern consumers' reservation values. The authors consider a setting where consumers interact once with a single monopolist. If she is not prevented from doing so, the monopolist can choose whether to discover consumers' reservation values for her product (with a specified probability). At personal cost, a consumer can eliminate the monopolist's ability to discern the consumer's personal reservation value. The authors show that consumers may be better off when they are unable to limit the monopolist's ability to discern reservation values.

of welfare-enhancing transactions.

In contrast, unsophisticated consumers benefit when the platform never shares transactions data with third parties. This privacy policy prevents merchants from exploiting unsophisticated consumers by charging them higher prices after they are observed to pay high prices to other merchants. Thus, the privacy policy that best serves unsophisticated consumers harms sophisticated consumers. Consequently, the formulation of privacy regulations for online platforms can be challenging even when the sole objective of the regulations is to maximize consumer welfare.

The formulation of opt-in and opt-out mandates can be similarly challenging. As others have recognized, the impact of such mandates often varies with the prevailing status quo (i.e., whether the platform's prevailing default policy is to share or not share transactions data with third parties).[7] The impact of opt-in and opt-out mandates also varies with the degree of consumer sophistication and with the magnitude of the costs that consumers must incur to opt in to or opt out of the platform's prevailing privacy policy.[8] Furthermore, findings that may seem counterintuitive can emerge. For instance, requiring explicit consumer consent before transactions data are shared with third parties can harm sophisticated consumers.

The effects of data breaches and willful violations of announced platform privacy policies also can vary with the prevailing status quo and the degree of consumer sophistication. To illustrate, when the platform announces it will implement the privacy policy that maximizes the welfare of unsophisticated consumers, these consumers are harmed both by a data breach and by a willful violation of the platform's announced privacy policy. However, not all consumers are harmed, and harm only arises under certain configurations of merchant costs.[9]

Different conclusions arise when consumers are sophisticated. In particular, when the platform adopts the policy that maximizes the welfare of sophisticated consumers, a data breach does not harm consumers (because their transactions data are already known to all

[7] For example, see Federal Trade Commission (2009).

[8] Athey, Catalini, and Tucker (2017) demonstrate that even small costs of opting in or opting out can greatly affect consumers' privacy choices.

[9] Specifically, consumers with low reservation values for the merchants' products are not harmed. A consumer with a high reservation value is harmed when the data from his transaction with a high-cost merchant is shared with a low-cost merchant (who would set a low price for her product in the absence of data sharing).

merchants on the platform). In contrast, a willful violation of this privacy policy can harm sophisticated consumers by foreclosing the signalling channel through which they can secure price concessions. Therefore, the effects of data breaches and violations of privacy policies can differ for sophisticated and unsophisticated consumers. In addition, the effects of data breaches can differ from the effects of willful violations of stated privacy policies.

We find that total social welfare, platform profits, and the welfare of sophisticated consumers are maximized when the platform provides transactions data to third parties. Consequently, under a *laissez faire* policy that permits the platform to implement its preferred privacy policy (and potentially charge fees for these data), the platform will adopt the privacy policy that maximizes the welfare of sophisticated consumers. This privacy policy is not ideal for unsophisticated consumers, however. It is also not the best policy for all merchants.

We also examine the impact of removing all information about a consumer's identity before transactions data are shared with third parties. The removal of such personal information can benefit unsophisticated consumers, but does not always do so. The removal generally harms sophisticated consumers by effectively closing the channel through which they might signal their low reservations values for the merchants' products.

Our analysis differs from the seminal work of Taylor (2004) and Acquisti and Varian (2005) primarily by analyzing the effects of several different elements of platform privacy policy (e.g., mandates to implement the privacy policy that best serves particular types of consumers, opt-in and opt-out mandates, and requirements to remove all information about a consumer's identity before transactions data are shared with third parties).[10] In contrast, Taylor (2004) focuses on the impact of limits on the ability of individual merchants (rather than the platform) to sell customer transactions data to other merchants.[11] We find

[10] Our analysis also differs in this respect with the important related work of Conitzer, Taylor, and Wagman (2012) (CTW). CTW analyze a setting where consumers interact with a monopolist in each of two periods. If the monopolist can track individual consumers, she can charge a higher price in period 2 to consumers who purchased her product in period 1. When consumers can preclude such tracking at low personal cost, they will do so to avoid exploitation in period 2. CTW demonstrate that the monopolist also can gain when she is unable to track consumers. This inability allows the monopolist to commit not to exploit consumers in period 2, which makes them willing to pay more for the monopolist's product in period 1, thereby increasing her two-period expected profit.

[11] Calzolari and Pavan (2006) analyze a related setting in which an agent (A) interacts sequentially with two principals (P1 and P2). A is privately informed about a personal characteristic that affects the value he derives from his interaction with the principals. The authors examine the policy (including the information disclosure policy) that maximizes P1's expected welfare in a setting where payments from P2 to P1 can reflect

that the incentive of the platform to disclose transactions data to its merchants can differ significantly from the incentive of an individual merchant to disclose its transactions data to other merchants. In our model, a merchant whose data are released to other merchants often suffers a reduction in profit (which it cannot recoup by charging a fee for the data in our model). In contrast, the platform maximizes its profit by disclosing all transactions data to its merchants.[12] As noted above, this disclosure also maximizes total welfare and the welfare of sophisticated consumers, but reduces the welfare of unsophisticated consumers.[13]

It is important to emphasize at the outset that our analysis only considers policies that pertain to the privacy of basic transactions data—price and quantity data and the customer's identity. In practice, transactions data can include additional information, including a consumer's credit card number or product descriptions that reveal sensitive information about the consumer (e.g., the consumer's health status or sexual orientation). Our analysis does not address the important additional considerations introduced by the disclosure of such information.

Our analysis proceeds as follows. Section 2 describes the key elements of our model. Section 3 characterizes equilibrium outcomes, both when consumers are sophisticated and when they are unsophisticated. Section 4 identifies the distinct privacy policies that maximize the welfare of sophisticated and unsophisticated consumers. Section 4 also explains how data breaches, violations of stated privacy policies, and opt-in or opt-out policies affect consumer welfare. Section 5 examines the impact of privacy policies on total (rather than consumer) welfare, identifies the privacy policy that maximizes platform profit, and explores the effects of requirements to remove all information about a consumer's identity before transactions

the policy that P1 adopts. The authors identify conditions under which P1 does not disclose any relevant information to P2. The authors also demonstrate that P1's optimal policy can entail some information disclosure and that such disclosure can secure Pareto gains.

[12]Like Taylor (2004), Kim and Wagman (2015) (KW) analyze a setting in which merchants may be permitted to sell information about their customers. In KW's model, consumers interact sequentially with two merchants, M1 and M2. Through costly effort, M1 can acquire more accurate information about the cost of serving individual consumers. KW identify conditions under which consumers are better off (and total welfare increases) when M1 is permitted to sell information about its customers to M2. The welfare gains arise in part because the potential to profit from the sale of information induces M1 to acquire better information about its customers, which helps to screen out consumers who are unduly costly to serve.

[13]Our model also differs from Taylor (2004)'s model by allowing merchants to have distinct production costs. Consequently, the welfare effects of privacy policies in our model vary with the configuration of firms' costs.

data are shared with third parties. Section 6 concludes.[14]

2 Elements of the Model

We consider the interaction between consumers and merchants on a platform (e.g., an online shopping platform). Merchants post prices on the platform and consumers decide whether to purchase the merchants' products at the posted prices. For simplicity, we assume a consumer interacts at most once with each merchant. Our formal analysis will focus on the interaction between a consumer and two merchants.[15] In this setting of primary interest, the consumer interacts first with Merchant 1 and subsequently with Merchant 2.

To isolate the key strategic effects associated with transactions data sharing, we abstract from direct competition among merchants by assuming that: (i) the consumer purchases either $n_i > 0$ units or 0 units from Merchant i ($i \in \{1, 2\}$); and (ii) the consumer derives a constant value from each unit of the product he purchases.[16] For simplicity, we assume this constant "reservation value" (r) is either low ($\underline{r}$) or high ($\overline{r}$) (with $\overline{r} > \underline{r}$). The consumer knows r from the outset of his interaction with the merchants. The merchants do not know r initially, but believe that $r = \overline{r}$ with probability $\phi \in (0, 1)$. Merchant i produces her product at constant average cost $c_i < \underline{r}$. Consequently, in principle, gains from trade are possible in every consumer-merchant interaction.

The price each merchant sets for her product depends in part on her beliefs about r. These beliefs, in turn, may be influenced by the platform's privacy policy. We focus on two such policies: one where the platform reveals all transactions data to other merchants on the platform ("third parties"), and one where the platform reveals no such data.[17] In the setting of primary interest, if the platform reveals all transactions data to third parties, then when Merchant 2 sets the price at which she will sell her product to the consumer, she knows the

[14] The proofs of the propositions in Sections 4 and 5 follow from Lemmas 1 - 6, which are presented in Section 3. The proofs of these lemmas are explained in the text. More detailed proofs are available in Baye and Sappington (2018).

[15] As explained in Section 4, the conclusions from this analysis generalize to settings with many consumers and merchants.

[16] We further assume that when he is indifferent between purchasing n_i units and purchasing 0 units, the consumer purchases n_i units. The ensuing analysis focuses exclusively on pure strategies.

[17] Section 5 allows for the possibility that merchants might learn some, but not all, details of other consumer-merchant interactions.

price that Merchant 1 set and whether the consumer purchased n_1 units or 0 units at this price. In contrast, if the platform reveals no transactions data to third parties, then at the time Merchant 2 sets her price, she does not know the price that Merchant 1 set (p_1) or whether the consumer purchased n_1 units of Merchant 1's product or declined to purchase any of Merchant 1's product.

Under *privacy*, i.e., when the platform reveals no transactions data to third parties, the consumer and Merchant 1 know that Merchant 2 will learn nothing about their interaction. Consequently, Merchant 1's considerations when she sets p_1 are straightforward. The merchant knows that if she sets $p_1 = \underline{r}$, she will sell n_1 units of her product and thereby secure payoff $n_1 [\, \underline{r} - c_1 \,]$.[18] Alternatively, if Merchant 1 sets $p_1 = \overline{r}$, the consumer will buy n_1 units of the merchant's product if and only if $r = \overline{r}$. Merchant 1's corresponding (expected) payoff is $\phi\, n_1 [\, \overline{r} - c_1 \,]$.[19] Therefore, Merchant 1 will set the lower price ($p_1 = \underline{r}$) and always sell to the consumer if and only if her unit cost of production is sufficiently low, i.e.

$$n_1 [\, \underline{r} - c_1 \,] \;\geq\; \phi\, n_1 [\, \overline{r} - c_1 \,] \;\Leftrightarrow\; c_1 \;\leq\; \widehat{c} \;\equiv\; \overline{r} - \frac{\overline{r} - \underline{r}}{1 - \phi}\,.$$

Merchant 2's considerations are identical to those of Merchant 1 under privacy. Therefore, $\widehat{c}$ is the unit cost for which a merchant's payoff is the same whether she sets price $\underline{r}$ or $\overline{r}$ under privacy.

Additional considerations arise in the absence of privacy, i.e., when all transactions data are revealed to third parties. In this event, Merchant 2 may infer something about r from the details of the consumer's interaction with Merchant 1. For example, if Merchant 2 learns the consumer bought Merchant 1's product at price $p_1 = \overline{r}$, Merchant 2 might infer that $r = \overline{r}$. Because he recognizes that Merchant 2 may draw such an inference, the consumer may decline to buy Merchant 1's product at price $p_1 = \overline{r}$ even when $r = \overline{r}$. Merchant 1 is aware that the consumer's action may be affected by his concerns about what Merchant 2 will infer from the action. This awareness, in turn, may influence the price that Merchant 1 sets for her product.

[18] Merchant 1 is also certain to sell n_1 units to the consumer if she sets $p_1 < \underline{r}$. However, she secures a lower payoff by setting $p_1 < \underline{r}$ than by setting $p_1 = \underline{r}$ (i.e., $n_1 [\, p_1 - c_1 \,] < n_1 [\, \underline{r} - c_1 \,]$).

[19] For expositional ease, we will employ the term "payoff" to denote "expected payoff" in the ensuing discussion. If Merchant 1 sets $p_1 \in (\underline{r}, \overline{r})$, her payoff is $\phi\, n_1 [\, p_1 - c_1 \,] < \phi\, n_1 [\, \overline{r} - c_1 \,]$. If Merchant 1 sets $p_1 > \overline{r}$, her payoff is $0 < \min \{\, n_1 [\, \underline{r} - c_1 \,],\, \phi\, n_1 [\, \overline{r} - c_1 \,] \,\}$.

We consider perfect Bayesian equilibria when transactions data are revealed to third parties. Formally, we consider equilibria in which: (i) each party's action is rational, given prevailing beliefs about r; and (ii) Merchant 2's beliefs about r reflect Bayes rule for all equilibrium actions. In a separating (pooling) equilibrium, the number of units the consumer buys from Merchant 1 varies (does not vary) with r.

The timing in the model is as follows. First, the platform announces whether transactions data will or will not be revealed to third parties. Then Merchant 1 sets price p_1 and the consumer purchases either 0 or n_1 units of Merchant 1's product. Merchant 2 learns this information when the platform reveals transactions data to third parties. Otherwise, Merchant 2 learns nothing about the consumer's interaction with Merchant 1. Finally, Merchant 2 sets price p_2 and the consumer purchases either 0 or n_2 units of Merchant 2's product.

3 Equilibrium Outcomes

The ensuing analysis will distinguish between sophisticated and unsophisticated consumers. As noted in the Introduction, an unsophisticated consumer considers only the interaction in which he is presently engaged when he decides whether to buy a merchant's product. Specifically, even when all transactions data are revealed to third parties, the consumer will purchase a merchant's product if and only if the merchant's price does not exceed r. In contrast, a sophisticated consumer considers all future interactions with other merchants each time he interacts with a merchant. In the absence of privacy in the setting of primary interest, the consumer recognizes that his decision to buy or not buy Merchant 1's product may affect Merchant 2's belief about r and therefore may affect the price that Merchant 2 sets for her product.

3.1 Unsophisticated Consumers

We first characterize equilibrium outcomes when the consumer is unsophisticated. Observe that under privacy, the consumer's behavior does not differ with his level of sophistication. In particular, the consumer will purchase n_1 units from Merchant 1 if and only if Merchant 1 sets a price that does not exceed r. When she faces an unsophisticated consumer, Merchant 1 will set $p_1 = \underline{r}$ (and thereby ensure the consumer always buys her product) when

c_1 is sufficiently low. In contrast, Merchant 1 will set $p_1 = \overline{r}$ (and sell the product to the consumer at this higher price only if $r = \overline{r}$) when c_1 is sufficiently high.

Lemma 1 *Suppose the consumer is unsophisticated. Then Merchant* 1*'s payoff is* $\phi\, n_1 [\, \overline{r} - c_1 \,]$ $(n_1 [\, \underline{r} - c_1 \,])$ *when* $c_1 > \widehat{c}$ $(c_1 \leq \widehat{c})$*, both under privacy and in the absence of privacy.*

Under privacy, Merchant 2 receives no information about the consumer's interaction with Merchant 1. Consequently, Merchant 2 acts exactly as Merchant 1 does. The consumer never has an opportunity to purchase at a price below $\underline{r}$. Consequently, his welfare (i.e., the difference between the value he derives from the products he purchases and the amount he pays for the products) is 0 when $r = \underline{r}$. When $r = \overline{r}$, the consumer secures strictly positive welfare when and only when a merchant has a relatively low cost and so sets price $\underline{r}$.

Lemma 2 *Suppose the consumer is unsophisticated. Then under privacy, Merchant* 2*'s payoff is* $\phi\, n_2 [\, \overline{r} - c_2 \,]$ $(n_2 [\, \underline{r} - c_2 \,])$ *when* $c_2 > \widehat{c}$ $(c_2 \leq \widehat{c})$*. The consumer's welfare is* 0 *when* $r = \underline{r}$*. His welfare when* $r = \overline{r}$ *is: (i)* $[\, n_1 + n_2 \,][\, \overline{r} - \underline{r} \,]$ *if* $c_1 \leq \widehat{c}$ *and* $c_2 \leq \widehat{c}$*; (ii)* $n_1 [\, \overline{r} - \underline{r} \,]$ *if* $c_1 \leq \widehat{c}$ *and* $c_2 > \widehat{c}$*; (iii)* $n_2 [\, \overline{r} - \underline{r} \,]$ *if* $c_1 > \widehat{c}$ *and* $c_2 \leq \widehat{c}$*; and (iv)* 0 *if* $c_1 > \widehat{c}$ *and* $c_2 > \widehat{c}$*.*

In the absence of privacy, if Merchant 1 sets $p_1 = \overline{r}$ (because $c_1 > \widehat{c}$), Merchant 2 learns that $r = \overline{r}$ $(r = \underline{r})$ if the unsophisticated consumer buys (does not buy) the product from Merchant 1. Merchant 2 sets p_2 equal to the revealed value of r in this case. In contrast, if Merchant 1 sets $p_1 = \underline{r}$ (because $c_1 \leq \widehat{c}$), the unsophisticated consumer always buys her product. Consequently, Merchant 2 learns nothing about r from observing the details of the consumer's interactions with Merchant 1, and so acts as she does under privacy.

Lemma 3 *Suppose the consumer is unsophisticated. Then in the absence of privacy, Merchant* 2*'s payoff is: (i)* $n_2 [\, \underline{r} - c_2 \,]$ *if* $c_1 > \widehat{c}$ *and* $r = \underline{r}$*; (ii)* $n_2 [\, \overline{r} - c_2 \,]$ *if* $c_1 > \widehat{c}$ *and* $r = \overline{r}$*; (iii)* $\phi\, n_2 [\, \overline{r} - c_2 \,]$ *if* $c_1 \leq \widehat{c}$ *and* $c_2 > \widehat{c}$*; (iv)* $n_2 [\, \underline{r} - c_2 \,])$ *if* $c_1 \leq \widehat{c}$ *and* $c_2 \leq \widehat{c}$*. The consumer's welfare is* 0 *when* $r = \underline{r}$ *or* $c_1 > \widehat{c}$*. His welfare when* $r = \overline{r}$ *and* $c_1 \leq \widehat{c}$ *is* $n_1 [\, \overline{r} - \underline{r} \,]$ $([\, n_1 + n_2 \,][\, \overline{r} - \underline{r} \,])$ *when* $c_2 > \widehat{c}$ $(c_2 \leq \widehat{c})$*.*

3.2 Sophisticated Consumers

Now consider equilibrium outcomes when the consumer is sophisticated. Under privacy, a sophisticated consumer knows that the details of his transaction with one merchant will not be revealed to another merchant. Therefore, the consumer's concern in each transaction is solely with the details of that transaction. Consequently, the sophisticated consumer acts exactly as an unsophisticated consumer acts, and the corresponding outcomes are precisely those that arise under privacy when the consumer is unsophisticated.

Lemma 4 *Suppose the consumer is sophisticated. Then equilibrium outcomes under privacy are as specified in Lemmas 1 and 2 (where the consumer is unsophisticated).*

Important differences arise in the absence of privacy when the consumer is sophisticated. In this setting, when $r = \overline{r}$, the consumer may decline to purchase from Merchant 1 at price $p_1 = \overline{r}$ to avoid having Merchant 2 raise her price to $\overline{r}$ after inferring that $r = \overline{r}$. In such a setting, Merchant 1 has to reduce p_1 below $\overline{r}$ (her preferred price when $c_1 > \widehat{c}$) to convince the consumer to purchase her product when $r = \overline{r}$ even though doing so causes Merchant 2 to infer that $r = \overline{r}$. In equilibrium, when Merchant 1's cost is sufficiently high (i.e., when $c_1 > c^* \equiv \widehat{c} + \phi \frac{n_2}{n_1} \left[\frac{\overline{r} - \underline{r}}{1 - \phi} \right]$), she will reduce p_1 to $\widehat{p}_1 \equiv \overline{r} - \frac{n_2}{n_1} [\, \overline{r} - \underline{r} \,]$, which is the highest price the consumer will pay when $r = \overline{r}$ to buy from Merchant 1 when doing so causes Merchant 2 to infer that $r = \overline{r}$ and so increase her price from $\underline{r}$ to $\overline{r}$.[20] Formally, $n_1 [\, \overline{r} - \widehat{p}_1 \,] = n_2 [\, \overline{r} - \underline{r} \,]$.

The price reduction that Merchant 1 is compelled to deliver to the sophisticated consumer in the absence of privacy reduces her payoff by $\phi\, n_1 [\, \overline{r} - \widehat{p}_1 \,]$. Merchant 2's payoff increases by the same amount, $\phi\, n_2 [\, \overline{r} - \underline{r} \,]$, because she sells her product to the consumer at price $p_2 = \overline{r}$ rather than at price $p_2 = \underline{r}$ when $r = \overline{r}$. The consumer's welfare is the same under privacy and in its absence because the welfare gain the consumer secures from the lower price he pays to Merchant 1 in the absence of privacy ($n_1 [\, \overline{r} - \widehat{p}_1 \,]$) is offset by the welfare reduction he incurs from the higher price he pays to Merchant 2 ($n_2 [\, \overline{r} - \underline{r} \,]$).

[20] c^* is the c_1 realization for which Merchant 1's payoff is the same when she: (i) sets price $\widehat{p}_1$ and sells to the consumer only when $r = \overline{r}$; and (ii) sets price $\underline{r}$ and always sells to the consumer.

Lemma 5 *In any separating equilibrium that arises in the absence of privacy when the consumer is sophisticated, Merchant 1's payoff is* $\phi\, n_1\,[\,\widehat{p}_1 - c_1\,]$, *Merchant 2's payoff is* $n_2\,[\,\phi\,(\overline{r} - c_2) + (1-\phi)\,(\underline{r} - c_2)\,]$, *and the consumer's welfare is* 0 *(*$n_1\,[\,\overline{r} - \widehat{p}_1\,] = n_2\,[\,\overline{r} - \underline{r}\,]$*) when* $r = \underline{r}$ *(*$r = \overline{r}$*).*[21]

When Merchant 1's cost is sufficiently low ($c_1 \in (\widehat{c}, c^*)$), she prefers to reduce her price to $\underline{r}$ and always sell to the sophisticated buyer than to set $p_1 = \widehat{p}_1$ and sell to him only when $r = \overline{r}$.

Lemma 6 *In any pooling equilibrium that arises in the absence of privacy when the consumer is sophisticated, Merchant* 1*'s payoff is* $n_1\,[\,\underline{r} - c_1\,]$, *Merchant* 2*'s payoff is* $n_2\,[\,\underline{r} - c_2\,]$, *and the consumer's welfare is* 0 *(*$[\,n_1 + n_2\,]\,[\,\overline{r} - \underline{r}\,]$*) when* $r = \underline{r}$ *(*$r = \overline{r}$*).*[22]

The separating equilibrium identified in Lemma 5 only arises when the consumer's interaction with Merchant 1 is relatively important in the sense that $n_1 > n_2$. If $n_1 \leq n_2$, the consumer sacrifices relatively little when $r = \overline{r}$ if he declines to purchase from Merchant 1 at a price $p_1 \in (\underline{r}, \overline{r}]$. Consequently, Merchant 2 will not conclude that $r = \underline{r}$ when she learns that the consumer did not purchase Merchant 1's product at such a price. Therefore, the only equilibrium that arises when $n_1 \leq n_2$ is the pooling equilibrium in which both merchants charge $\underline{r}$, as in Lemma 6.[23]

4 Privacy Regimes and Consumer Welfare

We now employ the findings in Section 3 to address several important privacy issues. Specifically, we examine: (a) how the platform's privacy policy affects consumer welfare; (b) whether data breaches or willful violations of a platform's privacy policy harm consumers; and (c) the effects of policies that require consumers to "opt in" before their transactions data with one merchant are provided to other merchants. We find that the effects of privacy regimes, privacy breaches, and opt-in policies vary with many factors, including the level of consumer sophistication. For example, privacy regimes that best serve unsophisticated

[21] This equilibrium arises when $n_1 > n_2$, $c_1 > c^*$, and $c_2 \leq \widehat{c}$.

[22] This equilibrium arises when $n_1 > n_2$, $c_1 \in (\widehat{c}, c^*)$, and $c_2 \leq \widehat{c}$.

[23] This equilibrium arises when $n_1 \leq n_2$, $c_1 > \widehat{c}$, and $c_2 \leq \widehat{c}$.

consumers can harm sophisticated consumers, and *vice versa*. Consequently, it typically is difficult to formulate Pareto-improving privacy policies, even if the exclusive concern is with consumer welfare.

Although the formal conclusions in Section 3 were derived in a setting with a single consumer and two merchants, the findings are readily extended. To illustrate, suppose there are many ($N > 1$) consumers, a fraction of whom ($\theta \in [\,0, 1\,]$) have high reservation values while the remaining fraction $(1 - \theta)$ have low reservation values. Each consumer arrives at the online platform at a random time during the day to purchase Merchant 1's product, and then returns at a random time during the evening to buy Merchant 2's product. At the end of the evening, these consumers retire and are replaced by a corresponding new generation of N consumers who start anew the next day. When merchants can engage in dynamic pricing (e.g., charge different consumers different prices), the behavior of individual consumers and merchants in this setting will match their behavior in the setting of Section 3. The ensuing discussion will focus on the setting with many consumers.

4.1 Optimal Privacy Regimes for Consumers

We first examine the implications of our formal conclusions for settings where consumers are unsophisticated. Recall that an unsophisticated consumer does not account for the possibility that information about his interaction with one merchant might affect his interaction with another merchant. Further recall from Lemmas 2 and 3 that the *ex post* welfare of an unsophisticated consumer is never higher, and is sometimes lower, when the platform reveals transactions data to other merchants than under privacy. Therefore, the *ex ante* expected welfare of unsophisticated consumers (i.e., their expected welfare before consumers learn their reservation values) is higher under privacy than in its absence.

Proposition 1 *Under the privacy policy that maximizes the welfare of unsophisticated consumers, the platform does not reveal any transactions data.*

Intuitively, under privacy, merchants must base their pricing decisions solely on the *ex ante* characteristics of a particular buyer, and these characteristics are the same for all consumers in our model. Consequently, low-cost merchants charge each customer a low price

and high-cost merchants charge each buyer a high price. In contrast, in the absence of privacy, a low-cost merchant who learns that a consumer is willing to pay a high price will charge that consumer a high price. Keeping transactions data private protects unsophisticated consumers by preventing low-cost merchants from using information from prior transactions to extract additional surplus from unsophisticated consumers with high reservation values.

Now consider the setting where consumers are sophisticated, so they recognize that if data pertaining to their interaction with Merchant 1 are revealed, the revelation may affect their transaction with Merchant 2.[24] Lemmas 4, 5, and 6 imply that in any equilibrium, a sophisticated consumer's payoff is never lower, and is sometimes higher, in the absence of privacy than in its presence.

Proposition 2 *Under the privacy policy that maximizes the welfare of sophisticated consumers, the platform reveals all transactions data.*

The welfare gain the sophisticated consumer experiences in the absence of privacy stems from the lower price (p_1) that Merchant 1 sets. In the absence of privacy, Merchant 1 must reduce p_1 to convince the consumer with a high reservation value ($r = \overline{r}$) to buy her product even though doing so leads Merchant 2 to infer that $r = \overline{r}$, and therefore to set a high price ($p_2 = \overline{r}$) for her product. In the absence of privacy, the sophisticated consumer with a high reservation value will intentionally sacrifice some surplus in his interaction with Merchant 1 if doing so convinces Merchant 2 that $r = \underline{r}$ and thereby leads her to set a low price ($p_2 = \underline{r}$) rather than a high price ($p_2 = \overline{r}$) for her product. Merchant 1 recognizes that the sophisticated consumer with $r = \overline{r}$ will only purchase her product if she sets p_1 so low that the surplus the consumer secures when he buys Merchant 1's product ($n_1 [\overline{r} - p_1]$) exceeds the reduction in surplus the consumer suffers ($n_2 [\overline{r} - \underline{r}]$) when Merchant 2 infers that $r = \overline{r}$ and so increases her price from $\underline{r}$ to $\overline{r}$. The resulting reduction in the price that Merchant 1 sets for her product increases the equilibrium welfare of the sophisticated consumer with a high reservation value ($r = \overline{r}$).[25]

[24] Transactions data include data from interactions in which a consumer purchases 0 units of a merchant's product.

[25] This discussion pertains to the settings characterized in Lemma 5, where Merchant 1's cost is sufficiently high ($c_1 > c^*$) that she prefers to set p_1 above $\underline{r}$ and sell her poduct to the consumer only when he has

Under privacy, Merchant 2 does not observe the details of the consumer's interaction with Merchant 1. Consequently, the consumer's strategic considerations (and the corresponding considerations of Merchant 1) no longer arise. In essence, even when consumers are sophisticated, privacy induces the behavior that prevails when consumers are unsophisticated. Expressed differently, concealing transactions data harms sophisticated consumers by limiting their ability to protect themselves.

Propositions 1 and 2 demonstrate that the optimal privacy regime can vary with the level of consumer sophistication. More strikingly, the propositions imply that in the present setting, the privacy regime that best serves unsophisticated consumers is the worst privacy regime for sophisticated consumers, and *vice versa*.

4.2 Consumer Harm from Data Breaches and Violations of Privacy Policies

We now consider how consumers are affected by "breaches" that can take one of two forms: either transactions data are revealed to third parties (perhaps because the platform's data servers were hacked) or the platform reneges on terms in its privacy disclosure (so the platform "deceives" consumers). To examine the effects of breaches, we focus on the setting where the platform announces the privacy policy that maximizes consumer welfare (perhaps to help attract consumers to its platform or in response to a government mandate, for example).[26]

Suppose initially that consumers are unsophisticated. Recall from Proposition 1 that consumer welfare is maximized in this setting under privacy, where transactions data are not revealed to third parties. If the platform's stated privacy policy is breached in this setting (either due to willful platform deception or a data breach caused by hackers), Merchant 2 will observe the details of prior transactions. Lemmas 2 and 3 imply that no unsophisticated consumers benefit from the breach, and some may be harmed (for the reasons discussed below). Consequently, the breach (weakly) harms unsophisticated consumers.

---

the high reservation value ($r = \overline{r}$). As Lemma 6 reports, when Merchant 1's costs are more moderate ($c_1 \in (\widehat{c}, c^*)$), she will reduce her price from $\overline{r}$ all the way to $\underline{r}$ in the pooling equilibrium that arises in the absence of privacy.

[26] Tsai et al. (2011) find that online shoppers tend to frequent retailers who announce policies that better protect shopper privacy.

Proposition 3 *Suppose the platform announces it will implement the privacy policy that maximizes the welfare of unsophisticated consumers. Then unsophisticated consumers are (weakly) harmed if transactions data are revealed, contrary to the platform's announcement.*

The magnitude of the harm a breach imposes on unsophisticated consumers varies with the consumers' reservation values and the merchants' costs. Furthermore, the realized harm may be zero. To explain these conclusions, observe from Lemmas 2 and 3 that unsophisticated consumers with low reservation values are not harmed by a breach because their welfare is zero both in the presence a breach and in its absence.[27] Likewise, unsophisticated consumers with high reservation values are not harmed by the breach when Merchant 1 has low costs.[28] In contrast, unsophisticated consumers with high reservation values are harmed by a breach that reveals Merchant 1's transactions data when she has high costs ($c_1 > \widehat{c}$) and Merchant 2 has low costs ($c_2 \leq \widehat{c}$). The breach harms these consumers in this instance because it reveals the consumers have a high reservation value ($r = \overline{r}$), a fact that Merchant 2 exploits in her interaction with these consumers.

To calculate the magnitude of the potential harm from a breach, suppose the transactions data at Merchant 1 that pertain to N transactions with N different unsophisticated consumers are breached. If these data are representative of the population, a fraction θ of the transactions involve consumers with high reservation values and the fraction $1 - \theta$ involve consumers with low reservation values. Because consumers with low reservation values are not harmed by the breach, θN is an upper bound for the expected (average) number of consumers who are harmed by the breach. Of these θN consumers who are potentially harmed, actual harm is zero if Merchant 1 has a low cost or if Merchant 2 has a high cost. In both these cases, Merchant 2's optimal price is not affected by the breach. Only when Merchant 1 has a high cost and Merchant 2 has a low cost are these θN consumers harmed by the breach. The dollar value of the harm to each of the θN consumers is the product of the price increase he faces ($\overline{r} - \underline{r}$, from Lemmas 2 and 3) and the number of units of

[27] Merchants never set a price below $\underline{r}$, so the consumer's welfare is never strictly positive when his reservation value is $\underline{r}$.

[28] When $c_1 \leq \widehat{c}$, Merchant 1 sets $p_1 = \underline{r}$, thereby inducing all unsophisticated consumers to buy her product. Consequently, Merchant 2 cannot identify consumers with high reservation values by observing Merchant 1's transactions data.

Merchant 2's product he buys (n_2). Formally, the welfare reduction experienced by each of these consumers is:

$$\Delta W \;=\; \begin{cases} n_2\,[\,\overline{r}-\underline{r}\,] & \text{if} \quad c_1 > \widehat{c} \text{ and } c_2 \leq \widehat{c} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

These calculations imply:

Proposition 4 *Suppose the platform announces it will implement the privacy policy that maximizes the welfare of unsophisticated consumers. Then the expected number of consumers harmed by a breach that involves N transactions between Merchant 1 and unsophisticated consumers is at most θN. Furthermore, the expected magnitude of consumer harm is*

$$H \;=\; \begin{cases} \theta N\, n_2\,[\,\overline{r}-\underline{r}\,] & \textit{if} \quad c_1 > \widehat{c} \textit{ and } c_2 \leq \widehat{c} \\ 0 & \textit{otherwise.} \end{cases}$$

Before proceeding to consider the impact of a breach on sophisticated consumers, we note that in principle, a willful violation of an announced privacy policy can actually benefit unsophisticated consumers. To see why, suppose the platform announces that it will implement the privacy policy that maximizes the welfare of sophisticated consumers, i.e., all transactions data will be revealed to third parties. Further suppose that, contrary to its announced policy, the platform does not reveal any transactions data. Proposition 1 implies that the resulting data concealment can benefit unsophisticated consumers with high reservation values ($r = \overline{r}$) by preventing Merchant 2 from learning that $r = \overline{r}$ when the consumers pay $p_1 = \overline{r}$ for Merchant 1's product. Of course, a data breach will not affect the welfare of unsophisticated consumers when the platform prominently discloses that it reveals all transactions data to third parties. Therefore, the impact of a data breach (by hackers, say) can differ from the impact of a platform's failure to honor its stated privacy policy.

We now consider how data breaches and violations of announced privacy policies affect sophisticated and unsophisticated consumers. The next two propositions illustrate that breaches due to hacking of the platform's data and a platform's failure to honor its stated privacy policy can affect consumers differently.

Proposition 5 *Suppose the platform implements the privacy policy that maximizes the welfare of sophisticated consumers. Then a transactions data breach harms neither sophisticated nor unsophisticated consumers.*

Proposition 5 (which is essentially a corollary to Proposition 2) might seem to suggest that the platform's willful violation of the privacy policy identified in the proposition would not harm consumers. However, this is not the case.

Proposition 6 *Suppose the platform announces it will implement the privacy policy that maximizes the welfare of sophisticated consumers. Then if the platform willfully violates this policy: (a) unsophisticated consumers are never harmed, and they secure strict gains when* $r = \overline{r}$, $c_1 > \widehat{c}$, *and* $c_2 \leq \widehat{c}$; *whereas (b) sophisticated consumers are harmed when* $r = \overline{r}$.

The different conclusions in Propositions 5 and 6 reflect the following considerations. Proposition 5 reflects the fact that when the platform routinely reveals all transactions data, a breach that makes the data available to all merchants will not affect the actions of any sophisticated consumer or merchant. Proposition 6 reflects the more subtle considerations that arise when the breach entails the platform violating its announced policy of revealing all transactions data. Recall that this violation of the announced privacy policy benefits unsophisticated consumers with high reservation values because when Merchant 2 does not observe data from the previous transaction, she cannot opportunistically charge a high price when she would otherwise charge a low price. In contrast, this same violation of the announced privacy policy harms sophisticated consumers with high reservation values. It does so because under the platform's announced privacy policy, these consumers would reject favorable prices from Merchant 1 in order to influence Merchant 2's beliefs about r. This behavior by sophisticated consumers would induce Merchant 1 to reduce the price she charges for her product. However, when Merchant 2 cannot observe data involving transactions with Merchant 1, consumers no longer have an incentive to reject favorable prices from Merchant 1, which leads Merchant 1 to set higher prices for her product.

Together, Propositions 5 and 6 provide two conclusions. First, the effects of data breaches and violations of privacy policies can differ for sophisticated and unsophisticated consumers. Second, the effects of data breaches can differ from the effects of willful violations of stated privacy policies.

4.3 Impact of Opt-in and Opt-out Policies on Consumer Welfare

We now consider the effects of mandated consumer "opt-in" and "opt-out" policies. An opt-in policy requires the platform to secure a consumer's explicit consent before it provides transactions data concerning the consumer to third parties. An opt-out policy requires the platform to allow a consumer to request, and thereby secure, a personal exemption from the platform's stated policy to provide transactions data to third parties. The ensuing focus is on whether mandated opt-in or opt-out polices enable consumers to secure the privacy regime that maximizes their welfare.

If consumers do not make utility-maximizing opt-in and opt-out decisions, then opt-in and opt-out policies are unlikely to improve their welfare. To focus on the "best-case" scenario for opt-in and opt-out policies when consumers are unsophisticated, we assume an unsophisticated consumer makes the opt-in or opt-out decision that maximizes his welfare.

It is well-known that the welfare effects of opt-in or opt-out policies can vary with the prevailing status quo.[29] We first consider the impact of opt-in and opt-out policies on unsophisticated consumers when the platform does not reveal transactions data to third parties under the initial status quo. Recall from Proposition 1 that this status quo maximizes the welfare of unsophisticated consumers. Consequently, a regulation that requires the platform to adopt an opt-in or opt-out policy cannot increase the welfare of unsophisticated consumers above the level they secure under the initial status quo. Moreover, if consumers must incur a positive (but possibly negligible) "hassle cost" to opt out, consumers are harmed when an opt-out policy is mandated because they cannot costlessly replicate the (welfare-maximizing) status quo. In summary:

Proposition 7 *Suppose the platform does not reveal transactions data to third parties under the initial status quo. Then a regulation that requires the platform to adopt an opt-in (or opt-out) policy does not improve the welfare of unsophisticated consumers.*

A different conclusion can arise under the same status quo when consumers are sophisticated. Recall from Proposition 2 that the welfare of sophisticated consumers is lowest under

[29] For example, see Federal Trade Commission (2009).

the postulated status quo, where the platform does not reveal transactions data to third parties. Consequently, sophisticated consumers cannot be harmed by the adoption of an opt-in policy or an opt-out policy. Furthermore, if it is costless for consumers to opt in or opt out of the status quo policy, sophisticated consumers will endogenously create the privacy regime that maximizes their welfare by authorizing the platform to reveal all transactions data to third parties.[30]

Proposition 8 *Suppose the platform does not reveal transactions data to third parties under the initial status quo. Then a regulation that requires the platform to adopt an opt-in (or opt-out) policy improves the welfare of sophisticated consumers in the absence of hassle costs.*[31]

We now consider the alternative status quo in which the platform reveals transactions data to third parties. Suppose a regulation is imposed in this setting that prohibits the platform from revealing a consumer's transactions data to third parties unless the consumer explicitly consents to such revelation by opting in. Proposition 1 implies that this regulation has the potential to increase the welfare of unsophisticated consumers by effectively replacing the status quo with the policy that maximizes the welfare of unsophisticated consumers (by not revealing transactions data to third parties). Also recall from Lemmas 1 and 3 that no unsophisticated consumer is harmed and unsophisticated consumers with high reservation values benefit when transactions data are concealed from, rather than revealed to, third parties. Therefore, if it is costless to opt out, all unsophisticated consumers will opt out of the platform's status quo privacy policy. In doing so, unsophisticated consumers endogenously move the platform's privacy policy to the policy that maximizes their welfare.

More subtle considerations arise if consumers must incur a positive (but possibly negligible) hassle cost to opt out. In this case, only the θN unsophisticated consumers with high reservation values have an incentive to opt out (when $c_1 > \widehat{c}$, $c_2 \leq \widehat{c}$, and the hassle cost

[30] We assume here and throughout the ensuing discussion that when an unsophisticated consumer is indifferent between the privacy policy that maximizes the *ex ante* expected welfare of unsophisticated consumers and some other privacy policy, the consumer will opt for the former policy.

[31] From Lemmas 4, 5, and 6, sophisticated consumers with high reservation values ($r = \overline{r}$) secure strict gains by opting out of the status quo privacy policy. The welfare of sophisticated consumers with low reservation values ($r = \underline{r}$) is the same whether they opt out of the status quo policy or decline to do so. Additional considerations arise in the presence of strictly positive hassle costs. See the discussion below, immediately preceding Proposition 9.

is less than the potential gain identified in equation (1), $n_2[\overline{r} - \underline{r}]$). However, by opting out, these consumers would effectively reveal their high reservation values by behaving differently than consumers with low reservation values (who do not opt out because they secure no strict gain by doing so).[32] Such revelation would eliminate the potential gain from opting out. If unsophisticated consumers recognize that opting out would not allow them to secure a welfare gain, they will not bear the cost of opting out. In this event, the mandated opt-out policy will not affect their welfare. If the unsophisticated consumers fail to recognize the inference that would be drawn from their opting out and so incur the hassle cost required to opt out, the mandated policy would harm them. To summarize:

Proposition 9 *Suppose the platform reveals transactions data to third parties under the initial status quo. Then: (a) a regulation that requires the platform to adopt an opt-in policy improves the welfare of unsophisticated consumers; and (b) a regulation that requires the platform to adopt an opt-out policy increases the welfare of unsophisticated consumers only if it is costless for consumers to opt-out.*

We now consider the corresponding considerations that arise when consumers are sophisticated. We continue to suppose the platform reveals transactions data to third parties under the initial status quo. In addition, a regulation is imposed that requires the platform to adopt an opt-in policy under which a consumer's explicit consent must be secured before his transactions data can be revealed to third parties. Proposition 2 implies that sophisticated consumers are worse off when transactions data are not revealed to third parties than when it is revealed. Therefore, the regulation cannot improve the welfare of sophisticated consumers. Furthermore, if consumers must incur a cost to opt in, then the cost required to ensure that the welfare-maximizing status quo is maintained harms sophisticated consumers.

Now consider an opt-out regulation that allows consumers to secure an exemption from the platform's status quo policy of revealing transactions data to third parties. Proposition 2 implies that the welfare of sophisticated consumers is maximized under the platform's status quo policy of revealing transactions data. Consequently, sophisticated consumers will not exercise their option to opt out of this policy, even if doing so is costless. In summary:

[32] Recall Lemmas 4, 5, and 6.

Proposition 10 *Suppose the platform reveals transactions data to third parties under the initial status quo. Then: (a) a regulation that requires the platform to adopt an opt-in policy harms sophisticated consumers unless it is costless for them to opt in; and (b) a regulation that requires the platform to adopt an opt-out policy does not affect the welfare of sophisticated consumers.*

In summary, the impact of mandated opt-in or opt-out policies can vary widely, depending on the prevailing status quo, the costs of opting in or opting out, and the level of consumer sophistication.

5 Social Welfare, Platform Incentives, and PII

The foregoing analysis has focused on consumer welfare, reflecting the primary concern of most antitrust and consumer protection agencies. However, consideration of total welfare (the sum of consumer and merchant welfare) is relevant for at least two reasons. First, even agencies that are charged with protecting consumer welfare consider the likely impact of proposed policies on total welfare.[33] Second, the incentives of two-sided platforms, such as the online shopping platform in our model, typically are not fully aligned with the welfare of the participants on just one side of the platform.[34] In particular, if the online shopping platform in our model sought to maximize its profit and could charge consumers and merchants to use the platform, the platform would adopt the privacy policy that maximizes the total welfare of all consumers and merchants on the platform and employ fixed fees to extract all rent.

The privacy policy that maximizes social welfare varies with the degree of consumer sophistication. When consumers are unsophisticated, total welfare is the same whether the platform reveals or does not reveal transactions data to third parties. This conclusion reflects two observations. First, recall from Lemma 1 that the welfare of Merchant 1 is the same under the two privacy regimes (privacy and the absence of privacy). Second, Lemmas 2 and 3 imply that any increase (or reduction) in welfare that unsophisticated consumers experience

[33] The mission of the U.S. Federal Trade Commission is to protect consumers. However, Section 5 cases (those alleging an "unfair business practice") require an accounting of countervailing benefits to consumers or to competition. See the Federal Trade Commission Act Incorporating U.S. SAFE WEB Act amendments of 2006 at § 45 (Section 5), available at https://www.ftc.gov/sites/default/files/documents/statutes/federal-trade-commission-act/ftc_act_incorporatingus_safe_web_act.pdf.

[34] See Baye and Morgan (2001), for instance.

under one of the privacy regimes is exactly offset by a reduction (or increase) in Merchant 2's payoff. Therefore, total welfare does not vary across privacy regimes when consumers are unsophisticated.

In contrast, Lemmas 4, 5, and 6 imply that when consumers are sophisticated, total welfare is highest under the privacy regime that maximizes the welfare of sophisticated consumers, i.e., when the platform reveals transactions data to third parties. The higher total welfare arises because surplus-enhancing sales are consummated more often when transactions data are revealed. The expanded sales arise in the absence of privacy because Merchant 1 reduces her price to account for the sophisticated consumer's incentive to reject an otherwise favorable price in an attempt to conceal information from Merchant 2.[35] To summarize:

Proposition 11 *Suppose the platform adopts the privacy policy that maximizes the welfare of sophisticated consumers by revealing transactions data to third parties. Then total welfare is maximized both when consumers are sophisticated and when they are unsophisticated.*

Proposition 11 implies that a laissez-faire policy which allows a profit-maximizing platform to implement its preferred privacy policy will ensure that the welfare of sophisticated consumers is maximized. However, this laissez-faire policy will not necessarily maximize the welfare of all relevant parties. In particular, the policy will leave unsophisticated consumers worse off than they are under a policy that prohibits the platform from providing transactions data to third parties. Furthermore, Lemmas 1, 4, 5, and 6 imply that Merchant 1 is worse off under the laissez-faire policy than when the platform cannot reveal transactions data, and she is strictly worse off when consumers are sophisticated.

We conclude by considering the effects of removing personally identifiable information (PII) from transactions data that are revealed to third parties. We take PII to include any information that would allow Merchant 2 to infer from Merchant 1's transactions data the identity of the consumer with whom Merchant 2 is presently interacting. Observe that PII is not limited to a consumer's name, address, and telephone number. PI might include, for

[35] This welfare improvement does not reflect the increased surplus that typically arises from a price reduction in the presence of a downward-sloping demand curve. Recall that a consumer's demand is completely price-inelastic below his reservation value in our model.

example, the consumer's IP address or identifying information that Merchant 2 gleans from cookies, for example.[36]

Removing PII from transactions data can benefit consumers in our model. Alternatively, it can harm consumers or not affect them at all. To illustrate these varied effects, first suppose that the $N > 1$ consumers are unsophisticated, so they are harmed when all transactions data (including PII) are revealed to third parties. This harm is eliminated when PII is removed from transactions data as long as consumers' reservation values are not perfectly correlated. In this case, Merchant 2 learns nothing about the reservation value of any particular consumer from observing (only) the price and quantity data from Merchant 1's N transactions. Consequently, if Merchant 2 has low costs ($c_2 \leq \widehat{c}$), she will set price $p_2 = \underline{r}$ for all N consumers. In contrast, if Merchant 2 could observe all of Merchant 1's transactions data (including PII), she would charge a high price ($p_2 = \overline{r}$) for her product to any unsophisticated consumer that paid $p_1 = \overline{r}$ for Merchant 1's product. Therefore, removing PII increases the welfare of unsophisticated consumers with high reservation values ($r = \overline{r}$) when Merchant 1 has high costs ($c_1 > \widehat{c}$) and Merchant 2 has low costs ($c_2 \leq \widehat{c}$).

To demonstrate that the removal of PII from transactions data may not affect consumer welfare, consider the extension of our model in which the reservation values of the N unsophisticated consumers are perfectly correlated. Suppose all PII is removed from transactions data before the data from transactions with Merchant 1 are revealed to Merchant 2 in this setting. Further suppose Merchant 1 has high costs ($c_1 > \widehat{c}$), and so sets price $p_1 = \overline{r}$. Then if Merchant 2 learns that some consumer purchased n_1 units of Merchant 1's product at this price, she knows that all consumers have high reservation values ($r = \overline{r}$) due to the presumed correlation in reservation values. Consequently, Merchant 2 will set $p_2 = \overline{r}$, which leaves unsophisticated consumers with the same welfare they achieve when PII is not removed from transactions data in the absence of privacy.

In contrast, the removal of PII can harm sophisticated consumers when their reservation values are not perfectly correlated. To see why, suppose again there are $N > 1$ consumers.

[36] The FTC (2009, footnote 47) observes that "Traditionally, PII has been defined as information that can be linked to a specific individual including, but not limited to, name, postal address, email address, Social Security number, or driver's license number...[but in online markets] the traditional notion of what constitutes PII versus non-PII is becoming less and less meaningful ...".

When PII is removed from transactions data in this setting, each sophisticated consumer knows that his actions with Merchant 1 will reveal nothing about his personal identity to Merchant 2. Consequently, the consumer acts precisely as an unsophisticated consumer acts. In particular, the sophisticated consumer gains nothing by rejecting a favorable price from Merchant 1 because Merchant 2 cannot link this rejection to the identity of any particular consumer. Consequently, when she has high costs ($c_1 > \widehat{c}$), Merchant 1 will charge sophisticated consumers a higher price when transactions data are "sanitized" by removing PII, so consumer welfare declines. In summary:

Proposition 12 *Removing PII before transactions data are revealed to third parties does not improve (and may reduce) the welfare of sophisticated consumers. Such removal of PII can benefit unsophisticated consumers, but does not necessarily do so.*

6 Conclusions

The growing prevalence of "big data" has raised serious concerns about the use of these data. We have employed a simple model in the spirit of important predecessors (especially Taylor (2004) and Acquisti and Varian (2005)) to examine the effects of sharing transactions (price and quantity) data on an online platform. We found that such sharing can have important effects on consumer, merchant, and platform welfare. Relatively subtle effects can arise because the sharing of transactions data opens a channel through which sophisticated consumers may attempt to signal or conceal their reservation values for merchants' products.

We found that total welfare, the welfare of sophisticated (fully rational) consumers, and platform profit are all maximized when the platform provides transactions data to all merchants. In contrast, the welfare of unsophisticated consumers is maximized when no transactions data are shared with third parties. Consequently, an important tension arises. Privacy policies that best protect unsophisticated consumers may do so at the expense of sophisticated consumers. These policies may also reduce social welfare (and platform profit).

This tension between policies that best serve different types of consumers raises subtle considerations in the formulation of platform privacy policies. For example, opt-in or opt-out requirements can benefit unsophisticated consumers but harm sophisticated consumers. In addition, data breaches and willful violations of platform privacy policies can have different

effects, and can affect sophisticated and unsophisticated consumers in different ways. Consequently, the most appropriate privacy policy for online shopping platforms typically will vary with the relevant social objective and with prevailing institutional features, including the status quo policy, the costs of opting into and out of a privacy policy, and the degree of consumer sophistication.[37]

Before concluding, we note that our model has potential implications for antitrust policy, as well as consumer protection policy. We found that information sharing though a third party (the platform in our model) can increase total welfare in part by promoting the consummation of welfare-enhancing transactions. This finding lends support to the current antitrust practice in the U.S. which recognizes that information exchanges are not necessarily anti-competitive. However, this finding also suggests that current requirements for information sharing to fall in an "antitrust safety zone" may be unduly restrictive in some settings (e.g., when firms do not compete directly). Under current policy,

> "...the agencies will not [generally] challenge a data exchange if: (1) the exchange is managed by a third-party, like a trade association; (2) the information provided by participants is more than three months old; and (3) at least five participants provide the data underlying each statistic shared, no single provider's data contributes more than 25% of the "weight" of any statistic shared, and the shared statistics are sufficiently aggregated that no participant can discern the data of any other participant."[38]

In our model, even the sharing of data that are current and that explicitly identify the specific data source can enhance welfare.

In concluding, we reiterate that we have only considered policies that pertain to the privacy of basic transactions data—price and quantity data and the customer's identity. We have not considered the additional considerations that arise when transactions data include potentially sensitive financial or personal information. Explicit analysis of the appropriate treatment of such additional information merits further study.

[37] Taylor and Wagman (2014, p. 81) similarly caution that "studies of consumer privacy must be understood within their individual context and industries, and that their conclusions depend on the specific competitive landscapes at play – and may not necessarily apply more broadly."

[38] Bloom (2014).

References

Acquisti, Alessandro, Curtis R. Taylor, and Liad Wagman, "The Economics of Privacy," *Journal of Economic Literature* (2016), 52(2), pp. 442-92.

Acquisti, Alessandro and Hal R. Varian, "Conditioning Prices on Purchase History," *Marketing Science* (2005), 24(3), pp. 367-381.

Athey, Susan, Christian Catalini, and Catherine Tucker, "The Digital Privacy Paradox: Small Money, Small Costs, Small Talk," National Bureau of Economic Research Working Paper w23488, June 2017.

Baye, Michael R. and John Morgan, "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets," *American Economic Review* (2001), 91(3), pp. 454-474.

Baye, Michael R. and David E. M. Sappington, "Technical Appendix to Accompany 'Revealing Transactions Data to Third Parties: Implications of Privacy Regimes for Welfare in Online Markets" (2018), available at http://nash-equilibrium.com/PDFs/Appendix.pdf.

Belleflamme, Paul and Wouter Vergote, "Monopoly Price Discrimination and Privacy: The Hidden Cost of Hiding," *Economics Letters* (2016), 149, pp. 141-144.

Blaug, Mark, "The Fundamental Theorems of Modern Welfare Economics, Historically Contemplated," *History of Political Economy* (2007), 39(2), pp. 185-207.

Bloom, Michael, "Information Exchange: Be Reasonable," Federal Trade Commission (2014), available at https://www.ftc.gov/news-events/blogs/competition-matters/2014/12/information-exchange-be-reasonable.

Calzolari, Giacomo and Alessandro Pavan, "On the Optimality of Privacy in Sequential Contracting," *Journal of Economic Theory* (2006), 130(1), pp. 168-204.

Campbell, James, Avi Goldfarb, and Catherine Tucker, "Privacy Regulation and Market Structure," *Journal of Economics & Management Strategy* (2015), 24(1), pp. 47-73.

Conitzer, Vincent, Curtis R. Taylor, and Liad Wagman, "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases," *Marketing Science* (2012), 31(2), pp. 277-292.

Evans, David S., "The Online Advertising Industry: Economics, Evolution, and Privacy," *Journal of Economic Perspectives* (2009), 23(3), pp. 37-60.

Federal Trade Commission, "Self-Regulatory Principles for Online Behavioral Advertising," FTC Staff Report (2009), February, available at https://www.ftc.gov/sites/default/files/docu

ments/reports/federal-trade-commission-staff-report-self-regulatory-principles-online-behavioral-advertising/p085400behavadreport.pdf.

Kim, Jin-Hyuk and Liad Wagman, "Screening Incentives and Privacy Protection in Financial Markets: A Theoretical and Empirical Analysis," *RAND Journal of Economics* (2015), 46(1), pp. 1-22.

Smith, Adam, "An Inquiry into the Nature and Causes of the Wealth of Nations," London: W. Strahan and T. Cadell (1776).

Stigler, George J., "A Theory of Oligopoly," *Journal of Political Economy* (1964), 72(1), pp. 44-61.

Taylor, Curtis R., "Consumer Privacy and the Market for Customer Information," *RAND Journal of Economics* (2004), pp. 631-650.

Taylor, Curtis and Liad Wagman, "Consumer Privacy in Oligopolistic Markets: Winners, Losers, and Welfare." *International Journal of Industrial Organization* (2014), 34, pp. 80-84.

Tsai, Janice Y., Serge Egelman, Lorrie Cranor, and Alessandro Acquisti, "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Information Systems Research* (2011), 22(2), pp. 254-268.

Tucker, Catherine E., "The Economics of Advertising and Privacy," *International Journal of Industrial Organization* (2012), 30(3), pp. 326-329.