Is voluntary profiling welfare enhancing?

Byungwan Koh[†]

Srinivasan Raghunathan[††]

Barrie R. Nault[†]

[†]Haskayne School of Business
University of Calgary
bkoh@ucalgary.ca; nault@ucalgary.ca

[††]Naveen Jindal School of Management
University of Texas at Dallas
sraghu@utdallas.edu

December 10, 2013

The authors gratefully acknowledge helpful comments from the participants at the 2009 BIG XII + MIS Research Symposium, the INFORMS 2010 annual meeting, LG CNS/KrAIS Post-ICIS 2010 workshop, and the Information Systems seminar at University of Texas at Dallas and Seoul National University. We thank the Natural Science and Engineering Research Council of Canada, the Social Science and Humanities Research Council of Canada, and the Robson Endowment and the IRC at the Haskayne School of Business at the University of Calgary for support.

Is voluntary profiling welfare enhancing?

Abstract

Although consumer profiling advocates tout benefits from personalization, consumer advocacy groups oppose profiling in online markets because of concerns about privacy and price discrimination. Policies such as "opt-out" or "opt-in" that provide consumers the option to voluntarily participate in profiling are the favored compromise. We compare voluntary profiling to no profiling and show that voluntary profiling leads to some counter-intuitive results. Consumers that do not participate in profiling and some that participate are worse off under voluntary profiling. Neither social welfare nor aggregate consumer surplus is necessarily higher under voluntary profiling; even when voluntary profiling leads to an increase in social welfare, it may come at the expense of consumer surplus. If the seller cannot price discriminate, then aggregate consumer surplus under voluntary profiling is higher and a reduction in privacy cost has a positive impact on all consumers as well as the seller. When price discrimination is possible, however, reducing privacy cost alone may reduce aggregate consumer surplus. The primary reason for these counter-intuitive results is that voluntary profiling allows the seller to identify high valuation consumers that have no incentive to participate and set a higher price for them (compared to no profiling) while simultaneously benefitting from the profile information of low valuation consumers that participate. However, a positive privacy cost mitigates the participation incentives of even low valuation consumers and hence sellers' ability to engage in price discrimination.

Keywords: Profiling, Electronic Marketplaces, Privacy, Price Discrimination, Social Welfare.

1 Introduction

Consumer profiling is ubiquitous in online markets because firms can address individual consumers and access consumers' purchase history, browsing behavior as well as demographic and psychographic information that consumers may provide when they register at online sites. Such consumer profile information allows firms to offer personalized web services, which have the potential to reduce consumers' search effort prior to purchase, as well as personalized pricing which may convert some consumers into buyers.

Surveys generally show that most consumers are uneasy about consumer profiling. Many studies attribute the negative consumer attitude towards profiling to privacy and price discrimination concerns (Hoffman et al. 1999; Volokh 2000; Earp and Baumer 2003; Acquisti and Grossklags 2005). Privacy concern relates to consumers' loss of control over their information – the individual's ability to choose the extent to which her personal information is used, protected, and shared with others (Chellappa and Shivendu 2007, 2008). Even with privacy policies that govern the use and sharing of the information by firms, privacy concerns persist. Consumers' awareness about and resentment towards price discrimination has also grown steadily (Turow et al. 2005; Edwards 2006). The most publicized negative reaction to price discrimination was when Amazon.com experimented with differential pricing for DVDs. Though it announced the termination of this practice (Streitfeld 2000; Edwards 2006), blog sites continue to cite cases where Amazon.com charges differential prices for those that register in profiling-like programs and those that do not.[1] Academic literature (Odlyzko 2003; Viswanathan et al. 2007; Schön 2010; Li and Dinlersoz 2012), news articles (Valentino-Devries et al. 2012) and online discussion forums provide anecdotal evidence about price discrimination practices in e-commerce. Empirical studies also suggest that online firms indeed practice price discrimination, often subtly (Clemons et al. 2002; Baye and Morgan 2002; Odlyzko 2003; Ellison and Ellison 2004; Viswanathan et al. 2007).

Although industry groups support the use of consumer profiling, consumer advocates

[1] see http://crookedtimber.org/2008/12/22/amazons-price-discrimination/; http://brightviolet.wordpress.com/2010/05/18/amazon-coms-price-discrimination-how-your-amazon-prime-membership-isnt-such-a-good-deal-after-all/; http://iterativepath.wordpress.com/2013/01/17/amazon-price-discrimination-done-well/

oppose them on privacy and price discrimination grounds (Chester 2009). As a compromise between a ban and an unfettered use of profiling, policy makers in the U.S. have proposed a permission-based approach for the collection and use of online consumer information. This policy, known as *Voluntary Profiling*, allows firms to collect and use consumer information if consumers voluntarily participate in profiling (FTC 2000).[2] One might expect that no consumer will be worse off under voluntary profiling compared to no profiling because only those consumers who expect to benefit from profiling will participate in it. One might also expect that voluntary profiling will increase the aggregate consumer surplus and social welfare because voluntary profiling reduces search cost, and may increase market size – the number of consumers that purchase.

We show that this intuition is not accurate. Voluntary profiling can have adverse implications for consumers and society unless privacy, and more importantly, price discrimination are addressed. The reason is that voluntary profiling allows the firm to exploit the heterogeneity in consumers' inherent incentives to participate in profiling. Our context is the following. An online monopolist – the seller – sells a large number of products from a product category to consumers each of whom has an ideal product. Further, consumers have different valuations for their ideal products. In this setup, consumers that have a low valuation – especially, those that are not served when there is no profiling and have a low privacy cost – have an incentive to participate in profiling and benefit from reduced search cost and possibly reduced price. Consumers that have a high valuation, however, do not have similar incentives to participate in profiling despite the reduced search cost because the seller may charge a higher price to them. The direct effect of low-valuation consumers' participation in profiling is that profile information reduces the seller's uncertainty about the valuation of these consumers. The indirect effect is that the seller infers that consumers that do not participate in profiling have a higher valuation than those that do. Hence, compared to no profiling, the seller charges a higher price to non-participants under voluntary profiling. Due to this indirect effect, some consumers that participate are worse off under voluntary profiling than no profiling. Similarly, every non-participating consumer under voluntary profiling

[2] Also known as the "Do not Track List" policy modeled after the "Do Not Call List" policy implemented for telemarketers.

is worse off compared to no profiling. Furthermore, the seller extracts the search cost saving enjoyed by participating consumers, and therefore, voluntary profiling may result in a net transfer of wealth from consumers to the seller.

We also examine the roles played by the two components of the profiling debate – price discrimination and privacy cost – in the adverse impact of voluntary profiling on consumers. We show that if price discrimination is not possible, then aggregate consumer surplus under voluntary profiling compared to no profiling is higher and a reduction in privacy cost has a positive impact on all consumers as well as the seller. However, when price discrimination is possible, reducing privacy cost may reduce consumer surplus. Interestingly, a positive privacy cost for at least some consumers moderates the ability of the seller to engage in price discrimination. These results highlight the economic tensions related to price discrimination and privacy cost that should be evaluated by policy makers before implementing voluntary profiling policy in online markets.

1.1 Relationship to Existing Literature

There is a growing literature that addresses electronic privacy, namely the ability of firms to track and collect information about individual consumers. Bulk of this literature examines the price discrimination component of customer addressability (Chen 1997; Fudenberg and Tirole 1998; Villas-Boas 1999; Chen and Iyer 2002; Villas-Boas 2004; Taylor 2004; Zhang and Krishnamurthi 2004; Iyer et al. 2005; Fudenberg and Villas-Boas 2007; Cachon et al. 2008; Aloysius et al. 2012). This stream of literature shows that with passive consumers, consumer addressability and personalized pricing benefits a monopoly, but competing firms may not choose full addressability because it may intensify competition. However, research on dynamic targeted pricing in online markets using multi-period models (Taylor 2004; Acquisti and Varian 2005; Calzolari and Pavan 2006; Chen and Zhang 2009) shows how strategic consumers' actions in the first period anticipating price in the second period can hurt the firm.

Chellappa and Shivendu (2007, 2008) analyze the welfare implications of regulatory regimes that control if a vendor can "buy" consumer information and can force consumers to accept collection of information. Challappa and Shivendu (2010) examine vendor strategies regarding the extent of personalization and provision of coupons when faced with consumers with heterogeneous privacy concerns. Although related, their models do not incorporate key features of ours, namely price discrimination and the externality that participation of one group of consumers may have on those that do not participate.

Closest to our paper is recent research that examined contexts in which consumers decide the information they want to reveal (Aron et al. 2006; Hann et al. 2008; Wagman et al. 2010). In Aron et al., a customer receives customized products in return for personalized information, and is faced with the trade-off between better product-fit and price discrimination. However, they do not consider a consumer's trade-off between better product-fit and privacy. In our model, a consumer participates in profiling only if the reduction in search cost and possible price discounts exceeds the privacy cost and potential price increases of participating. Moreover, in their model, information provided by a consumer is used to infer only her valuation. In contrast, in our model, the seller uses information provided by participating consumers to estimate the valuation of these consumers as well as to update the distribution of non-participating consumers valuation.

Hann et al. introduce marketing avoidance – consumer strategies such as concealment (by opting out of telemarketing lists) and deflection (using filters at the consumer end) to avoid being targeted by sellers – and analyze welfare implications of these strategies. They assume that every marketing solicitation imposes privacy harm to consumers. In addition, sellers cannot address individual consumers and hence their solicitations are randomly distributed, and that price discrimination is not allowed.

Wagman et al. extend the dynamic pricing literature by considering welfare implications of an opt-out model in a repeat purchase context. In their two-period model, consumers that buy in the first period can choose to delete purchase history at a cost at the end of the first period. The seller decides the pricing strategies by anticipating consumer action. Wagman et al. do not consider privacy cost. The seller cannot address individual consumers, offer personalized services, nor engage in price discrimination at the individual level. In their model a consumer becomes addressable to the seller at the end of the first period and hence consumers can wait in the first period, whereas we consider a single period game in which consumers' opt-in decision is made before any purchase.

We extend the existing literature by modeling a richer context in which each consumer voluntarily allows herself to be identified by the seller by considering the tradeoff between her individual benefit (reduction in search cost and possible price discounts) and cost (privacy cost and possible price increase) from participating in profiling. Further, we model both price discrimination and privacy, two distinct dimensions of the profiling debate. Consequently, we derive new results regarding the impact of voluntary profiling on individual consumers, impact of privacy cost and (separately) price discrimination, and the roles of privacy and price discrimination in how voluntary profiling affects consumer and social welfare.

2 Model Description

An online monopolist – the seller – sells a large number of products from a product category. A consumer visiting the seller's site searches and finds her "ideal" product – the one that is closest to her preference – at the end of search. If the seller does not profile a consumer (or the consumer does not participate in profiling), the consumer searches without any support from the seller. If the consumer participates in profiling, the seller supports search using profile information. We do not assume any specific form of support but we assume that if the seller provides support then the search cost of a consumer to find her ideal product is reduced by the amount of support that the seller provides.

Consumers We denote consumers by their valuation from consuming their ideal product: $v \in [v_L, v_H]$ where probability distribution $f(v)$ is positive over its support. We take v as uniformly distributed as valuation can be scaled as needed. Without providing profile information to the seller, a consumer incurs a fixed search cost $c \in R^+$ to find her ideal product. We use α to denote the reduction in search cost a consumer receives from participating, and reflect the amount of search support that seller provides. Thus, we refer to α as the *search support* from the seller. A consumer participating in profiling reduces the search cost by proportion $\alpha \in [0,1]$ such that the effective search cost is $[1-\alpha]\,c$. In our setting the consumer always finds her ideal product at the end of the search.[3] Further, the consumer preference is diffuse so that *ex ante* each product offered by the seller is equally likely to be preferred and *ex ante* demand is the same for each product so that the seller offers all products at the same price. Therefore, the purpose of search is only to identify the ideal product.

Consumers that participate in profiling incur privacy costs. This privacy cost includes any cost resulting from unwanted marketing solicitations (Hann et al. 2008), information sharing, privacy leakage, and security breaches (Milne and Rohm 2000). However, it does not include potential loss (or gain) in surplus from price discrimination (Chellappa and Shivendu 2007, 2008). Consumers are either *privacy-sensitive (S)* or *privacy-non-sensitive (NS)* – our results extend in a qualitatively similar way if there are more than two groups. The privacy cost of a consumer in S is $r \in R^+$, and in NS is normalized to zero. The proportion of consumers that belong to S is $\lambda \in [0,1]$. The valuation and privacy cost are independent, and profile information does not reveal a consumer's privacy cost. Furthermore, the privacy cost is sunk after a consumer makes her participation decision; hence, the seller does not take the privacy cost into account in its pricing decision.

The Seller The seller employs a profiler – specialized Business Intelligence (BI) or Business Analytics (BA) tool – to utilize consumer information. We take the fixed cost to develop a profiler with search support α to be increasing at an increasing rate in α. Without loss of generality as any convex function of α yields qualitatively similar results, we take the functional form of fixed costs to be $t\,\alpha^2$. There is no variable cost to use the profiler.

The profile information of a participating consumer provides the seller a signal $\hat{v}$ about their valuation. The signal reveals the consumer's true valuation with probability β and no new information with probability $[1-\beta]$. We refer to β as the *valuation accuracy* of the profiler. If the prior probability distribution function of valuation for a participating

[3]We can extend the model to when the consumer identifies a less-than-ideal product with some positive probability at the end of search. In that case, the expected utility rather than the utility from the ideal product determines our results.

consumer is $g(v)$, the conditional probability distribution function of the signal given true valuation is given by:

$$q(\hat{v}|v) = \begin{cases} \beta & \text{if } \hat{v} = v \\ [1-\beta]g(v) & \text{otherwise} \end{cases}$$

Clearly, $g(v)$ depends on $f(v)$ and consumers' participation decisions. For instance, if consumers with valuation in the range $[v_a, v_b]$ participate in profiling, $g(v)$ is uniform with support $[v_a, v_b]$. We derive the exact functional form for $g(v)$ in Section 3.2.

Under voluntary profiling, we consider when the valuation accuracy is sufficiently high so that the seller prefers to use the signal when it sets prices. If the seller finds it profitable to ignore the signal, then it charges a uniform price. In Section 5 we examine how eliminating price discrimination under voluntary profiling affects our results.

Timing of the Game The sequence of the game is depicted in Figure 1. In stage 1, the seller decides the search support. In stage 2 after observing the search support, each consumer decides whether to participate in profiling.[4] A consumer indifferent between participating and non-participating participates in profiling. In stage 3, after it receives the signal from the profiler, the seller sets prices for participating and non-participating consumers. Finally, in stage 4, consumers that expect a non-negative surplus search and purchase.

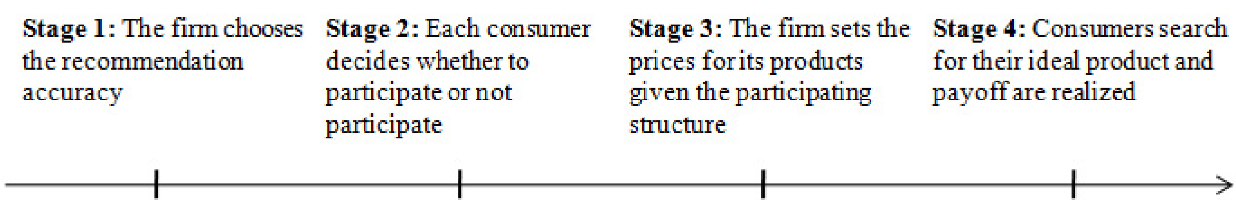


Figure 1: Timeline of the game

In summary, in our model, the seller does not know either the valuation or the privacy cost

[4] We assume that in stage 2 consumers can observe the search support (i.e., the benefit in terms of reduction in search cost they would get if they participate). We also analyzed a model in which consumers are not able to observe the true search support when they make their participation decisions in stage 2. Consumers have a belief about the search support, modeled as a probability distribution, and they make participation decisions based on this belief. If all consumers have the same belief (i.e., expected search support), then the model yields qualitatively identical results to what we report.

of a consumer, and a consumer does not know the signal the seller has about her valuation. Thus, we have a sequential game with private information, and consequently, we use the *Perfect Bayesian Equilibrium (PBE)* as our solution concept. A PBE in our model consists of the following: (i) seller's strategy composed of search support (in stage 1), and the prices for participating and non-participating consumers (in stage 3), and (ii) consumers' strategies composed of participation decision (in stage 2) and search and purchase decision (in stage 4). These constitute a PBE if all strategies are sequentially rational given the beliefs and all beliefs are consistent given the strategies (Fudenberg and Tirole 1991).

3 Model Analysis

3.1 Benchmark: No profiling

We use the no profiling scenario as the benchmark and denote it using superscript b. Under no profiling there are two stages. In stage 1 the seller sets a uniform price, and in stage 2 consumers search and then purchase. Let p^b be the uniform price in the benchmark scenario. Consumers whose valuation v exceeds $p^b + c$ search and purchase their ideal products. The seller maximizes the expected profit given by:

$$\max_{p^b} \pi^b(p^b) = p^b \int_{p^b+c}^{v_H} f(v)dv,$$

and it yields:

$$p^{b^*} = \frac{v_H - c}{2}. \tag{1}$$

3.2 Voluntary profiling

We use superscript ν to denote the voluntary profiling scenario, and subscripts 1 and 0, respectively, to denote a participating and a non-participating consumer. In stage 3, after knowing whether a consumer has participated in profiling, the seller sets prices. For a participating consumer, the seller sets price considering the posterior distribution of the consumer's valuation conditional on participation as well as the signal it receives about their valuation. For a non-participating consumer, the seller sets a uniform price. Consumers purchase their ideal products in stage 4 if their expected surpluses are non-negative.

To determine a PBE, we first propose the seller's belief about which consumers participate in profiling. Given this belief, we compute the posterior probability distribution of consumer valuation. We then derive sequentially rational strategies for the seller and consumers, and show that their strategies and the belief are consistent. We first introduce a key Lemma that helps in deriving the seller's belief about consumers' participation decisions. {Proofs for the Lemma and Theorems are provided in Appendix A}.

Lemma 1. *Consider consumers whose valuations are greater than $p_0^\nu + c$ for a given p_0^ν in each privacy group. Among these consumers,*
(i) if a consumer with valuation v participates in profiling, then any consumer whose valuation is less than v in the same privacy group participates in profiling, and
(ii) if a consumer with valuation v does not participate in profiling, any consumer whose valuation is greater than v in the same privacy group does not participate in profiling.

Lemma 1 is effectively a separation result segmenting consumers that participate and not. In Lemma 1 because the expected value of the signal increases in consumer valuation, the expected price for a participating consumer also increases in consumer valuation. On the other hand, with a uniform price to non-participating consumers, the relative benefit from participating decreases in consumer valuation. Therefore, Lemma 1 shows that among those consumers whose valuations are higher than $p_0^\nu + c$ for a given p_0^ν, if there are both participating and non-participating consumers in each privacy group, non-participating consumers have greater valuations than participating consumers. Of course, for some parameter values, it is possible that we only have either participating or non-participating consumers in each privacy group.

Using Lemma 1, we derive the seller's belief about consumers' participation decisions. We provide the intuition here and the details of the derivation in Appendix B.

The seller's belief. *The seller's belief about consumers' participation decisions consists of either one of the following two participation structures:*

(a) High-privacy-cost equilibrium: A subset of privacy-non-sensitive (NS) consumers but no privacy-sensitive (S) consumers participate in profiling.

(b) Low-privacy-cost equilibrium: All privacy-non-sensitive (NS) consumers as well as a subset of privacy-sensitive (S) consumers participate in profiling.

Figure 2 illustrates these two participation structures. The privacy-non-sensitive consumer that is indifferent between participating and not is denoted by $\bar{v}_{NS}$, and the privacy-sensitive consumer that is indifferent between participating and not is denoted by $\bar{v}_S$. While the expected surplus of the consumer with valuation $\bar{v}_{NS}$ is positive, the expected surplus of the consumer with valuation $\bar{v}_S$ is zero.

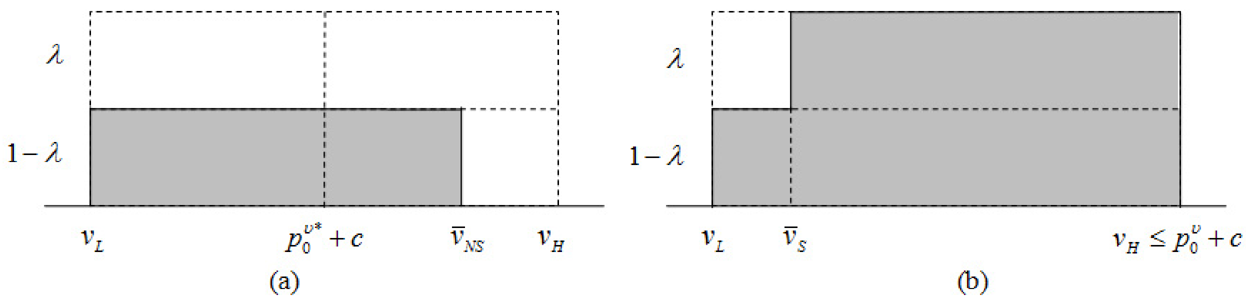


Figure 2: The seller's belief about consumers' participation decision if the privacy cost is (a) high and (b) low

Consider a privacy-non-sensitive consumer with a valuation less than $p_0^{\nu}+c$ for a given p_0^{ν}. If they do not participate in profiling, then their expected surplus is zero. If they participate, then their expected surplus is positive (or zero if their valuation is v_L) as there is a positive probability that the signal is less than (or equal to) their true valuation. Therefore, all privacy-non-sensitive consumers with valuations less than $p_0^{\nu}+c$ participate in profiling. Similarly, if privacy cost is sufficiently low, then all privacy-sensitive consumers with valuations less than $p_0^{\nu}+c$ also participate in profiling except those that have valuations below a threshold ($\bar{v}_S$ in Figure 2(b)). The expected surplus of these consumers with valuations less than $\bar{v}_S$ is not large enough to offset the privacy cost. Therefore, if privacy cost is sufficiently low (i.e., $\bar{v}_S$ is sufficiently small), then the seller has an incentive to set a high price (i.e., any price not less than $v_H - c$) for a non-participating consumer which induces all privacy-non-sensitive consumers to participate in profiling, and we have low-privacy-cost equilibrium (Figure 2(b)).

As privacy cost increases, more privacy-sensitive consumers choose to not participate in profiling (i.e., $\bar{v}_S$ increases) and the seller's incentive to charge a high price for a non-participating consumer decreases. If the privacy cost is sufficiently high (i.e., $\bar{v}_S > p_0^{\nu *} + c$), then the seller sets the price for a non-participating consumer $p_0^{\nu *}$ maximizing its expected profit from non-participating consumers and no privacy-sensitive consumer participates (Figure 2(a)). Privacy-sensitive consumers with valuations less than $p_0^{\nu *} + c$ do not participate in profiling because the expected surplus does not offset the privacy cost, and those with valuations greater than $p_0^{\nu *} + c$ do not participate in profiling because their expected surplus when they do not participate is higher. Given $p_0^{\nu *}$, only some privacy-non-sensitive consumers with valuations greater than $p_0^{\nu *} + c$ participate in profiling. As we discussed in the Lemma, if both participating and non-participating consumers exist, then there is a threshold ($\bar{v}_{NS}$ in Figure2(a)) that places all non-participating consumers to the right of all participating consumers on the valuation dimension. We derive $\bar{v}_{NS}$, $\bar{v}_S$, and the equilibrium condition for each of these two structures in Section 3.2.2.

From the seller's belief about consumers' participation decisions, the probability distribution function of valuation of a participating consumer $g(v)$ is written as:

(a) In a high-privacy-cost equilibrium

$$g(v) = \frac{1}{\bar{v}_{NS} - v_L} \text{ for } v_L \leq v \leq \bar{v}_{NS} \tag{2}$$

(b) In a low-privacy-cost equilibrium

$$g(v) = \begin{cases} \dfrac{1 - \lambda}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} & \text{for } v_L \leq v \leq \bar{v}_S \\ \dfrac{1}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} & \text{for } \bar{v}_S < v \leq v_H \end{cases} \tag{3}$$

and, the posterior distribution function of valuation of a participating consumer given the signal $\hat{v}$ is computed as:

$$s(v = y|\hat{v} = x) = \\ \frac{q(\hat{v} = x|v = y)g(y)}{q(\hat{v} = x|v = x)g(x) + \int_{z \neq x} q(\hat{v} = x|v = z)g(z)dz} = \begin{cases} \beta & \text{if } y = x \\ [1 - \beta]g(y) & \text{if } y \neq x \end{cases} \tag{4}$$

where $g(v)$ for a high-privacy-cost equilibrium and a low-privacy-cost equilibrium is given in (2) and (3) respectively.

3.2.1 Stage 3: The optimal price

When setting prices the seller knows whether a particular consumer has participated in profiling and sets the price accordingly. Thus, the seller's profit from a participating consumer depends on the posterior probability of their valuation conditional on the signal. The seller's expected profit for a participating consumer is given by:

$$\pi_1^\nu(p_1^\nu) = \begin{cases} p_1^\nu \left[s(v=\hat{v}|\hat{v}) + \displaystyle\int_{v>p_1^\nu+[1-\alpha]c} s(v \neq \hat{v}|\hat{v})dv \right] - t\alpha^2 & \text{if } p_1^\nu \leq \hat{v} - [1-\alpha]c \\ p_1^\nu \displaystyle\int_{v>p_1^\nu+[1-\alpha]c} s(v \neq \hat{v}|\hat{v})dv - t\alpha^2 & \text{if } p_1^\nu > \hat{v} - [1-\alpha]c \end{cases} \tag{5}$$

where $s(v|\hat{v})$ is given in (4). We only consider when the seller prefers to use the signal when it sets the price for a participating consumer. We show in Appendix C that if (i) $v_H - v_L$ is not too large:

$$v_L > \frac{v_H + [1-\alpha]c}{2}, \tag{6}$$

and (ii) the valuation accuracy is sufficiently high:

$$\beta > \begin{cases} \dfrac{\bar{v}_{NS} - [1-\alpha]c}{2\bar{v}_{NS} - v_L[1-\alpha]c} & \text{for a high-privacy-cost equilibrium} \\ \dfrac{v_H - [1-\alpha]c}{2v_H - [1-\lambda]v_L - \lambda\bar{v}_S - [1-\alpha]c} & \text{for a low-privacy-cost equilibrium,} \end{cases} \tag{7}$$

then it is optimal for the seller to use the signal when it sets the price for a participating consumer. Maximizing (5), yields the optimal price for a participating consumer:

$$p_1^{\nu*} = \hat{v} - [1-\alpha]c. \tag{8}$$

We now derive the optimal price for a non-participating consumer. In a high-privacy-cost equilibrium, (i) privacy-sensitive consumers whose valuations are greater than $p_0^\nu + c$ and (ii) privacy-non-sensitive consumers whose valuations are greater than $\bar{v}_{NS}$ search and purchase their ideal products at price p_0^ν. Therefore, the seller maximizes the expected profit from a non-participating consumer given by:

$$\max_{p_0^\nu} \pi_0^\nu(p_0^\nu) = p_0^\nu \left[\lambda \int_{p_0^\nu+c}^{v_H} f(v)dv + [1-\lambda] \int_{\bar{v}_{NS}}^{v_H} f(v)dv \right],$$

and it yields:

$$p_0^{\nu *} = \frac{1}{2}\left[\frac{v_H - [1-\lambda]\bar{v}_{NS}}{\lambda} - c\right]. \tag{9}$$

As we showed in the seller's belief about consumers' participation decisions, in a low-privacy-cost equilibrium for a non-participating consumer the seller charges a price not less than $v_H - c$ and no consumer purchases their ideal product at this price. The details of the derivation are available in Appendix B.

3.2.2 Stage 2: The equilibrium participation decisions

In this section, we analyze stage 2 and derive the optimal participation decisions of consumers given the seller's pricing decisions. We denote the expected surplus of a consumer with valuation v in privacy group $i \in \{S, NS\}$ as $U(v, i)$.

In a high-privacy-cost equilibrium, no privacy-sensitive consumer participates in profiling. Hence, we analyze the participation decisions of privacy-non-sensitive consumers only and show the condition under which no privacy-sensitive consumer participates in profiling later in this Section. Consider a privacy-non-sensitive consumer whose valuation is v. If they do not participate in profiling, their expected surplus is:

$$U_0^\nu(v, NS) = \begin{cases} 0 & \text{for } v_L \le v < \frac{1}{2}\left[\frac{v_H - [1-\lambda]\bar{v}_{NS}}{\lambda} + c\right] \\ v - \frac{1}{2}\left[\frac{v_H - [1-\lambda]\bar{v}_{NS}}{\lambda} + c\right] & \text{for } \frac{1}{2}\left[\frac{v_H - [1-\lambda]\bar{v}_{NS}}{\lambda} + c\right] \le v \le v_H. \end{cases} \tag{10}$$

Otherwise if they participate in profiling, then their expected surplus is:

$$U_1^\nu(v, NS) = q(\hat{v} = v|v) \cdot 0 + \int_{v_L}^{\nu} q(\hat{v} \neq v|v)[v - \hat{v}]d\hat{v} = \frac{[1-\beta][v - v_L]^2}{2[\bar{v}_{NS} - v_L]}. \tag{11}$$

Using (10) and (11), by solving $U_1^\nu(v = \bar{v}_{NS}, NS) = U_0^\nu(v = \bar{v}_{NS}, NS)$, we have:

$$\bar{v}_{NS} = \frac{v_H - \lambda[1-\beta]v_L + \lambda c}{1 + \lambda\beta}. \tag{12}$$

In a low-privacy-cost equilibrium, the seller charges any price not less than $v_H - c$ for a non-participating consumer, and hence, we have: $U_0^\nu(v, i) = 0$ for all v and $i \in \{S, NS\}$. If a privacy-sensitive consumer with valuation v participates in profiling, then their expected surplus is:

$$U_1^\nu(v,S) = \frac{[1-\beta]\left[[v-\lambda\bar{v}_S-[1-\lambda]v_L]^2+\lambda[1-\lambda][\bar{v}_S-v_L]^2\right]}{2\left[[v_H-v_L]-\lambda[\bar{v}_S-v_L]\right]} - r. \tag{13}$$

If a privacy-non-sensitive consumer with valuation v participates in profiling, her expected surplus is:

$$U_1^\nu(v,NS) = \begin{cases} \dfrac{[1-\beta][1-\lambda][v-v_L]^2}{2\left[[v_H-v_L]-\lambda[\bar{v}_S-v_L]\right]} & \text{for } v_L \le v \le \bar{v}_S \\ \dfrac{[1-\beta]\left[[v-\lambda\bar{v}_S[1-\lambda]v_L]^2+\lambda[1-\lambda][\bar{v}_S-v_L]^2\right]}{2\left[[v_H-v_L]-\lambda[\bar{v}_S-v_L]\right]} & \text{for } \bar{v}_S \le v \le v_H. \end{cases} \tag{14}$$

Using (13), by solving $U_1^\nu(v=\bar{v}_S,S) = U_0^\nu(v=\bar{v}_S,S) = 0$, we have:

$$\bar{v}_S = \frac{[1-\lambda][1-\beta]v_L-\lambda r+\sqrt{2r[1-\lambda][1-\beta][v_H-v_L]+\lambda^2r^2}}{[1-\lambda][1-\beta]}. \tag{15}$$

As we discussed in the seller's belief about consumers' participation decisions, no privacy-sensitive consumer participates in profiling if $\bar{v}_S > p_0^{\nu*}+c$. From (9) and (15) we can rewrite the condition as:

$$r > \frac{[1-\lambda][1-\beta]\left[v_H-[1-\lambda]\bar{v}_{NS}-2\lambda v_L+\lambda c\right]^2}{4\lambda^2\left[v_H+[1-\lambda]\bar{v}_{NS}-2[1-\lambda]v_L-\lambda c\right]}. \tag{16}$$

We now verify that consumers' participation strategies do not deviate from those implied by equilibrium structures (i.e., the seller's belief about consumers' participation decisions). If condition in (16) is satisfied, $U_1^\nu(v,S)$ given in (13) is negative for any v. Thus, we have:

$$U_1^\nu(v,S) < U_0^\nu(v,S) \text{ for all } v.$$

Further, by substituting (12) in (10) and (11), we show:

$$U_1^\nu(v,NS) - U_0^\nu(v,NS) \begin{cases} \ge 0 & \text{for } v_L \le v \le \bar{v}_{NS} \\ < 0 & \text{for } \bar{v}_{NS} < v \le v_H. \end{cases}$$

If condition in (16) is not satisfied, for a privacy-sensitive consumer, by substituting (15) in (13), we show:

$$\begin{cases} U_1^\nu(v,S) < U_0^\nu(v,S) = 0 & \text{for } v_L \le v < \bar{v}_S \\ U_1^\nu(v,S) \ge U_0^\nu(v,S) = 0 & \text{for } \bar{v}_S \le v \le v_H \end{cases}$$

and, for a privacy-non-sensitive consumer, by substituting (15) in (14), we show:

$$U_1^{\nu}(v, NS) \geq U_0^{\nu}(v, NS) = 0 \text{ for all } v.$$

Therefore, we confirm that the equilibrium structures characterized by (12), (15), and (16) constitute a PBE. Our next theorem shows the properties of the equilibrium.

Theorem 1. *In each equilibrium, the proportion of consumers that choose to participate is decreasing in the (i) privacy cost, (ii) valuation accuracy, and (iii) fraction of privacy-sensitive consumers.*

Theorem 1 is the outcome of the tradeoff faced by consumers when they make participation decisions. A participating consumer benefits in two ways: a reduction in search cost and a possibility of receiving a low signal and a low price. A participating consumer pays in two ways: privacy cost and a possibility of receiving a high signal and a high price. A decrease in privacy cost or the valuation accuracy increases any consumer's benefit from participation relative to the cost, inducing more consumers to participate. On the other hand, an increase in the fraction of privacy-sensitive consumers induces the seller to reduce the price for non-participating consumers that reduces the proportion of consumers (even privacy-non-sensitive) that choose to participate.

3.2.3 Stage 1: The optimal search support

Given the seller's price strategies and consumers' participation strategies, in a high-privacy-cost equilibrium, the seller maximizes the expected profit given by:

$$\begin{aligned}\max_{\alpha} \pi^{\nu}(\alpha) = p_0^{\nu *} &\left[\lambda \int_{p_0^{\nu *}+c}^{v_H} f(v)dv + [1-\lambda] \int_{v_{NS}}^{v_H} f(v)dv \right] \\ &+ [1-\lambda] \int_{v \in \Omega} \left[q(\hat{v} = v|v) p_1^{\nu *} + \int_{\hat{v}<v} q(\hat{v} \neq v|v) p_1^{\nu *} d\hat{v} \right] f(v)dv - t\alpha^2,\end{aligned}$$

where Ω is the set of valuations of consumers who participate in profiling. It yields:

$$\alpha^* = \frac{[1+\beta][1-\lambda][\bar{v}_{NS} - v_L]c}{4t[v_H - v_L]}.$$

In a low-privacy-cost equilibrium, the seller maximizes the expected profit given by:

$$\max_{\alpha} \pi^{\nu}(\alpha) = \int_{v\in\Omega} \left[q(\hat{v} = v|v)p_1^{\nu*} + \int_{\hat{v}<v} q(\hat{v} \neq v|v)p_1^{\nu*} d\hat{v} \right] f(v)dv - t\alpha^2,$$

and it yields:

$$\alpha^* = \frac{[1+\beta][[v_H - v_L] - \lambda[\bar{v}_S - v_L]]c}{4t[v_H - v_L]}.$$

Our next theorem shows the properties of the optimal search support in the high- and low-privacy cost equilibria.

Theorem 2. *In each equilibrium, the optimum level of search support is (i) decreasing in the fraction of privacy-sensitive consumers, (ii) decreasing in the privacy cost in a low-privacy-cost equilibrium but independent of the privacy cost in a high-privacy-cost equilibrium, and (iii) increasing in the valuation accuracy.*

The intuition for Theorem 2(i) is as follows. The price for a participating consumer increases in the search support (i.e., the amount of support the seller provides to a participating consumer to reduce her search cost) (see (8)). Therefore, if the proportion of participating consumers increases, then the marginal benefit from improving search support increases, and the seller has more incentive to improve the search support. As a decrease in the fraction of privacy-sensitive consumers increases the proportion of participating consumers (Theorem 1(iii)), we obtain Theorem 2(i). Similarly, for Theorem 2(ii) in a low-privacy-cost equilibrium, as the privacy cost increases, fewer privacy-sensitive consumers participate in profiling, and the seller invests less in the search support. On the other hand, in a high-privacy-cost equilibrium, no privacy-sensitive consumer participates in profiling. Hence, the number of participating consumers, thereby, the search support is independent of privacy cost. For Theorem 2(iii), although an increase in the valuation accuracy decreases the proportion of participating consumers (Theorem 1(ii)), it increases the search support. It is because an increase in the valuation accuracy increases the expected proportion of participating consumers that purchase their ideal products (i.e., the demand) which increases the marginal benefit from improving search support, and the seller has more incentive to improve the preferenc accuracy.[5]

[5]Note that the signal can be lower or higher than the true valuation of a consumer with some probability.

3.2.4 The impact of privacy sensitivity and price discrimination on the seller, consumer surplus, and social welfare under voluntary profiling

We show a set of results that provides key insights into how privacy sensitivity (as measured by the fraction of privacy-sensitive consumers and privacy cost) and price discrimination (using valuation accuracy as a proxy) impact the seller, consumer surplus, and social welfare under voluntary profiling. As our analytical model leads to some complex expressions, we provide analytical results for tractable cases and confirm these results using numerical analyses for more general cases.

Theorem 3. *(i) If there are no privacy-sensitive consumers, then consumer surplus is decreasing and both seller's profit and social welfare are increasing in the valuation accuracy. (ii) If all consumers are privacy sensitive, then the seller's profit, consumer surplus, and social welfare are independent of both the valuation accuracy and privacy cost. (iii) If the valuation accuracy is perfect, then consumer surplus is increasing and seller's profit is decreasing in the fraction of privacy-sensitive consumers; and social welfare is non-monotonic in the fraction of privacy-sensitive consumers but is higher when there are no privacy sensitive consumers than when all consumers are privacy sensitive.*

In Theorem 3, if there are no privacy-sensitive consumers, then every consumer participates in profiling and an increase in the valuation accuracy allows the seller to extract more of consumer surplus. Further, an increase in the valuation accuracy decreases the possibility that the signal is higher than the true valuation of a consumer, and hence, it increases the number of participating consumers who actually purchase their ideal products. Consequently, although price discrimination hurts consumers, it benefits the seller and society. If all consumers are privacy-sensitive, then no consumer participates in profiling, and the seller's profit, consumer surplus and social welfare are independent of both valuation accuracy and privacy cost. Theorem 3(iii) offers an important insight: if profiling enables the seller to infer the true valuation of consumers through perfect valuation accuracy, then a decrease in the fraction of privacy-sensitive consumers does not benefit consumer surplus.

If the signal is higher than the true valuation of a consumer, she does not purchase. The increase in the valuation accuracy decreases the possibility that the signal is higher than the true valuation of a consumer.

The decrease in the fraction of privacy-sensitive consumers increases the proportion of participating consumers and the seller's incentive to charge a high price for a non-participating consumer, thus hurting consumers in aggregate and benefitting the seller. Social welfare, on the other hand, is higher when more consumers participate in profiling compared to when no consumer participates in profiling because if valuation accuracy is perfect, then all participating consumers purchase their ideal products but only a subset of non-participating consumers purchases their ideal products.

Extensive numerical analysis show that the essential insights provided by Theorem 3 hold for other cases. A representative set of results is shown in Figure 3. The following parameter values are used: $V_H = 100$, $V_L = 70$, $c = 90$, $t = 45$, $\lambda = 0.3$, $\beta \in \{0.85, 0.9, 0.95\}$, and $r \in [0, 2.5]$. In Figure 3, the discontinuity in each of the graphs corresponds to the point at which the equilibrium structure switches from one to the other.

A decrease in privacy cost affects consumer surplus only when some privacy-sensitive consumers find it beneficial to participate in profiling (i.e., low-privacy-cost equilibrium). Consumer surplus decreases when the participation structure switches from a high-privacy-cost equilibrium to a low-privacy-cost equilibrium, as privacy cost decreases. However, once it switches to a low-privacy-cost equilibrium, further decrease in privacy cost increases consumer surplus. It is worth noting that consumer surplus when privacy cost is low is not necessarily higher than when it is high. If the valuation accuracy is high ($\beta = 0.95$), the consumer surplus in a high-privacy-cost equilibrium is higher than that at any privacy cost (even when the privacy cost is zero) in a low-privacy-cost equilibrium. The increase in valuation accuracy, on the other hand, always reduces consumer surplus.

A decrease in privacy cost always increases both the seller's profit and social welfare. An increase in the valuation accuracy, however, does not necessarily increase the seller's profit and social welfare. The increase in the valuation accuracy decreases the number of non-participating consumers who purchase their ideal products, whereas increases the number of participating consumers who purchase their ideal products. Therefore, the increase in the valuation accuracy hurts the seller and the society when the former effect is greater than the latter effect. Figure 3 shows that if privacy cost r is in the range of $[0.25, 1.50]$, the seller's

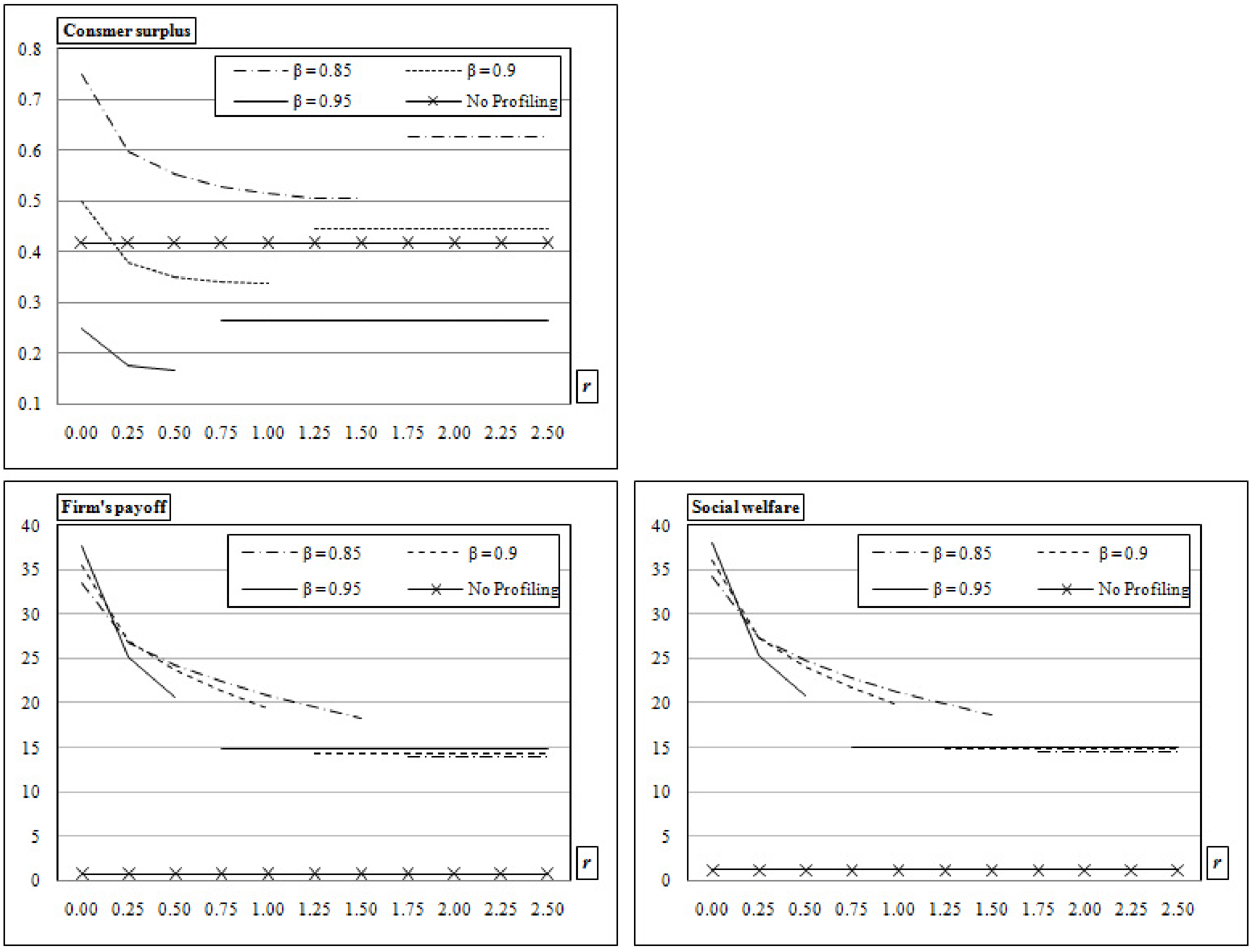


Figure 3: Impact of r and β on consumer surplus, seller's payoff and social welfare

profit and social welfare when valuation accuracy $\beta = 0.85$ is higher than when $\beta = 0.95$.

4 Comparison of Voluntary Profiling and No Profiling

Here we show key results regarding how voluntary profiling affects the price paid by consumers, consumer surplus, and social welfare, relative to no profiling. The seller's profit is higher under voluntary profiling than under no profiling because the seller can always choose to not use consumer profile information. The sufficient conditions for the seller's use of profile information are given in (6) and (7).

4.1 Price paid by consumers

The price offered to a participating consumer depends on the signal obtained by the seller for her. Further, a participating consumer will not purchase if the signal is higher than her true valuation. Therefore, the expected average price paid by a participating consumer is computed as:

$$\begin{aligned} &E_v(E_{\hat{v}}(I(\hat{v},v)\; p_1^{\nu}(\hat{v}|v))) \\ &= \int_v g(v) \left[q(\hat{v}=v|v)[v-[1-\alpha]c] + \int_{\hat{v}<v} q(\hat{v}\neq v|v)[\hat{v}-[1-\alpha]c]d\hat{v} \right] dv \end{aligned} \tag{17}$$

where $I(\hat{v},v)$ is an indicator function that is equal to 1 if $\hat{v}\leq v$ and 0 otherwise. In a high-privacy-cost equilibrium, the price for a non-participating consumer is given in (9). In a low-privacy-cost equilibrium, the seller charges a price not less than $v_H - c$ for a non-participating consumer, and no non-participating consumer purchases at this price. Therefore, there is no price at which a non-participating consumer purchases in a low-privacy-cost equilibrium. Comparing the prices under voluntary profiling (given in (9) and (17)) and the price under no profiling (given in (1)), we have the following theorem.

Theorem 4. *Compared to no profiling, under voluntary profiling: (i) the expected price paid by a non-participating consumer is higher, and (ii) if the search cost is sufficiently high, then the expected price paid by a participating consumer is higher.*

Because voluntary profiling skews the distribution of valuations of non-participating consumers to the right, on the average, the valuation of non-participating consumers is higher than that of the consumer population. This induces the seller to charge a higher price for non-participating consumers under voluntary profiling than for the population under no profiling. Further, although participating consumers include low valuation consumers (those not served under no profiling), the expected average price paid by a participating consumer under voluntary profiling is higher than the price paid by a consumer under no profiling if the search cost is sufficiently high. This is because under voluntary profiling the seller is able to charge a higher price and extracts the participating consumer's search cost saving from the personalized search process.

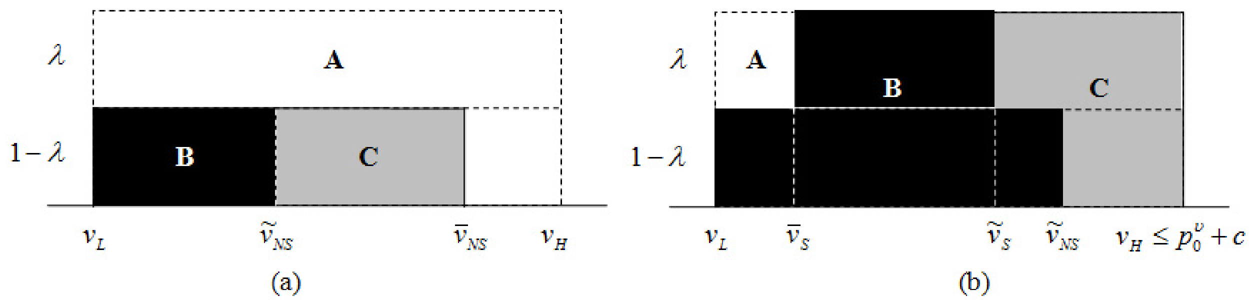
By comparing (9) and (17), we find that the expected price paid by a participating consumer can be even higher than the price paid by a non-participating consumer under voluntary profiling if the search cost is sufficiently high (i.e., $c > \frac{3[v_H - \bar{v}_{NS}] + 2\lambda[1-\beta]\bar{v}_{NS} - \lambda[2+\beta]v_L + 3\lambda\beta c}{3\lambda[\alpha[1+\beta]-\beta]}$). This is surprising because conventional reasoning is that the seller rewards those that voluntarily participate in profiling. Despite the counter-intuitive nature of this result, it provides important theoretical support to the anecdotal observations reported in blog sites about prices paid by consumers that participate in profiling. For instance, it has been reported that some members of *Prime membership* of Amazon.com pay about 20% more than non-members for the same item.

4.2 Consumer surplus

Let $\tilde{v}_i$, $i \in \{S, NS\}$ denote the valuation of the participating consumer in privacy group i that is indifferent between voluntary profiling and no profiling. We have the following theorem about the surplus of individual consumers.

Theorem 5. *Compared to no profiling, under voluntary profiling (i) the surplus of a non-participating consumer is not larger, and (ii) the surplus of a participating consumer whose valuation is greater than $\tilde{v}_i$ is smaller.*

Figure 4 illustrates Theorem 5. Region A includes all non-participating consumers, region B includes participating consumers that enjoy a higher surplus under voluntary profiling than under no profiling, and region C includes participating consumers that have a smaller surplus under voluntary profiling than under no profiling. Theorem 5 shows that contrary to intuition, not all consumers are better off under voluntary profiling compared to no profiling. In fact, no non-participating consumer is better off under voluntary profiling (Region A). More importantly, some consumers (Region C) may choose to participate in profiling although they are worse off under voluntary profiling compared to no profiling. This unintended consequence of voluntary profiling arises from the externality imposed by those low-valuation consumers in Region B that are better off under voluntary profiling. As these consumers prefer to participate, the distribution of the valuations of non-participating consumers is skewed to the right under voluntary profiling. Although the seller does not collect any information from non-participating consumers, as shown in Theorem 4(i), the seller charges non-participating consumers a higher price than what it would have charged under no profiling. Hence, consumers in Region C face the problem of deciding between participating and not participating, both of which are worse than no profiling, and find that participating is better.



Note: A. $U_0^{\nu}(v) \leq U^b(v)$, B. $U_1^{\nu}(v) > U^b(v)$, and C. $U_1^{\nu}(v) \leq U^b(v)$

Figure 4: The equilibrium under the voluntary profiling if the privacy cost is (a) high and (b) low

Although not all consumers are better off under voluntary profiling than under no profiling, there are consumers that are clearly better off (Region B). Hence, it is possible that aggregate consumer surplus is higher under voluntary profiling than under no profiling. To see if this is the case, we compare the aggregate consumer surplus under voluntary profiling and no profiling. The aggregate consumer surplus under no profiling is

$$TU^b = \int_{p^{b*}+c}^{v_H} f(v)[v - p^{b*} - c]dv.$$

The aggregate consumer surplus under voluntary profiling for a high-privacy-cost equilibrium is

$$\begin{aligned} TU^{\nu} = {} & \lambda \int_{p_0^{\nu*}+c}^{v_H} f(v) U_0^{\nu}(v, S) dv \\ & + [1-\lambda] \left[\int_{v_L}^{\bar{v}_{NS}} f(v) U_1^{\nu}(v, NS) dv + \int_{\bar{v}_{NS}}^{v_H} f(v) U_0^{\nu}(v, NS) dv \right], \end{aligned}$$

and for a low-privacy-cost-equilibrium is

$$TU^{\nu} = \lambda \int_{\bar{v}_S}^{v_H} f(v) U_1^{\nu}(v, S) dv + [1-\lambda] \int_{v_L}^{v_H} f(v) U_1^{\nu}(v, NS) dv.$$

Using above expressions for aggregate consumer surplus, we illustrate with a numerical example in Appendix A that the aggregate consumer surplus maybe higher or lower under voluntary profiling than no profiling. Consequently, the benefit offered by voluntary profiling to consumers in Region B may or may not offset the loss it inflicts on consumers in regions A and C in Figure 4.

In order to provide additional insights into when voluntary profiling increases (or decreases) aggregate consumer surplus, we present sample results from our numerical analysis in Figure 5 (parameter values used: $v_H = 100$, $v_L = 70$, $t = 45$, $\lambda = 0.3$ for both panels and $r \geq 1.5$ for panel a and $r = 0.2$ for panel b). Figure 5 suggests that voluntary profiling will likely lead to a higher consumer surplus compared to no profiling when the search cost savings is high (i.e., which is the case when the search cost c is high) and the valuation accuracy is low. Furthermore, comparing the two panels of the figure, we find that a reduction in privacy cost under voluntary profiling may harm consumers compared to no profiling. This occurs in the region where it is gray in panel (a) and white in panel (b) of the figure.

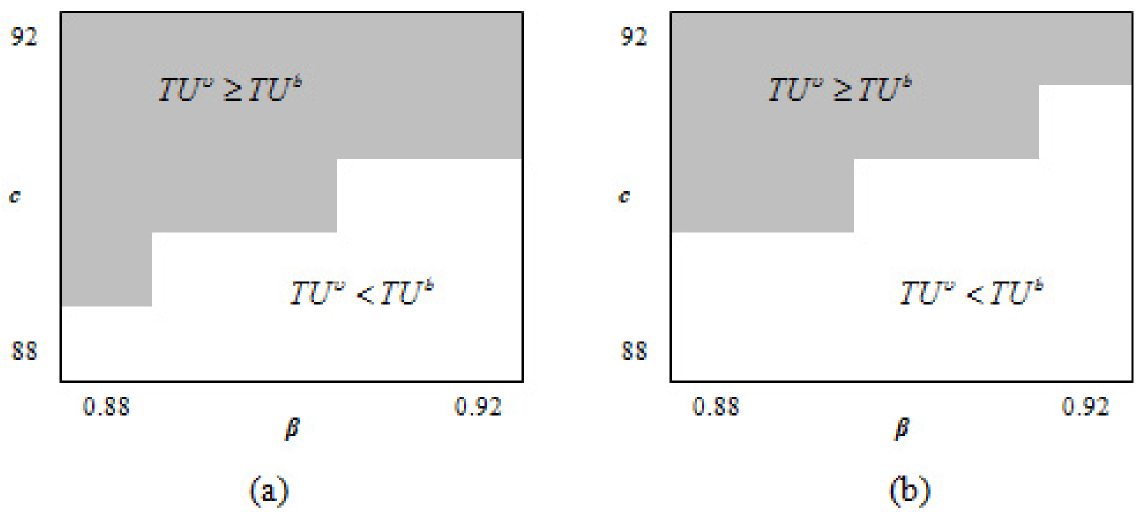


Figure 5: Comparison of consumer surplus under no profiling and voluntary profiling if the privacy cost is (a) high and (b) low

4.3 Social welfare

Our interest in analyzing the impact of voluntary profiling on social welfare stems from the fact that a social planner regulates profiling policy. We define social welfare (SW) as *the seller's expected profit + expected aggregate consumer surplus.* Social welfare under no profiling is

$$SW^b = \int_{p^{b*}+c}^{v_H} f(v)[v-c]dv.$$

Social welfare under voluntary profiling for a high-privacy-cost equilibrium is

$$\begin{aligned} SW^\nu = \lambda & \int_{p_0^{\nu *}+c}^{v_H} f(v)[v-c]dv \\ &+ [1-\lambda] \int_{v_L}^{\bar{v}_{NS}} f(v) \left[q(\hat{v}=v|v)[v-[1-\alpha]c] + \int_{\hat{v}<v} q(\hat{v}\neq v|v)[v-[1-\alpha]c]d\hat{v} \right] dv \\ &+ [1-\lambda] \int_{\bar{v}_{NS}}^{v_H} f(v)[v-c]dv - t\alpha^2, \end{aligned}$$

and for a low-privacy-cost-equilibrium is

$$\begin{aligned} SW^\nu = \lambda & \int_{\bar{v}_S}^{v_H} f(v) \left[q(\hat{v}=v|v)[v-[1-\alpha]c] + \int_{\hat{v}<v} q(\hat{v}\neq v|v)[v-[1-\alpha]c]d\hat{v} - r \right] dv \\ &+ [1-\lambda] \int_{v_L}^{v_H} f(v) \left[q(\hat{v}=v|v)[v-[1-\alpha]c] + \int_{\hat{v}<v} q(\hat{v}\neq v|v)[v-[1-\alpha]c]d\hat{v} \right] dv - t\alpha^2. \end{aligned}$$

Using above expressions for social welfare, we illustrate with a numerical example in Appendix A that social welfare maybe higher or lower under voluntary profiling than no profiling. Profiling benefits society in two ways: it reduces search cost that is a deadweight loss to the society and allows the seller to serve some low valuation consumers that would not be served under no profiling. Profiling may also harm society in two ways. Privacy-sensitive consumers incur privacy cost if they participate in profiling and some privacy-sensitive consumers that would be served under no profiling may not be served under voluntary profiling. If the latter negative effect dominates the former positive effect, the society is worse off under voluntary profiling compared to no profiling.

We present results from our numerical analysis in Figure 6 (parameter values used: $v_H = 100$, $v_L = 70$, $t = 45$, $\lambda = 0.3$ for both panels, $r \geq 0.67$ for panel (a) and $r = 0.1$ for panel (b)). Figure 6 suggests that voluntary profiling will likely lead to higher social welfare compared to no profiling when the search cost is high. Voluntary profiling leads to a significant reduction in the deadweight loss due to wasteful search when the search cost is high.

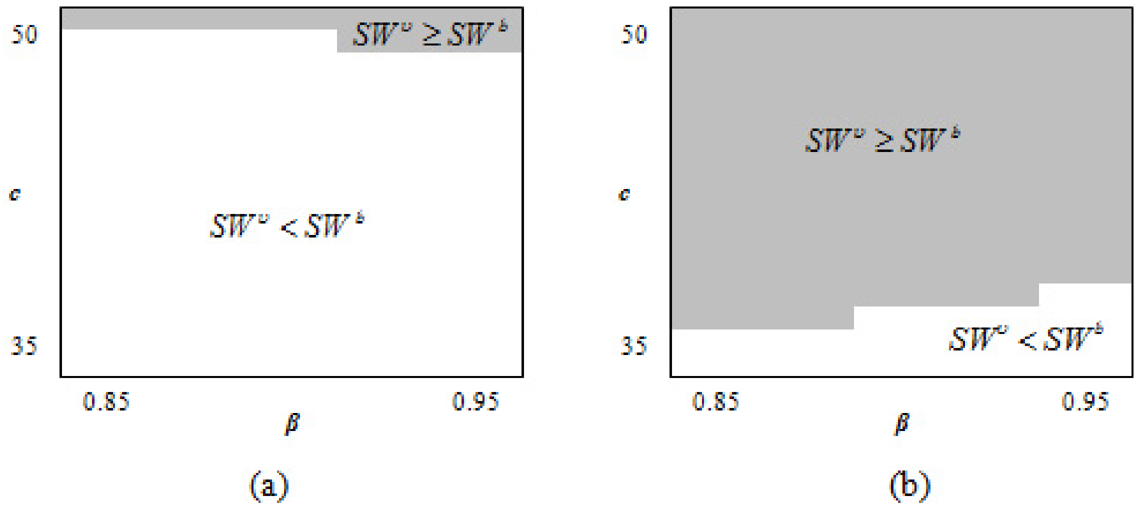


Figure 6: Comparison of social welfare under no profiling and voluntary profiling if the privacy cost is (a) high and (b) low

5 Social policy

Although voluntary profiling always benefits the seller, it does not always benefit consumers and society. Even if voluntary profiling benefits society, it is possible that the seller may benefit at the expense of consumers. Therefore, a policy maker has to balance the interests of industry and consumers before choosing between voluntary profiling and no profiling. One approach that a policy maker may follow is to adopt voluntary profiling only when it is Pareto optimal: neither the seller nor the consumers are worse off under voluntary profiling. In addition, as consumer advocates oppose voluntary profiling on price discrimination and privacy grounds, a mitigation of these concerns may also help a policy maker gain the support of consumer advocates for voluntary profiling. However, our theoretical results showed that reducing or eliminating privacy concerns may in fact harm consumers (Theorem 3(iii)), so reducing or eliminating privacy concerns alone would not necessarily lead to a Pareto improvement. Furthermore, price discrimination may favor the seller at the expense of consumers (Theorem 3(i)). Therefore, in this section, we analyze whether removing the price discrimination under voluntary profiling may result in a Pareto improvement or even Pareto optimality.

Suppose the policy maker could preclude the seller using price discrimination. This could happen through legal means, through social ones such as online forums creating negative publicity for the seller if it practices price discrimination so that the seller does not engage in price discrimination (e.g., consumer forums on the Internet caused Amazon.com to stop its price discrimination experiment for DVDs), or through self-regulation where the seller credibly commits not to price discriminate. Thus, the seller would charge a uniform price for all consumers under voluntary profiling – what we term *price discrimination-free voluntary profiling*. In this scenario, we have the following theorem.

Theorem 6. *Compared to no profiling, under price discrimination-free voluntary profiling (i) the surplus of every individual privacy-non-sensitive consumer is higher, (ii) aggregate consumer surplus is higher, and (iii) social welfare is higher.*

Since the seller's profit under voluntary profiling is always higher compared to no profiling, we have Theorem 6(iii) from (ii). Nonetheless, even though the benefits of price discrimination-free voluntary profiling to privacy-non-sensitive consumers outweigh the effects on privacy-sensitive consumers, this latter group may still be worse off. As we show in the following corollary, this depends on the magnitude of the privacy cost.

Corollary 1. *Compared to no profiling, the surplus of every privacy-sensitive consumer is higher under price discrimination-free voluntary profiling if and only if the privacy cost is sufficiently low.*

The aggregate consumer surplus, and thereby, social welfare is higher under voluntary profiling if price discrimination is eliminated. Further, if the privacy cost is sufficiently low, voluntary profiling policy is Pareto optimal under price discrimination-free voluntary profiling. This suggests that price discrimination rather than privacy cost may hinder consumers from reaping the potential benefits from voluntary profiling. Therefore, while the intuition that voluntary profiling makes every one better off compared to no profiling is valid if price discrimination and privacy cost are completely eliminated, our analysis provides theoretical support to consumer advocates' concern that consumers are harmed by voluntary profiling when price discrimination and privacy harm exist.

6 Presence of an Outside Option

We assumed a monopoly online seller in the preceding analysis. Consumer profiling often allows a monopoly seller to practice price discrimination, which as we showed may harm consumers and reduce social welfare. In Section 5, we showed that a policy maker can achieve Pareto improvement or maybe Pareto optimality by restricting price discrimination along with privacy concerns under voluntary profiling. It is also likely that competition among sellers will restrict price discrimination.

In this section, we study voluntary profiling when consumers have an outside option (i.e., a substitute product) available at a fixed price p^o. This outside option price, p^o, can be interpreted as a proxy for the intensity of price competition in the market; a smaller value for p^o suggests more intense price competition. In this case a consumer will purchase from the seller only if the price is lower than p^o, and as such p^o bounds the upper limit of the seller's prices. We take p^o to be exogenously given and known to the seller as well as all consumers. We restrict our analysis to the high privacy-cost equilibrium and when the search support is exogenously specified. In this scenario, we have the following theorem.

Theorem 7. *When consumers have an outside option at p^o, compared to no profiling, under voluntary profiling, there is a threshold value for p^o: (i) if p^o is less than the threshold value, then no non-participating consumer is worse off and all participating consumers are better off; hence, aggregate consumer surplus as well as social welfare are higher, otherwise (ii) if p^o is greater than or equal to the threshold value, then the results are qualitatively similar to those in Section 3 and 4; all non-participating consumers as well as some participating consumers are worse off, and aggregate consumer surplus and social welfare can be higher or lower.*

We illustrate Theorem 7 regarding how the outside option price affects the consumer surplus and social welfare under voluntary profiling using numerical examples in Figure 7 (parameter value used: $v_H = 100$, $v_L = 70$, $\lambda = 0.3$, $\beta \in \{0.85, 0.95\}$, $p^o \in [46, 60]$ for both panels, $c = 90$, $\alpha = 0.5$ for panel a, and $c = 40$, $\alpha = 0.45$ for panel b). Figure 7(a) shows that as p^o decreases, consumers benefit more from voluntary profiling. When p^o is sufficiently high, aggregate consumer surplus under voluntary profiling can be either higher or lower than no profiling. However, when p^o is lower than a certain threshold, aggregate consumer surplus under voluntary profiling is always higher compared to no profiling. Similar results hold for social welfare (Figure 7(b)).

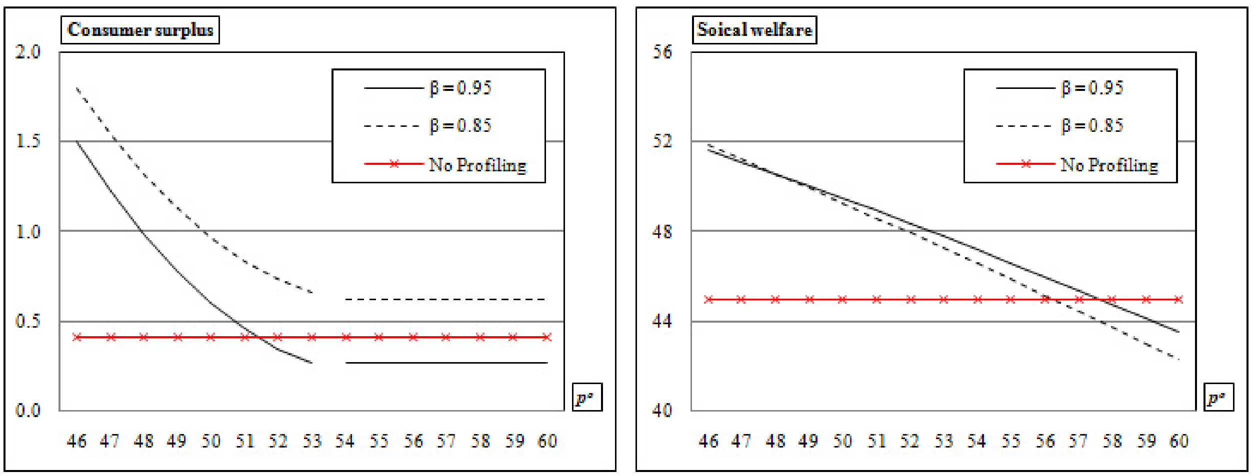


Figure 7: The effect of p^o on (a) consumer surplus and (b) social welfare under voluntary profiling

Our results provide an important insight for a policy maker. If the market is competitive in prices, then the social planner may not need to intervene in the market and restrict price discrimination and address privacy concerns. Even when consumers' privacy concerns are severe, price competition among sellers will preclude the seller from customizing prices for high-valuation consumers, and voluntary profiling will likely benefit all consumers and society. However, if the market is differentiated or segmented in some way (e.g., regionally separated (Dana and Spier 1994) or quality differentiated (Judd 1985; Nault 1997)) such that price competition is mitigated, then voluntary profiling can still harm consumers and reduce social welfare. If this is the case, the social planner may need to restrict the use of price discrimination to ensure that voluntary profiling benefits consumers as well as society.

7 Conclusion

In a model with privacy-sensitive and privacy-non-sensitive consumers, and a seller that may practice price discrimination, we studied the impact of the voluntary profiling policy on consumer surplus, seller's profit, and social welfare. We showed that every consumer that chooses to not participate in profiling and, more importantly, some consumers that choose to participate in profiling are worse off under voluntary profiling than under no profiling. Furthermore, neither the social welfare nor the aggregate consumer surplus is necessarily higher under voluntary profiling relative to no profiling. If price discrimination is restricted, then voluntary profiling leads to a Pareto-optimal outcome, but eliminating privacy cost alone does not guarantee a Pareto-optimal outcome. The primary reason for these results is that voluntary profiling allows the seller to exploit the heterogeneity in consumers' inherent incentives to participate in profiling and force participating consumers to impose a negative externality on non-participating consumers, but privacy cost moderates the impact of price discrimination on consumers. Without profiling, the seller is unable to exploit the consumer heterogeneity.

There are both managerial and policy implications of our results. Taking into account that sellers will choose to use profile information and to price discriminate when it is profit maximizing, sellers have control over their level of search support, and this in turn depends on consumers' privacy cost and the fraction of privacy-sensitive consumers. To the degree that sellers can reassure consumers about the security of the information they provide online, and consequently reduce the fraction of privacy-sensitive consumers and the privacy costs faced by those consumers, sellers can be more profitable. In this sense, sellers' focus should be on consumers' privacy concerns.

In contrast, a social planner should be more concerned about price discrimination under voluntary profiling than about privacy. The ability to price discriminate based on profile information is what social planners and regulators are typically concerned about, but as we show it is not primarily because profile information is used to charge higher prices to participating consumers. Rather it is because of the self-selection of consumers participating in voluntary profiling – lower valuation consumers participate, and with this knowledge the seller can charge a higher price to those higher valuation consumers that do not participate. Thus, the social planner may attempt to restrict price discrimination, which is hard to do in practice with special offers being easily directed to participating consumers, or encourage competition as a way to mitigate price discrimination.

Limitations We used a stylized model to perform our analysis, and in order to test the robustness of our results, we analyzed several extensions. First we considered an extension in which a participating consumer can misrepresent her identity at a cost and purchase at the non-participant price. We found that our qualitative results hold if the cost of misrepresentation is not too low. Next, we examined a model in which there are homogenous competing sellers. In this setup, a consumer can search for her ideal product using the recommendation features of a seller that has her profile information, but can switch to a competitor at a cost, such as the effort required to enter required information such as address, credit card details, and shipping preferences, and purchase the product as a non-participant from the competitor. Again, our qualitative results hold if the switching costs are not too low.

As with any analytical model based on a series of assumptions, our model has some limitations. We assumed that consumers incur a common fixed search cost to find their ideal product. In addition, we assumed that all privacy-sensitive consumers have the same fixed privacy cost. In reality, however, each consumer may have a different search cost and a different privacy cost. An extension of our research may be analyzing consumers' participation decisions under a heterogeneous search and privacy cost setup, where these costs are related to consumers' valuations. Although not directly an extension of our model, a richer description for how consumer profile information affects the seller's support for search process and consumers search cost could provide more comprehensive insights about ways in which profiling is implemented, and this might be a fruitful area of future research.

8 References

Acquisti, A., and Grossklags, J. 2005. Privacy and Rationality in Individual Decision Making. *IEEE Security & Privacy*. 3(1): 26-33.

Acquisti, A. and H. Varian 2005. Conditioning Prices on Purchase History. *Marketing Science*. 24(3): 367-381.

Aloysius, J., Deck, C., and Farmer, A. 2012. Sequential Pricing of Multiple Products: Leveraging Revealed Preferences of Retail Customers Online and with Auto-ID Technologies. *Information Systems Research*. Articles in Advance: 1-22.

Aron, R., Sundararajan, A., and Viswanathan, S. 2006. Intelligent Agents in Electronic Markets for Information Goods: Customization, Preference Revelation and Pricing. *Decision Support Systems*. 41(4): 764-786.

Baye, M.R., and Morgan, J. 2002. Information Gatekeepers and Price Discrimination on the Internet. *Economics Letters*. 76(1): 47-51.

Cachon G., Terwiesch, C., and Xu, Y. 2008. On the Effects of Consumer Search and Firm Entry in a Multiproduct Competitive Market. *Marketing Science*. 27(3): 461-473.

Calzolari, G. and A. Pavan 2006. On the Optimality of Privacy in Sequential Contracting. *Journal of Economic Theory*. 130: 168204.

Chellappa, R., and Shivendu, S. Winter 2007-8. An Economic Model of Privacy: A Property Rights Approach to Regulatory Choices for Online Personalization. *Journal of Management Information Systems*. 24(3): 193-225.

Chellappa, R., and Shivendu, S. 2010. Mechanism Design for Free but No Free Disposal Services: The Economics of Personalization under Privacy Concerns. *Management Science*. 56(10): 1766-1780.

Chen, Y. 1997. Paying Customers to Switch. *Journal of Economics and Management Strategy*. 6: 877897.

Chen, Y., and Iyer, G. 2002. Consumer Addressability and Customized Pricing. *Marketing*

Science. 21(2): 197-208.

Chen, Y. and J. Zhang. 2009. Dynamic targeted pricing with strategic consumers. *International Journal of Industrial Organization*. 27: 4350.

Chester, J. 2009. Testimony on Behavioral Advertising: Industry Practices and Consumers' Expectations. Center for Digital Democracy. Available at http://www.democraticmedia.org/doc/cdd-testimony-20090618. (18 June 2009).

Clemons E.K., Hann, I., and Hitt, L.M. 2002. Price Dispersion and Differentiation in Online Travel: An Empirical Investigation. *Management Science*. 48(4): 534-549.

Conitzer, V., Taylor, C., and Wagman, L. 2011. Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases. *Marketing Science*. 31: 277-292.

Dana, J., and K.E. Spier. 1994. Designing a private industry : Government auctions with endogenous market structure. *Journal of Public Economics.*. 53(1): 127-147.

Earp, J.D., and Baumer, D. 2003. Innovative Web Use to Learn About Consumer Behavior and Online Privacy. *Communications of The ACM*. 46(4): 81-83.

Edwards, M. 2006. Price and Prejudice: The Case Against Consumer Equality in the Information Age. *Lewis & Clark Law Review*.

Ellison, G., and Ellison, S.F. 2004. Search, Obfuscation, and Price Elasticities on the Internet. *Econometrica*. 77(2): 427-452.

Federal Trade Commission. 2000. Online Profiling: A Report to Congress. Available at http://www.ftc.gov/reports/index.shtm#2000.

Fudenberg, D. and J. Tirole, 1991. Perfect Bayesian Equilibrium and Sequential Equilibrium. *Journal of Economic Theory*. 53(2): 236-260.

Fudenberg, D. and J. Tirole, 1998. Upgrades, Tradeins, and Buybacks. *RAND Journal of Economics*. 29: 238258.

Fudenberg, D., and Villas-Boas J.M. 2007. Behavior-Based Price Discrimination and Customer Recognition. Economics and Information Systems. Volume 1. Elsevier Science.

Hann, I., Hui, K., Lee, S.T., and Png., I. 2008. Consumer Privacy and Marketing Avoidance: A Static Model. *Management Science*. 54(6): 1094-1103.

Hoffman, D., Novak, T.P., and Peralta, M. 1999. Building Con Trust Online. *Communications of the ACM*. 42(4): 80-85.

Iyer G., Soberman, D., and Villas-Boas, J.M. 2005. The Targeting of Advertising. *Marketing Science*. 24(3): 461-476.

Judd, K.L. 1985. Credible Spatial Preemption. *RAND Journal of Economics*. 16(2): 153-166.

Li, H., and Dinlersoz, E. 2012. Quality-based Price Discrimination: Evidence from Internet Retailers Shipping Options. *Journal of Retailing*. 88(2): 276-290.

Milne, G.R., and Rohm., A.J. 2000. Consumer Privacy and Name Removal across Direct Marketing Channels, Exploring Opt-In and Opt-Out Alternatives. *Journal of Public Policy & Marketing*. 19(2): 238-249.

Nault, B.R. 1997. Quality differentiation and adoption costs: The case for interorganizational information system pricing. *Annals of Operations Research*. 71: 115-142.

Odlyzko, A. 2003. Privacy, Economics, and Price Discrimination on the Internet. *ACM International Conference Proceeding Series*. 50: 255-266.

Schön, C. 2010. On the Optimal Product Line Selection Problem with Price Discrimination. *Management Science*. 56(5): 896-902.

Streitfeld, D. 2000. On the Web, Price Tags Blur: What You Pay Could Depend on Who You Are. Available at http://www.washingtonpost.com/ac2/wp-dyn/A15159-2000Sep25. (27 September 2000).

Taylor, C. 2004. Consumer Privacy and the Market for Customer Information. *RAND Journal of Economics*. 35(5): 631-650.

Turow, J., Feldman, L., and Meltzer, K. 2005. Open to Exploitation: America's Shoppers Online and Offline. A Report from the Annenberg Public Policy Center of the University of

Pennsylvania.

Valentino-Devries, J., J. Singer-Vine, A. Soltani. 2012. Websites Vary Prices, Deals Based on Users Information. *The Wall Street Journal*. Available at: http://online.wsj.com/news/articles/SB1000142
(24 December 2012).

Villas-Boas, J. M. 1999. Dynamic Competition with Customer Recognition. *RAND Journal of Economics*. 30: 604631.

Villas-Boas, J. M. 2004. Price Cycles in Markets with Customer Recognition. *RAND Journal of Economics* 35: 486501.

Viswanathan, S., Kuruzovich J., Gosain, S., and Agarwal, R. 2007. Online Infomediaries and Price Discrimination: Evidence from the Automotive Retailing Sector. *Journal of Marketing*. 71(3): 89-107.

Volokh, E. 2000. Personalization and Privacy. *Communications of The ACM*. 43(8): 84-88.

Zhang, J., and Krishnamurthi, L. 2004. Customizing Promotions in Online Stores. *Marketing Science*. 23(4): 561-578.

9 Appendix A

A.1. Proof of Lemma 1

Consider a consumer whose valuation, v, is greater than $p_0^\nu + c$ for a given p_0^ν in privacy group $i \in \{S, NS\}$. Her expected surplus if she participates in profiling is given by:

$$U_1^\nu(v,i) = v - E_{\hat{v}|v}\left(p_1^\nu(\hat{v})|v\right) - [1-\alpha]c - r_i \tag{18}$$

and, if she does not participate in profiling is given by:

$$U_0^\nu(v,i) = v - p_0^\nu - c \tag{19}$$

where, $r_S = r$, $r_{NS} = 0$, and

$$E_{\hat{v}|v}\left(p_1^\nu(\hat{v})|v\right) = q(\hat{v}=v|v)p_1^\nu(\hat{v}=v) + \int_{\hat{v}\neq v} q(\hat{v}|v)p_1^\nu(\hat{v})d\hat{v} \tag{20}$$

Let there is v_1 such that $U_1^{\nu}(v_1, i) < U_0^{\nu}(v_1, i)$. Then, from (18) and (19), we have:

$$E_{\hat{v}|v}\left(p_1^{\nu}(\hat{v})|v = v_1\right) > p_0^{\nu} + \alpha c - r_i \tag{21}$$

Further, from (20), for any $v_2 > v_1$, we have:

$$E_{\hat{v}|v}\left(p_1^{\nu}(\hat{v})|v = v_2\right) > E_{\hat{v}|v}\left(p_1^{\nu}(\hat{v})|v = v_1\right) \tag{22}$$

because: (i) $q(\hat{v} = v_2|v_2) = q(\hat{v} = v_1|v_1)$, (ii) $\int_{\hat{v}\neq v_2} q(\hat{v}|v_2)p_1^{\nu}(\hat{v})d\hat{v} = \int_{\hat{v}\neq v_1} q(\hat{v}|v_1)p_1^{\nu}(\hat{v})d\hat{v}$, and (iii) $p_1^{\nu}(\hat{v} = v_2) > p_1^{\nu}(\hat{v} = v_1)$. From (21) and (22), therefore, we show:

$$U_1^{\nu}(v_2, i) < U_0^{\nu}(v_2, i) \text{ for any } v_2 > v_1$$

Now suppose: $U_1^{\nu}(v_1, i) > U_0^{\nu}(v_1, i)$. Analogously, we show:

$$U_1^{\nu}(v_2, i) > U_0^{\nu}(v_2, i) \text{ for any } v_2 < v_1. \ \square$$

A.2. Proof of Theorem 1

Using $\bar{v}_{NS}$ given (12) and $\bar{v}_S$ given (15), we show:

$$\frac{\partial \bar{v}_{NS}}{\partial \beta} = -\frac{\lambda\left[[v_H - v_L] - \lambda[v_L - c]\right]}{[1+\lambda\beta]^2} \leq 0,\ \ \frac{\partial \bar{v}_{NS}}{\partial \lambda} = -\frac{\beta[v_H - v_L] + [v_L - c]}{[1+\lambda\beta]^2} \leq 0,$$

$$\frac{\partial \bar{v}_S}{\partial r} = \frac{[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2 - \lambda\sqrt{2r[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2}}{[1-\lambda][1-\beta]\sqrt{2r[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2}} \geq 0,$$

$$\frac{\partial \bar{v}_S}{\partial \beta} = \frac{\left[[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r - \lambda\sqrt{2r[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2}\right] r}{[1-\lambda][1-\beta]^2\sqrt{2r[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2}} \geq 0, \text{ and}$$

$$\frac{\partial \bar{v}_S}{\partial \lambda} = \frac{\left[[1-\lambda][1-\beta][v_H - v_L] + \lambda r - \lambda\sqrt{2r[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2}\right] r}{[1-\lambda]^2[1-\beta]\sqrt{2r[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2}} \geq 0. \ \square$$

A.3. Proof of Theorem 2

Using α^* in a high privacy-cost equilibrium given in Section 3.2.3, we show:

$$\frac{\partial \alpha^*}{\partial \beta} = \frac{[1-\lambda]\left[[\bar{v}_{NS} - v_L] + [1+\beta]\frac{\partial \bar{v}_{NS}}{\partial \beta}\right]}{4t[v_H - v_L]} \geq 0,\ \ \frac{\partial \alpha^*}{\partial r} = 0, \text{ and}$$

$$\frac{\partial \alpha^*}{\partial \lambda} = -\frac{[1+\beta]\left[[\bar{v}_{NS} - v_L] + [1-\lambda]\frac{\partial \bar{v}_{NS}}{\partial \beta}\right]}{4t[v_H - v_L]} \leq 0.$$

Using α^* in a low privacy-cost equilibrium given in Section 3.2.3, we show:

$$\frac{\partial \alpha^*}{\partial \beta} = \frac{\left[[v_H - v_L] - \lambda\left[[\bar{v}_S - v_L] + [1+\beta]\frac{\partial \bar{v}_S}{\partial \beta}\right]\right]c}{4t[v_H - v_L]} \geq 0, \quad \frac{\partial \alpha^*}{\partial r} = -\frac{\lambda[1+\beta]\frac{\partial \bar{v}_S}{\partial r}c}{4t[v_H - v_L]} \leq 0, \text{ and}$$

$$\frac{\partial \alpha^*}{\partial \lambda} = -\frac{[1+\beta]\left[[\bar{v}_S - v_L] + \lambda\frac{\partial \bar{v}_S}{\partial \lambda}\right]c}{4t[v_H - v_L]} \leq 0. \quad \square$$

A.4. Proof of Theorem 3

(i) If $\lambda = 0$, all consumers participate in profiling. Hence, we have:

$$\pi^\nu = \int_v f(v)\left[\beta[v - [1-\alpha]c] + \int_{\hat{v}<v}[1-\beta]f(v)[\hat{v} - [1-\alpha]c]d\hat{v}\right]dv - t\alpha^2 \tag{23}$$

$$TU^\nu = \int_v \int_{\hat{v}<v}[1-\beta]f(v)[v-\hat{v}]d\hat{v}dv \tag{24}$$

$$SW^\nu = \int_v f(v)\left[\beta[v - [1-\alpha]c] + \int_{\hat{v}<v}[1-\beta]f(v)[v - [1-\alpha]c]d\hat{v}\right]dv - t\alpha^2 \tag{25}$$

Using (23) - (25), we show:

$$\frac{\partial \pi^\nu}{\partial \beta} = \frac{[8[v_H - c] + 4[v_L - c]]\,t + 3[1+\beta]c^2}{24t} \geq 0, \quad \frac{\partial TU^\nu}{\partial \beta} = -\frac{v_H - v_L}{6} \leq 0, \text{ and}$$

$$\frac{\partial SW^\nu}{\partial \beta} = \frac{[4[v_H - c] + 8[v_L - c]]\,t + 3[1+\beta]c^2}{24t} \geq 0.$$

(ii) If $\lambda = 1$, no consumer participates in profiling, and hence, no profiling and voluntary profiling are identical.

(iii) If $\beta = 1$, those non-privacy sensitive consumers whose valuations are less than or equal to $p_0^\nu + c$ for a given p_0^ν participate in profiling, and the expected surplus of participating consumers is zero. Hence, we have:

$$\pi^\nu = [1-\lambda]\int_{v \leq p_0^\nu + c} f(v)[v - [1-\alpha]c]dv + \int_{v > p_0^\nu + c} f(v)p_0^\nu dv - t\alpha^2 \tag{26}$$

$$TU^\nu = \int_{v > p_0^\nu + c} f(v)[v - p_0^\nu - c]dv \tag{27}$$

Using (26) and (27), we show:

$$\frac{\partial \pi^{\nu}}{\partial \lambda} = -\frac{[p_0^{\nu} + c - v_L][p_0^{\nu} - [1-2\alpha]c + v_L]}{2[v_H - v_L]} \leq 0,$$

$$\frac{\partial TU^{\nu}}{\partial \lambda} = -\frac{\partial p_0^{\nu}}{\partial \lambda} \cdot \frac{v_H - p_0^{\nu} - c}{v_H - v_L} \geq 0,$$

and, we show:

$$SW^{\nu}(\lambda = 0) - SW^{\nu}(\lambda = 1) = \int_{v \leq p^b + c} f(v)(v - (1-\alpha)c)dv + \frac{(v_L - c)c^2}{4t(v_H - v_L)} \geq 0. \ \square$$

A.5. Proof of Theorem 4

(i) In a high-privacy-cost equilibrium, comparing the price under no profiling (given in (1)) and the price for a non-participating consumer under voluntary profiling (given in (9)), we show:

$$p_0^{\nu*} - p^{b*} = \frac{[1-\lambda][v_H - \bar{v}_{NS}]}{2\lambda} \geq 0$$

In a low-privacy-cost equilibrium, $p_0^{\nu*} > v_H - c$. Hence, there is no price at which a non-participating consumer purchases.

(ii) Comparing the price under no profiling (given in (1)) and the expected average price paid by a participating consumer under voluntary profiling (given in (17)), we show:

$$E_v(E_{\hat{v}}(I(\hat{v}, v)\, p_1^{\nu}(\hat{v}|v))) - p^{b*} \begin{cases} > 0 & \text{if } c > c^* \\ \leq 0 & \text{otherwise} \end{cases}$$

where, $c^* = \frac{3v_H - [1+2\beta]\bar{v}_{NS} - [2+\beta]v_L}{3[\alpha[1+\beta]-\beta]}$ for a high-privacy-cost equilibrium, and $c^* = \frac{3v_H - [2+\beta]v_L}{3[\alpha[1+\beta]-\beta]} - \frac{[1+2\beta][v_H - \lambda\bar{v}_S]}{3[\alpha[1+\beta]-\beta][1-\lambda]} + \frac{\lambda[v_H - \bar{v}_S]^2[[1+2\beta][v_H - \lambda\bar{v}_S] - [1-\lambda][[1-\beta]\bar{v}_S + 3\beta v_L]]}{3[\alpha[1+\beta]-\beta][1-\lambda][[v_H - v_L] - \lambda[\bar{v}_S - v_L]]^2}$ for a low-privacy-cost equilibrium. $\square$

A.6. Proof of Theorem 5

(i) From $p_0^{\nu*} \geq p^{b*}$ (Theorem 4(i)), we show: $U_0^{\nu}(v) \leq U^b(v)$ for all v.

(ii) Further: $U_0^{\nu}(v) \parallel U^b(v)$ when $U_0^{\nu}(v) \neq 0$ and $U^b(v) \neq 0$. Hence, from $U_1^{\nu}(\bar{v}_{NS}, NS) = U_0^{\nu}(\bar{v}_{NS}, NS) \leq U^b(\bar{v}_{NS}, NS)$ and $U_1^{\nu}(v, NS) > U_0^{\nu}(v, NS)$ for $v \leq \bar{v}_{NS}$, we show:

there is a $\tilde{v}_{NS} \leq \bar{v}_{NS}$ such that $U_1^{\nu}(\tilde{v}_{NS}, NS) = U^b(\tilde{v}_{NS}, NS)$.

and, from $U_1^{\nu}(v, S) > U_1^{\nu}(v, S)$; we show:

there is a $\tilde{v}_S \le \tilde{v}_{NS}$ such that $U_1^\nu(\tilde{v}_S, S) = U^b(\tilde{v}_S, S)$. $\square$

A.7. Illustration of aggregate consumer surplus under voluntary profiling versus no profiling

In a high-privacy-cost equilibrium (i.e., $r \ge 3.5$), for $v_H = 100$, $v_L = 60$, $c = 80$, $t = 20$, and $\lambda = 0.5$, we have:

$$TU^\nu = 1.57 > TU^b = 1.25 \text{ when } \beta = 0.7$$

whereas,

$$TU^\nu = 0.88 < TU^b = 1.25 \text{ when } \beta = 0.9.$$

Similarly, in a low-privacy-cost equilibrium (i.e., $r = 0.001$), for $v_H = 100$, $v_L = 80$, $c = 92.5$, $t = 50$, and $\lambda = 0.5$, we have:

$$TU^\nu = 0.48 > TU^b = 0.35 \text{ when } \beta = 0.85$$

whereas,

$$TU^\nu = 0.16 < TU^b = 0.35 \text{ when } \beta = 0.95. \ \square$$

A.8. Illustration of social welfare under voluntary profiling versus no profiling

In a high-privacy-cost equilibrium ($r \ge 0.57$), for $v_H = 100$, $v_L = 68$, $c = 50$, $t = 20$, and $\lambda = 0.5$, we have:

$$SW^\nu = 29.43 > SW = 29.30 \text{ when } \beta = 0.9$$

whereas,

$$SW^\nu = 29.20 < SW^b = 29.30 \text{ when } \beta = 0.8.$$

Similarly, in a low-privacy-cost equilibrium (i.e., $r = 0.001$), for $v_H = 100$, $v_L = 80$, $c = 48$, $t = 50$, and $\lambda = 0.5$, we have:

$$SW^\nu = 50.74 > SW^b = 50.70 \text{ when } \beta = 0.9$$

whereas,

$$SW^\nu = 49.54 < SW^b = 50.70 \text{ when } \beta = 0.8. \ \square$$

A.9. Proof of Theorem 6

(i) The seller can always choose to ignore the profile information, and hence, $\pi^\nu \ge \pi^b$.

(ii) - (iii) The seller charges an uniform price, $p^{\nu *}$, for all consumers under voluntary profiling. Hence, from:

$$U_1^\nu = v - p^{\nu *} - [1-\alpha]c > U_0^\nu = v - p^{\nu *} - c,$$

privacy-non-sensitive consumers always participate in profiling, and from

$$U_1^\nu = v - p^{\nu *} - [1-\alpha]c - r > U_0^\nu = v - p^{\nu *} - c,$$

privacy-sensitive consumers participate in profiling if $r < \alpha c$. Therefore, the seller's expected profit is given by:

$$\pi^V = \begin{cases} \dfrac{p^\nu [v_H - p^\nu - [1-[1-\lambda]\alpha]c]}{v_H - v_L} - t\alpha^2 & \text{if } r \geq \alpha c \\ \dfrac{p^\nu [v_H - p^\nu - [1-\alpha]c - \lambda r]}{v_H - v_L} - t\alpha^2 & \text{if } r < \alpha c \end{cases}$$

and, maximizing it yields:

$$p^{\nu *} = \begin{cases} \dfrac{1}{2}\left[v_H - [1-[1-\lambda]\alpha]c\right] & \text{if } r \geq \alpha c \\ \dfrac{1}{2}\left[v_H - [1-\alpha]c - \lambda r\right] & \text{if } r < \alpha c \end{cases}$$

and, we have:

$$\alpha^* = \begin{cases} \dfrac{[1-\lambda][v_H - c]c}{4t[v_H - v_L] - [1-\lambda]^2 c^2} & \text{if } r \geq \dfrac{[1-\lambda][v_H - c]c^2}{4t[v_H - v_L] - [1-\lambda]^2 c^2} \\ \dfrac{[v_H - c - \lambda r]c}{4t[v_H - v_L] - c^2} & \text{if } r < \dfrac{[1-\lambda][v_H - c]c^2}{4t[v_H - v_L] - [1-\lambda]^2 c^2} \end{cases}$$

Therefore, we show:

$$U^\nu(v, NS) - U^b(v, NS) = \begin{cases} \dfrac{1}{2}[1+\lambda]\alpha c \geq 0 & \text{if } r \geq \dfrac{[1-\lambda][v_H - c]c^2}{4t[v_H - v_L] - [1-\lambda]^2 c^2} \\ \dfrac{1}{2}[\alpha c + \lambda r] \geq 0 & \text{if } r < \dfrac{[1-\lambda][v_H - c]c^2}{4t[v_H - v_L] - [1-\lambda]^2 c^2} \end{cases}$$

$$U^\nu(v, S) - U^b(v, S) = \begin{cases} -\dfrac{1}{2}[1-\lambda]\alpha c \leq 0 & \text{if } r \geq \dfrac{[1-\lambda][v_H - c]c^2}{4t[v_H - v_L] - [1-\lambda]^2 c^2} \\ \dfrac{1}{2}[\alpha c - [2-\lambda]r] \begin{cases} \leq 0 & \text{if } \frac{\alpha c}{2-\lambda} \leq r < \frac{[1-\lambda][v_H - c]c^2}{4t[v_H - v_L] - [1-\lambda]^2 c^2} \\ > 0 & \text{if } r < \frac{\alpha c}{2-\lambda} \end{cases} \end{cases}$$

Aggregate consumer surplus is given by:

$$TU^\nu = \begin{cases} \dfrac{\lambda[v_H - p^\nu - c]^2 - [1-\lambda][v_H - p^\nu - [1-\alpha]c]^2}{2[v_H - v_L]} & \text{if } r \geq \alpha c \\ \dfrac{[v_H - p^\nu - (1-\alpha)c)^2 - \lambda r^2}{2[v_H - v_L]} & \text{if } r < \alpha c \end{cases}$$

Therefore, we show:

$$TU^{\nu}-TU^{b}=\begin{cases}\dfrac{[1-\lambda][2[v_H-c]+[3\lambda+c]\alpha]\alpha c}{8[v_H-v_L]}\geq 0 & \text{if } r\geq \alpha c\\ \dfrac{[\alpha c+\lambda r]^2+2[v_H-c][\alpha c+\lambda r]-4\lambda r^2}{8[v_H-v_L]}\geq 0 & \text{if } r<\alpha c\end{cases}$$

(iv) From (i) and (ii), we have $SW^{\nu}\geq SW^{b}$. □

A.10. Proof of Corollary 1

It is shown in A.10. Proof of Theorem 6(ii). □

A.11. Proof of Theorem 7

Given that a consumer can purchase a product elsewhere at p^o, the seller's price under no profiling (given as (1) in the monopoly model) is now written as:

$$p^b=\min\left\{\frac{v_H-c}{2},p^o\right\},$$

and, the seller's price for a participating and a non-participating consumer (given as (8) and (9) in the monopoly model) under voluntary profiling is written as:

$$p_1^{\nu}=\begin{cases}\hat{v}-[1-\alpha]c & \text{if } v_L\leq\hat{v}\leq p^o+[1-\alpha]c,\\ p^o & \text{if } p^o+[1-\alpha]c<\hat{v}\leq\bar{v}_{NS},\end{cases}$$

$$p_0^{\nu}=\min\left\{\frac{1}{2}\left[\frac{v_H-[1-\lambda]\bar{v}_{NS}}{\lambda}-c\right],p^o\right\}.$$

We identify three cases depending on p^o as the following:

Case 1: $p^o\geq\bar{v}_{NS}-[1-\alpha]c$

The seller's and consumers' decisions are not affected by p^o. Hence, all results are identical to those for a monopoly seller.

Case 2: $\frac{1}{2}\left[\frac{v_H-[1-\lambda]\bar{v}_{NS}}{\lambda}-c\right]<p^o<\bar{v}_{NS}-[1-\alpha]c$

Consider a privacy non-sensitive consumer whose valuation is v. If she does not participate in profiling, her expected surplus is:

$$U_0^{\nu}(v,NS)=\begin{cases}0 & \text{if } v<\frac{1}{2}\left[\frac{v_H-[1-\lambda]\bar{v}_{NS}}{\lambda}+c\right],\\ v-\frac{1}{2}\left[\frac{v_H-[1-\lambda]\bar{v}_{NS}}{\lambda}+c\right] & \text{if } v\geq\frac{1}{2}\left[\frac{v_H-[1-\lambda]\bar{v}_{NS}}{\lambda}+c\right],\end{cases}$$

and if she participates in profiling, her expected surplus is:

$$U_1^\nu(v, NS) = \begin{cases} q(\hat{v} = v|v) \cdot 0 + \displaystyle\int_{v_L}^{v} q(\hat{v} \neq v|v)[v - \hat{v}]d\hat{v} & if\ v_L \leq v \leq p^o + [1-\alpha]c, \\ q(\hat{v} = v|v)[v - p^o - [1-\alpha]c + \displaystyle\int_{v_L}^{p^o+[1-\alpha]c} q(\hat{v} \neq v|v)[v - \hat{v}]d\hat{v} & \\ \quad + \displaystyle\int_{p^o+[1-\alpha]c}^{\bar{v}_{NS}} q(\hat{v} \neq v|v)[v - p^o - [1-\alpha]c]d\hat{v} & if\ p^o + [1-\alpha]c < v \leq \bar{v}_{NS}. \end{cases}$$

By solving $U_1^\nu(v = \bar{v}_{NS}, NS) = U_0^\nu(v = \bar{v}_{NS}, NS)$, we have:

$$\bar{v}_{NS} = \frac{\psi - \sqrt{\psi^2 + 4[1-\lambda]\left[[1-\beta]\lambda[p^o + [1-\alpha]c - v_L]^2 - [v_H - [2p^o + [1-2\alpha]c]\lambda]\, v_L\right]}}{2[1-\lambda]}$$

where $\psi = v_H + [1-\lambda]v_L - [2p^o + [1-2\alpha]c]\lambda$, and we have:

$$U_1^\nu(v, NS) - U_0^\nu(v, NS) \begin{cases} \geq 0 & \text{if } v_L \leq v \leq \bar{v}_{NS}, \\ < 0 & \text{if } \bar{v}_{NS} < v \leq v_H. \end{cases}$$

The price for a non-participating consumer is not affected by p^o. Further, given that $\bar{v}_{NS} \leq v_H$ exist, we have: all non-participating consumers as well as some participating consumers are worse off, and aggregate consumer surplus and social welfare under voluntary profiling can be higher or lower compared to no profiling. Proofs are analogous to those for a monopoly seller.

Case 3: $p^o < \frac{1}{2}\left[\frac{v_H - [1-\lambda]\bar{v}_{NS}}{\lambda} - c\right]$

Since $p_1^\nu \leq p_0^\nu$ regardless of the signal, we have: $U_1^\nu(v, NS) \geq U_0^\nu(v, NS)$ for all v. Therefore, all privacy-non-sensitive consumers participate in profiling (i.e., $\bar{v}_{NS} = v_H$). Further, since we have: $p^b = p_0^\nu \geq p_1^\nu(\hat{v})$ for all $\hat{v}$, no non-participating consumer is worse off and all participating consumers are better off under voluntary profiling compared to no profiling.

From Case 1 - 3, we show: if p^o is sufficiently low (Case 3), then under voluntary profiling, no non-participating consumer is worse off, all participating consumers are better off, and social welfare is higher compared to no profiling. However, if p^0 is not sufficiently low (Case 1 and 2), then all our results that we reported in Section 3 and 4 hold. □

10 Appendix B

(a) Let $X_1 = v - E_{\hat{v}|v}(p_1^\nu(\hat{v})|v) - [1-\alpha]c$ for $v_L \leq v \leq p_0^\nu + c$ and $X_2 = p_0^\nu - E_{\hat{v}|v}(p_1^\nu(\hat{v})|v) + \alpha c$ for $p_0^\nu + c < v \leq v_H$. From (30) and (31), if $r > \max\{X_1, X_2\}$, then we have: $U_1^\nu(v, S) < U_0^\nu(v, S)$ for all v.

For a participating consumer, since there is always a positive probability that the signal is equal to or lower than her true valuation, we have: $U_1^\nu(v, NS) \geq 0$. On the other hand, for a given p_0^ν, we have: $U_0^\nu(v, NS) = 0$ for $v_L \leq v \leq p_0^\nu + c$. Therefore, we have: $U_1^\nu(v, NS) \geq U_0^\nu(v, NS)$ for $v_L \leq v \leq p_0^\nu + c$. Further, from Lemma 1, we have: $U_1^\nu(v, NS) \geq U_0^\nu(v, NS)$ for $p_0^\nu + c < v \leq \bar{v}_{NS}$; and $U_1^\nu(v, NS) < U_0^\nu(v, NS)$ for $\bar{v}_{NS} < v \leq v_H$, where $U_1^\nu(v = \bar{v}_{NS}, NS) = U_0^\nu(v = \bar{v}_{NS}, NS)$.

(b) If $r \leq \left[v - E_{\hat{v}|v}(p_1^\nu(\hat{v})|v) - [1-\alpha]c\right]$, then we have $\bar{v}_S$ such that $U_1^\nu(v = \bar{v}_S, S) = 0$, and from (30) and (31), for a given p_0^ν, we have: $U_1^\nu(v, S) \geq U_0^\nu(v, S)$ for $\bar{v}_S \geq v \geq p_0^\nu + c$. Further, from Lemma 1, we have: $U_1^\nu(v, S) \geq U_0^\nu(v, S)$ for $p_0^\nu + c < v \leq Z$; and $U_1^\nu(v, S) < U_0^\nu(v, S)$ for $Z < v \leq v_H$, where $U_1^\nu(v = Z, S) = U_0^\nu(v = Z, S)$.

Analogous to what we showed in (a) we have: $U_1^\nu(v, NS) \geq U_0^\nu(v, NS)$ for $v_L < v \leq \bar{v}_{NS}$; and $U_1^\nu(v, NS) < U_0^\nu(v, NS)$ for $\bar{v}_{NS} < v \leq v_H$. Therefore, for a non-participating consumer, the seller maximizes the expected profit given by: $\pi_0^\nu(p_0^\nu) = p_0^\nu \left[\lambda \int_Z^{v_H} f(v)dv + [1-\lambda] \int_{\bar{v}_{NS}}^{v_H} f(v)dv\right]$, and it yields $p_0^\nu \geq v_H - c$, and thereby, $Z = \bar{v}_{NS} = v_H$. $\square$

11 Appendix C

The seller's expected profit from participating consumers in a high-privacy-cost equilibrium can be written as:

$$\pi_1^\nu(p_1^\nu) = \begin{cases} p_1^\nu \left[\beta + \dfrac{[1-\beta][\bar{v}_{NS} - p_1^\nu - [1-\alpha]c]}{\bar{v}_{NS} - v_L}\right] & \text{if } p_1^\nu \leq \hat{v} - [1-\alpha]c \\ \dfrac{p_1^\nu[1-\beta][\bar{v}_{NS} - p_1^\nu - [1-\alpha]c]}{\bar{v}_{NS} - v_L} & \text{if } p_1^\nu > \hat{v} - [1-\alpha]c \end{cases}$$

For the seller to always use the signal, it is sufficient if the following are satisfied:

$$\frac{\partial \pi^{\nu}(p_1^{\nu})}{\partial p_1^{\nu}} = \begin{cases} \beta + \dfrac{[1-\beta][\bar{v}_{NS} - 2p_1^{\nu} - [1-\alpha]c]}{\bar{v}_{NS} - v_L} > 0 \\ \dfrac{[1-\beta][\bar{v}_{NS} - 2p_1^{\nu} - [1-\alpha]c]}{\bar{v}_{NS} - v_L} < 0 \end{cases}$$

Therefore:

$$\beta > \frac{\bar{v}_{NS} - [1-\alpha]c}{2\bar{v}_{NS} - v_L - [1-\alpha]c}$$

and

$$v_L > \frac{1}{2}[\bar{v}_{NS} + [1-\alpha]c] \text{ which holds if } v_L > \frac{1}{2}[v_H + [1-\alpha]c]$$

The seller's expected profit from participating consumers in a low-privacy-cost equilibrium can be written as:

If $\hat{v} \in [v_L, \bar{v}_S]$,

$$\pi_1^{\nu} = \begin{cases} p_1^{\nu}\left[\beta + \dfrac{[1-\beta][[v_H - \lambda\bar{v}_S] - [1-\lambda][p_1^{\nu} + [1-\alpha]c]]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]}\right] & \text{if } p_1^{\nu} \leq \hat{v} - [1-\alpha]c \\ \dfrac{p_1^{\nu}[1-\beta][[v_H - \lambda\bar{v}_S] - [1-\lambda][p_1^{\nu} + [1-\alpha]c]]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} & \text{if } \hat{v} - [1-\alpha]c < p_1^{\nu} \leq \bar{v}_S - [1-\alpha]c \\ \dfrac{p_1^{\nu}[1-\beta][v_H - p_1^{\nu} - [1-\alpha]c]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} & \text{if } p_1^{\nu} > \bar{v}_S - [1-\alpha]c \end{cases}$$

and, if $\hat{v} \in (\bar{v}_S, v_H]$,

$$\pi_1^{\nu}(p_1^{\nu}) = \begin{cases} p_1^{\nu}\left[\beta + \dfrac{[1-\beta][[v_H - \lambda\bar{v}_S] - [1-\lambda][p_1^{\nu} + [1-\alpha]c]]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]}\right] & \text{if } p_1^{\nu} \leq \bar{v}_S - [1-\alpha]c \\ p_1^{\nu}\left[\beta + \dfrac{[1-\beta][v_H - p_1^{\nu} - [1-\alpha]c]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]}\right] & \text{if } \bar{v}_S - [1-\alpha]c < p_1^{\nu} \leq \hat{v} - [1-\alpha]c \\ \dfrac{p_1^{\nu}[1-\beta][v_H - p_1^{\nu} - [1-\alpha]c]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} & \text{if } p_1^{\nu} > \hat{v} - [1-\alpha]c \end{cases}$$

For the seller to always use the signal, it is sufficient if the following are satisfied:

$$\frac{\partial \pi^{\nu}(p_1^{\nu})}{\partial p_1^{\nu}} = \begin{cases} \beta + \dfrac{[1-\beta][[v_H - \lambda\bar{v}_S] - [1-\lambda][2p_1^{\nu} + [1-\alpha]c]]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} > 0 \\ \dfrac{[1-\beta][[v_H - \lambda\bar{v}_S] - [1-\lambda][2p_1^{\nu} + [1-\alpha]c]]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} < 0 \\ \dfrac{[1-\beta][v_H - 2p_1^{\nu} - [1-\alpha]c]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} < 0 \\ \beta + \dfrac{[1-\beta][v_H - 2p_1^{\nu} - [1-\alpha]c]}{[v_H - v_L] - \lambda[\bar{v}_S - v_L]} > 0 \end{cases}$$

Therefore:

$$\beta > \frac{v_H - [1-\alpha]c}{2v_H - [1-\lambda]v_L - \lambda\bar{v}_S - [1-\alpha]c}$$

and

$$v_L > \frac{1}{2}\left[\frac{v_H - \lambda\bar{v}_S}{1-\lambda} + [1-\alpha]c\right] \text{ which holds if } v_L > \frac{1}{2}[v_H + [1-\alpha]c]$$

Under the above sufficient condition, $p_1^{\nu *} = \hat{v} - [1-\alpha]c$. $\square$