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

June 15, 2015

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, 2014 Theory in Economics of Information Systems (TEIS) 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. We also thank Jeanette Burman for editing assistance.

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 and charge only a uniform price for everyone or the seller can only charge different prices based on the consumers participation status, 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. However, when personalized pricing is possible, reducing privacy cost alone may reduce aggregate consumer surplus. The primary reason for these 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: Voluntary Profiling, Privacy, Price Discrimination, Social Welfare.

1 Introduction

"*There's going to come a point where our shelf pricing is pretty irrelevant because we can be so personalized in what we offer people*" – Steven Burd, Safeway CEO (Tuttle, 2013).

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

An example of profiling both online and offline is the Safeway Rewards card. Those customers that sign up either online or through forms at the store give up private identity information and optionally other information. In return for discounted prices on items in the store, they also give up their detailed purchase histories by swiping their Safeway Rewards card during the checkout process or entering it during their online grocery order. In addition, customized coupons are provided through web applications, and occasionally through the mail. Finally, Safeway has developed a mobile application that can guide customers through the store, potentially simplifying the shopping process but also tracking shopping patterns and selling through push marketing. Those customers that do not sign up for the Safeway Rewards card and are not profiled pay the higher posted prices across the board.

Surveys show, however, that not all consumers are comfortable with sellers collecting their profile information. These studies attribute the negative consumer attitude towards profiling to concerns about privacy and price discrimination (Hoffman et al., 1999; Volokh, 2000; Earp and Baumer, 2003; Acquisti and Grossklags, 2005). Privacy concerns relate to consumers' loss of control over their information – the individual's ability to choose the extent to which their personal information is used, protected, and shared with others (Chellappa and Shivendu, 2007-8). Even with policies that govern the use and sharing of the information by sellers, 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, offering different prices to consumers based on their profile information (Streitfeld, 2000). Since then many articles and online blogs have cited cases where online sellers charge different prices for identical products based on consumers' location (Valentino-Devries et al., 2012) and web browser (Villarreal, 2010).[1] Furthermore, as we describe in our example above, grocers such as Safeway and Kroger offer individualized offers/prices based on data collected from shoppers' purchases (Clifford, 2012). The academic literature also provides anecdotal evidence about price discrimination practice in e-commerce (Odlyzko, 2003; Viswanathan et al., 2007; Schön, 2010; Li and Dinlersoz, 2012). A number of empirical studies suggest that online sellers 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 oppose it on the grounds of aforementioned privacy and price discrimination (Chester, 2009). As a compromise, policy makers in the U.S. have proposed a policy known as *Voluntary Profiling* that allows sellers to collect and use consumer information if consumers voluntarily participate in profiling (Federal Trade Commission, 2000, 2012). This is also known as the "Do Not Track List" policy modeled after the "Do Not Call List" policy implemented for telemarketers. Complying with this policy, most online sellers do not mandate that a consumer participates in profiling (e.g., sign up at their websites) for online shopping, yet if they do then sellers provide them with benefits such as personalized offers and express check out in exchange for personal information. One might expect that no consumer is worse off under voluntary profiling compared to no profiling because only those consumers that expect to benefit from profiling will participate. One might also expect that voluntary profiling increases aggregate consumer surplus and social welfare because voluntary profiling reduces search cost and other transaction costs, and may increase market size – the number of consumers that purchase. We show, however, that this intuition is not accurate. Volun-

[1] More examples are available at: https://www.privacyrights.org/online-shopping-tips-e-commerce-and-you#dynamic and https://www.privacyrights.org/what-personal-information-should-you-give-merchants.

tary profiling can have adverse implications for consumers and society unless privacy, and more importantly, price discrimination (personalized pricing) are addressed. The reason is that voluntary profiling allows the seller to exploit the heterogeneity in consumers' inherent incentives to participate in profiling.

In our context an online monopolist – the seller – sells a large number of products from a product category to consumers each of whom has an ideal product. Consumers have different valuations for their ideal products. In addition to the price charged by the seller, a consumer incurs other transaction costs. These costs include the effort related to searching for the ideal product and completing the transaction by providing shipping and payment details, among others. Consumers have the option to voluntarily allow the seller to collect their personal data and use that data – consumer profiles – to reduce these transaction costs, which we hereafter refer to as *search cost*. Participation in profiling can take various forms such as registering at the seller's website, participating in rewards or loyalty programs, and providing information through surveys. By using consumer profiles, the seller supports the purchase process. The support – hereafter referred to as *search support* – includes recommendations, custom search results, promotions and other alerts, and providing 1-click ordering, among others. The profile information also provides a signal about a consumer's valuation of the product. Consequently, the seller may charge different prices for an identical product based on consumer profiles. This personalized pricing can be implemented directly or indirectly through targeted coupons and discounts.

In this setup we find that consumers with low valuations – especially those that are not served when there is no profiling – and have low privacy cost have an incentive to participate in profiling and benefit from reduced search cost and possibly reduced price. On the other hand, consumers with high valuations do not have similar incentives to participate in profiling despite the reduced search cost because the seller may charge them a higher price. 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, under voluntary profiling the seller charges a weakly higher price to non-participating consumers. Due to this indirect effect, a *negative externality*, no non-participating consumers are better off and some participating consumers are worse off under voluntary profiling than 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.

Interestingly, consumers' online shopping behaviors reported in articles in the popular press are consistent with the above findings, which suggest that consumers are aware of sellers' use of consumer profiles to offer personalized pricing and consequently consumers act strategically. For instance, consumers with low valuations persistently leave items they want in the shopping cart and abandon it because sellers often give discounts to these consumers to induce them to purchase (Tuttle, 2012). On the other hand, in order to avoid paying premium price those consumers with higher valuations often hide their information by deleting cookies or use a different machine or web browser, despite the search support they could otherwise receive.[2]

We also examine the different roles played by price discrimination and privacy cost in the adverse impact of voluntary profiling on consumers. We show that if price discrimination is not possible and the seller can only charge a single price, or if only third-degree price discrimination is possible whereby the seller can charge only one price for participating consumers and another for non-participating consumers, then aggregate consumer surplus under voluntary profiling compared to no profiling is not lower and a reduction in privacy cost has a positive impact on all consumers as well as the seller. However, when price discrimination in the form of personalized pricing is possible, reducing privacy cost may reduce consumer surplus. Thus, the existence of some consumers with a positive privacy cost 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 the voluntary profiling policy in online markets.

[2]See: http://online.wsj.com/news/articles/SB10001424052702304458604577488822667325882 and http://www.theguardian.com/money/blog/2010/aug/07/computer-cookies-booking-online.

1.1 Relationship to Existing Literature

There is growing literature that addresses electronic privacy, namely the ability of sellers to track and collect information about individual consumers. The 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; Cachon et al., 2008; Aloysius et al., 2012). This literature shows that with passive consumers, consumer addressability and personalized pricing benefits a monopoly, but competing sellers may not choose full addressability because it may intensify competition. Choudhary et al. (2005) show that personalized pricing aggravates price competition and it may or may not benefit competing sellers. Research on dynamic targeted pricing in online markets using multi-period models shows how consumers acting strategically in the first period anticipating price in the second period can hurt the seller (Taylor, 2004; Acquisti and Varian, 2005; Calzolari and Pavan, 2006; Chen and Zhang, 2009).

Chellappa and Shivendu (2007-8) analyze the welfare implications of regulatory regimes that control if a seller can "buy" consumer information and can force consumers to accept such information collection. Chellappa and Shivendu (2010) examine vendor strategies regarding the extent of personalization and provision of coupons when faced with consumers with heterogeneous privacy concerns. In contrast to that work our models incorporate price discrimination and the externality that participation of one group of consumers may have on those that do not participate.

Closest to our work is research that examines contexts in which consumers decide the information they want to reveal (Aron et al., 2006; Hann et al., 2008; Conitzer et al., 2011). 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. To this we add 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 their 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' valuations.

Hann et al. (2008) 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 price discrimination is not allowed.

Conitzer et al. (2011) consider welfare implications of an opt-out model in a repeat purchase context. In their two-period model, consumers that purchase in the first period can choose to delete their purchase history at a cost at the end of the first period and the seller decides the pricing strategies by anticipating consumer actions. They do not consider privacy cost, and 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 purchase model in which the 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 through considering the tradeoff between their 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 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) of price discrimination, and the roles of privacy and price discrimination in how voluntary profiling affects consumer and social welfare.

2 Model Description

Overview An online monopolist – the seller – sells a large number of products from a product category. A consumer visiting the seller's website searches and finds their ideal product – the one that is closest to their preference – at the end of search. Under voluntary profiling a consumer can choose to participate in profiling. If the consumer participates in profiling, then the seller supports the transaction process using profile information. We do not assume any specific form of support but we assume that if the seller provides support then the consumer's search cost to find and purchase their ideal product is reduced by the amount of support that the seller provides. We use the term search support to refer to all support provided by the seller to decrease the consumers' transaction costs. If the consumer does not participate in profiling, then the consumer does not receive any support from the seller in the transaction process. In our extensions section we consider a positive level of generic search support to non-participating consumers. For our base model, we normalize such generic search support to zero.

Consumer Valuations We denote consumers by their valuation from consuming their ideal product: $v \in [v_L, v_H]$ where the probability distribution $f(v)$ is positive over its support. We take v as uniformly distributed as valuation can be scaled as needed, $f(v) = \frac{1}{v_H - v_L}$ for $v_L \leq v \leq v_H$ and 0 otherwise. We take the consumer preferences to be sufficiently 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; prices only differ due to price discrimination based on valuation for participating consumers and non-participating consumers are offered a uniform price.

Consumer Search Cost and Support When not participating in profiling a consumer incurs a constant search cost $c \in R^+$ to find her ideal product. A constant search cost is often assumed in the search literature (Bakos, 1997; Arbatskaya, 2007). When participating in profiling we use $\alpha \in [0, 1]$ to denote the *search support* from the seller. It reflects the proportional reduction in search cost a consumer receives. As α is often constrained by the technology, we assume it is exogenously given. We relax the assumption in our extensions section in which α is a function of the number of participating consumers on the basis that the profiler has a better predictive power when more consumers participate. The extension yields qualitatively similar results. The search support reduces the search cost for a participating consumer such that the effective search cost is $[1-\alpha]\,c$. In our setting the consumer always finds their ideal product at the end of the search.[3] Therefore, the purpose of search is only to identify the ideal product.

Consumer Privacy Costs Consumers that participate in profiling incur privacy costs. These costs include unwanted marketing solicitations (Hann et al., 2008), information sharing, privacy leakage, and security breaches (Milne and Rohm, 2000). However, privacy cost does not include potential loss (or gain) from price discrimination (Chellappa and Shivendu, 2007-8). Consumers are in one of two groups, 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 is uninformative regarding a consumer's privacy cost.

As noted in the previous three paragraphs, in our model consumers face search cost, privacy cost, and price discrimination (consumer valuation) aspects. Because privacy cost and price discrimination are key dimensions of the profiling debate, we develop a model that has heterogeneity on the privacy cost and valuation dimensions. However, as discussed in Section 6.3, a model that has heterogeneity on the search cost and privacy cost dimensions leads to qualitatively identical results.

The Seller The seller employs a profiler – a specialized business analytics tool – to develop consumer profiles. The profiles are used to provide search support and a signal $\hat{v}$ about a consumer's valuation. The signal reveals the consumer's true valuation with probability

[3] We can extend the model to a case where 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.

$\beta \in [0, 1]$ and no new information with probability $[1 - \beta]$. We refer to β as the *valuation accuracy* of the profiler.

Timing of the Game The sequence of the game is depicted in Figure 1. In stage 1, the seller sets the price for a non-participating consumer and each consumer decides whether to participate in profiling. In stage 2, the seller sets the price for a participating consumer based on the signal it receives from the profile. Finally, in stage 3, a consumer that expects a non-negative surplus searches and purchases after observing the offered price. We note that consumers make their participation decisions in stage 1 using expected prices, and their purchasing decisions in stage 3 after they observe the offered prices.

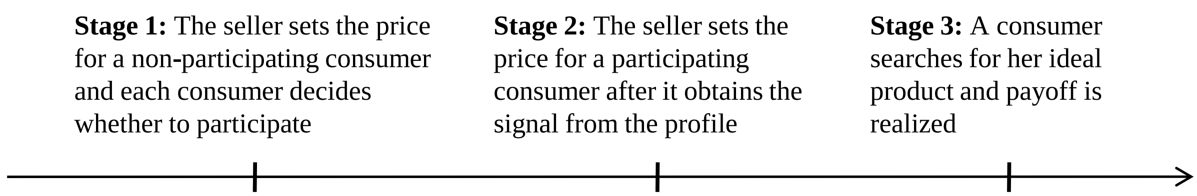


Figure 1: Timeline of the game

Equilibrium In our model, the seller does not know a consumer's (true) valuation and privacy cost, and a consumer does not know the signal the seller has about her valuation. Every other parameter is common knowledge. Thus, we have a sequential game with incomplete information and consequently we use the *Perfect Bayesian Equilibrium (PBE)* as our solution concept. A PBE in our model consists of the seller's strategy (the prices for participating and non-participating consumers), consumers' strategies (participation and purchase decisions), and a system of beliefs. These constitute a PBE if all strategies are sequentially rational given the beliefs and those beliefs are consistent with the strategies (Fudenberg and Tirole, 1991; Conitzer et al., 2011).

3 Model Analysis

3.1 Benchmark: No Profiling

We use no profiling as our benchmark and denote it using superscript b. The sequence of the game under no profiling is as follows: the seller sets a uniform price in stage 1 and a consumer that expects a non-negative surplus searches and purchases her ideal product in stage 2. We make the following technical assumption to ensure an interior solution for the seller's problem.

Assumption 1. $v_H > 2v_L - c$.

The seller maximizes profit given by

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

It yields the optimal price under no profiling

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

Please note that given Assumption 1 we have $p^{b*} + c > v_L$, which implies that not all consumers are served by the seller under no profiling.

3.2 Voluntary Profiling

We employ the following standard procedure to compute a PBE for the game under voluntary profiling. First, we propose a belief about consumers' participation structure. Second, based on this belief, using backward induction, we determine the consumers' purchase decisions in stage 3, compute the seller's optimal price for a participating consumer in stage 2, followed by the price for a non-participating consumer and consumers' participation strategies in stage 1. Lastly, we show that the seller's and consumers' optimal strategies are consistent with the proposed belief and all strategies are sequentially rational based on Bayesian belief updating. We use superscript ν to denote voluntary profiling, and subscripts 1 and 0, respectively, to denote a participating and a non-participating consumer.

3.2.1 The belief about consumers' participation structure

We propose two *ex ante* beliefs about consumers' participation structure as it is defined in Definition 1 and illustrate it using Figure 2. We denote the expected surplus of a participating consumer with valuation v in privacy group $i \in \{NS, S\}$ by $U_1^{\nu}(v, i)$ and surplus of a non-participating consumer by $U_0^{\nu}(v, i)$. The gray (white) areas in the figure indicate the valuations of (non-)participating consumers.

Definition 1.
(a) High-privacy-cost belief: The privacy cost r is greater than a threshold $\bar{r}$ such that no privacy-sensitive (S) consumer participates in profiling. Privacy-non-sensitive (NS) consumers whose valuations are not greater than $\bar{v}_{NS}$ participate in profiling, where $U_1^{\nu}(v = \bar{v}_{NS}, NS) = U_0^{\nu}(v = \bar{v}_{NS}, NS)$.
(b) Low-privacy-cost belief: The price cost r is less than or equal to $\bar{r}$ such that some privacy-sensitive (S) consumers – those whose valuations are between $\bar{v}_S$ and $\bar{v}'$ – participate in profiling, where $U_1^{\nu}(v = \bar{v}_S, S) = U_0^{\nu}(v = \bar{v}_S, S) = 0$ and $U_1^{\nu}(v = \bar{v}'_S, S) \geq U_0^{\nu}(v = \bar{v}'_S, S)$. Privacy-non-sensitive (NS) consumers whose valuations are not greater than $\bar{v}_{NS}$ participate in profiling, where $U_1^{\nu}(v = \bar{v}_{NS}, NS) \geq U_0^{\nu}(v = \bar{v}_{NS}, NS)$.

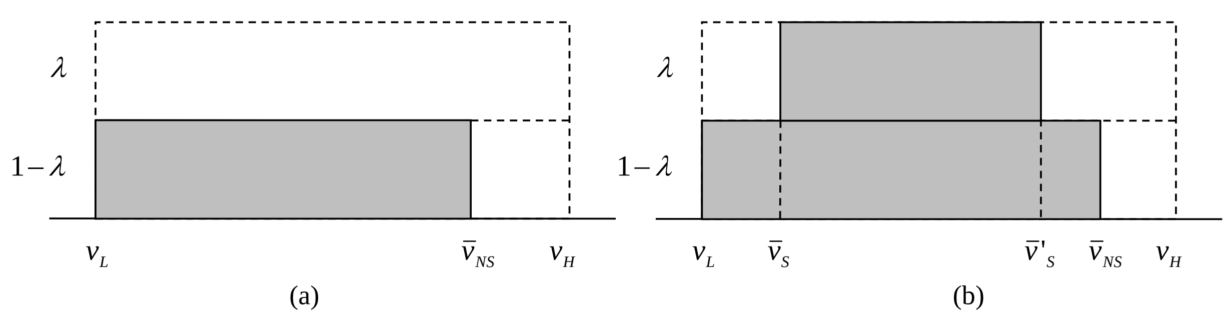


Figure 2: The illustration of (a) high- and (b) low-privacy-cost beliefs

The beliefs can be formally defined as the following posterior probability density functions for a participating consumer's valuation $g(v)$:

(a) High-privacy-cost belief

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

(b) Low-privacy-cost belief

$$g(v) = \begin{cases} \dfrac{1-\lambda}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} & \text{for } v_L \leq v \leq \bar{v}_S \text{ and } \bar{v}'_S \leq v \leq \bar{v}_{NS}, \\ \dfrac{1}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} & \text{for } \bar{v}_S < v \leq \bar{v}'_S. \end{cases}$$

Then the conditional probability distribution function of the signal given true valuation v is written as

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

and the posterior density function for a participating consumer's valuation conditional on the signal is computed as

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

We identify a PBE only for the case when voluntary profiling is decision-relevant and make the following assumption.

Assumption 2. *The valuation accuracy β is sufficiently high so that the seller does not ignore the signal when it sets the price for a participating consumer.*

The mathematical conditions for Assumption 2 are available in Appendix A.0.[4] If Assumption 2 does not hold then voluntary profiling does not provide useful information for pricing, which in turn implies the seller simply charges a uniform price. In Section 5.2 we examine how restricting price discrimination (personalized pricing) under voluntary profiling affects our results.

3.2.2 Stage 3: Consumers' purchase decisions

Let $p_1^\nu(\hat{v})$ and p_0^ν denote the price offered to a participating and a non-participating consumer respectively. In stage 3 of the game, a participating consumer with valuation v purchases

[4] As we noted in Appendix A.0 we assume in our analysis that the lower bound for β is an interior value between 0 and 1.

her ideal product if $v - p_1^\nu(\hat{v}) - [1-\alpha]c \geq 0$, noting that the privacy cost is sunk after a consumer makes her participation decision. A non-participating consumer with valuation v purchases her ideal product if $v - p_0^\nu - c \geq 0$.

3.2.3 Stage 2: The optimal price for a participating consumer

Based on the belief about consumers' participation structure the seller maximizes profit from a participating consumer conditional on the signal given by

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

Under Assumption 2 in both high- and low-privacy-cost beliefs it yields the optimal price for a participating consumer

$$p_1^{\nu*} = \hat{v} - [1-\alpha]c. \tag{2}$$

The detailed derivation of optimal solutions for problems in stages 1 and 2 is available in Appendix A.1.

3.2.4 Stage 1: The optimal price for a non-participating consumer and consumers' participation decisions

If a consumer with valuation v does not participate in profiling her surplus is

$$U_0^\nu(v,i) = \begin{cases} v - p_0^\nu - c & \text{if } v \geq p_0^\nu + c, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

for $i = NS$ and S, and if she participates in profiling her surplus is

$$U_1^\nu(v,i) = \begin{cases} q(\hat{v}=v|v) \cdot 0 + \int_{v_L}^{v} q(\hat{v} \neq v|v) \, [v - \hat{v}] \, d\hat{v} & \text{if } i = NS, \\ q(\hat{v}=v|v) \cdot 0 + \int_{v_L}^{v} q(\hat{v} \neq v|v) \, [v - \hat{v}] \, d\hat{v} - r & \text{if } i = S. \end{cases} \tag{4}$$

High-privacy-cost belief Under the high-privacy-cost belief the seller maximizes profit from a non-participating consumer given by

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

Further, from the belief we have $U_1^\nu(v = \bar{v}_{NS}, NS) = U_0^\nu(v = \bar{v}_{NS}, NS)$. Hence, it yields the optimal price for a non-participating consumer

$$p_0^{\nu *} = \frac{[1+\beta]v_H + [1-\lambda][1-\beta]v_L - [2-[1-\beta]\lambda]c}{2[1+\lambda\beta]}, \tag{5}$$

and the consumer indifferent between participating and not has the valuation

$$\bar{v}_{NS} = \frac{v_H - \lambda[1-\beta]v_L + \lambda c}{1+\lambda\beta}. \tag{6}$$

Low-privacy-cost belief Under the low-privacy-cost belief the seller maximizes profit given by

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

Further, from the belief we have $U_1^\nu(v = \bar{v}_{NS}, NS) \geq U_0^\nu(v = \bar{v}_{NS}, NS)$, $U_1^\nu(v = \bar{v}_S, S) = U_0^\nu(v = \bar{v}_S, S) = 0$, and $U_1^\nu(v = \bar{v}_S', S) \geq U_0^\nu(v = \bar{v}_S', S)$. Hence, it yields the optimal price for a non-participating consumer

$$p_0^{\nu *} > v_H - c, \tag{7}$$

and the consumer indifferent between participating and not has valuation

$$\bar{v}_S = \frac{[1-\lambda][1-\beta]v_L - \lambda r + \sqrt{2r[1-\lambda][1-\beta][v_H - v_L] + \lambda^2 r^2}}{[1-\lambda][1-\beta]}, \tag{8}$$

and

$$\bar{v}_S' = \bar{v}_{NS} = v_H. \tag{9}$$

3.2.5 The perfect Bayesian equilibrium

In this section, we first derive the condition for r under which each belief about the participation structure can be sustained and verify that consumers' participation strategies do not deviate from those implied by the belief.

Comparing (3) and (4), if $U_1^\nu(v, S) < 0$ for $v \leq p_0^\nu + c$ then we have $U_1^\nu(v, S) < U_0^\nu(v, S)$ for all v (i.e., no privacy-sensitive consumers participate). Hence, solving

$$\int_{v_L}^{v} q(\hat{v} \neq v | v) \, [v - \hat{v}] \, d\hat{v} - r < 0 \text{ for } v \leq p_0^{\nu *} + c,$$

after substituting $p_0^{\nu *}$ given in (5), we obtain

$$\bar{r} = \frac{[1-\lambda][1-\beta][1+\beta]^2[v_H - [1+\lambda]v_L + \lambda c]^2}{4[1+\lambda\beta]\,[[2-[1-\beta]\lambda]v_H - [1-\lambda][2+[1+\beta]\lambda]v_L - [1+\beta]\lambda^2 c]}. \tag{10}$$

If $r > \bar{r}$ then we have the high-privacy-cost equilibrium; otherwise, we have the low-privacy-cost equilibrium.

Now we verify consumers' participation strategies. If $r > \bar{r}$, comparing (3) and (4) after substituting (5) and (6), we verify that

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

and

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

If $r \leq \bar{r}$, comparing (3) and (4) after substituting (7) – (9), we verify that

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

and

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

We characterize the PBE we derive for the voluntary profiling game as the following Lemma and illustrate it using Figure 2. The gray (white) areas in the figure indicate the valuations of (non-)participating consumers.

Lemma 1. *The following constitutes the PBE under voluntary profiling:*
(a) If r is greater than (10), then we have the high-privacy-cost equilibrium in which (i) the seller's optimal price for a non-participating consumer is given in (5), (ii) no privacy-sensitive consumers participate in profiling, and (iii) privacy-non-sensitive consumers whose valuations are less than (6) participate in profiling.
(b) If r is less than or equal to (10), then we have the low-privacy-cost equilibrium in which (i) the seller's optimal price for a non-participating consumer is given in (7), (ii) privacy-sensitive consumers whose valuations are greater than (8) participate in profiling, and (iii) all privacy-non-sensitive consumers participate in profiling.
(c) In both equilibria the seller's optimal price for a participating consumer is given in (2).

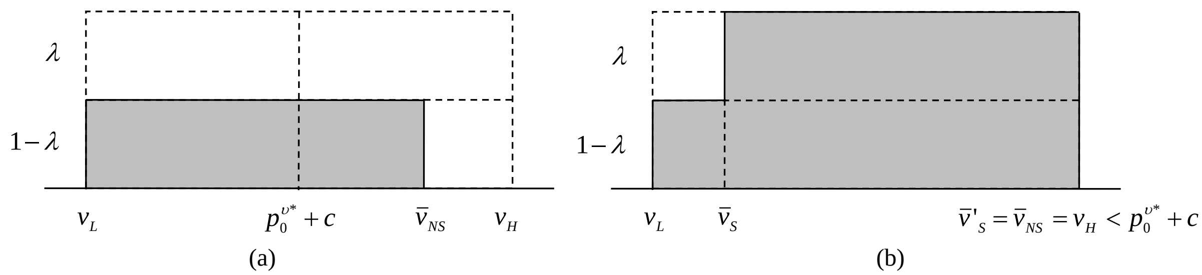


Figure 3: The illustration of (a) high- and (b) low-privacy-cost equilibria

Our next theorem shows the properties of the equilibrium. Proofs for the Theorems are provided in Appendix A.

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

Theorem 1 is the result 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 the seller receiving a low signal and thereby offering a low price. A participating consumer pays in two ways: privacy cost and a possibility of a seller receiving a high signal and thereby offering a high price. A decrease in privacy cost or the valuation accuracy increases all consumers' benefit from participation, 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 a non-participating consumer, and that reduces the proportion of consumers (even privacy-non-sensitive) that participate.

4 Voluntary Profiling Versus 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. It is straightforward that the seller's profit is higher under voluntary profiling than under no profiling because the seller can exploit additional information about a participating consumer in setting prices.

4.1 Price Paid by Consumers

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

$$\begin{aligned} & E_v(E_{\hat{v}}(I(\hat{v},v)\, p_1^{\nu}(\hat{v}|v))) \\ & \quad = \int_v g(v) \left[q(\hat{v}=v|v)[v-[1-\alpha]c] + \int_{\hat{v}<v} q(\hat{v}\neq v|v)[\hat{v}-[1-\alpha]c] d\hat{v} \right] dv, \end{aligned} \tag{11}$$

where $I(\hat{v},v)$ is an indicator function that is equal to 1 if $\hat{v} \leq v$ and 0 otherwise. In the high-privacy-cost equilibrium, the price for a non-participating consumer is given in (5). In the low-privacy-cost equilibrium, no non-participating consumer purchases (see (7)). Comparing the prices under voluntary profiling (given in (5) and (11)) and the price under no profiling (given in (1)), we have the following theorem.

Theorem 2. *Compared to no profiling, under voluntary profiling: (i) the expected price paid by a non-participating consumer is not lower and (ii) if the search cost is sufficiently high, then the expected price paid by a participating consumer is higher.*

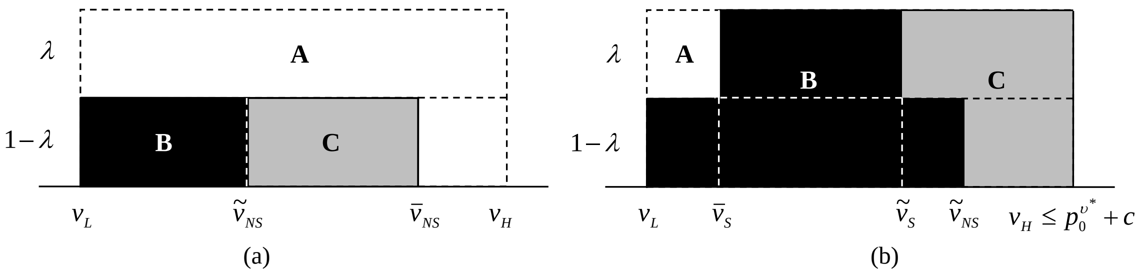
Because voluntary profiling skews the distribution of valuations of non-participating consumers to the right on the valuation dimension, on average, the valuation of non-participating consumers is higher than that of the consumer population. This induces the seller to charge a higher price for non-participating consumers under voluntary profiling than for the consumer population under no profiling. If the search cost is sufficiently high, then the expected price paid by a participating consumer under voluntary profiling is higher than the price paid by a consumer under no profiling. This is true even though participating consumers include low valuation consumers (those not served under no profiling), and 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. Moreover, comparing (5) and (11), 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., $c > \frac{[2+\beta]v_H+[[1-\beta][1-2\lambda]-3\beta[1+\lambda\beta]]v_L}{3-[2\lambda+6\beta]+[2\lambda-6\lambda\beta]\beta+3\alpha[1+\beta][1+\lambda\beta]}$).

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. Then, we have the following theorem about the surplus of individual consumers.

Theorem 3. *Compared to no profiling, under voluntary profiling (i) the surplus of no non-participating consumer is larger and (ii) the surplus of a participating consumer whose valuation is greater than* $\tilde{v}_i$ *is smaller.*

Figure 4 illustrates Theorem 3. Region A includes all non-participating consumers, region B includes participating consumers that enjoy a higher surplus under voluntary profiling than under no profiling (valuations less than $\tilde{v}_i$), and region C includes participating consumers that have a smaller surplus under voluntary profiling than under no profiling (valuations greater than $\tilde{v}_i$). Theorem 3 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 on the valuation dimension under voluntary profiling. Although the seller does not collect any information from non-participating consumers, as shown in Theorem 2(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: Consumer surplus under voluntary profiling versus no profiling for the (a) high- and (b) low-privacy-cost equilibria

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 (CS) 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

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

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

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

and for the low-privacy-cost-equilibrium is

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

We illustrate with a numerical example in Appendix A.5 that the aggregate consumer surplus may be higher or lower under voluntary profiling than no profiling. Consequently, the benefit offered by voluntary profiling to consumers in region B may 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: $v_H = 100$, $v_L = 70$, $\lambda = 0.6$, and $\alpha = 0.9$ for both panels and $r > 2$ for panel (a) and $r = 0.1$ for panel (b)). Consistent with intuition, Figure 5 suggests that voluntary profiling likely leads to a higher aggregate consumer surplus compared with no profiling when search cost savings αc are high (i.e., when the search cost c is high) and the valuation accuracy β is low.

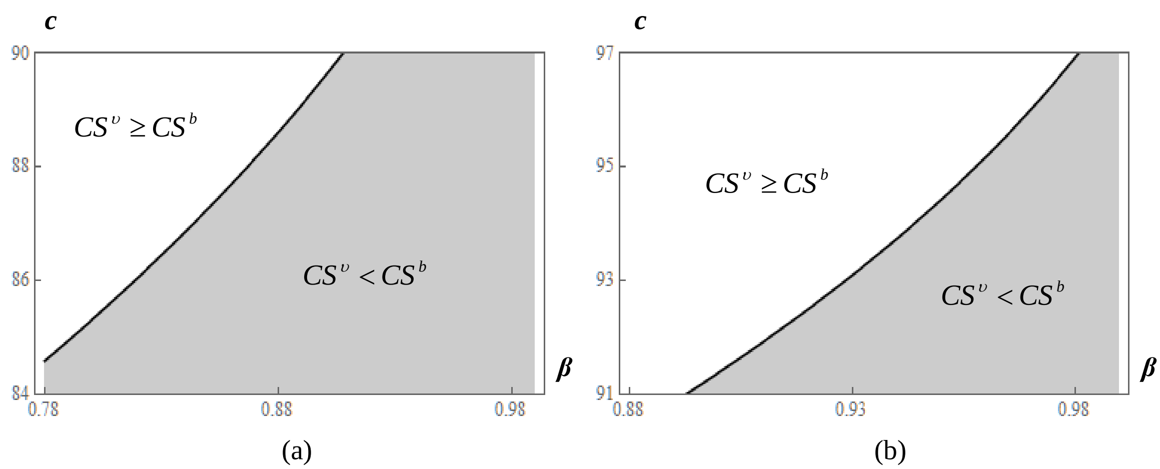


Figure 5: Comparison of aggregate consumer surplus under no profiling and voluntary profiling for the (a) high- and (b) low-privacy-cost equilibria[5]

4.3 Social Welfare

Analyzing the impact of voluntary profiling on social welfare is important because a policy-maker has the potential to regulate profiling policy. We define social welfare (SW) as *the seller's profit + aggregate consumer surplus*. Social welfare under no profiling is

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

[5]We note that given other parameter values Assumption 2 holds when $\beta > 0.77$ and $\beta > 0.87$ in the high- and low-privacy-cost equilibrium respectively. We discuss the case where Assumption 2 does not hold and thereby the seller ignores the signal in Section 5.2.

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

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

and for the low-privacy-cost-equilibrium is

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

Using above expressions for social welfare, we illustrate with a numerical example in Appendix A.6 that social welfare may be higher or lower under voluntary profiling compared with no profiling. This accounts for the fact that price is a transfer between consumers and the seller, only affecting social welfare through consumer choices. Voluntary profiling benefits society in two ways: it reduces search cost, a deadweight loss to 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, society is worse off under voluntary profiling compared to no profiling.

We present results from our numerical analysis in Figure 6 (parameter values: $v_H = 100$, $v_L = 70$, $\lambda = 0.1$, and $\alpha = 0.1$ for both panels, $r > 2$ for panel (a) and $r = 0.4$ for panel (b)). Figure 6 suggests that voluntary profiling likely leads to higher social welfare compared to no profiling when the search cost c and the valuation accuracy β are high. Voluntary profiling leads to a significant reduction in the deadweight loss due to wasteful search when the search cost is high.

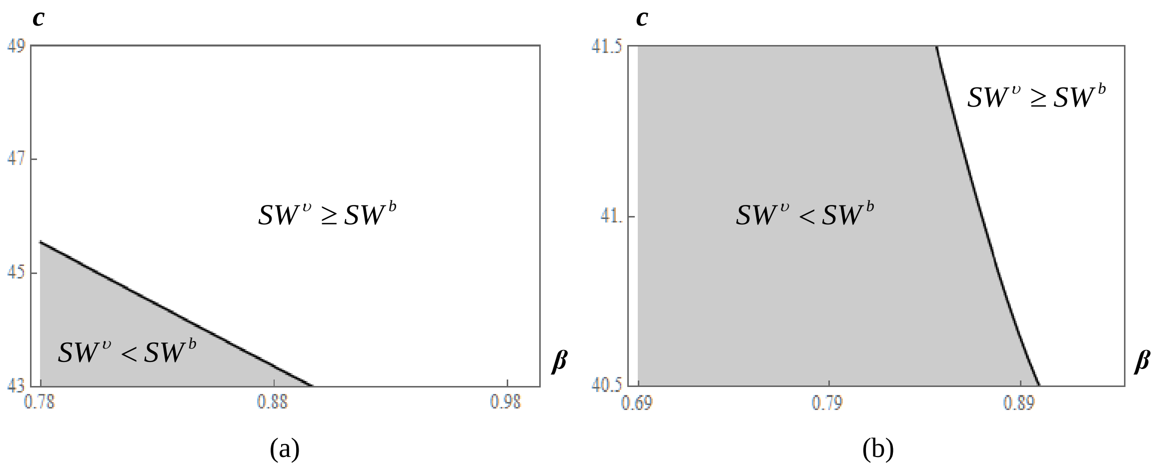


Figure 6: Comparison of social welfare under no profiling and voluntary profiling for the (a) high- and (b) low-privacy-cost equilibria[6]

5 Social Policy

Although voluntary profiling always benefits the seller, it does not always benefit consumers and society. Therefore, a policy-maker has to balance the interests of industry and consumers before choosing voluntary profiling. Here we show the different roles played by privacy sensitivity and price discrimination in the adverse impact of voluntary profiling on consumers and society, and derive policy implications.

5.1 The Impact of Privacy Sensitivity and Valuation Accuracy on the Seller, Consumers, and Society

Our next theorem shows how privacy sensitivity (as measured by the fraction of privacy-sensitive consumers and privacy cost) and valuation accuracy impact the seller's profit, aggregate consumer surplus, and social welfare under voluntary profiling. As our model leads to some complex expressions, we provide analytical results for tractable cases and confirm these results using numerical analyses for more general cases.

[6]We note that given other parameter values Assumption 2 holds when $\beta > 0.77$ and $\beta > 0.68$ in the high- and low-privacy-cost equilibrium respectively. We discuss the case where Assumption 2 does not hold and thereby the seller ignores the signal in Section 5.2.

Theorem 4. *If valuation accuracy is perfect then aggregate consumer surplus is increasing and the 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.*

Theorem 4 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 increase aggregate consumer surplus. A decrease in the fraction of privacy-sensitive consumers increases the proportion of participating consumers along with 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 but only a subset of non-participating consumers purchases.

Extensive numerical analysis shows that the essential insights provided by Theorem 4 hold for other cases. A representative set of results is shown in Figure 7. The following parameter values are used: $v_H = 100$, $v_L = 70$, $c = 90$, $\lambda = 0.3$, $\alpha = 0.7$, $\beta \in \{0.85, 0.9, 0.95\}$, and $r \in [0, 2.5]$. In Figure 7, the discontinuity seen in each of the graphs corresponds to the point at which the equilibrium switches from low-privacy-cost to high-privacy-cost or vice versa.

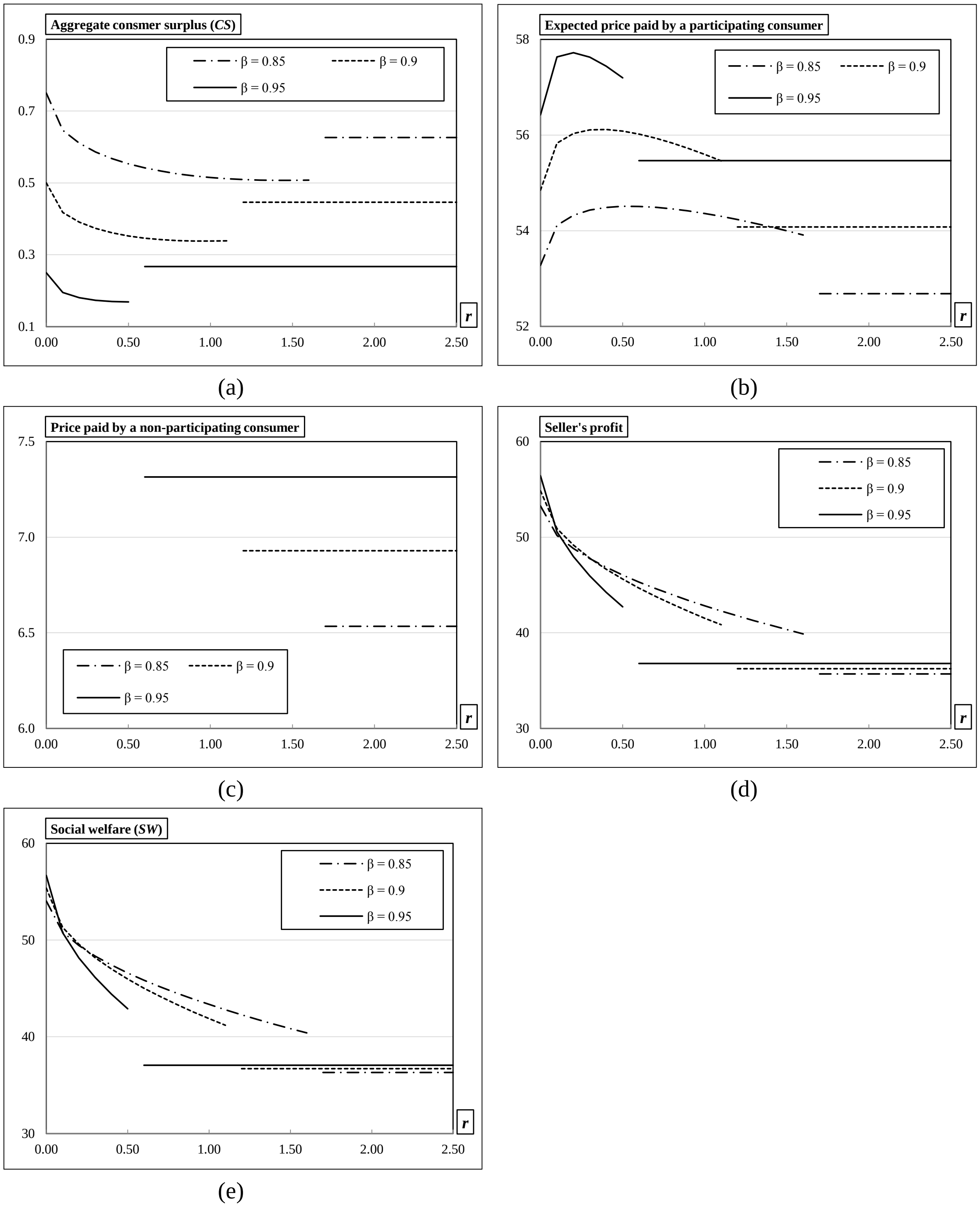


Figure 7: Impact of r and β on (a) aggregate consumer surplus (CS), (b) expected price paid by a participating consumer, (c) price paid by a non-participating consumer, (d) seller's profit, and (e) social welfare (SW)[7]

Figure 7(a) reveals that a decrease in privacy cost affects aggregate consumer surplus only

[7] The mathematical expression of each economic measures used in the figure is given in Section 4.

when privacy cost is lower than a threshold such that some privacy-sensitive consumers find it beneficial to participate in profiling (i.e. the low-privacy-cost equilibrium). As privacy cost decreases aggregate consumer surplus decreases when the participation structure switches from the high-privacy-cost equilibrium to the low-privacy-cost equilibrium; however, once it switches to the low-privacy-cost equilibrium, a further decrease in privacy cost increases aggregate consumer surplus. It is worth noting that aggregate consumer surplus when privacy cost is low is not necessarily higher than when privacy cost is high. For instance, when valuation accuracy is high such as when $\beta = 0.95$, the aggregate consumer surplus in the high-privacy-cost equilibrium is higher than that at any privacy cost (even when privacy cost is zero) in the low-privacy-cost equilibrium.

The expected price paid by a participating consumer is higher in the low-privacy-cost equilibrium than in the high-privacy-cost equilibrium because privacy-non-sensitive consumers with high valuations find it beneficial to participate in profiling when some privacy-sensitive consumers participate in profiling (Figure 7(b)). In the low-privacy-cost equilibrium, as privacy cost decreases more privacy-sensitive consumers, starting from those have high valuations, participate in profiling; hence, a decrease in privacy cost initially increases the expected price paid by a participating consumer yet once low valuation privacy-sensitive consumers also start to participate, a further decrease in privacy cost decreases the expected price paid by a participating consumer. Figure 7(c) suggests that the price paid by a non-participating consumer increases as β increases. Notice that there is no price paid by a non-participating consumer in the low-privacy-cost equilibrium (see (7)).

In Figure 7(d) and Figure 7(e) we find that a decrease in privacy cost never decreases the seller's profit nor social welfare. An increase in valuation accuracy, however, does not necessarily increase the seller's profit and social welfare. For instance, when r is in the range of $[0.25, 1.50]$, the seller's profit and social welfare are higher when valuation accuracy is lower: $\beta = 0.85$ versus $\beta = 0.95$. The increase in valuation accuracy increases the number of participating consumers that purchase but decreases the number of non-participating consumers that purchase. Therefore, if the latter effect is greater than the former effect, then the total number of consumers that purchase decreases and the increase in the valuation accuracy hurts the seller and society.

5.2 Price Discrimination and Pareto Optimality

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 compared to no profiling. However, our theoretical results in the previous subsection show that reducing (or even eliminating) privacy concerns may in fact harm consumers and reducing privacy concerns alone would not necessarily lead to a Pareto improvement. In this subsection, we analyze whether restricting the price discrimination under voluntary profiling may result in a Pareto improvement or even Pareto optimality.

5.2.1 Price discrimination-free voluntary profiling

Suppose the policy-maker could preclude the seller from using price discrimination. This could happen through legal regulation, through social ones such as online forums creating negative publicity for the seller if it practices 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. Then, 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 5. *Compared to no profiling, under price discrimination-free voluntary profiling (i) the surplus of every individual privacy-non-sensitive consumer is not lower, (ii) aggregate consumer surplus is not lower, and (iii) social welfare is not lower.*

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 not lower under price discrimination-free voluntary profiling if the privacy cost is sufficiently low.*

Together Theorem 5 and Corollary 1 show that the price discrimination-free voluntary profiling indeed results in a Pareto optimal outcome if the privacy cost is sufficiently low.

5.2.2 Group-pricing voluntary profiling

Alternatively, the seller could be allowed to charge two different prices: one for participating consumers and the other for non-participating consumers – what we term *group-pricing voluntary profiling*. This group-pricing (third-degree price discrimination) is less restrictive for the seller compared to price discrimination-free but more restrictive than personalized pricing analyzed in Sections 3 and 4. We have the following theorem for group-pricing voluntary profiling.

Theorem 6. *There is a threshold for privacy cost such that (a) if the privacy cost is lower than the threshold, all consumers that are served participate in profiling and group-pricing voluntary profiling is identical to price discrimination-free voluntary profiling and (b) if the privacy cost is not lower than the threshold, under group-pricing voluntary profiling compared to no profiling (i) the surplus of every individual privacy-non-sensitive consumer is not lower, (ii) privacy-sensitive consumers are indifferent, (iii) aggregate consumer surplus is not lower, and (iv) social welfare is not lower.*

Theorem 6 reveals that if the privacy cost is sufficiently low, no consumer purchases her ideal product at the price for a non-participating consumer; hence, effectively group-pricing voluntary profiling is identical to price discrimination-free voluntary profiling. If the privacy cost is sufficiently high, voluntary profiling is weakly Pareto optimal when the seller is allowed to charge two different prices.

The analysis in this subsection shows that voluntary profiling can be Pareto optimal as long as the use of participating consumers' individual information for pricing is successfully restricted. In sum, our findings show that the intuition that no consumer is worse off under voluntary profiling compared to no profiling is valid if the seller's use of individual consumer information for price discrimination (personalized pricing) is restricted and privacy cost is reduced.

6 Extensions

6.1 Presence of an Outside Option

In addition to legal, social, and self-regulation means, it is possible that competition among sellers restricts price discrimination. In this subsection, we study voluntary profiling when consumers have an outside option (i.e., a substitute product) available at a fixed price. This fixed 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 purchases from the seller only if the price is lower than p^o, and as such p^o provides an upper bound of the seller's prices. We take p^o to be exogenous and known to the seller as well as all consumers, as might be the case with a competitor's posted price. As our objective is to show that our main results about voluntary profiling possibly being welfare reducing can be generalized, we restrict our analysis to the high-privacy-cost equilibrium. In the outside option setting 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 such that: (a) if p^o is less than the threshold value, then no non-participating consumer is worse off and all participating consumers are weakly better off; hence, aggregate consumer surplus as well as social welfare are weakly higher and (b) if p^o is greater than or equal to the threshold value, then the results are qualitatively similar to those in Sections 3 and 4: no non-participating consumers are better off, some participating consumers are worse off, and aggregate consumer surplus and social welfare can be higher or lower.*

The result provides an important insight for policy. If the market is competitive in prices, then the policy-maker may not need to intervene in the market and to restrict price discrimination and address privacy concerns. Even when consumers' privacy concerns are severe, price competition among sellers precludes the seller from customizing prices for high-valuation consumers and voluntary profiling likely benefits 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. In this case, the policy-maker may need to restrict price discrimination to ensure that voluntary profiling benefits consumers as well as society.

6.2 Search Support

We demonstrate the generalizability of our results by relaxing some of the assumptions we made earlier regarding search support. We provide a brief description of the results below and leave the details to Appendix B.

6.2.1 Network effects in the profiler

We assume in our base model that search support α is exogenous and independent of the size of the participant group. However, a profiler is likely to be more efficient and effective as information from more consumers becomes available; hence, the seller may be able to provide better search support as more consumers participate in profiling. In order to capture this relationship between the search support and the number of participating consumers we suppose α increases in proportion to the number of participating consumers (i.e., $\alpha = a\,\bar{v}_{NS}$). As our objective is to show that our main results about voluntary profiling possibly being welfare reducing can be generalized to different treatments of search support, we restrict our analysis to the high-privacy-cost equilibrium and the other aspects of the model remain the same.

We find that the optimal price for a participating consumer increases in $\bar{v}_{NS}$. Knowing that a participating consumer enjoys increased savings in search cost as more consumers participate in profiling, the seller charges a higher price and extracts more savings in search cost as more consumers participate in profiling. Therefore, the effect of information from more consumers to develop search support is offset by the effect on price for a participating consumer; all results become qualitatively identical to those we have in Sections 3 and 4. That is, briefly, even in the presence of a network effect on the search support, no non-participating consumer is better off and some participating consumers are worse off under voluntary profiling compared to no profiling. The aggregate consumer surplus and social welfare under voluntary profiling can be higher or lower than under no profiling.

6.2.2 Generic search support for a non-participating consumer

In our base model, we normalize the search support such that the seller's search support without individual consumer information is zero. In this subsection, we suppose the seller provides generic search support to those consumers that do not participate in profiling, i.e., the search cost of a non-participating consumer under voluntary profiling and a consumer under no profiling is $[1-\alpha_0]c$, where $\alpha_0 = \alpha - \epsilon$. Again, as our objective is to show that our main results about voluntary profiling possibly being welfare reducing can be generalized to different treatments of search support, we restrict our analysis to the high-privacy-cost equilibrium and the other aspects of the model remain the same.

We find that with generic search support, as compared to when the seller's search support without individual consumer information is zero (i.e., $\alpha_0 = 0$), the seller charges a higher price under no profiling and to a non-participating consumer under voluntary profiling. However, because the seller does not have information about the individual consumers that do not participate in profiling the seller extracts only a part of savings in search cost, thus the benefit from participation relative to non-participation decreases due to generic search support to non-participating consumers. Consequently, if there is generic search support, then more consumers choose to not participate in profiling. However, although more consumers do not participate, no non-participating consumer is better off under voluntary profiling compared to no profiling. Further, some participating consumers are worse off, and thereby, aggregate consumer surplus and social welfare can be higher or lower under voluntary profiling compared to no profiling.

6.3 Search Cost Heterogeneity

To model consumer heterogeneity we allowed consumers to differ in their valuations for their ideal product. This is consistent with anecdotal evidence that consumers with different profiles have different willingness-to-pay. Our key insight from our model is that low willingness-to-pay consumers or consumers with lower valuations, participate in profiling and consumers with higher valuations do not. In our model we take search cost as identical across consumers.

In an alternative model we consider search costs that differ across consumers but all consumers have identical valuations. This setup is isomorphic to our current model as net valuation – valuation less search cost – is the same in both models. Consequently, all of the main results from the alternative model with heterogeneous search cost are qualitatively the same as the current model, although the interpretations in some cases are slightly different. For example, in the alternative model low net valuation consumers (i.e., those with higher search cost) participate in profiling as they expect greater benefits (i.e., possibly a lower price as well as greater saving in search cost). We have included our analysis of the alternative model as an online supplement.

7 Conclusion

In a model with privacy-sensitive and privacy-non-sensitive consumers, and the seller that may practice price discrimination, we study the impact of the voluntary profiling policy on the seller's profit, consumer surplus, and social welfare. We show 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 social welfare nor 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 reducing (or even eliminating) privacy cost alone does not guarantee a Pareto optimal outcome. The primary reason is that voluntary profiling allows the seller to exploit the heterogeneity in consumers' inherent incentives to participate in profiling and forces 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 the seller chooses to use profile information and to price discriminate when it is profit maximizing, the seller's profit depends on consumers' privacy cost and the fraction of privacy-sensitive consumers. To the degree that the seller 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, the seller can be more profitable. In this sense, the seller's focus should be on consumers' privacy concerns.

In contrast, a policy-maker should be more concerned about price discrimination under voluntary profiling than about privacy. The ability to price discriminate based on individual consumers' information is what policy-makers 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 policy-maker 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 As with any analytical model based on a series of assumptions, our model has some limitations. We assume that consumers incur a common fixed search cost to find their ideal product. In addition, we assume 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 could be 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 regarding ways in which profiling is implemented. This is a potentially fruitful area of future research.

References

Acquisti, A., and J. Grossklags. 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., C. Deck, and A. Farmer. 2012. Sequential Pricing of Multiple Products: Leveraging Revealed Preferences of Retail Customers Online and with Auto-ID Technologies. *Information Systems Research*. **24**(2): 372-393.

Arbatskaya, M. 2007. Ordered Search. *The RAND Journal of Economics*. **38**(1): 119-126.

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

Bakos, J.Y. 1997. Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science*. **43**(12): 1676-1692.

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

Cachon, G.P., C. Terwiesch, and Y. Xu. 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**(1): 168-204.

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

Choudhary, V., A. Ghose, T. Mukhopadhyay, and U. Rajan. 2005. Personalized Pricing and Quality Differentiation. *Management Science*. **51**(7): 1120-1130.

Chen, Y. 1997. Paying Customers to Switch. *Journal of Economics and Management Strategy*. **6**(4): 877-897.

Chen, Y., and G. Iyer. 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**(1): 43-50.

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

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

Clifford, S. 2012. Shopper Alert: Price May Drop for You Alone. *The New York Times*. August 9. Available at: http://www.nytimes.com/2012/08/10/business/supermarkets-try-customizing-prices-for-shoppers.html?pagewanted=all&_r=1&_r=1&.

Conitzer, V., C. Taylor, and L. Wagman. 2011. Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases. *Marketing Science*. **31**(2): 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 D. Baumer. 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 and Clark Law Review*. **10**(Fall): 559-562.

Ellison, G., and S.F. Ellison. 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.

Federal Trade Commission. 2012. Recommendations for Businesses and Policymakers. Available at: http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

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**(2): 238-258.

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

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

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

Iyer G., D. Soberman, and J.M. Villas-Boas. 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 E. Dinlersoz. 2012. Quality-based Price Discrimination: Evidence from Internet Retailers Shipping Options. *Journal of Retailing*. **88**(2): 276-290.

Milne, G.R., and A.J. Rohm. 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. *Proceedings of the 5th international conference on Electronic commerce*. **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. *Washington Post* September 27. Available at: http://www.washingtonpost.com/ac2/wp-dyn/A15159-2000Sep25.

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

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

Tuttle, B. 2012. The Passive-Aggressive Way to Haggle Online: Abandon Your Shopping Cart. *Time* September 27. Available at: http://business.time.com/2012/09/27/the-passive-aggressive-way-to-haggle-online-abandon-your-shopping-cart/.

Tuttle, B. 2013. A Disloyalty Movement? Supermarkets and Customers Drop Loyalty Card Programs. *Time* July 11. Available at: http://business.time.com/2013/07/11/a-disloyalty-movement-supermarkets-and-customers-drop-loyalty-card-programs/.

Valentino-Devries, J., J. Singer-Vine, A. Soltani. 2012. Websites Vary Prices, Deals Based on Users Information. *The Wall Street Journal* December 24. Available at: http://online.wsj.com/news/articles/SB10001424127887323777204578189391813881534.

Villas-Boas, J.M. 1999. Dynamic Competition with Customer Recognition. *RAND Journal of Economics*. **30**(4): 604-631.

Villas-Boas, J.M. 2004. Price Cycles in Markets with Customer Recognition. *RAND Journal of Economics* **35**(3): 486-501.

Villarreal, P. 2010. Capital One Made Me Different Loan Offers Depending on Which Browser I used. *Consumerist* November 1. Available at: http://consumerist.com/2010/11/01/capital-one-made-me-different-loan-offers-depending-on-which-browser-i-used/.

Viswanathan, S., J. Kuruzovich, S. Gosain, and R. Agarwal. 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 L. Krishnamurthi. 2004. Customizing Promotions in Online Stores. *Marketing Science*. **23**(4): 561-578.

Appendix A

A.0. The conditions for Assumption 2

The mathematical conditions for Assumption 2 under each of the high-privacy-cost and low-privacy-cost beliefs are

(a) For the high-privacy-cost belief,

$$Max\left\{\frac{\sqrt{X^2+4\lambda[v_L-[1-\alpha]c][v_H-\lambda v_L-[1-\alpha-\lambda]c]}-X}{2\lambda[v_L-[1-\alpha]c]}, \frac{v_H-[2+\lambda]v_L+[1-\alpha+\lambda]c}{\lambda[v_L-[1-\alpha]c]}\right\} < \beta \leq 1,$$

(b) For the low-privacy-cost belief,

$$Max\left\{\frac{[v_H-[1-\alpha]c]\left[Y+\lambda^2 r+\sqrt{[2Y+\lambda^2 r]\lambda^2 r}\right]}{[1-\lambda][2v_H-v_L-[1-\alpha]c]^2}, 1-\frac{2[v_L-[1-\alpha]c]\lambda^2 r}{[v_H-[2-\lambda]v_L+[1-\lambda][1-\alpha]c]^2}\right\} < \beta \leq 1,$$

where $X = 2v_H - [1+3\lambda]v_L - [1-[1-\lambda]\alpha - 3\lambda]c$ and $Y = [1-\lambda][2v_H - v_L - [1-\alpha]c]$.

We assume in our analysis that the lower bound for β is an interior value between 0 and 1.

A.1. The derivation of optimal solutions for problems in stages 1 and 2

Stage 2: The price for a participating consumer

The seller maximizes profit from a participating consumer in the high-privacy-cost belief given by

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

It yields the following first order condition

$$\frac{\partial \pi^\nu(p_1^\nu)}{\partial p_1^\nu} = \begin{cases} [1-\lambda]\left[\beta + \dfrac{[1-\beta][\bar{v}_{NS} - 2p_1^\nu - [1-\alpha]c]}{\bar{v}_{NS} - v_L}\right] = 0 & \text{for } p_1^\nu \leq \hat{v} - [1-\alpha]c, \\ \dfrac{[1-\lambda][1-\beta][\bar{v}_{NS} - 2p_1^\nu - [1-\alpha]c]}{\bar{v}_{NS} - v_L} = 0 & \text{for } p_1^\nu > \hat{v} - [1-\alpha]c. \end{cases}$$

If the following conditions hold: $\beta - \frac{[1-\beta][\bar{v}_{NS}-[1-\alpha]c]}{\bar{v}_{NS}-v_L} > 0$ and $\bar{v}_{NS} - 2v_L + [1-\alpha]c < 0$, then we have

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

Under Assumption 2, substituting $\bar{v}_{NS}$ given in (6), we verify that the above conditions are satisfied.

The seller maximizes profit from a participating consumer in the low-privacy-cost belief given by

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

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

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

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

and if $\hat{v} \in (\bar{v}'_S, \bar{v}_{NS}]$,

$$\pi_1^\nu(p_1^\nu) = \begin{cases} p_1^\nu \left[\beta + \dfrac{[1-\beta][\bar{v}'_S - \bar{v}_S + [1-\lambda][\bar{v}_S - p_1^\nu - [1-\alpha]c + \bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] \\ \qquad\qquad \text{if } p_1^\nu \leq \bar{v}_S - [1-\alpha]c, \\ p_1^\nu \left[\beta + \dfrac{[1-\beta][\bar{v}'_S - p_1^\nu - [1-\alpha]c + [1-\lambda][\bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] \\ \qquad\qquad \text{if } \bar{v}_S - [1-\alpha]c < p_1^\nu \leq \bar{v}'_S - [1-\alpha]c, \\ p_1^\nu \left[\beta + \dfrac{[1-\beta][1-\lambda][\bar{v}_{NS} - p_1^\nu - [1-\alpha]c]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] \\ \qquad\qquad \text{if } \bar{v}'_S - [1-\alpha]c < p_1^\nu \leq \hat{v} - [1-\alpha]c, \\ \dfrac{p_1^\nu[1-\beta][1-\lambda][\bar{v}_{NS} - p_1^\nu - [1-\alpha]c]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} \qquad \text{if } \hat{v} - [1-\alpha]c < p_1^\nu \leq \bar{v}_{NS} - [1-\alpha]c. \end{cases}$$

It yields the following first order condition

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

$$\frac{\partial \pi^\nu(p_1^\nu)}{\partial p_1^\nu} = \begin{cases} \left[\beta + \dfrac{[1-\beta][\bar{v}'_S - \bar{v}_S + [1-\lambda][\bar{v}_S - 2p_1^\nu - [1-\alpha]c + \bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] = 0 \\ \qquad \text{for } p_1^\nu \leq \hat{v} - [1-\alpha]c, \\ \dfrac{[1-\beta][\bar{v}'_S - \bar{v}_S + [1-\lambda][\bar{v}_S - 2p_1^\nu - [1-\alpha]c + \bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} = 0 \\ \qquad \text{for } \hat{v} - [1-\alpha]c < p_1^\nu \leq \bar{v}_S - [1-\alpha]c, \\ \dfrac{[1-\beta][\bar{v}'_S - 2p_1^\nu - [1-\alpha]c + [1-\lambda][\bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} = 0 \\ \qquad \text{for } \bar{v}_S - [1-\alpha]c < p_1^\nu \leq \bar{v}'_S - [1-\alpha]c, \\ \dfrac{[1-\beta][1-\lambda][\bar{v}_{NS} - 2p_1^\nu - [1-\alpha]c]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} = 0 \quad \text{for } \bar{v}'_S - [1-\alpha]c < p_1^\nu \leq \bar{v}_{NS} - [1-\alpha]c, \end{cases}$$

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

$$\frac{\partial \pi^\nu(p_1^\nu)}{\partial p_1^\nu} = \begin{cases} \left[\beta + \dfrac{[1-\beta][\bar{v}'_S - \bar{v}_S + [1-\lambda][\bar{v}_S - 2p_1^\nu - [1-\alpha]c + \bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] = 0 \\ \qquad \text{for } p_1^\nu \leq \bar{v}_S - [1-\alpha]c, \\ \left[\beta + \dfrac{[1-\beta][\bar{v}'_S - 2p_1^\nu - [1-\alpha]c + [1-\lambda][\bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] = 0 \\ \qquad \text{for } \bar{v}_S - [1-\alpha]c < p_1^\nu \leq \hat{v} - [1-\alpha]c, \\ \dfrac{[1-\beta][\bar{v}'_S - 2p_1^\nu - [1-\alpha]c + [1-\lambda][\bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} = 0 \\ \qquad \text{for } \hat{v} - [1-\alpha]c < p_1^\nu \leq \bar{v}'_S - [1-\alpha]c, \\ \dfrac{[1-\beta][1-\lambda][\bar{v}_{NS} - 2p_1^\nu - [1-\alpha]c]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} = 0 \quad \text{for } \bar{v}'_S - [1-\alpha]c < p_1^\nu \leq \bar{v}_{NS} - [1-\alpha]c, \end{cases}$$

and if $\hat{v} \in (\bar{v}'_S, \bar{v}_{NS}]$,

$$\frac{\partial \pi^\nu(p_1^\nu)}{\partial p_1^\nu} = \begin{cases} \left[\beta + \dfrac{[1-\beta][\bar{v}'_S - \bar{v}_S + [1-\lambda][\bar{v}_S - 2p_1^\nu - [1-\alpha]c + \bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] = 0 \\ \qquad \text{for } p_1^\nu \leq \bar{v}_S - [1-\alpha]c, \\ \left[\beta + \dfrac{[1-\beta][\bar{v}'_S - 2p_1^\nu - [1-\alpha]c + [1-\lambda][\bar{v}_{NS} - \bar{v}'_S]]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] = 0 \\ \qquad \text{for } \bar{v}_S - [1-\alpha]c < p_1^\nu \leq \bar{v}'_S - [1-\alpha]c, \\ \left[\beta + \dfrac{[1-\beta][1-\lambda][\bar{v}_{NS} - 2p_1^\nu - [1-\alpha]c]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]}\right] = 0 \\ \qquad \text{for } \bar{v}'_S - [1-\alpha]c < p_1^\nu \leq \hat{v} - [1-\alpha]c, \\ \dfrac{[1-\beta][1-\lambda][\bar{v}_{NS} - p_1^\nu - [1-\alpha]c]}{\lambda[\bar{v}'_S - \bar{v}_S] + [1-\lambda][\bar{v}_{NS} - v_L]} = 0 \quad \text{for } \hat{v} - [1-\alpha]c < p_1^\nu \leq \bar{v}_{NS} - [1-\alpha]c. \end{cases}$$

If the following conditions hold: $\beta - \frac{[1-\beta][v_H-[1-\alpha]c]}{[v_H-v_L]-\lambda[\bar{v}_S-v_L]} > 0$ and $\lambda[v_H - \bar{v}_S] + [1-\lambda][v_H - 2v_L + [1-\alpha]c] < 0$, then we have

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

Under Assumption 2, substituting $\bar{v}_S$, $\bar{v}'_S$, and $\bar{v}_{NS}$ given in (8) and (9), we verify that the above conditions are satisfied.

Stage 1: The price for a non-participating consumer and the valuation of consumer indifferent between participating and not

Under the high-privacy-cost belief the seller maximizes profit from a non-participating consumer given by

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

The first order condition is

$$\frac{\partial \pi_0^\nu(p_0^\nu)}{\partial p_0^\nu} = \frac{v_H - [1-\lambda]\bar{v}_{NS} - 2\lambda p_0^\nu - \lambda c}{v_H - v_L} = 0.$$

Further, using (3) and (4), from $U_0^\nu(v = \bar{v}_{NS}, NS) = U_1^\nu(v = \bar{v}_{NS}, NS)$, we have

$$\bar{v}_{NS} - p_0^\nu - c = \int_{v_L}^{\bar{v}_{NS}} q(\hat{v} \neq \bar{v}_{NS}|\bar{v}_{NS}) \left[\bar{v}_{NS} - \hat{v}\right] d\hat{v}.$$

Hence, we obtain

$$p_0^{\nu *} = \frac{[1+\beta]v_H + [1-\lambda][1-\beta]v_L - [2 - [1-\beta]\lambda]c}{2[1+\lambda\beta]},$$

and

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

Under the low-privacy-cost belief the seller maximizes profit from a non-participating consumer given by

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

The first order condition is

$$\frac{\partial \pi_0^\nu(p_0^\nu)}{\partial p_0^\nu} = \frac{v_H - \lambda\bar{v}'_S - [1-\lambda]\bar{v}_{NS}}{v_H - v_L} = 0.$$

Further, using (3) and (4), from $U_0^\nu(v=\bar{v}_{NS}, NS) \leq U_1^\nu(v=\bar{v}_{NS}, NS)$, we have

$$\bar{v}_{NS} - p_0^\nu - c \leq \int_{v_L}^{\bar{v}_{NS}} q(\hat{v} \neq \bar{v}_{NS}|\bar{v}_{NS})\,[\bar{v}_{NS} - \hat{v}]\,d\hat{v},$$

from $U_0^\nu(v=\bar{v}'_S, S) \leq U_1^\nu(v=\bar{v}'_S, S)$, we have

$$\bar{v}'_S - p_0^\nu - c \leq \int_{v_L}^{\bar{v}'_S} q(\hat{v} \neq \bar{v}'_S|\bar{v}'_S)\,[\bar{v}'_S - \hat{v}]\,d\hat{v} - r,$$

and from $U_1^\nu(v=\bar{v}_S, S) = U_0^\nu(v=\bar{v}_S, S) = 0$, we have

$$\int_{v_L}^{\bar{v}_S} q(\hat{v} \neq \bar{v}_S|\bar{v}_S)\,[\bar{v}_S - \hat{v}]\,d\hat{v} - r = 0.$$

Hence, we obtain

$$p_0^{\nu*} > v_H - c,$$

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

and

$$\bar{v}'_S = \bar{v}_{NS} = v_H.$$

A.2. Proof of Theorem 1

Using $\bar{v}_{NS}$ given in (6) and $\bar{v}_S$ given in (8), we show

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

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

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

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

A.3. Proof of Theorem 2

(i) In the 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 (5)), we show

$$p_0^{\nu*} - p^{b*} = \frac{[1-\lambda][\beta v_H + [1-\beta]v_L - c]}{2[1+\lambda\beta]} \geq 0.$$

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

(ii) Comparing the price under no profiling (given in (1)) and the expected price paid by a participating consumer under voluntary profiling (given in (11)), 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, $c^* = \frac{[2[1-\beta]+3\lambda\beta]v_H - \left[2+\beta[1+\lambda]-[1-3\beta^2]\lambda\right]v_L}{[1+2\beta]\lambda - 3\beta[1+\lambda\beta]+3\alpha[1+\beta][1+\lambda\beta]}$ for the high-privacy-cost equilibrium, and $\frac{3v_H-[2+\beta]v_L}{3[\alpha[1+\beta]-\beta]} - \frac{[1+2\beta][v_H-\lambda\bar{v}_S]}{3[\alpha[1+\beta]-\beta][1-\lambda]} + \frac{\lambda[v_H-\bar{v}_S]^2[[1+2\beta][v_H-\lambda\bar{v}_S]-[1-\lambda][[1-\beta]\bar{v}_S+3\beta v_L]]}{3[\alpha[1+\beta]-\beta][1-\lambda][[v_H-v_L]-\lambda[\bar{v}_S-v_L]]^2}$ for the low-privacy-cost equilibrium. □

A.4. Proof of Theorem 3

(i) From $p_0^{\nu *} \geq p^{b*}$ (Theorem 2(i)), we show $U_0^{\nu}(v) \leq U^b(v)$ for all v.

(ii) Further, using $U_1^{\nu}(\bar{v}_{NS}, NS) = U_0^{\nu}(\bar{v}_{NS}, NS) \leq U^b(\bar{v}_{NS}, NS)$ 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 $U_1^{\nu}(\tilde{v}_{NS}, NS) = U^b(\tilde{v}_{NS}, NS)$.

From $U_1^{\nu}(v, NS) > U_1^{\nu}(v, S)$ we show

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

A.5. Illustration of aggregate consumer surplus under voluntary profiling versus no profiling

In the high-privacy-cost equilibrium (i.e., $r > 2$), for $v_H = 100$, $v_L = 70$, $c = 87$, $\lambda = 0.6$, and $\alpha = 0.9$, we have

$$CS^{\nu} = 0.79 > CS^b = 0.70 \text{ when } \beta = 0.80,$$

whereas

$$CS^{\nu} = 0.63 < CS^b = 0.70 \text{ when } \beta = 0.88.$$

Similarly, in the low-privacy-cost equilibrium (i.e., $r = 0.1$), for $v_H = 100$, $v_L = 70$, $c = 93$, $\lambda = 0.6$, and $\alpha = 0.9$, we have:

$$CS^{\nu} = 0.31 > CS^b = 0.20 \text{ when } \beta = 0.90,$$

whereas

$$CS^{\nu} = 0.16 < CS^{b} = 0.20 \text{ when } \beta = 0.94. \ \square$$

A.6. Illustration of social welfare under voluntary profiling versus no profiling

In the high-privacy-cost equilibrium ($r > 2$), for $v_H = 100$, $v_L = 70$, $c = 44$, $\lambda = 0.1$, and $\alpha = 0.1$, we have

$$SW^{\nu} = 38.54 < SW^{b} = 39.20 \text{ when } \beta = 0.80,$$

whereas

$$SW^{\nu} = 40.47 > SW^{b} = 39.20 \text{ when } \beta = 0.95.$$

Similarly, in the low-privacy-cost equilibrium (i.e., $r = 0.4$), for $v_H = 100$, $v_L = 70$, $c = 41.2$, $\lambda = 0.1$, and $\alpha = 0.1$, we have

$$SW^{\nu} = 43.07 < SW^{b} = 43.22 \text{ when } \beta = 0.87,$$

whereas

$$SW^{\nu} = 43.81 > SW^{b} = 43.22 \text{ when } \beta = 0.92. \ \square$$

A.7. Proof of Theorem 4

If $\beta = 1$, no privacy-sensitive consumer participates in profiling and privacy-non-sensitive consumers whose valuations are not greater than $p_0^{\nu*} + c$ participate in profiling. Hence, from $p_0^{\nu*} + c = \bar{v}_{NS}$, we have: $p_0^{\nu*} = \frac{v_H - c}{1+\lambda}$, and from

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

and

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

we show

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

$$\frac{\partial CS^{\nu}}{\partial \lambda} = -\frac{\partial p_0^{\nu*}}{\partial \lambda} \cdot \frac{v_H - p_0^{\nu*} - c}{v_H - v_L} \geq 0, \text{ and}$$

$$SW^{\nu}(\lambda = 0) - SW^{\nu}(\lambda = 1) = \int_{v \leq \frac{v_H + c}{2}} f(v)[v - c]dv + \int_{v > \frac{v_H + c}{2}} f(v) \alpha c \, dv \geq 0. \ \square$$

A.8. Proof of Theorem 5

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

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

all privacy-non-sensitive consumers 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 $r \leq \alpha c$.

The seller's profit is given by

$$\pi^{\nu} = \begin{cases} \dfrac{p^{\nu}[v_H - p^{\nu} - [1-[1-\lambda]\alpha]c]}{v_H - v_L} & \text{if } r > \alpha c, \\ \dfrac{p^{\nu}[v_H - p^{\nu} - [1-\alpha]c - \lambda r]}{v_H - v_L} & \text{if } r \leq \alpha c, \end{cases}$$

and maximizing it yields

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

Using $p^{\nu *}$ obtained above, we show

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

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

(ii) Aggregate consumer surplus is given by

$$CS^{\nu} = \begin{cases} \dfrac{\lambda[v_H - p^{\nu} - c]^2 - [1-\lambda][v_H - p^{\nu} - [1-\alpha]c]^2}{2[v_H - v_L]} & \text{if } r > \alpha c, \\ \dfrac{[v_H - p^{\nu} - [1-\alpha]c]^2 - \lambda r^2}{2[v_H - v_L]} & \text{if } r \leq \alpha c. \end{cases}$$

Hence, we show

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

(iii) As the seller can always choose to ignore the profile information, we have $\pi^{\nu} \geq \pi^{b}$. Hence, together with (ii) we have $SW^{\nu} \geq SW^{b}$. □

A.9. Proof of Corollary 1

As is shown in Proof of Theorem 5(i), $U^{\nu}(v, NS) \geq U^{b}(v, NS)$ for all v, and $U^{\nu}(v, S) > U^{b}(v, S)$ for all v if $r < \frac{\alpha c}{2-\lambda}$. □

A.10. Proof of Theorem 6

From $U_1^{\nu}(v, NS) = v - p_1^{\nu} - [1-\alpha]c \geq U_0^{\nu}(v, NS) = v - p_0^{\nu} - c$, privacy-non-sensitive consumers participate in profiling if $\alpha c - [p_1^{\nu} - p_0^{\nu}] \geq 0$. Similarly, privacy-sensitive consumers participate in profiling if $\alpha c - r - [p_1^{\nu} - p_0^{\nu}] \geq 0$.

(a) If $r \leq \alpha c - [p_1^{\nu} - p_0^{\nu}]$, then both privacy-sensitive and privacy-non-sensitive consumers (whose surplus are positive) participate in profiling. Hence, there is only one price (for participating consumers) and results are identical to those for price discrimination-free voluntary profiling.

(b) If $r > \alpha c - [p_1^{\nu} - p_0^{\nu}]$, then privacy-sensitive consumers do not participate in profiling. Suppose privacy-non-sensitive consumers participate in profiling. Then, maximizing the seller's profit from participating consumers given by

$$\max_{p_1^{\nu}} \pi_1^{\nu}(p_1^{\nu}) = p_1^{\nu}\,[1-\lambda] \int_{p_1^{\nu}+[1-\alpha]c}^{v_H} f(v)dv,$$

we have

$$p_1^{\nu *} = \frac{1}{2}\left[v_H - [1-\alpha]c\right],$$

and maximizing the seller's profit from non-participating consumers given by

$$\max_{p_0^{\nu}} \pi_0^{\nu}(p_0^{\nu}) = p_0^{\nu}\,\lambda \int_{p_0^{\nu}+c}^{v_H} f(v)dv,$$

we have

$$p_0^{\nu *} = \frac{1}{2}\left[v_H - c\right].$$

Using $p_1^{\nu *}$ and $p_0^{\nu *}$, from $\alpha c - [p_1^{\nu *} - p_0^{\nu *}] = \frac{\alpha c}{2} \geq 0$, we confirm that privacy-non-sensitive consumers participate. Further, we have: (i) $U_1^{\nu} - U^b = \frac{\alpha c}{2} \geq 0$ and from $p_0^{\nu *} = p^{b*}$ (ii) $U_0^{\nu} = U^b$. Therefore, we have: (iii) $CS^{\nu} \geq CS^b$ and (iv) $SW^{\nu} \geq SW^b$. □

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

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

and the seller's price for a participating and a non-participating consumer (given as (2) and (5) in the monopoly model) under voluntary profiling is

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

$$p_0^{\nu} = \min\left\{\frac{[1+\beta]v_H + [1-\lambda][1-\beta]v_L - [2-[1-\beta]\lambda]c}{2[1+\lambda\beta]}, p^o\right\}.$$

We identify three cases depending on p^o as follows:

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: $\frac{[1+\beta]v_H + [1-\lambda][1-\beta]v_L - [2-[1-\beta]\lambda]c}{2[1+\lambda\beta]} \leq p^o < \bar{v}_{NS} - [1-\alpha]c$

Consider a privacy non-sensitive consumer with valuation v. If this consumer participates in profiling, their surplus is

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

and if this consumer does not participate in profiling, their surplus is

$$U_0^\nu(v, NS) = \begin{cases} 0 & \text{if } v < p_0^\nu + c, \\ v - p_0^\nu - c & \text{if } v \geq p_0^\nu + c, \end{cases}$$

Hence, solving $\frac{\partial \pi_0^\nu(p_0^\nu)}{\partial p_0^\nu} = \frac{v_H - [1-\lambda]\bar{v}_{NS} - 2\lambda p_0^\nu - \lambda c}{v_H - v_L} = 0$ and $U_1^\nu(v = \bar{v}_{NS}, NS) = U_0^\nu(v = \bar{v}_{NS}, NS)$ together we have

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

and

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

where $\psi = v_H + [1-\lambda]v_L - [2p^o + [1-2\alpha]c]\lambda$. Therefore, we show

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

We have $p_0^{\nu *} \geq p^{b*}$. Further, given that there is a $\bar{v}_{NS} \leq v_H$, 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: $p^o < \frac{[1+\beta]v_H + [1-\lambda][1-\beta]v_L - [2-[1-\beta]\lambda]c}{2[1+\lambda\beta]}$

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

Appendix B

B.1. Network effects in the profiler

We restrict our analysis to the high-privacy-cost equilibrium and assume Assumption 2 continues to hold. Suppose α increases in proportion to the number of participating consumers (i.e., $\alpha = a\,\bar{v}_{NS}$). Then the seller maximizes profit from a participating consumer given by

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

It yields the following first order condition:

$$\frac{\partial \pi^{\nu}(p_1^{\nu})}{\partial p_1^{\nu}} = \begin{cases} [1-\lambda]\left[\beta + \dfrac{[1-\beta][\bar{v}_{NS} - 2p_1^{\nu} - [1-a\bar{v}_{NS}]c]}{\bar{v}_{NS} - v_L}\right] = 0 & \text{for } p_1^{\nu} \leq \hat{v} - [1-a\bar{v}_{NS}]c, \\ \dfrac{[1-\lambda][1-\beta][\bar{v}_{NS} - 2p_1^{\nu} - [1-a\bar{v}_{NS}]c]}{\bar{v}_{NS} - v_L} = 0 & \text{for } p_1^{\nu} > \hat{v} - [1-a\bar{v}_{NS}]c. \end{cases}$$

If (i) $\beta - \frac{[1-\beta][\bar{v}_{NS} - [1-a\bar{v}_{NS}]c]}{\bar{v}_{NS} - v_L} > 0$ and (ii) $\bar{v}_{NS} - 2v_L + [1-a\bar{v}_{NS}]c < 0$, then

$$p_1^{\nu *} = \hat{v} - [1-a\bar{v}_{NS}]c.$$

Under Assumption 2, substituting $\bar{v}_{NS}$, we verify that the above conditions are satisfied.

Further, solving $\frac{\partial \pi_0^{\nu}(p_0^{\nu})}{\partial p_0^{\nu}} = \frac{v_H - [1-\lambda]\bar{v}_{NS} - 2\lambda p_0^{\nu} - \lambda c}{v_H - v_L} = 0$ and $U_0^{\nu}(v = \bar{v}_{NS}, NS) = U_1^{\nu}(v = \bar{v}_{NS}, NS)$ together we have

$$p_0^{\nu *} = \frac{[1+\beta]v_H + [1-\lambda][1-\beta]v_L - [2 - [1-\beta]\lambda]c}{2[1+\lambda\beta]},$$

and

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

We show $p_0^{\nu *} - p^{b*} = \frac{[1-\lambda][\beta v_H + [1-\beta]v_L - c]}{2[1+\lambda\beta]} \geq 0$. Hence, compared to no profiling, under voluntary profiling, no non-participating consumers are better off. Also, given that there is a $\bar{v}_{NS}$ such that $U_1^{\nu}(v, NS) \geq U_0^{\nu}(v, NS)$ if $v \leq \bar{v}_{NS}$ and $U_1^{\nu}(v, NS) < U_0^{\nu}(v, NS)$ otherwise, some participating consumers are worse off. Therefore, aggregate consumer surplus and social welfare under voluntary profiling can be higher or lower. Proofs are analogous to those for the base model.

B.2. Generic search support for a non-participating consumer

We restrict our analysis to the high-privacy-cost equilibrium. Suppose the search cost of non-participating consumer is $[1-\alpha_0]c$, where $\alpha_0 = \alpha - \epsilon$. Under no profiling the seller maximizes profit given by

$$\max_{p^b} \pi^b(p^b) = p^b \int_{p^b+[1-\alpha_0]c}^{v_H} f(v)dv.$$

It yields

$$p^{b^*} = \frac{1}{2}\left[v_H - [1-\alpha_0]c\right].$$

Under voluntary profiling, $p_1^{\nu *}$ is given in (2). Further, solving $\frac{\partial \pi_0^\nu(p_0^\nu)}{\partial p_0^\nu} = \frac{v_H - [1-\lambda]\bar{v}_{NS} - 2\lambda p_0^\nu - \lambda[1-\alpha_0]c}{v_H - v_L} = 0$ and $U_1^\nu(v = \bar{v}_{NS}, NS) = U_0^\nu(v = \bar{v}_{NS}, NS)$ together, we have

$$p_0^{\nu *} = \frac{[1+\beta]v_H + [1-\beta][1-\lambda]v_L - [2-[1-\beta]\lambda][1-\alpha_0]c}{2[1+\lambda\beta]},$$

and

$$\bar{v}_{NS} = \frac{v_H - \lambda[1-\beta]v_L + \lambda[1-\alpha_0]c}{1+\lambda\beta}.$$

We show $p_0^{\nu *} - p^{b^*} = \frac{[1-\lambda][\beta v_H + [1-\beta]v_L - [1-\alpha_0]c]}{2[1+\lambda\beta]} \geq 0$. Hence, compared to no profiling, under voluntary profiling, no non-participating consumers are better off. Also, given that there is a $\bar{v}_{NS}$ such that $U_1^\nu(v, NS) \geq U_0^\nu(v, NS)$ if $v \leq \bar{v}_{NS}$ and $U_1^\nu(v, NS) < U_0^\nu(v, NS)$ otherwise, some participating consumers are worse off. Therefore, aggregate consumer surplus and social welfare under voluntary profiling can be higher or lower. Proofs are analogous to those for the base model