Available online at www.sciencedirect.com
ScienceDirect
European Journal of Operational Research 187 (2008) 857–870
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/ejor

Privacy and pricing personal information

Jeevan Jaisingh [a,*], Jack Barron [b], Shailendra Mehta [b], Alok Chaturvedi [b]

[a] HKUST, Clear Water Bay, Kowloon, Hong Kong
[b] Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47906, USA

Received 31 August 2005; accepted 22 March 2006
Available online 1 November 2006

Abstract

The issues we address here are – How should a firm (e.g. Internet service provider (ISP)) that is capable of collecting personal information (browsing information, purchase history, etc.) about consumers, price its service, given that consumers vary in their valuation for privacy, and also vary in terms of the value of their personal information to a third party (firms that need consumer information)? Should the firm have a blanket policy of never collecting, or a policy of always collecting and revealing information? Surprisingly we find that in some cases the collector of information may be no worse off in the asymmetric information case than in the full information case. The paper provides a justification for the strategy of some firms such as ISP's which never collect information and also for the strategy of other firms, like grocery stores that do. We also find that it is non-optimal for the firm to design contracts where the consumer can choose an intermediate level of privacy.

Keywords: Information privacy; Pricing; Personal information; Profiles

1. Introduction

The technological developments that have made e-commerce possible have also enhanced the ability of companies to collect, store, transfer, and analyze vast amounts of data, from and about the consumers who visit their store on the World Wide Web [6].

> "Data Collection is the dominant activity of commercial web sites. Nearly 92 percent of them collect personal data from web users, which they can aggregate, sort and use." – [11]

These developments have increased privacy concerns among consumers, especially online consumers. Privacy is defined as the moral claim of the individuals to be left alone and to control the flow of information about themselves [5]. The loss of privacy in the Internet context comes from the inherent disutility associated with firms collecting browsing behavior, or tracking websites visited. According to a Business Week/Harris poll [2], privacy is the number one consumer issue facing the Internet. Another poll by the Wall Street Journal [21] on what Americans feared most in the next millennium, concerns regarding privacy, rated highest.

* Corresponding author.
E-mail addresses: jeevan@ust.hk (J. Jaisingh), barron@mgmt.purdue.edu (J. Barron), mehta@mgmt.purdue.edu (S. Mehta), alok@mgmt.purdue.edu (A. Chaturvedi).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2006.03.062

The other side of the story is that personal information is a valuable asset to private and governmental institutions, which use it to reduce their costs of operation [10]. The increasing thirst for personal information, and innovations in targeting technique, have made targeted marketing the hottest form of marketing, growing at twice the rate of America's GNP. It has been suggested by several authors [10,14] that the personal information, required for targeted marketing, can be treated as a commodity, and that firms who require this information, would be ready to pay a price for it in an 'information market'. There are several examples of firms providing an incentive to consumers to reveal personal information. Software firms provide free versions of their software which are advertising-supported (commonly called adware[1]). Customers view personalized advertisements from marketers based on the information collected while using the adware. Advertising free version of the same software can be bought at a price. Catalina Marketing Corporation offers monetary incentives to induce consumers to provide personal information. For example, a person's zip code and preferred supermarket is worth $40 of coupons, and a personal shopping card number garners free products [12,18]. Grocery stores offer discounts on goods purchased using grocery store cards. Consumers accept discounts, with the implicit assumption that information on purchases collected, will be collected, and then sold to third parties by the grocery store [3]. The third parties then target the consumers with one to one marketing, based on the consumer profiles formed using the purchase history.

This collection of information by firms is leading to increased privacy concerns among consumers. These concerns about privacy need to be addressed to enable the growth of E-commerce. A proper balance needs to be achieved between the consumers' concerns for privacy and the needs of firms for personal information. This paper seeks to address the sharp increase in public debate about privacy issues, particularly on the issues of Internet privacy and the value of personal information.

In this paper the problem that we model is that of a collector of information (Internet service provider (ISP), adware seller, grocery store, etc.). It is technically feasible for the collector to track consumer behavior – browsing profiles in case of the ISP and purchasing information in the case of the grocery store. If allowed to collect data, the collector can potentially build up a profile of the consumer, which may include but is not limited to information about preferences, income segment, etc. This information may be of great use to a third party who wants to target the consumers with a personalized product or service. By providing the third party access to this information, the collector enables the matching of the third party with the right consumers. This collection and release of consumer information, is what is termed as privacy violation in our context.

The research questions we are addressing here are – How should the collector of information price its service given that the consumers vary in their valuation for privacy and also vary in terms of the value of their personal information to the third party? Should the collector have a blanket policy of never collecting, or a policy of always collecting and revealing information? Is it better off providing its customers with a choice of several pricing options, each varying in the amount of information collected and revealed?

The results offer insights into when collection of information is optimal/non-optimal and provide a justification for the strategy of firms such as ISP's which never collect information, and other firms like grocery stores that do. Contrary to intuition, collection of information is not always optimal from the perspective of the firm collecting information. The environment that we have, is characterized by asymmetric information, where the consumers know their type while the collector and third parties do not know the type of the consumer. Surprisingly we find that in some cases the collector is no worse off than in the first best case, with no asymmetry of information. We also find that it is non-optimal for the firm to design contracts where the consumer can choose an intermediate level of privacy. Insights into conditions under which different pricing strategies are optimal are provided.

The rest of the paper is organized as follows: Section 2 covers the background literature on privacy. We specify our assumptions and the model framework in Section 3 – subsections look at the public, full and asymmetric information cases. We extend our basic framework in Section 4. Section 5 summarizes the results and the managerial implications. We also look at the limitations of our results and possible extensions.

[1] http://en.wikipedia.org/wiki/Adware.

2. Background

Privacy is defined as the moral claim of the individuals to be left alone and to control the flow of information about themselves [5,22]. Coasian logic suggests that given sufficient flexibility in contracting, stringent privacy laws are unnecessary to achieve efficiency in concealment or revelation of information [8]. For a background on the evolution of privacy concerns see [4].

Sovern [14] addresses the issue of opting in or opting out of databases and mailing lists. The article argues that businesses have both the incentive and the ability to increase consumers' transaction costs in protecting privacy. Faced with this, and other constraints, many consumers decide ultimately not to protect their privacy. The paper proposes several ways by which consumers' transaction costs can be reduced or eliminated. Kang [9] is more relevant to our research question, since it focuses on the problem of personal data generated in cyberspace transactions. The paper proposes solving the problem by treating personal information as a commodity that interested parties should contract for in the course of negotiating a cyberspace transaction. The proposal for treating information as a commodity, is equivalent to putting a price to ones personal information or privacy. Laudon [10] proposes a framework for an 'Information Market' where cybermediaries would aggregated and sell consumer profiles as commodities. Consumers would be compensated for the loss of privacy by money or personalized services.

Stewart and Segars [16] develop an instrument that identifies and measures the consumers concerns for information privacy. Other similar empirical examinations of privacy include Culnan [4] and Smith et al. [13]. Smith et al. [13] classifies privacy costs along four dimensions: collection, error, secondary use and improper access. Hann et al. [7] conduct a conjoint analysis to explore individuals' tradeoffs between the benefits and costs, of providing personal information to websites. The paper finds monetary values for the last three dimensions of privacy – particularly disallowance of secondary use of personal information was found to be worth between \$39.83 and \$49.78. Tam et al. [19] offers a framework of seven motivators that can induce consumers to disclose their personal information. These include monetary savings, time savings and pleasure. They find that although all seven motivators are generally desirable to consumers, the consumers may differ in terms of their inclination towards specific motivators.

Kahn et al. [8] look at efficiency gains under different privacy regimes. They analyze the problem of a social planner of when and when not transactional information should be made public knowledge. Our framework is similar, although we look at completely different research questions. Other papers that look at privacy from an analytical framework include [20,1]. Both papers look at the issue of using/selling a customer list (purchasing decisions in previous period) which would enable firms to classify customers according to their valuations. Two settings are studied, an anonymity regime in which sale of the customer information is not possible, and a recognition regime in which a firm may compile and sell/use a customer list. Although there are similarities in the issues that we study – privacy, sale of customer information, etc.; a major difference is that we consider that individuals are inherently different in their privacy preferences, while they focus on another important reason for wishing to remain anonymous, discrimination in the form of dynamic pricing. In this paper customer information could be any information – browsing information, demographic information, etc., whereas in [1,20], customer information is their purchasing decision in the previous period.

3. Model

There are three sets of agents in the market – the consumers (A), the collector of information (B) and the third parties (C). A third party is a firm that requires the consumer's information to target her with a personalized product or service. There are a large number of consumers and third parties but just a single collector (firm). All consumers are assumed to have the same valuation for the product/service provided by B. This simplifying assumption was made since our primary interest is in the pricing of a homogeneous service or product given heterogeneity in consumers privacy valuations, and value of their information to a third party. Assuming that consumer differ in their value for the service, would only complicate the framework. Some of the consumers can be potential customers of a third party C. However without prior information about consumer A's buying profile, it is prohibitively expensive for a third party C to target a consumer A. An example of targeting could be sending a sample product. If the third party has no information about the consumers' interests then sending samples to the entire set of consumers would be very expensive.

If A allows B to collect information on her browsing behavior and subsequently release that information to C, then C can identify if A can be a potential customer or not. The consumers differ in their valuation for privacy, and also in their value to a third party based on their interests. Consumers who have a higher privacy valuation are consumers who associate a higher disutility with the revelation of their information. For example high privacy valuation consumers could be those consumers who put up more sophisticated filters to avoid junk mail. The cost associated with putting up these filters both in effort terms and in monetary terms is their cost associated with loss of privacy. Monetary costs associated with loss of privacy can be estimated using conjoint analysis as shown by [7].

Consumers who have a high value of information, are consumers who could potentially buy a personalized product or service from the third party. For example if the third party is a firm that sells Jazz music, then consumers who listen to jazz will be of value to the third party. Let Consumer type be a vector $\theta = (\alpha, \beta)$. α represents the valuation for privacy (L-LOW, H-HIGH).[2] β represents whether the buyer's information is useful to the third party and affiliates of B (H-HIGH, L-LOW). Thus there are four consumer types:

$$\theta_1 = (\mathrm{H}, \mathrm{L}); \quad \theta_2 = (\mathrm{H}, \mathrm{H}); \quad \theta_3 = (\mathrm{L}, \mathrm{H}); \quad \theta_4 = (\mathrm{L}, \mathrm{L}). \tag{1}$$

Consumers know their type while the collector and third parties do not know the type of the consumer. Thus there is asymmetry in information among market participants. All agents are risk-neutral and the reservation utility of all consumers is assumed to be zero. The firm B is a profit-maximizing monopolist.[3] We consider a two period model. In the first period the consumer A chooses to subscribe to the product/service provided by B. Based on the option chosen by A (to allow B to collect information or not), B collects information on A and then releases the information to a third party C. Now C knows A's type perfectly and targets her with a personalized service or product. An example of a personalized product or service could be a CD on 1960s jazz. In actual practice C would update her prior belief on the type of A based on the consumer profile (signal), formed from the browsing information. Here we make the simplifying assumption that the signal received is a perfect one.[4] As a result of this targeting, the utility of agents A, B and C change. An example of this change in utility could be if the consumer ends up buying Jazz CD's from the third party – the consumer gets some utility from the purchase and pays a price to the third party. Since B enables the matching between A and C, B can extract a rent from C. We summarize our key assumptions below:

Key assumptions:

A1: Three types of agents in the market: consumers (A), collector (B), third parties (C).
A2: Monopoly market for product/service sold by B.
A3: Four types of consumers (2×2) differing in privacy valuation (HIGH/LOW) and value of information to third party (HIGH/LOW).
A4: Asymmetry of information on consumer type between the consumers and the collector, and the consumer and the third parties.
A5: Prohibitively expensive for third parties to target consumers, without prior information on consumer type.
A6: Cost of collecting information negligible.
A7: Risk Neutral agents.
A8: Perfect signal (browsing information) to identify consumer type.
A9: Perfect information on cost and benefit of revealing information.
A10: Consumers have property right over their information. There is sufficient flexibility in contracting and negligible negotiation costs.

Notation: v is the reservation value of all consumers for product/service provided by B, i.e. the maximum price they are willing to pay for the product/

[2] In actual practice, consumer valuation for privacy will be a continuum of values in a range. However for ease of modeling we restrict ourselves to only two values. This type of assumption is standard in screening literature [19;16].

[3] In practice we know that the collector's market may not be monopolistic in nature. However this simplifying assumption is crucial to avoiding analysis of strategic interaction among firms. The study of the monopoly setting is important from a benchmarking perspective and is a first step before studying oligopolistic and competitive markets.

[4] We discuss the issue of better posteriors vs. perfect signals in greater detail in the limitations of the paper.

service. p_j is the price charged for the service for consumer type θ_j, $j = 1,2,3,4$. v_i is the utility of agent i ($i = \mathrm{A,B,C}$) due to agent A subscribing to the service provided by B. Note that only the utility of agents A and B is affected if A subscribes to the service provided by B. w_i is the change in utility of agent i ($i = \mathrm{A,B,C}$) due to release of information and subsequent match between A and C. If A is of value to a third party, then C targets the consumer with a personalized product or service, thus affecting the utility of both A and C. Since B enables this match, it can extract a rent from C. Let S be the gains from trade of personalized product or service enabled due to the matching between a third party and a consumer. This trade would not have been possible if the third party did not have access to the consumer's profile. The surplus from trade of the personalized product or service is distributed between the consumer and the third party. Specifically the consumer gets a surplus w, and the third party gets a surplus u, where $S = u + w$. c_{H} and c_{L} are the cost associated with loss of privacy for consumers who have a high and low valuation for privacy, respectively. By definition $c_{\mathrm{H}} > c_{\mathrm{L}}$. p^{I} is the rent that B extracts from C for the information collected on A. λ_j is B and C's prior that the consumer is of type θ_j. A summary of the notation used appears in Table 1.

As is clear from our setting, our environment is characterized by asymmetric information. Both B and C do not know A's type in period 1 while A knows her type. In period 2 if A allows B to collect information, then B gets to know A's type. B charges a rent from C for the right to know A's type. Before we analyze this asymmetric information setting we look at two benchmark cases – public information and full information. We describe these environments in the following Sections 3.1 and 3.2 and then compare the results with the asymmetric information case in Section 3.3.

3.1. Public information

As before A knows her type. Also, B knows A's type in period 1 (hypothetically). Information is always collected and is always revealed to C. Thus C always gets to know A's type in period 2. B cannot prevent C from knowing A's type, since information is publicly available. The surplus/profits of the agents appear in the table below (see Table 2):

The monopolist firm will extract the entire consumer surplus. Hence the schedule of prices would be:

$$p_1 = v - c_{\mathrm{H}};\quad p_2 = v - c_{\mathrm{H}} + w;$$
$$p_3 = v - c_{\mathrm{L}} + w;\quad p_4 = v - c_{\mathrm{L}}. \tag{2}$$

Table 1
Notation

Notation	Description
i	Type of agent – consumer (A)/ISP (B)/third party (C)
$\theta = (\alpha, \beta)$	Type of consumer
j	Index for each consumer type $j = 1$–4
α	Valuation of privacy – HIGH or LOW
β	Value of consumers information to third party – HIGH or LOW
v	Common reservation utility of all consumers for Internet service
v_i	Change in utility of agent i because of agent A subscribing to the Internet service
w_i	Change in utility of agent i because of release of A's information to C
π_i	Net profit/consumer surplus of agent i
Parameters	
λ_j	Firm B's prior on distribution of consumer type θ_j
N	Total number of consumers
M	Total number of third parties
S	Total gains from trade of a personalized product or service enabled by revelation of information
w	Consumer's surplus from trade of a personalized product or service
u	Third party's surplus from trade of a personalized product or service
c_{H}	Privacy loss of consumer having a high valuation for privacy
c_{L}	Privacy loss of consumer having a low valuation for privacy.
Variables	
P_j	Price charged for Internet service for consumer type θ_j
p^{I}	Price charged by B for information on type of A
k	Consumer choice variable to allow B to collect information or not. $k = 1$, A allows B to collect information; $k = 0$ otherwise

Table 2
Profits and surplus under public information

Consumer type	v_A	v_B	v_C	w_A	w_B	w_C	$\pi_A = v_A + w_A$	$\pi_B = v_B + w_B$	$\pi_C = v_C + w_C$
(H, L)	$v - p_1$	p_1	0	$-c_H$	0	0	$v - p_1 - c_H$	p_1	0
(H, H)	$v - p_2$	p_2	0	$-c_H + w$	0	u	$v - p_2 - c_H + w$	p_2	u
(L, H)	$v - p_3$	p_3	0	$-c_L + w$	0	u	$v - p_3 - c_L + w$	p_3	u
(L, L)	$v - p_4$	p_4	0	$-c_L$	0	0	$v - p_4 - c_L$	p_4	0

Prices charged will have the following relationship:

$$p_3 > p_4; \quad p_3 > p_2; \quad p_2 > p_1; \quad p_4 > p_1. \tag{3}$$

3.2. Full Information

B knows A's type in period 1 (hypothetically) and can commit not to reveal the information to C. In period 1, A can choose whether to allow B to reveal information to C or not. A distinction has been made between the public information case and the full information case, because when comparing with the asymmetric information case, there are two crucial considerations:

- The asymmetry in information on the type of consumer.
- Whether B can prevent C from accessing A's information.

Let k be A's choice variable regarding whether to allow B to collect information, and reveal it to C or not. $k = 1$ corresponds to release of information to C and $k = 0$ corresponds to no release of information. B sets two sets of prices depending on A's type and her choice. Profits and surpluses of the different agents, under the different two choices made by A, appear below (see Tables 3 and 4):

Case 1: $c_L < w < c_H$

B will offer the service to A at a contract that specifies not just the price but also whether information will be collected and released; i.e. the pricing contract is conditioned on the consumer choice variable k. Each contract is thus a tuple (p_j, k). Contract $(p_1 = v, k = 0)$ means that consumer Type 1 would be charged a price v if she decides not to reveal information $(k = 0)$. The contract $(p_1 = v - c_H + \varepsilon, k = 1)$ (see (4)) means that the consumer Type 1 would be charged a price $v - c_H + \varepsilon$, where $\varepsilon > 0$, if she decides to allow B to collect information and reveal it to C. These two contracts together are the menu of contract offered to consumer type 1, and the consumer chooses one contract from the menu. Since B wants A to pick the pricing contract that maximizes her profit, B will provide the following pricing options:

$$\{(p_1^* = v; k^* = 0); (p_1 = v - c_H + \varepsilon, k = 1)\}, \tag{4}$$

$$\left\{p_2^* = \left(\begin{cases} v - c_H + w & \text{if } S > c_H; k^* = 1 \\ v & \text{if } S \leqslant c_H; k^* = 0 \end{cases}\right);\right.$$
$$\left. p_2 = \left(\begin{cases} v + \varepsilon & \text{if } S > c_H; k = 0 \\ v - c_H + w + \varepsilon & \text{if } S \leqslant c_H; k = 1 \end{cases}\right)\right\}, \tag{5}$$

$$\{(p_3^* = v - c_L + w; k^* = 1); (p_3 = v + \varepsilon, k = 0)\}, \tag{6}$$

$$\{(p_4^* = v, k^* = 0); (p_4 = v - c_L + \varepsilon, k = 1)\}. \tag{7}$$

In all cases $\varepsilon > 0$ ensures that the A will pick the option (*) that maximizes B's profit. The pricing contract with (*) is what is termed as the first best contract or full information contract, i.e. the contract that B would offer each consumer type, if she knew her type, and if she could prevent C from accessing A's information. Notice that although B knows A's type, she cannot force A to reveal information or otherwise. What B can do, is to distort the non-profit maximizing contract, so that each consumer type is better off picking the profit maxi-

Table 3
Profits/surplus of agents under full information if A chose to reveal information

Consumer type	v_A	v_B	v_C	w_A	w_B	w_C	$\pi_A = v_A + w_A$	$\pi_B = v_B + w_B$	$\pi_C = v_C + w_C$
(H, L)	$v - p_1$	p_1	0	$-c_H$	0	0	$v - p_1 - c_H$	p_1	0
(H, H)	$v - p_2$	p_2	0	$-c_H + w$	p^I	u	$v - p_2 - c_H + w$	p_2	u
(L, H)	$v - p_3$	p_3	0	$-c_L + w$	p^I	u	$v - p_3 - c_L + w$	p_3	u
(L, L)	$v - p_4$	p_4	0	$-c_L$	0	0	$v - p_4 - c_L$	p_4	0

Table 4
Profits of agents under full information if A chooses not to reveal information

Consumer type	v_A	v_B	v_C	w_A	w_B	w_C	$\pi_A = v_A + w_A$	$\pi_B = v_B + w_B$	$\pi_C = v_C + w_C$
$\theta_1 = (H, L)$	$v - p_1$	p_1	0	0	0	0	$v - p_1$	p_1	0
$\theta_2 = (H, H)$	$v - p_2$	p_2	0	0	0	0	$v - p_2$	p_2	0
$\theta_3 = (L, H)$	$v - p_3$	p_3	0	0	0	0	$v - p_3$	p_3	0
$\theta_4 = (L, L)$	$v - p_4$	p_4	0	0	0	0	$v - p_4$	p_4	0

mizing contract. *Case 2:* $c_L < c_H < w$ (see Appendix A.1). *Case 3:* $w < c_L < c_H$ (see Appendix A.2).

3.3. Asymmetric information

B and C do not know A's type. A knows her valuation for privacy, and also knows that her information is going to be useful to one of the third parties or in other words the personalized service provided by one of the C's will be useful to her. An example could be if the consumer listens to Jazz music, then she knows that information on her tastes, would be of use to a third party that sells Jazz CD's. B provides a set of contracts (for the service she provides) – one where information is not collected and one where information is collected and potentially revealed. Note that in the asymmetric information case, B cannot condition the contracts on the type of the consumer. Let p_P be the price of the contract in which information is kept private, and p_R be the price of the contract where information is revealed. Each of these contracts could be offered alone or in combination with the other contract, and A chooses from this set of contracts. Thus contracts that could be offered by B are:

$$\{(p_P, k = 0\} \quad \text{or} \quad \{p_R, k = 1\} \quad \text{or} \quad \{(p_P, k = 0); (p_R, k = 1\}. \tag{8}$$

The first two are pooling contracts, i.e. contracts in which all agents are offered the same price. The intuition for these types of contracts is that the firms could have a blanket policy regarding collection of information and release: they could either follow a policy of never collecting information or they could have a policy of always collecting and releasing information. In the first contract, B provides service to only those consumers who do not want B to collect information and then reveal it to C. The second contract specifies that B only provides service to consumers who are willing to allow B to collect information and then reveal it to C. In the third menu of contracts, information is collected and revealed to B based on A's choice. This menu of contracts is a partial pooling contract. In a partial pooling contract two or more consumer types choose the same contract.

Pooling–non-revealing: If B has a policy of never collecting and releasing information – $\{p_P, k = 0\}$. B will charge a price $p_P = v$ and extract the entire consumer surplus of all types.

Pooling–revealing: If B has a policy of always collecting and releasing information – $\{p_R, k = 1\}$. B has a number of pricing options here. Consider the case $c_L < w < c_H$.[5] Also assume that $c_H - c_L < w$. Thus we have

$$v - c_H < v - c_L < v - c_H + w < v - c_L + w. \tag{9}$$

In the price range $[0, v - c_H]$ all consumer types subscribe to the service. However the firm would leave the least on the table if it charges the price $p_R = v - c_H$. In the price range $(v - c_H, v - c_L]$B can sell to all consumer types except type θ_1. The firm is better off picking the price $p_R = v - c_L$ than any other price in this range. In the price range $(v - c_L, v - c_H + w]$B sells to types θ_2 and θ_3 and is better off charging the price $p_R = v - c_H + w$. Similarly in the price range $(v - c_H + w, v - c_L + w]$B only sells to types θ_3 and is better off charging the price $p_R = v - c_L + w$. B cannot charge a price greater than $v - c_L + w$ since no consumer type would buy the service at that price.

Partial pooling strategy: The only two partial pooling strategies (menu of contracts) that B needs to consider are PP1 – $\{(p_P = v, k = 0); (p_R = v - c_H + w, k = 1)\}$ and PP2 – $\{p_P = v; p_R = v - c_L + w\}$. All other Partial pooling strategies are strictly dominated by PP1 and/or PP2. If B chooses partial pooling strategy PP1, then types θ_1 and θ_4 choose not to reveal information and types θ_2 and θ_3 reveal information by choosing $(p_R = v - c_H + w, k = 1)$. Type θ_3 ends up with a surplus of $c_H - c_L$. If B chooses the other partial pooling strategy PP2, then types θ_1, θ_2 and θ_4 choose the contract $(p_P = v, k = 0)$. Type θ_3 chooses

[5] Results do not change for other relative magnitudes, i.e. Cases 2 and 3.

the contract ($p_R = v - c_L + w, k = 1$). All consumer types get zero surplus.

B's profits under the different pooling and partial pooling strategies appear below:

Proposition 1. *It is never optimal for the firm to have a pure* revealing *pooling pricing strategy.*

Proof. For the case $c_L < w < c_H$ Partial pooling strategy PP1 (Strategy 6) dominates strategies 1, 2, and 3 (comparing profits). Partial pooling strategy PP2 (Strategy 7) dominates strategy 4. The same is true for $c_L < c_H < w$ and $w < c_L < c_H$. □

From Proposition 1 we can now concentrate on the two Partial Pooling Strategies PP1 (Strategy 6) and PP2 (Strategy 7) and the pooling non-revealing strategy P3 (Strategy 5) for all three cases. The following result is useful for calculating the optimal strategy of the firm.

Proposition 2. *PP1 dominates PP2 if* $S > c_L[r_{HL} + r_{32}(r_{HL} - 1)]$, *where* $r_{32} = \frac{\lambda_3}{\lambda_2}$ *and* $r_{HL} = \frac{c_H}{c_L}$, *else PP2 dominates PP1.*

Proof. Let $\Delta\Pi$ to be the difference in profits between the PP1 (Strategy 6 Table 5) and the PP2 (Strategy 7 Table 5).

$$\Delta\Pi = N[\lambda_2(S - c_H) - \lambda_3(c_H - c_L)].$$

Here S is the total gain from trade between a third party and a consumer. If $\Delta\Pi > 0$, then B prefers PP1 to PP2. The reverse is true if $\Delta\Pi < 0$, and the firm is indifferent between the two strategies if $\Delta\Pi = 0$. B prefers PP1 to PP2 if

$$S > c_H + \frac{\lambda_3}{\lambda_2}(c_H - c_L). \tag{10}$$

Eq. (10) can be written as

$$S > c_L[r_{HL} + r_{32}(r_{HL} - 1)], \tag{11}$$

where $r_{HL} = \frac{c_H}{c_L} > 1$ and $r_{32} = \frac{\lambda_3}{\lambda_2}$. r_{HL} is the ratio of privacy loss of HIGH privacy valuation type to privacy loss of LOW privacy valuation type. r_{32} is the ratio of the proportion of type θ_3 to the proportion of type θ_2. This basically says that if the gain from trade (of a personalized product or service) is greater than the expression on the right of (11) then choosing PP1 is optimal for B. A plot of the expression on the right side of (11) for $1 < r_{HL} < \infty$, $0 < r_{32} < \infty$ and for a given value of c_L appears in Fig. 1. For a gain from trade S which lies above the plane shown in Fig. 1, strategy PP1 dominates PP2. Conversely for any S that lies below the plane, PP2 dominates PP1. □

It is clear from Fig. 1 that when either of the two ratios r_{32} or r_{HL} is small, then even a small surplus would lie above the surface of the plane in Fig. 1 and hence PP1 will dominate PP2. For higher values of r_{32} and r_{HL}, the gains form trade S would have to be very high for PP1 to dominate. For moderate and low values of S, PP2 will dominate PP1.

Case 1: $c_L < w < c_H$

Proposition 3. *Optimal strategy for the firm for the case* $c_L < w < c_H$ *is PP2, if* $S \leqslant c_H$. *If* $S > c_H$, *then the optimal strategy is PP1, if* $S > c_L[r_{HL} + r_{32}(r_{HL} - 1)]$, *and PP2 otherwise.*

Proof. When the gains from trade of personalized product or service (S) is less than the privacy loss of the HIGH privacy valuation consumer, i.e. $S \leqslant c_H$, then the first-best (full information) contracts are ($p_P = v, k = 0$) for types θ_1, θ_2 and θ_4 and ($p_R = v - c_L + w, k = 1$) for type θ_3 (PP2). In the second best case (asymmetric information on types), with $S \leqslant c_H$, and for any prior on the distribution of types, there is no incentive for any agent to deviate from the first best contract (PP2). To see this, the Individual rationality and Incentive compatibility constraints of the consumer types appear below:

Table 5
Profits under the different pooling/partial-pooling pricing strategies

Strategy	Prices	Firm's (B) surplus/profits
1	$p_R = v - c_H$	$N[(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)(v - c_H) + (\lambda_2 + \lambda_3)u]$
2	$p_R = v - c_L$	$N[(\lambda_2 + \lambda_3 + \lambda_4)(v - c_L) + (\lambda_2 + \lambda_3)u]$
3	$p_R = v - c_H + w$	$N[(\lambda_2 + \lambda_3)(v - c_H + w) + (\lambda_2 + \lambda_3)u]$
4	$p_R = v - c_L + w$	$N[\lambda_3(v - c_L + w) + \lambda_3 u]$
5	$p_P = v$	Nv
6	$\{p_P = v; p_R = v - c_G + w\}$	$N[(\lambda_1 + \lambda_4)v + (\lambda_2 + \lambda_3)(v - c_H + w) + (\lambda_2 + \lambda_3)u]$
7	$\{p_P = v; p_R = v - c_L + w\}$	$N[(\lambda_1 + \lambda_2 + \lambda_4)v + \lambda_3(v - c_L + w) + \lambda_3 u]$

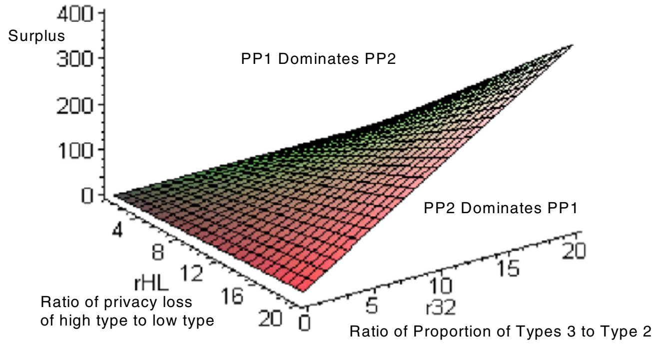

Fig. 1. Regions where the two partial pooling strategies dominate.

$$v - p_{\mathrm{P}} \geqslant 0, \tag{12}$$

$$v - p_{\mathrm{R}} - c_{\mathrm{L}} + w \geqslant 0, \tag{13}$$

$$v - p_{\mathrm{R}} - c_{\mathrm{L}} + w \geqslant v - p_{\mathrm{P}}, \tag{14}$$

$$v - p_{\mathrm{P}} \geqslant v - p_{\mathrm{R}} - c_{\mathrm{H}}, \tag{15}$$

$$v - p_{\mathrm{P}} \geqslant v - p_{\mathrm{R}} - c_{\mathrm{H}} + w, \tag{16}$$

$$v - p_{\mathrm{P}} \geqslant v - p_{\mathrm{R}} - c_{\mathrm{L}}. \tag{17}$$

The reader can verify that the first best contracts satisfy the above constraints. When the total gains from trade of a personalized product or service is greater that the privacy loss of the HIGH privacy valuation consumer, i.e.$S > c_{\mathrm{H}}$ then in the second best case, the types have an incentive to deviate if charged the first best contracts which are $(p = v, k = 0)$ for types θ_1 and θ_4, $(p = v - c_{\mathrm{H}} + w, k = 1)$ for type θ_2 and $(p = v - c_{\mathrm{L}} + w, k = 1)$ for type θ_3. Specifically type θ_3 will pick the contract meant for type θ_2. This problem comes up because the pricing cannot be made conditional on the binary decision variable k alone.[6] The best available option then is to allow type θ_3 to choose the contract $(p = v - c_{\mathrm{H}} + w, k = 1)$ and leave type θ_3 some surplus (PP1) or to force type θ_2 to pick the contract $(p = v, k = 0)$ by offering this contract in addition to the contract $(p = v - c_{\mathrm{L}} + w, k = 1)$ (PP2). Note that with $S > c_{\mathrm{H}}$, PP2 (Strategy 7) dominates pure pooling non-revealing strategy P3 (Strategy 5). The conditions where PP1 and PP2 are optimal come from Proposition 2. □

Case 2: $c_{\mathrm{L}} < c_{\mathrm{H}} < w$

Proposition 4. *Optimal strategy for the firm for the case* $c_{\mathrm{L}} < c_{\mathrm{H}} < w$ *is PP1, if* $S > c_{\mathrm{L}}[r_{\mathrm{HL}} + r_{32}(r_{\mathrm{HL}} - 1)]$, *and PP2 otherwise.*

Proof. The first best (Full Information) contracts (see Appendix) are to offer types θ_1 and θ_4 the contract $(p_{\mathrm{P}} = v, k = 0)$ and to offer types θ_2 and θ_3 the contracts $(p_{\mathrm{R}} = v - c_{\mathrm{H}} + w, k = 1)$ and $(p_{\mathrm{R}} = v - c_{\mathrm{L}} + w, k = 1)$, respectively. However under asymmetric information type θ_3 would always pick the contract designed for type θ_2. Thus in Case 2 the first best prices cannot partially separate the types as under certain conditions in Case 1. The only option for B then is to offer one of the partial pooling contracts PP1 or PP2. Note that both PP1 and PP2 dominate the pure pooling non-revealing strategy (Strategy 5). The conditions where PP1 and PP2 are optimal come from Proposition 2. □

Case 3: $w < c_{\mathrm{L}} < c_{\mathrm{H}}$

Proposition 5. *Optimal strategy for the firm for the case* $w < c_{\mathrm{L}} < c_{\mathrm{H}}$ *is PP2, if* $c_{\mathrm{L}} < S \leqslant c_{\mathrm{H}}$, *and P3 if* $S > c_{\mathrm{L}}$. *If* $S > c_{\mathrm{H}}$, *then the optimal strategy is PP1, if* $S > c_{\mathrm{L}}[r_{\mathrm{HL}} + r_{32}(r_{\mathrm{HL}} - 1)]$, *and PP2 otherwise.*

Proof. When $c_{\mathrm{L}} < S \leqslant c_{\mathrm{H}}$, then the first best contract (see Appendix) is to offers types θ_1, θ_2 and θ_4 the contract $(p_{\mathrm{P}} = v, k = 0)$ and θ_3 the contract $(p_{\mathrm{R}} = v - c_{\mathrm{L}} + w, k = 1)$. None of the consumer types has an incentive to deviate (all incentive compatibility constraints are satisfied) even when offered the first best contracts. When $S \leqslant c_{\mathrm{L}}$ then the first best contract is to offer types θ_1, θ_2, θ_3 and θ_4 the contract P3 $(p_{\mathrm{P}} = v, k = 0)$. Thus when the total gain from trade is less than the privacy loss of the HIGH privacy valuation consumer, then the optimal strategy for the firm is to offer the first best contract. When the total gains from trade of

[6] In Section 4 we show how this problem can be solved by considering the decision variable of the consumer k to be continuous.

Table 6
Optimal Strategy for firm under different conditions

	$S \leqslant c_H$	$S > c_H$
Case 1 $c_L < w < c_H$	PP2[a]	PP1 if $S > c_L[r_{HL} + r_{32}(r_{HL} - 1)]$ PP2 if $S \leqslant c_L[r_{HL} + r_{32}(r_{HL} - 1)]$
Case 2 $c_L < c_H < w$	NA	PP1 if $S > c_L[r_{HL} + r_{32}(r_{HL} - 1)]$ PP2 if $S \leqslant c_L[r_{HL} + r_{32}(r_{HL} - 1)]$
Case 3 $w < c_L < c_H$	PP2[a] if $S > c_L$ P3[a] if $X \leqslant c_L$	PP1 if $S > c_L[r_{HL} + r_{32}(r_{HL} - 1)]$ PP2 if $S \leqslant c_L[r_{HL} + r_{32}(r_{HL} - 1)]$

Partial Pooling PP1: ($p_P = v, k = 0$); ($p_R = v - c_H + w, k = 1$).
Partial Pooling PP2: ($p_P = v, k = 0$); ($p_R = v - c_L + w, k = 1$).
P3: ($p_P = v; k = 0$) – Pooling private.
[a] First Best.

a personalized product or service is greater that the privacy loss of the HIGH privacy valuation consumer, i.e. $S > c_H$ then in the second best case, the types have an incentive to deviate if charged the first best contracts which are ($p = v, k = 0$) for types θ_1 and θ_4, ($p = v - c_H + w, k = 1$) for type θ_2 and ($p = v - c_L + w, k = 1$) for type θ_3. Specifically type θ_3 will pick the contract meant for type θ_2. Thus in Case 3 with $S > c_H$, the first best prices cannot partially separate the types. The only option for B then is to offer one of the partial pooling contracts PP1 or PP2. Note that with $S > c_H$, PP2 dominates P3 (strategy 5). The conditions where PP1 and PP2 are optimal come from Proposition 2. □

The results of Propositions 3–5, which specify the optimal strategy for the firm B are summarized in Table 6. As is clear from the table, when the gains from trade is greater than the privacy loss of the HIGH privacy valuation consumer, i.e. when $S > c_H$, then the choice of pricing strategy is independent of the relative magnitudes of the privacy loss of the HIGH and LOW privacy valuation types and the value to the consumer due to a personalized product or service, i.e. Cases 1, 2 or 3. The choice between the two partial pooling strategies only depends on whether the gains from trade of a personalized product or services are greater than or less than $c_L[r_{HL} + r_{32}(r_{HL} - 1)]$. When $S \leqslant c_H$, the results are very different. For Case 1, $c_L < w < c_H$, the optimal strategy is the partial pooling strategy PP2, which is also the first-best pricing strategy (i.e. the firm is no worse than in the first best full information case). Case 2, $c_L < c_H < w$ with $S \leqslant c_H$ is not a valid case since the two inequalities cannot be true at the same time. For Case 3, $w < c_L < c_H$, the optimal strategy is PP2 if $S > c_L$ and P3 (to never collect information and charge the uniform price $p_P = v$) if $S \leqslant c_L$. Both these strategies are also the first best pricing strategies.

4. Varying degrees of privacy

Our analysis till now assumes that the privacy losses are binary in nature. Consumers if they reveal information, face a known privacy loss, and if they choose not to reveal information, they face a loss of zero. Thus the losses are assumed to be two extreme values and there are no privacy losses in between these two values. In this section we relax that assumption, and instead look at a continuum of privacy losses. A can now choose the number of third parties to whom information is released. Releasing information to M third parties corresponds to maximum privacy loss and not releasing information to any third party, corresponds to the zero privacy loss. Any value in between, specifies an intermediate valuation for privacy. The consumer's utility from personalized service due to revealing information to n different C's is given by nw.[7] c_H and c_L continue to be the marginal cost of privacy of types $\theta_2 \equiv \theta_H$ and $\theta_3 \equiv \theta_L$, respectively. The marginal cost of releasing information is more for high privacy consumer θ_H than for type θ_L ($c_H > c_L$). Let p_H and p_L be the price charged by B for the service provided. The utility function of each type due to revealing information to n different third parties is given by

$$v + n_H(w - c_H) - p_H, \tag{18}$$

$$v + n_L(w - c_L) - p_L. \tag{19}$$

$w - c_H$ and $w - c_L$ can be interpreted as the net marginal benefit to revealing information for the high and low types: $w'_L \equiv w - c_L > w - c_H \equiv w'_H$. The firm B in order to screen the consumer types offers two contracts (n_H,p_H); (n_L,p_L). The first contract is for type θ_H and the second contract for type θ_L. The contracts for this case under both full and asymmetric information are shown in Fig. 2.

The vertical axis is the price charged and the horizontal axis is the number of third parties to whom the information is revealed. The slopes of the indifference curves are w'_H and w'_L, respectively for the high and low types – $w'_L > w'_H$. Firm B under full

[7] Implicit in this assumption is that types θ_2 and θ_3 receive the same utility from personalized services and that the utility is increasing linearly in n, the number of third parties to whom information is revealed.

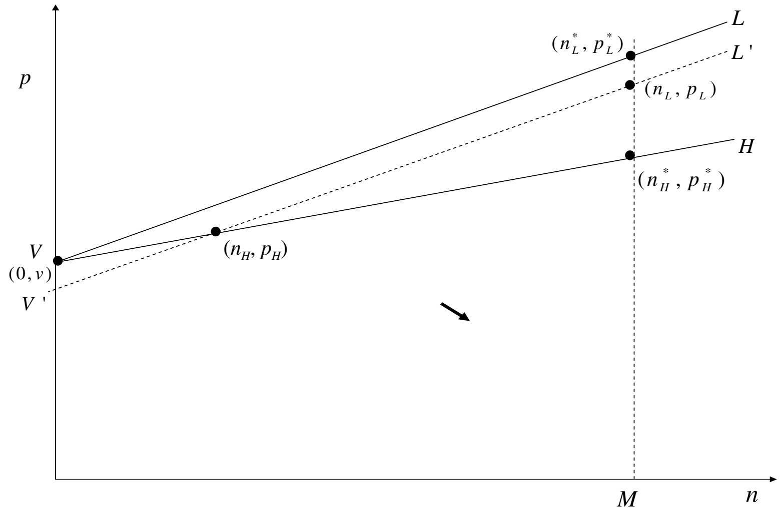

Fig. 2. Indifference curves and contracts under full and asymmetric information when $c_L < c_H < w$.

information, would charge prices along the indifference curves – lines VH and VL, respectively. These indifference curves both pass through the point $(0,v)$, which corresponds to zero privacy loss. Both consumer types are better off in the direction of the arrow-lower prices and more net benefits from personalized services. The firm is best off, if it offers the first best contracts (n_L^*, p_L^*) and (n_H^*, p_H^*) for the low and high type consumers, respectively, where $n_H^* = n_L^* = M$. In the second best case, i.e. with asymmetric information, the low privacy type prefers the contract (n_H^*, p_H^*) to (n_L^*, p_L^*) and has an incentive to lie about her type. So under asymmetric the contracts should be restricted and must satisfy the individual rationality and incentive compatibility constraints. The only contracts that achieve separation are ones where the individual rationality constraint of the high privacy type and the incentive compatibility constraint of the low privacy type bind. The firm B provides the low privacy type a subsidy. The low privacy type is thus better off in the asymmetric information case and enjoys a positive surplus. The new indifference curve is the line $V'L'$. The contract offered to the low privacy type is (n_L, p_L). The contract for the high type is (n_H, p_H) and lies at the point of intersection of the new indifference curve of the low type $V'L'$ and the indifference curve of the high type. If the subsidy to the low type is increased, then the firm can extract rent from more third parties for the information of the high type consumer. However the tradeoff is that the firm has to leave more surplus for the low type consumer. Thus the monopolist needs to choose the optimal amount of subsidy to give to the low privacy type, in order to maximize her total profits. Given that firm can design contracts that allow consumers to choose intermediate values of privacy, we get the following result.

Proposition 6. *Given that firms can design contract where consumer can choose from a continuum of privacy level, it is always optimal for the firm to offer contracts where information is revealed to either all third parties, or to none at all.*

Proof. Denoting the subsidy provided to the low privacy type by s, the contract offered to the low privacy type is $(n_L, p_L) \equiv (M, v + Mw'_L - s)$.[8] The contract offered to the high privacy type if the low privacy type is offered a surplus of s is: $(n_H, p_H) \equiv \left(\frac{s}{c_H - c_L}, v + s\frac{w'_H}{w'_L - w'_H}\right)$.[9] The total profit function for the monopolist is:

$$\Pi = \lambda_2[p_H + n_H u] + \lambda_3[p_L + n_L u]. \quad \square \tag{20}$$

Substituting the values of p_H, n_H, p_L and n_L into the above expression, and simplifying we get

[8] Note that the contract offered to the low privacy type in the first-best case is $(n_L^*, p_L^*) \equiv (M, v + Mw'_L)$.

[9] This contract is obtained by finding the point of intersection of the indifference curve of the high privacy type and the indifference curve of the low privacy type who is given a subsidy of s.

$$\begin{aligned}\Pi &= \lambda_2\left[v + s\left(\frac{S-c_{\rm H}}{c_{\rm H}-c_{\rm L}}\right)\right] + \lambda_3[v + M(S-c_{\rm L}) - s] \\ &= C + s\left[\lambda_2\left(\frac{S-c_{\rm H}}{c_{\rm H}-c_{\rm L}}\right) - \lambda_3\right],\end{aligned} \tag{21}$$

where C is a constant. To see how the profits of the monopolist varies with the subsidy, note that

$$\frac{\partial\Pi}{\partial s} = \left[\lambda_2\left(\frac{S-c_{\rm H}}{c_{\rm H}-c_{\rm L}}\right) - \lambda_3\right] = \left[\frac{(S/c_{\rm L})-1}{r_{\rm HL}-1}\right] - r_{32}. \tag{22}$$

The monopolist should increase the subsidy provided to the low privacy type if $\frac{\partial\Pi}{\partial s} > 0$, i.e. when $S > c_{\rm L}[r_{\rm HL} + r_{32}(r_{\rm HL}-1)]$. When $S \leqslant c_{\rm L}[r_{\rm HL} + r_{32}(r_{\rm HL}-1)]$, the monopolist should offer the minimum possible subsidy to the low privacy type. The optimal strategy for the firm is

$$\begin{aligned}&\{(n_{\rm H}, p_{\rm H}) \equiv (0, v); \{(n^*_{\rm L}, p^*_{\rm L}) \equiv (M, v + Mw'_{\rm L})\} \\ &\quad \text{if } S \leqslant c_{\rm L}[r_{\rm HL} + r_{32}(r_{\rm HL}-1)], \\ &\{(n^*_{\rm H}, p^*_{\rm H}) = (n_{\rm L}, p_{\rm L}) \equiv (M, v + Mw'_{\rm H})\} \\ &\quad \text{if } S > c_{\rm L}[r_{\rm HL} + r_{32}(r_{\rm HL}-1)].\end{aligned} \tag{23}$$

Thus if $S \leqslant c_{\rm L}[r_{\rm HL} + r_{32}(r_{\rm HL}-1)]$, the high privacy type choose the contract in which information is neither collected nor revealed and the low privacy type choose the first best contract. For the case $S > c_{\rm L}[r_{\rm HL} + r_{32}(r_{\rm HL}-1)]$, the low privacy type gets a subsidy of $M[c_{\rm H} - c_{\rm L}]$. This analysis shows that it is non-optimal for the firms to provides contracts where the consumers can choose an intermediate level of privacy. The indifference curves and contracts under full and asymmetric information for $c_{\rm L} < w < c_{\rm H}$ and $w < c_{\rm L} < c_{\rm H}$ can be drawn similar to Fig. 2.[10] Analysis for both these case are similar to Case 1 and is left out of the paper for brevity.

5. Discussion

The results obtained in the previous section and the managerial implications from them are summarized in this section. We also discuss the limitations and directions for future research. When the total gains from trade of a personalized product or service is less than the privacy loss of the low privacy type, results suggest that it is optimal for the firm not to collect information about consumers. With the recent articles on privacy issues on the Internet, in trade journals, and in the general media [2,3,21], there could be an increasing tendency for consumer to over value their loss of privacy, if they reveal information to firms. Because of increased uncertainty regarding how the collected information would be used, the benefits from personalized services are valued much smaller, in relative terms, than the privacy costs. Under these circumstances, when the total gains from the trade of a personalized product or service is less than the privacy loss of the low privacy type, collection of information by ISP's is non-optimal. A quick survey of the Privacy policies of the major ISP's revealed that, they explicitly state that they do not collect information on consumers. This is in spite of the fact that there are no specific laws, as of now, that prevent the ISP's from legally collecting information.

In most cases except the one mentioned above, partial pooling contracts (revealing and non-revealing) dominate the pure pooling contracts, i.e. the firm is better off offering a menu of contracts: one contract where information is collected at a certain price, and the other contract were information is not collected, at a different price. Evidence of this form of contracts exists in grocery stores where the consumers receive discounts on goods purchased if a grocery store card is used. The contract at the time of signing up for the grocery card explicitly states that information will be collected at the point of sale, on goods purchased using the grocery card. However this line is often buried in a footnote, and often consumers are not aware of this clause. Others who are aware of this clause, possibly do not overvalue their privacy losses, since there has been much less coverage of privacy issues in the grocery store setting as opposed to the Internet setting.[11] For this reason, in the grocery store example, the magnitude of the value from personalized service relative would be either greater than the privacy losses of both high privacy and low privacy types, or at least greater than the privacy loss of low privacy type. In both cases partial pooling contracts are optimal as predicted by our model, and as seen in practice in the pricing strategies of these stores.

Even if the ISP (or any other firm that is capable of collecting information) has an option to design contracts that allow consumers to choose an intermediate level of privacy, i.e. choose the number of third parties to whom information can be revealed,

[10] The figures are available upon request from the authors.

[11] The grocery stores are in some ways benign compared to the Internet, where firms are always tracking behavior, and privacy is a big concern.

it would always be optimal for the firm to choose contracts where privacy losses are extreme. In the menu of contracts offered, one contract reveals information to all third parties, while the other contract does not reveal information to any third party at all. The first contract is meant for the low privacy type while the second contract is meant for the high privacy type. This could be one possible reason why there is no evidence of contracts in practice, where the consumers can choose the number of third parties to whom information is revealed.[12] In some cases under asymmetric information, the collector of information is able to charge the first best prices and there is no incentive for any agent type to deviate. Thus under these cases the firm is no worse off because of the asymmetry of information, and can extract the entire consumer surplus. These cases are marked by a (*) in Table 6. This typically happens when the total gains from trade are less than the privacy loss of the high privacy type.

This paper does a formal analysis of the privacy construct and contrasts it with the value of personal information to firms. The model formalizes the incentives of the different agents, and how they are related to each other. To keep things tractable, we made a number of simplifying assumptions, and the managerial implications of the model must be seen in the light of the assumptions made:

We consider a monopoly setting for the product or service sold by the collector of information. It is evident that in the real world, the market for the collector's product or service is not monopolistic, and it would have been more realistic to consider an oligopolistic setting or a competitive one. However solving the monopoly case is a first step, which helps to formalize the interactions between the various agents, and is important from a benchmarking perspective. The next step would be to consider a competitive or an oligopolstic setting, and compare the results obtained here, with those.

We assume that the collector of information enables the matching between the consumers and the third parties, and extracts a rent from the third party, for the matching service provided. Implicit in the assumption is that collector has perfect knowledge on the information requirements of the third parties, i.e. it knows which consumer profiles would be of interest to which third parties. If we assume that collector has no prior information on the possibility of a match between a third party and a consumer, then the other option would be for the collector to charge the same price for all consumer profiles irrespective of whether there is a match or not. This modification would not change the nature of results obtained here. The same total rent that it extracts from the third parties for the profiles of consumer who have high information value would now be obtained for the profiles of all consumer types. The decision problem of the consumers remains unaffected.

All agents are assumed to risk neutral. This is not a standard assumption, as in screening problems such as labor markets [15] and insurance markets [17], where the informed agents are assumed to be risk averse. There are several reasons behind our choice of risk neutral agents. First, in analytical terms the analysis becomes much simpler since we have linear utility functions, as opposed to a more general functional form. Second, the choice of a concave utility function for the consumers (risk averse) does not really affect the nature of the results; it only changes the shape of the indifference curves. This of course was not known before hand, but since the intuition behind the results does not change with the choice of functional form, there does not seem to be much value in pursuing this alternate modeling approach, as it would just be a mathematical exercise.

We assume that the interests of the consumers change every period. This is equivalent to saying that a consumers who is of interest to a firm in the current period, may not be of interest in the next period. For example a consumer who has changing tastes, and who listens to jazz music in the current period, may not to be a jazz fan in the next period. A consumer who has just taken a home loan, would fall out of the set of consumers who are potential customers of home loan lenders in the next period. An alternate way of modeling this would be to consider that the consumers have a base set of interests that never change, and a random component that changes every period. Once a consumer has revealed their profile to the third party, since the base set of interests never changes, the third parties are only interested in the random component in a future period. The analysis for this case would be similar to

[12] Financial institutions such as banks and insurance companies often share customer information with their business affiliates, and customers are aware of this arrangement since they agree to it when they decide to transact with these organizations. However, sharing information with affiliates is equivalent to keeping information within a closed network of affiliates with similar privacy policies.

the one considered here, where we assume that tastes could change every period, and there is no constant component. Another interesting line of future research could be where consumers have different valuations of privacy based on the type of information collected.

For the mechanism designed here to be feasible, all agents should follow the actions specified by the contract. It is possible that the collector might release information to third parties even though the contract specifies otherwise. To tackle such problems we propose that Trusted Third Parties (TTP) should audit firm practices, with a severe penalty for non-conformance. This should again take care of the problem of secondary use of personal information that has not been addressed by the paper.

Acknowledgement

This research was partially funded by National Science Foundation Grant Nos. EIA-0075506 and DMI-0122214.

Appendix A

A.1. Full information

Case 2: $c_{\mathrm{L}} < c_{\mathrm{H}} < w$

$$\{(p_1^* = v; k^* = 0); (p_1 = v - c_{\mathrm{H}} + \varepsilon, k = 1),$$
$$\{(p_2^* = v - c_{\mathrm{H}} + w, k = 1); (p_2 = v + \varepsilon, k = 0)\},$$
$$\{(p_3^* = v - c_{\mathrm{L}} + w, k^* = 1); (p_3 = v + \varepsilon, k = 0)\},$$
$$\{(p_4^* = v, k^* = 0); (p_4 = v - c_{\mathrm{L}} + \varepsilon, k = 1)\}.$$

A.2. Full information

Case 3: $w < c_{\mathrm{L}} < c_{\mathrm{H}}$

$$\{(p_1^* = v; k^* = 0); (p_1 = v - c_{\mathrm{H}} + \varepsilon, k = 1),$$
$$\left\{ p_2^* = \left(\begin{cases} v - c_{\mathrm{H}} + w & \text{if } S > c_{\mathrm{H}}; k^* = 1 \\ v & \text{if } S \leqslant c_{\mathrm{H}}; k^* = 0 \end{cases} \right); \right.$$
$$\left. p_2 = \left(\begin{cases} v + \varepsilon & \text{if } S > c_{\mathrm{H}}; k = 0 \\ v - c_{\mathrm{H}} + w + \varepsilon & \text{if } S \leqslant c_{\mathrm{H}}; k = 1 \end{cases} \right) \right\},$$
$$\left\{ p_3^* = \left(\begin{cases} v - c_{\mathrm{L}} + w & \text{if } S > c_{\mathrm{L}}; k^* = 1 \\ v & \text{if } S \leqslant c_{\mathrm{L}}; k^* = 0 \end{cases} \right); \right.$$
$$\left. p_3 = \left(\begin{cases} v + \varepsilon & \text{if } S > c_{\mathrm{L}}; k = 0 \\ v - c_{\mathrm{L}} + w + \varepsilon & \text{if } S \leqslant c_{\mathrm{L}}; k = 1 \end{cases} \right) \right\},$$
$$\{(p_4^* = v, k^* = 0); (p_4 = v - c_{\mathrm{L}} + \varepsilon, k = 1)\}.$$

References

[1] A. Acquisti, H.R. Varian, Conditioning prices on purchase history, Forthcoming Marketing Science 24 (3) (2005) 367–381.
[2] Business Week. A little net privacy, please. Business Week. (Mar 16, 1998). Available from: www.businessweek.com.
[3] Business Week. How Grocery Stores are Feeding Fears. Business Week, June 2002. Available from: http://www.businessweek.com/bwdaily/dnflash/jun2002/nf20020620_7007. html.
[4] M.J. Culnan, How did they get my name: An exploratory investigation of consumer attitudes toward secondary information use, MIS Quarterly (1993) 341–361.
[5] R.H. Coase, The problem of social cost, Journal of Law and Economics 3 (1960) 1–44.
[6] FTC, 2000. Privacy Online: Fair Information Practices in the Electronic Marketplace. A Report to Congress. Federal Trade Commission.
[7] Hann, I.H., Hui, K.L., Lee, T.S., Png, I.P.L., 2002. Online information privacy: Measuring the cost-beneifit trade-off. In: Proceedings of International Conference on Information Systems (ICIS).
[8] Kahn, C.M., McAndrews, J., Roberds, W., 2000. A theory of transactions privacy. Federal Reserve Bank of Atlanta, Working Paper.
[9] J. Kang, Information privacy in cyberspace transactions, Stanford Law Review 50 (1998) 192–1294.
[10] K.C. Laudon, Markets and privacy, Communications of the ACM 39 (9) (1996) 92–104.
[11] Lessig, L. Code and other laws of cyberspace. Basic Books, June 2000.
[12] R. Quick, On-line groups are offering up privacy plans, Wall Street Journal 22 (1998) B1.
[13] Smith, H.J., Milberg, S.J., Burke, S.J. Information privacy: Measuring individuals' concerns about organizational practices. MIS Quarterly 20(2), 167–196.
[14] J. Sovern, Opting in, opting out, or no options at all: The fight for control of personal information, Washington Law Review 74 (1999) 1033–1118.
[15] M. Spence, Job market signaling, Quarterly Journal of Economics 87 (1973) 355–379.
[16] K.A. Stewart, A.H. Segars, An empirical examination of the concern for information privacy instrument, Information Systems Research 3 (1) (2002) 36–49.
[17] J. Stiglitz, Monopoly, non-linear pricing and imperfect information: The insurance market, The Review of Economic Studies 44 (3) (1977) 407–430.
[18] P. Thomas, Clicking coupons on-line has a cost: Privacy, Wall Street Journal 18 (1998) B1–B8.
[19] Tam, E.C., Hui, K.L., Tan, B.C., 2002. What do they want? Motivating consumers to disclose personal information to internet businesses. In: Proceedings of International Conference on Information Systems (ICIS).
[20] Taylor, C.R., 2003. Private demands and demands for privacy: Dynamic pricing and the market for customer information. Working paper, Duke University.
[21] Wall Street Journal, 1999. American opinion (a special report): Optimum out-duels pessimism. The Wall Street Journal, September 1999, A 10.
[22] A.F. Westin, Privacy and Freedom, Atheneum, New York, 1967.