Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2002 Proceedings

Americas Conference on Information Systems (AMCIS)

December 2002

PRIVACY ON THE INTERNET: AN ECONOMIC ANALYSIS

Jeevan Jaisingh
Purdue University

Jack Barron
Purdue University

Alok Chaturvedi
Purdue University

Shailendra Mehta
Purdue University

PRIVACY ON THE INTERNET: AN ECONOMIC ANALYSIS[1]

Jeevan Jaisingh, Jack Barron, Alok Chaturvedi, and Shailendra Mehta
Krannert Graduate School of Management
Purdue University
jeevan_jaisingh@mgmt.purdue.edu barron@mgmt.purdue.edu
alok@mgmt.purdue.edu mehta@mgmt.purdue.edu

Abstract

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. The research questions we are addressing here are – How should an Internet Service Provider (ISP) 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 a third party? Should the ISP have a blanket policy of never collecting, or a policy of always collecting and revealing information? We calculate the separating and pooling strategies for the ISP under asymmetric information and compare them with the full information benchmarks. We find that in some cases the ISP may be no worse off than in the full information case while in other cases it may have to restrict the set of contracts so that they are incentive compatible and individually rational.

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 (FTC 2000). These developments have increased privacy concerns among consumers, especially online consumers. According to a recent Business Week/Harris poll (Business Week 1998), privacy is the number one consumer issue facing the Internet. These concerns about privacy need to be addressed to enable the growth of E-commerce.

Te 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 (Laudon 1996). 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. In 1995, direct marketing resulted in $600 billion in sales (Solove 2000).

It is evident that 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. Past research (Laudon 1996, Kang 1998, Sovern 1999) has suggested an information market-based solution where consumers reveal potentially useful personal information in return for money and personalized services. However to the best of our knowledge, there has not been much research addressing the economic analysis of transactions privacy and information markets. It is the intention of this research to fill that void.

In this paper the problem that we model is that of an Internet Service Provider (ISP). It is technically feasible for the ISP to track the browsing habits of its consumers. If allowed to collect data on browsing habits, the ISP 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

[1]This research was partially funded by National Science Foundation grants #EIA-0075506 and DMI-0122214.

party access to this information, the ISP enables the matching of the third party with the right consumers. This collection and release of consumer information if done without the permission of the consumer is what is termed as privacy violation in our context. However the release of information to a third party if done with the consent of the consumer could lead to overall increase in social welfare.

The research questions we are addressing here are – How should the ISP 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 ISP 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? Which of these options is Pareto Optimal?

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. Section 4 summarizes the results and the managerial implications. In Section 5 we look at the limitations of our results and possible extensions.

Background

Privacy is defined as the moral claim of the individuals to be left alone and to control the flow of information about themselves (Coase 1960; Westin 1967). Coasian logic suggests that given sufficient flexibility in contracting, stringent privacy laws are unnecessary to achieve efficiency in concealment or revelation of information (Kahn et al 2000). For a background on the evolution of privacy concerns see Culnan (1993). More recently Wang et al (1998) have provided a taxonomy of consumer privacy concerns specific to the Internet marketing area.

Past research on privacy can be classified into three different streams – Legal, Information Systems and Economic. Sovern (1999) 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 (1998) is more relevant to our research question since it focuses on the problem of personal data generated in cyberspace transactions. Kang 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.

Literature in the Information Systems stream has concentrated on the technical aspects of privacy. The World-Wide Web Consortium (W3C)'s Platform for Privacy Preference Project (P3P) provides a framework for informed online interactions. P3P applications will allow users to be informed about Web Site practices, delegate decisions to their computer agent when they wish, and tailor relationships with specific sites (Reagle and Cranor 1999). Tools available for privacy protection include Crowds (Reiter and Rubin 1997), Onion Routing (Goldschlag et al 1999), Anonymizer (www.anonymizer.com), and anonymous re-mailers (Bacard 1996). Laudon (1996) proposes a framework for an 'Information Market' where consumer profiles would aggregated and sold as commodities. Consumers would be compensated for the loss of privacy by money or personalized services.

Economic studies of Privacy include –Hirshleifer (1980), Stigler (1980) and Posner (1981). Most of these were written at a time when technology was in its infancy – Internet was not even in existence. More recently Kahn et al (2000) have looked at efficiency gains under different privacy regimes. Their environment is characterized with complete information and they look at the social planners problem of when and when not transactional information should be made public knowledge.

Model

There are three sets of agents in the market - the consumers (A), the ISP (B) and the third parties (C). There are a large number of consumers and third parties but just a single ISP (firm). 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*.[2] 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. Let Consumer type be a vector $\theta = (\alpha, \beta)$. α represents the valuation for privacy (*L*-Low, *H*-High). β represents whether the buyer information is useful to the third party and affiliates of the ISP (*H*- High Value, *L*-Low Value). All consumers are assumed to have the same valuation for the Internet service provided by *B*. Thus there are four consumer types: $\theta_1 = (H, L); \theta_2 = (H, H); \theta_3 = (L, H); \theta_4 = (L, L)$. All agents are risk-neutral[3] and the reservation utility[4] of all consumers is assumed to be zero. The firm *B* is a profit-maximizing monopolist. We consider a two period model similar to Kahn et al (2000). In the first period the consumer *A* chooses to subscribe to the Internet Service provided by *B*. Based on the option chosen by *A* (to allow *B* to collect information or not), *B* collects information on the browsing behavior of *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. As a result of this targeting, the utility of agents *A*, *B* and *C* change. Since *B* enables the matching between *A* and *C*, *B* can extract a rent from *C*.

Notation: v is the reservation utility of all consumers for Internet Service provided by *B*. p_j is the price charged for the service for consumer type $\theta_j, j = 1,2,3,4$. v_i is the utility of agent *i* ($i = A, B, C$) due to agent *A* subscribing to the service provided by *B*. w_i is the change in utility of agent *i* ($i = A, B, C$) due to release of information and subsequent match between *A* and *C*. *V*,*P* and *c* are the reservation utility, price and marginal cost respectively of the personalized service provided by the third party *C*. Assume that $V > P$ i.e. all consumers would want to buy the personalized service at the price *P*. c_H and c_L are the disutility due to loss of privacy for consumers who have a high and low valuation for privacy respectively $(c_H > c_L)$. 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.

As is clear from our setting, our environment is characterized by asymmetric information. Both *B* and *C* don't know *A's* type in period 1 while *A* knows his 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 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.

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.

Table 1. Profits/Surplus for Agents Under Public Information

Cons. 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	$-c_H+(V-P)$	0	$P-c$	$v-p_2-c_H+(V-P)$	p_2	$P-c$
(L,*H*)	$v-p_3$	p_3	0	$-c_L+(V-P)$	0	$P-c$	$v-p_3-c_L+(V-P)$	p_3	$P-c$
(*L*,*L*)	$v-p_4$	p_4	0	$-c_L$	0	0	$v-p_4-c_L$	p_4	0

---

[2]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.

[3]Risk Neutrality means that the agents utility function is linear in wealth.

[4]Reservation utility is the utility that the agent gets by not subscribing to the ISP.

The monopolist firm will extract the entire consumer surplus. Hence the schedule of prices would be $p_1 = v - c_H; p_2 = v - c_H + (V - P); p_3 = v - c_L + (V - P); p_4 = v - c_L$. Prices charged will have the following relationship $p_3 > p_4; p_3 > p_2; p_2 > p_1; p_4 > p_1$.

Full Information

B knows *A*'s type in period 1 (hypothetically) and can commit not to reveal the information to *C* in period. *A* can choose whether to allow *B* to reveal information to *C* or not. Let *k* be *A*'s choice variable. *k* = 1corresponds 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.

Table 2. Profits/surplus of Agents Under Full Information if *A* Chooses *k* = 1

Cons. 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	$-c_H$+(V-*P*)	p^I	*P*-*c*-p^I	$v - p_2 - c_H + (V - P)$	p_2+p^I	*P*-*c*-p^I
(L,*H*)	v-p_3	p_3	0	$-c_L$+(V-*P*)	p^I	*P*-*c*-p^I	$v - p_3 - c_L + (V - P)$	p_3+p^I	*P*-*c*-p^I
(*L*,*L*)	v-p_4	p_4	0	$-c_L$	0	0	$v - p_4 - c_L$	p_4	0

Table 3. Profits/Surplus of Agents Under Full Information if *A* chooses (*k* = 0)

Cons. Type	v_A	v_B	v_C	w_A	w_B	w_C	$\pi_{A=} v_{A+}w_A$	$\pi_{B=} v_{B+} w_B$	$\pi_{C=} v_{C+}w_C$
θ_1 =(*H*,*L*)	v-p_1	p_1	0	0	0	0	$v - p_1$	p_1	0
θ_2 = (*H*,*H*)	v-p_2	p_2	0	0	0	0	$v - p_2$	p_2	0
θ_3 = (L,*H*)	v-p_3	p_3	0	0	0	0	$v - p_3$	p_3	0
θ_4 = (*L*,*L*)	v-p_4	p_4	0	0	0	0	$v - p_4$	p_4	0

Case 1: $c_L < V - P < c_H$

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. Since *B* wants *A* to pick the option that maximizes her profit, *B* will provide the following pricing options:

$$\{(p^*{}_1 = v; k^* = 0); (p_1 = v - c_H + \varepsilon, k = 1)\}$$

$$\left\{ p^*{}_2 = (\begin{cases} v - c_H + (V - P).if.(P - c) > c_H - (V - P); k^* = 1 \\ v......................if.(P - c) \le c_H - (V - P); k^* = 0 \end{cases}); \right.$$

$$\left. p_2 = (\begin{cases} v + \varepsilon........................if.(P - c) > c_H - (V - P); k = 0 \\ v - c_H + (V - P) + \varepsilon.if.(P - c) \le c_H - (V - P); k = 1 \end{cases}) \right\}$$

$$\{(p^*{}_3 = v - c_L + (V - P); k^* = 1); (p_3 = v + \varepsilon, k = 0)\}$$

$$\{(p^*{}_4 = v, k^* = 0); (p_4 = v - c_L + \varepsilon, k = 1)\}$$

In all cases $\varepsilon > 0$ ensures that the *A* will pick the option (*) that maximizes *B*'s profit

Case 2: $c_L < c_H < V - P$ ***(see Appendix Section A.1)***

Case 3: $V - P < c_L < c_H$ ***(see Appendix Section A.2)***

Asymmetric Information

B doesn't 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. *B* provides a set of contracts – one where information won't be collected and one where information will be collected and potentially revealed. Let p_P be the price of the first contract (information is kept private) and p_R be the price of the second contract (information is revealed). *A* chooses from this set of contracts. The contracts that could be offered by *B* are $\{(p_p, k = 0\}$ or $\{p_R, k = 1\}$ or $\{(p_P, k = 0); (p_R, k = 1\}$. The first two are pooling contracts. The intuition for these types of contracts could be that the firms have a blanket policy regarding collection of information and release: they could either follow a policy of never collecting information or they could have policy of always collecting and releasing information.[5] If $(k = 0)$, then the surplus and profits of agents are as follows:

Table 4. Profits/Surplus if All Agents Choose the Contract $\{p_P, k = 0\}$

Cons. Type	v_A	v_B	v_C	w_A	w_B	w_C	$\pi_{A=} v_{A+} w_A$	$\pi_{B=} v_{B+} w_B$	$\pi_{C=} v_{C+} w_C$
θ_1 =(*H*,*L*)	v-p_P	p_P	0	0	0	0	$v - p_P$	p_P	0
θ_2 = (*H*,*H*)	v-p_P	p_P	0	0	0	0	$v - p_P$	p_P	0
θ_3 = (L,*H*)	v-p_P	p_P	0	0	0	0	$v - p_P$	p_P	0
θ_4 = (*L*,*L*)	v-p_P	p_P	0	0	0	0	$v - p_P$	p_P	0

If $(k = 1)$, then the surplus and profits of agents are as follows:

[5]In the first contract *B* provides service to only those consumers who don't 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 contract information is collected and revealed to *B* based on *A*'s choice.

Table 5. Profits/Surplus if All Agents Choose the Contract $\{p_R, k=1\}$

Cons. Type	v_A	v_B	v_C	w_A	w_B	w_C	$\pi_{A=} v_{A+}w_A$	$\pi_{B=} v_{B+} w_B$	$\pi_C = v_{C+}w_C$
θ_1 =(*H*,*L*)	v-p_R	p_R	0	$-c_H$	0	0	$v-p_R-c_H$	p_R	0
θ_2 = (*H*,*H*)	v-p_R	p_R	0	$-c_H+(V-P)$	p^I	P-c-p^I	$v-p_R-c_H+(V-P)$	p_R+p^I	P-c-p^I
θ_3 = (L,*H*)	v-p_R	p_R	0	$-c_L+(V-P)$	p^I	P-c-p^I	$v-p_R-c_L+(V-P)$	p_R+p^I	P-c-p^I
θ_4 = (*L*,*L*)	v-p_R	p_R	0	$-c_L$	0	0	$v-p_R-c_L$	p_R	0

Pooling-Non-Revealing: If *B* has a policy of never collecting and releasing information -$\{p_P, k=0\}$. *B* will charge a price $p_p = v$. *B* will 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 $c_L < V-P < c_H$. Also assume that $v-c_L < v-c_H+V-P$. Thus we have $v-c_H < v-c_L < v-c_H+V-P < v-c_L+V-P$. 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 charged 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 $(v-c_L, v-c_H+V-P]$ *B* sells to types θ_2 and θ_3 and is better off charging the price $p_R = v-c_H+V-P$. Similarly in the price range $(v-c_H+V-P, v-c_L+V-P]$ *B* only sells to types θ_3 and is better off charging the price $p_R = v-c_H+V-P$. *B* cannot charge a price greater than $v-c_H+V-P$ since no consumer type would buy the service at that price.

Table 6. Surplus of Under the Different Pooling-Revealing Prices

	$p_R = v-c_H$	$p_R = v-c_L$	$p_R = v-c_H+V-P$	$p_R = v-c_L+V-P$
θ_1 =(*H*,*L*)	0	$-(c_H-c_L)$ 0	$-(V-P)$ 0	$-(c_H-c_L)-(V-P)$ 0
θ_2 = (*H*,*H*)	$(V-P)$	$[(V-P)$ $-(c_H-c_L)]$	0	$-(c_H-c_L)$ 0
θ_3 = (L,*H*)	c_H-c_L+V-P	$V-P$	c_H-c_L	0
θ_4 = (*L*,*L*)	$V-P$	0	$-[(V-P)$ $-(c_H-c_L)]$ 0	$-(V-P)$ [a]

[a]Since Consumer surplus is negative no trade takes place at this price.

Table 7. Firm Profits Under the Different Pooling-Revealing Prices

Prices	Firm's (*B*) surplus/profits
$p_R = v - c_H$	$N[(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)(v - c_H) + (\lambda_2 + \lambda_3)(P - c)]$
$p_R = v - c_L$	$N[(\lambda_2 + \lambda_3 + \lambda_4)(v - c_L) + (\lambda_2 + \lambda_3)(P - c)]$
$p_R = v - c_H + V - P$	$N[(\lambda_2 + \lambda_3)(v - c_H + V - P) + (\lambda_2 + \lambda_3)(P - c)]$
$p_R = v - c_L + V - P$	$N[\lambda_3(v - c_L + V - P) + \lambda_3(P - c)]$

The optimal price p_R^* (i.e. price that maximizes profits when information is always revealed) would depend on the values of the exogenous parameters and the firm's prior on types.

***Case 1*: $\lambda_1 = \lambda_2 = 0; \lambda_3 = \lambda_4$**

Firm profits $$\Pi = \begin{cases} N[(v-c_H)+\frac{1}{2}(P-c)] \\ N[\frac{1}{2}(v-c_H+V-P)+\frac{1}{2}(P-c)] \\ N[\frac{1}{2}(v-c_L+V-P)+\frac{1}{2}(P-c)] \\ N[(v-c_L)+\frac{1}{2}(P-c)] \end{cases}$$ under the different pooling revealing pricing options. The optimal pricing strategy for *B* if it always collects and reveals information is $p_R^* = \begin{cases} v-c_L \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots if \ldots\ldots v-c_L > V-P \\ v-c_L+(V-P)\ldots\ldots\ldots if \ldots\ldots v-c_L \le V-P \end{cases}$.

If *B* has a policy of never collecting information then *B*'s profits = $\Pi = N[v]$. Thus *B*'s optimal policy:

$$\begin{array}{l} (p=v, k=0)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots if..(v-c_L > V-P)..\&..c_L > \frac{1}{2}(P-c) \\ (p=v-c_L, k=1)\ldots\ldots\ldots\ldots\ldots if..(v-c_L > V-P)..\&..c_L \le \frac{1}{2}(P-c) \\ (p=v-c_L+V-P, k=1)\ldots.if..(v-c_L \le V-P)..\&..v > V-c-c_L \\ (p=v, k=0)\ldots\ldots\ldots\ldots\ldots\ldots\ldots if..(v-c_L \le V-P)..\&..v \le V-c-c_L \end{array}$$

Given conditions 1 and 4, *B* should have policy of never collecting and revealing information. Given conditions 2 and 3, it is optimal for the firm to always collect and reveal information to the third party.

Table 8. Profits and Surpluses under Full and Asymmetric Information and under Different Strategies

	Pricing Option	θ_3 = (L,*H*)	θ_4 = (*L*,*L*)	Profits
Asymmetric Information (Pooling)	$(p = v, k = 0)$	0	0	Nv
Asymmetric Information (Pooling)	$(p = v - c_L, k = 1)$	$V - P$	0	$N[(v - c_L) + \frac{1}{2}(P - c)]$
Asymmetric Information (Pooling)	$(p = v - c_L + V - P, k = 1)$	0	$-(V - P)$ (does not buy so 0)	$N[\frac{1}{2}(v - c_L + V - P) + \frac{1}{2}(P - c)]$
Full Information	$(p^*{}_3 = v - c_L + (V - P); k^* = 1)$ $(p^*{}_4 = v, k^* = 0)$	0	0	$N[\frac{1}{2}(2v - c_L + V - P) + \frac{1}{2}(P - c)]$
Asymmetric Information (Separating)	$(p^*{}_3 = v - c_L + (V - P); k^* = 1)$ $(p^*{}_4 = v, k^* = 0)$	0	0	$N[\frac{1}{2}(2v - c_L + V - P) + \frac{1}{2}(P - c)]$

The firm is thus always better off in the full information case. In the Asymmetric information case, pooling strategy, type θ_3 will be better off under one of the pricing policies ($(p = v - c_L, k = 1)$; i.e. where *B* has a policy of always collecting and revealing information and charges a price of $p_R = v - c_L$).

Separating Strategy: Since $\lambda_1 = \lambda_2 = 0$ there are only types $\theta_3 = (L, H); \theta_4 = (L, L)$ - low privacy valuation but differing in value of information. If ($(p_3, k = 1), (p_4, k = 0)$ are the pricing options for each type, then the prices should satisfy the following constraints:

$$v - p_3 - c_L + V - P \geq 0; v - p_4 \geq 0; v - p_3 - c_L + V - P \geq v - p_4; v - p_4 \geq v - p_3.$$ [5]

The full information prices $p_3 = v - c_L + V - P$ and $p_4 = v$, satisfies all the above constraints and so successfully separates type θ_3 and θ_4. The IR constraints of both types and the incentive compatibility constraint of type θ_3 bind. The reason the full information prices can separate the two types is because none of the types has an incentive to lie: type θ_4 is not interested in the personalized service type θ_3 gets and type θ_3 is no worse revealing that her information is going to be useful to at least one *C*. It should be noted that that type θ_3 is indifferent between pricing options $(v - c_L + V - P, k = 1)$ and $(v, k = 0)$ [6]. *B* is thus better off adopting a separating strategy than adopting a pooling pricing strategy. The interesting result here is that even with asymmetric information, the firms may be no worse than in the full information case.

Case 2: $\lambda_1 = \lambda_2; \lambda_3 = \lambda_4 = 0$ Analysis and results similar to Case 1.

---

[5]Individual Rationality (IR) and Incentive Compatibility (IC) constraints.

[6] Assuming θ_3 always chooses option $(v - c_L + V - P, k = 1)$, the option where he gets personalized service and which is more profitable to *B*.

Case 3: $\lambda_1 = \lambda_4 = 0; \lambda_2 = \lambda_3$ **see Appendix B.1.**

When $(P - c) > c_H - (V - P)$ then no separating strategy exists. For $(P - c) \le c_H - (V - P)$ the full information prices separate the two types, thus in this case *B* is no worse than under full information.

Case 4: $\lambda_1 = \lambda_4; \lambda_2 = \lambda_3 = 0$

It is easy to show that the pooling strategy $\{p_P = v, k = 0\}$ is the optimal strategy.

Case 5: $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \frac{1}{4}$

Outline of the results are given, details and pooling strategy prices are left to the interested reader.

Separating Strategy: If $((p_P, k = 0), (p_R, k = 1)$ are the separating pricing options, then firm *B* wants the consumer types θ_1, θ_2 and θ_4 to pick the pricing option $(p_P, k = 0)$ and for type θ_3 to pick the pricing option $(p_R, k = 1)$. The prices should satisfy the following constraints:

$$v - p_P \ge 0; v - p_R - c_L + V - P \ge 0; v - p_R - c_L + V - P \ge v - p_P$$
$$; v - p_P \ge v - p_R - c_H; v - p_P \ge v - p_R - c_H + V - P; v - p_P \ge v - p_R - c_L.$$

The full information prices satisfy these constraints for $(P - c) \le c_H - (V - P)$. However for $(P - c) > c_H - (V - P)$, as in Case 3 there are no prices that separate the types.

It is clear from the above analysis that full information prices are able to separate the consumer types in some of the cases. However with certain parameters, e.g. $c_L < V - P < c_H; \&. (P - c) > c_H - (V - P)$, the full information prices cannot separate types θ_2 and θ_3. Note that with these parameters under full information they should be charged prices $(v - c_H + V - P, k = 1)$ and $(v - c_L + V - P, k = 1)$ respectively. However type θ_3 will have an incentive to lie and say that she has a high valuation for privacy and hence be charged a lower price. Another example is where $c_L < c_H < V - P$. The full information prices cannot separate types θ_2 and θ_3 since as before the type θ_3 has an incentive to lie about her type. In order to find a separating mechanism to handle these cases, we modify our setting. We consider the case $c_L < c_H < V - P$ with $\lambda_1 = \lambda_4 = 0; \lambda_2 = \lambda_3 = \frac{1}{2}$. We assumed in our setting before that the information on customer would be released to only one third party *C* as a result of which each consumer *A* gets a utility $V - P = w$. We now relax this assumption – *B* can now release information (with permission from *A*) to more than one third party. In the base model setting, privacy was a binary variable – either information was released or not released. In the modified setting, there are varying degrees of privacy depending on the number of third parties to whom information is revealed. Thus if there are *M* third parties, then privacy could taken on any integer values in [0,*M*]. Releasing information to *M* third parties corresponds to zero valuation for privacy and 0 corresponds to the highest valuation for privacy. Any value in between specifies an intermediate valuation for privacy. The consumer's utility from personalized service is 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 $v + wn - nc_H - p_H = v + n_H(w - c_H) - p_H; v + n_L(w - c_L) - p_L$. $w - c_H$ and $w - c_L$ can be interpreted as the 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

[7]Implicit in this assumption is that types θ_2 and θ_3 receive the same utility from personalized services.

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.

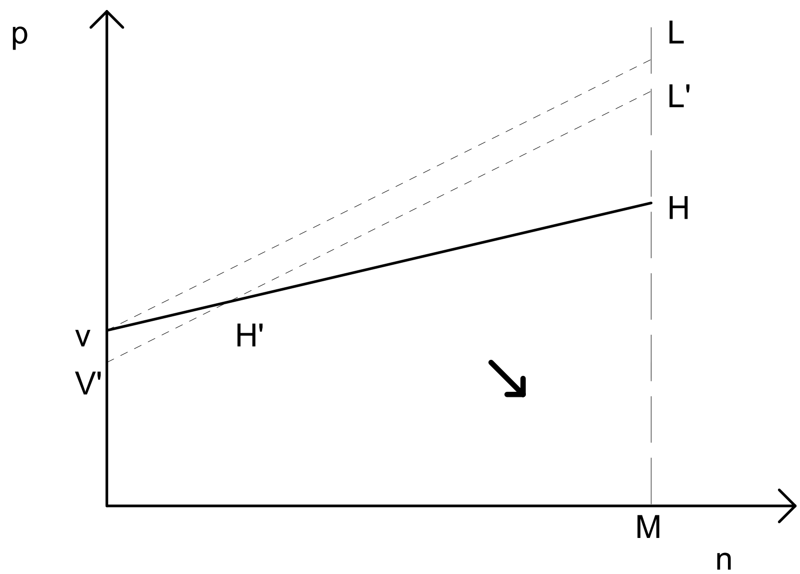

Figure 1. Indifference Curves and Contracts under Full and Asymmetric Information $c_L < c_H < w$

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- $w'_L > w'_H$. The firms under full information would charge prices along the indifference curves $v + n_H w'_H - p_H = 0; v + n_L w'_L - p_L = 0$ - lines VH and VL respectively. These indifference curves both pass through the point (0,*v*), which corresponds to highest level of privacy. The firm is best off if it offers the contracts at *L* and *H* for the low and high type consumers respectively which corresponds to the release of information to *M* third parties. Both consumer types are better off in the direction of the arrow-lower prices and more net benefits from personalized services. Under asymmetric information however the low privacy type has an incentive to lie. So under asymmetric the contracts should be restricted and must satisfy the individual rationality and incentive compatibility constraints:

$$v + n_H w'_H - p_H \geq 0; v + n_L w'_L - p_L \geq 0;$$
$$v + n_H w'_H - p_H \geq v + n_L w'_H - p_L; v + n_L w'_L - p_L \geq v + n_H w'_L - p_H$$

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 $V'L'$. There is a set of contracts available to the low privacy type while there is only one contract for the high privacy type. The contract for the high type is at the point of intersection H' of the new indifference curve of the low type $V'L'$ and the indifference curve of the high type. All the feasible contracts for the low type lie on the segment $H'L'$. The firm *B* is best off if the low types choose L'. Thus under asymmetric information the firm is best off offering contracts L' and H' compared with the full information contracts L and H. 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 is loses out on some part of the surplus it gets from the low type consumer. The Figs. for $c_L < V - P < c_H$ and $V - P < c_L < c_H$ are shown in the appendix section C.

Results and Implications

We have calculated the prices under asymmetric information only for the case $c_L < V - P < c_H$. The asymmetric information prices for the other cases can be calculated on similar line and have been left out of paper for brevity. For all cases, it is optimal for the firm to have a separating strategy. The full information prices successfully separate types except in cases where types θ_2 and θ_3 are present simultaneously and the rent that the firm can extract from the third party is less than the subsidy he has to give the high privacy type to participate. In the pooling-non-revealing strategy, the firm has much lower profits while in the pooling-revealing strategy; the firm leaves something on the table for type θ_3. We modify the setting later in order to successfully separate types θ_2 and θ_3 by allowing the revelation of information to more than one third party. By providing a subsidy to the low privacy type consumer and by restricting the set of contracts that are available to the high privacy consumer, successful separation of types can be achieved. The firm can extract the entire surplus from the high-risk types but leaves some surplus for the low risk types. 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 it loses out on some part of the surplus it gets from the low type consumer.

The results found here have important implications. They suggest that an ISP or any firm that can collect information and form consumer profiles for that matter is better off having a separating strategy for high privacy and low privacy types. This is contrary to some evidence that we see in practice – ISP's like AOL have a blanket policy of never collecting information and revealing it to third parties.[8] There is the other spectrum of firms too who do not have a privacy policy in place and presumably collect information and reveal it to third parties. Our theoretical results suggest that offering a set of polices designed for the different types of consumers and allowing them to self-select is optimal. We have shown here that in some cases the firm is no worse off than in the full information case and can extract the entire consumer surplus. However in cases where full information prices cannot separate types, a slight modification in the setting allows the firm to separate types, however loosing out on some profits. The results from our paper suggest that it is possible for firms to design more sophisticated mechanisms that allow the firms to reap higher profits.

Limitations and Directions for Future Research

In this paper we consider a monopolist *B* who prices the service based on the different privacy valuations and value of information to third parties. It would have been more realistic to consider an oligopolistic setting or a competitive one and to compare the results with those obtained here. Solving the monopoly case is a first step before solving the other cases and is important from a theoretical perspective. The extension to a two firm case could be done by consider the problem to be game where each firm decides whether to follow a separating strategy or a pooling strategy given the strategy of the other player. Firm *B* here enabled the matching between the consumers and the third parties and extracts a rent from the third party for the service provided. Implicit in the assumption is that *B* has perfect knowledge on the information requirements of the third parties. We assumed risk neutrality among all agents; there could be a qualitative change in the nature of the results if risk aversion is taken into account. For the mechanism designed here to be feasible, all agents should follow the actions specified by the contract. It is possible that the firm 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 the 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.

References

Baase, S. "Privacy and Information," in *A Gift of Fire: Social Legal and Ethical issues in Computing*, Prentice Hall, 1997, pp50.
Bacard, A. "Anonymous Remailer FAQ," 1996. http://www.well.com/user/abacard/remail.htm
Business Week. "A little net privacy, please," *Business Week*. (Mar 16, 1998); www.businessweek.com.
Branscomb, A.E. "Who Owns Information?" *Basic Books*, pp 9-10.

[8]There is some evidence however which provides partial support for our findings – grocery stores have a differential pricing of goods based on whether customers use grocery cards. Use of the grocery cards provides discounts on goods purchased, with the implicit assumption that the information collected by the grocery store can then be sold to third parties who then target the consumers with one to one marketing. On the Internet, the example of "Free PC.com" comes to mind – the company offered free PC's to consumer who would allow the company to collect information. Both the grocery store and "Free PC" are examples where the consumers were ready to put a price to their privacy in terms of reduced prices or other products.

Culnan, M.J. "How did they get my name: An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use," *MIS Quarterly*, September 1993 , pp 341-361.
Coase, R.H. "The problem of social cost," *Journal of Law and Economics*, 3, Oct 1960, pp1-44.
Goldschlag, D., Reed, M., and Syverson, P. "Onion Routing," *Communications of the ACM* Feb 1999, pp 39-41.
Kahn, C.M., McAndrews, J., and Roberds, W. "A Theory of Transactions Privacy," *Federal Reserve Bank of Atlanta, Working Paper,* Nov 2000.
Kang, J. "Information Privacy in Cyberspace Transactions," *Stanford Law Review,* Vol. 50, 1998, pp 192-1294.
Laudon, K.C. "Markets and Privacy," *Communications of the ACM*, Vol 39 (9), Sep 1996, pp 92-104.
FTC. "Privacy Online: Fair Information Practices in the Electronic Marketplace, A report to Congress," *Federal Trade Commission*, 2000
Reagle, J., and Cranor, L.F. "The Platform for Privacy Preferences," *Communications of the ACM*, 42(2) Feb 1999, pp 48-55.
Reiter, M.K., and Rubin, A.D. "Crowds: Anonymity for Web Transactions," *ACM Transactions on Information and System Security*, 1997
Shapiro, A.L., "Privacy for Sale: Peddling data on the Internet", *The Nation*, June 23, 1997, pp 11-12.
Solove, D.J. "Privacy and Power: Computer Databases and Metaphors for Information Privacy," *Working Paper.Seton Hall Law School* 2000.
Sovern, J. "Opting in, opting out, or no options at all: The fight for control of personal information," *Washington Law Review,* Vol. 74, 1999, pp 1033-1118.
Westin, A.F. "Privacy and Freedom," *Atheneum*, New York, 1967.

Appendix

A.1 Full Information

Case 2: $c_L < c_H < V - P$

$$\{(p^*{}_1 = v; k^* = 0); (p_1 = v - c_H + \varepsilon, k = 1)$$

$$\{(p^*{}_2 = v - c_H + (V - P), k = 1); (p_2 = v + \varepsilon, k = 0)\}$$

$$\{(p^*{}_3 = v - c_L + (V - P); k^* = 1); (p_3 = v + \varepsilon, k = 0)\}$$

$$\{(p^*{}_4 = v, k^* = 0); (p_4 = v - c_L + \varepsilon, k = 1)\}$$

A.2 Full Information

Case 3: $V - P < c_L < c_H$

$$\{(p^*{}_1 = v; k^* = 0); (p_1 = v - c_H + \varepsilon, k = 1)$$

$$\left\{ p^*{}_2 = (\begin{cases} v - c_H + (V - P).if.(P - c) > c_H - (V - P); k^* = 1 \\ v........................if.(P - c) \le c_H - (V - P); k^* = 0 \end{cases}); \right.$$

$$\left. p_2 = (\begin{cases} v + \varepsilon.........................if.(P - c) > c_H - (V - P); k = 0 \\ v - c_H + (V - P) + \varepsilon.if.(P - c) \le c_H - (V - P); k = 1 \end{cases}) \right\}$$

$$\left\{ p^*{}_3 = (\begin{cases} v - c_L + (V - P).if.(P - c) > c_L - (V - P); k^* = 1 \\ v........................if.(P - c) \le c_L - (V - P); k^* = 0 \end{cases}); \right.$$

$$\left. p_3 = (\begin{cases} v + \varepsilon.........................if.(P - c) > c_L - (V - P); k = 0 \\ v - c_L + (V - P) + \varepsilon.if.(P - c) \le c_L - (V - P); k = 1 \end{cases}) \right\}$$

$$\{(p^*{}_4 = v, k^* = 0); (p_4 = v - c_L + \varepsilon, k = 1)\}$$

B.1

Case 3: $\lambda_1 = \lambda_4 = 0; \lambda_2 = \lambda_3$

B's optimal policy:

$$(p = v, k = 0)........................if..v + V - c > 2c_H - c_L..\&..c_H > (V - c)$$
$$(p = v - c_H + V - P, k = 1)...if..v + V - c > 2c_H - c_L..\&..c_H \le (V - c)$$
$$(p = v - c_L + V - P, k = 1)....if..v + V - c \le 2c_H - c_L..\&.v \le V - c - c_L$$
$$(p = v, k = 0).......................if..v + V - c \le 2c_H - c_L..\&..v > V - c - c_L$$

Table A1. Profits and Surpluses under Full and Asymmetric Information and Under Different Strategies

	Optimal Pricing Option	$\theta_2 = (H,H)$	$\theta_3 = (L,H)$	Profits
Asymmetric Information (Pooling)	$(p = v, k = 0)$	0	0	Nv
Asymmetric Information (Pooling)	$(p = v - c_H + V - P, k = 1)$	0	$c_H - c_L$	$N[(v - c_H + V - P) + (P - c)]$
Asymmetric Information (Pooling)	$(p = v - c_L + V - P, k = 1)$	$-(c_H - c_L)$ does not buy	0	$N[\frac{1}{2}(v - c_L + V - P)$ $+ \frac{1}{2}(P - c)]$
Full Information	1. If $(P - c) > c_H - (V - P)$ $\{(p^*_2 = v - c_H + (V - P); k^* = 1)$ $(p^*_3 = v - c_L + V - P, k^* = 1)\}$; 2. If $(P - c) \le c_H - (V - P)$ $\{(p^*_2 = v; k^* = 1)$ $(p^*_3 = v - c_L + V - P, k^* = 1)\}$;	0 0	0 0	$N[\frac{1}{2}(v - c_L + V - P)$ $+ \frac{1}{2}(v - c_H + V - P)$ $+ (P - c)]$ $N[\frac{1}{2}(2v - c_L + V - P)$ $+ \frac{1}{2}(P - c)]$
Asymmetric Information (Separating)	1. If $(P - c) > c_H - (V - P)$ No Separating Strategy 2. If $(P - c) \le c_H - (V - P)$ $\{(p^*_2 = v; k^* = 0)$ $(p^*_3 = v - c_L + V - P, k^* = 1)\}$;	N/A	N/A	N/A $N[\frac{1}{2}(2v - c_L + V - P)$ $+ \frac{1}{2}(P - c)]$

C.

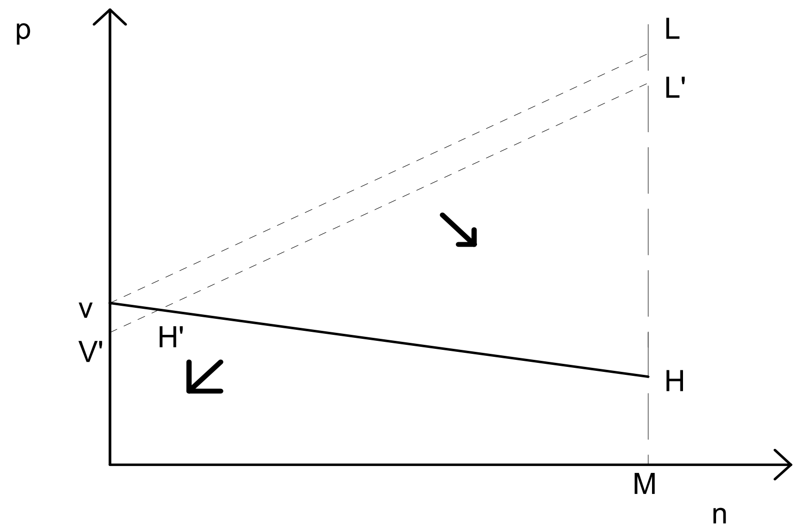

Figure 2. Indifference Curves and Contracts Under Full and Asymmetric Information $c_L < w < c_H$

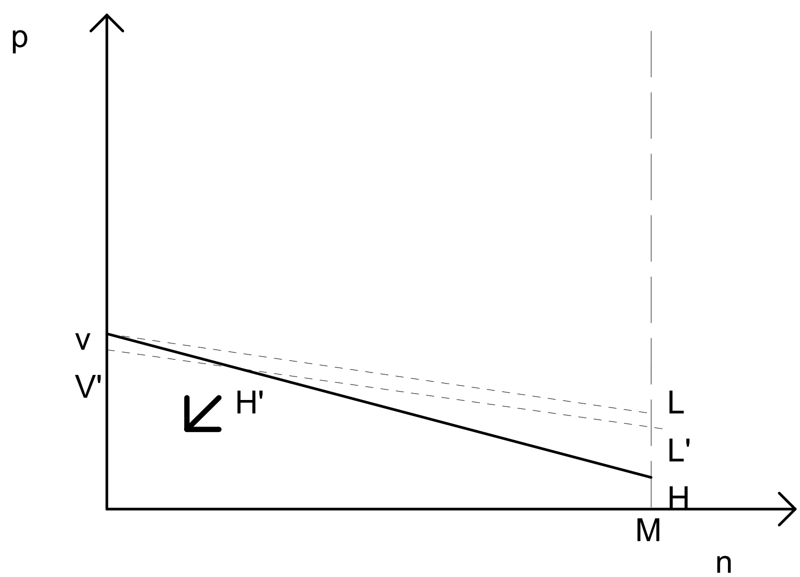

Figure 3. Indifference Curves and Contracts Under Full and Asymmetric Information $w < c_L < c_H$