Wharton Research Scholars Journal

Wharton School

9-1-2010

Individual Privacy and Online Services

Minh Chau
University of Pennsylvania, minh.chau.wh11@wharton.upenn.edu

Eric K. Clemons
University of Pennsylvania, clemons@wharton.upenn.edu

Individual Privacy and Online Services

Minh Chau | The Wharton School | mih@wharton.upenn.edu
Eric K. Clemons | The Wharton School | clemons@wharton.upenn.edu

Abstract

We explore consumer trade-offs between better performance through tailoring of online services to their individual needs and greater privacy as a result of reduced disclosure of personal information. We show that individuals have different willingness to accept loss of privacy that is a function of (1) the individual and his/her preferences, because the variation in demands for privacy is not uniform across individuals, (2) the service Domain, because individuals demand more privacy in some Domains than they do in others and (3) these differences themselves differ among consumers as well.

Keywords: Tradeoffs between privacy and enhanced web services, online privacy, online personalization.

1. Introduction

Despite claims to the contrary, we find that privacy is not yet dead. Our study shows that privacy remains an important issue to our study population (primarily college students), and that concerns vary among individuals and across different Domains. Here we define Domains[1] as areas of online services that are distinct from one another with respect to the content of services provided to the user, the information needed from the user in order to personalize the service (e.g., locating a coffee shop vs. providing driving directions), and our classification of the specific area of potential sensitivity (e.g., political preferences, sexual behavior, potential medical problems).

2. Background

2.1. Privacy Concerns

Privacy is defined on three dimensions. One dimension of privacy, *secrecy*, concerns concealing information to prevent it from being released accidentally or discovered by information seekers [19, 29, 32]. A second dimension of privacy, *seclusion*, refers to a state of peace, with freedom from unwanted interruption [19, 29]. Finally, the third dimension of privacy, *autonomy*, "concerns freedom from observation," leading to freedom of action [19, 29]. While all three dimensions of privacy – secrecy, seclusion, and autonomy – are important considerations in evaluating privacy regulations [19], we are concerned primarily with the first definition of privacy – the concealment of information. This is similar to the definition of privacy used in psychology literature [38]. Following Awad and Krishnan and Stone et al., we define information privacy as "the ability of the individual to personally control information about one's self" [4, 33]. When individuals choose to provide their personal information in return for enhanced online services, they may believe that they are handing over information to a party whose use of that information is now outside their control.

Firms' customization and personalization of online services has undeniably experienced significant technological progress [12, 26, 31, 35, 36]. In return for using these online services, consumers are confronted with advertisements on the website where the services are located. Through the use of cookies as well as voluntary disclosure, a range of consumer information may be collected through these services – from a simple email address to a complex set of search term histories, site preferences, past purchases, credit card information, physical mailing addresses, and so forth. Companies' possession of such rich datasets can benefit consumers through tailored services that can speed up purchase transactions, provide better search results, and other services. However, the potential for consumer harm exists, as in the case of Amazon's alleged differential pricing of DVDs based on customers' purchase histories and demographics [5]. For firms, such information can serve as a competitive advantage in environments where differentiation is key to their survival, or as a public relations nightmare as in Facebook's attempts at privacy policy changes, which has caused a rise in concern over privacy (see, for example, [8, 9, 15, 23, 32, 33, 39, 47, 49]). This motivated the study of consumers' tradeoffs between better performance through tailoring of websites to their individual needs and greater privacy as a result of reduced disclosure of personal information.

In the following section, we review prior literature. The third section discusses this paper's theoretical model and hypotheses. The fourth section reviews data methodology. The fifth section presents results and analyses. Section six presents a discussion of results and their business and policy implications, as well as potential areas for further research. We conclude by

[1] As with Domain, all variables and terms used in this paper that are used with precise meanings in our research will be capitalized to alert the reader.

addressing some of the limitations of our study.

2.2. Prior Literature

As firms, employers, advertisers, governments, and other parties increase their technological ability to track individuals' behavior and information, privacy concerns become more important for politicians and academics of all fields, including law, economics, political science, psychology, and others [19].

Even with legislative protections [13, 19, 28] in place to protect individual privacy, consumers still have reservations in providing more and more details of their Internet habits to firms. Studies by Pricewaterhouse-Coopers and Harris showed that online shopping would increase if consumers knew that retailers would not misuse their personal information [1, 3, 16].

It is necessary to find the proper balance between privacy and customization. Quantification of privacy valuations has been attempted in various contexts, as in privacy with regards to location data [11], marketers' call lists [37], and online activity [15]. In a conjoint analysis of students in United States and Singapore universities, Hann *et al.* identified three segments of the population: privacy guardians (the majority) who value privacy highly, information sellers who provide personal information in return for monetary rewards, and convenience seekers who provide personal information in return for increased convenience [15].

However, other researchers have identified factors that can increase individuals' willingness to provide private information. These results do not appear to apply only to "information sellers" or "convenience seekers." For example, in both field and laboratory experiments, Hui *et al*. showed that if provided with small monetary incentives, individuals became more willing to disclose private information [20, 21]. In his 2006 review of economic analyses of privacy, Hui concluded that consumers are not as sensitive to the sharing of personal information as previously believed and as previously reported in the literature (see surveys and experiments referenced by Hui [21] for example, [1, 7, 22, 35, 36, 54, 56]), and that economic solutions, such as monetary incentives, increased convenience, or the provision of resources, can mediate the exchange of information [14, 19, 23, 27]. In addition Chellappa and Sin showed that vendor trust-building activities, such as improving their brand image, can reduce consumers' concerns for privacy [10]. They showed that privacy is not absolute, and that consumers may give up more information in return for benefits [10]. Again, these results do not appear to apply only to "information sellers" or "convenience seekers" as the authors did not identify a group of consumers (i.e., "privacy guardians") in their study population (n=243) who were particularly insensitive to trust as a mediating factor. Furthermore, Chellappa showed that consumers' perceptions of privacy, security, and trust vary between online and offline transactions—even if the transactions occur through the same company (e.g., Best Buy online and Best Buy brick and mortar) [9].

We believe that the model of Hann, while supported by his data, was incomplete. Just as monetary incentives, trust, online vs. offline interactions, and other factors can affect all individuals', and not just "information sellers'" or "convenience seekers'," willingness to provide information in return for enhanced services, we believe that with a large sample size and a wider range of Domains people will, on average, seem to be less polarized and less different. Some will protect privacy in some Domains while some will protect privacy in others, but few individuals will be consistently more concerned than average or consistently less concerned than average. Indeed, as mentioned above, the results of other researchers have shown different results.

3. Theory and Hypotheses

3.1. Theory

Following Awad and Krishnan [4], we use utility maximization theory to study consumers' tradeoffs between better performance through tailoring of websites to their individual needs and greater privacy as a result of reduced disclosure of personal information. While Rust et al. [30] and other researchers have used utility maximization theory to examine consumer privacy, there are limitations to the use of the theory. For example, although utility maximization theory argues that consumers make decisions to maximize their economic utility, the majority of consumers do not make explicit financial calculations of the costs and benefits of social exchanges [6, 17]; however, previous research has shown that consumers do, to a certain extent, consider the tradeoffs involved [5, 17, 18, 19, 38, 40, 51].

The "privacy calculus" theory argues that individuals' decisions to reveal or not reveal their information is dependent on the potential benefits vs. potential negative consequences of providing their information [22, 24, 25, 34]. Factors that contribute to the privacy calculus include "personality and culture-based privacy attitudes, the type of information to be disclosed and its [deviation] from the average, the recipient, the value that is being assigned to personalization benefits, the extent to which users know what information has been disclosed and can control its usage, and various trust-establishing factors" [22]. Further, Huberman *et al.* showed that the price consumers demanded for that information varied with the sensitivity (broadly construed of that information: consumers who were significantly overweight or underweight demanded a higher price for releasing information on their weight) [18]. However, it should be noted that consumers often lack proper information or motivation to make privacy decisions, often having incorrect beliefs about whether certain pieces of information can be linked back to their identities or lacking the expertise to interpret privacy policies [22].

Our study focuses on whether enhanced online services (as measured through time savings via personalization of the online service to the consumer) have an effect on consumer willingness to disclose their personal information. Providing more information should represent a greater cost, while more time savings represents a greater benefit. If sensitivities vary as we expect, consumers will differ in the savings they demand before sharing information, and their demands will vary across Domains. Since subjects' demographics may affect their responses, we control for gender and age.

3.2. Hypotheses

We hypothesized that individuals have different willingness to accept a loss of privacy that is a function of (1) the individual and his or her preferences, because preferences for privacy are not uniform across individuals, (2) the service Domain because demand for privacy varies with the sensitivity of the Domain, and (3) variation among Domains that also vary across individuals. We explored the following, demonstrating:

Hypothesis 0: Individuals value their time and they value more time more.

We study consumer tradeoffs between enhanced online services (as demonstrated through time saved) and decrease in privacy by first determining that individuals do value their time and that they value more time more. While this hypothesis does not address privacy explicitly, it (1) assesses whether individual experimental subjects do value their time and (2) it is critical to our experimental design and if it were not supported it would be impossible for us to assess any of the hypotheses that follow.

Hypothesis 1: Individuals are not the same (e.g., they do not value their time or privacy equally).

Hypothesis 2: Individuals are not consistent across Domains (e.g., the variance of their responses to the Domain questions is not zero (0) and they value privacy in some Domains more than in others).

Hypothesis 3: Individuals are not uniformly more or less sensitive than average across Domains; individuals consider some Domains more sensitive and are less willing than average to share information in one or more Domains, and these same individuals are more willing than average to share information in other Domains.

Hypothesis 4: Individuals can be grouped with respect to their sensitivity over privacy across certain Domains; while individuals do vary, there are groups of individuals who are clustered and whose concerns for privacy are largely in agreement across Domains.

4. Research Methodology

We develop a set of experiments to test our hypotheses using a survey instrument.

4.1. Variable Definitions

Domain—areas of online services that are distinct from one another with respect to the type of services provided, the information needed from the user in order to personalize the service, and individuals' perceptions of the sensitivity of the topics addressed. Medical evaluation, advice for the unemployed, descriptions of aberrant sexual behavior, and restaurant reviews are all likely to be viewed very differently by different individuals, as providing home address or income levels might be. Given our research hypotheses, we choose Domains that we expected would reveal differences among individuals—e.g., Domains that we expected to be sensitive to some, to all, or to none of the experimental subjects. We used some Domains from previous researchers (e.g., financial services and travel services [14, 25]). We used the following Domains: medical symptom checker, sexual support, violent games, sexual fantasy games, children's games, financial help, map directions, relationship counseling, political commentaries, tutoring, finding restaurants, movie ratings, and finding coffee shops.

Feelings Towards Privacy (FTP)—As a proxy measure of the costs to individuals for revealing their private information, we asked subjects for their feelings towards having information being recorded about their internet usage through a five-point Likert-scaled item (Table 1).

Click Equivalents (Click(s))—As will be described, we asked subjects questions regarding tradeoffs between enhanced services in search and revelation of private information from the thirteen Domains. For each Domain, subjects received a hypothetical situation and then asked to choose between two services – one service tracked the individual's behavior but saved him or her X amount of time (X ranged from 0 to 60 minutes), the other did not track the individual's behavior but did not save him or her time. Subjects who chose the service that tracked his or her behavior indicated how much time, at a minimum, the service must save him or her. (Options ranged from 0 to 60 minutes at different increments.) In order to control for differences in individuals' absolute preferences for time, responses were standardized into "Click Equivalents," where each Click corresponds to 5 minutes. We do not use person-specific valuation of time due to concerns for unobserved heterogeneity (e.g., wealth), which could bias the results or make averages difficult to compare.

Above Average—For each subject, we calculated the Above Average, defined as the number of times the subject's Click value exceeded the average value by at least half the standard deviation of the responses for each Domain question. We followed this procedure to widen the range in which an individual's response would be counted as average in order to account for the unlikelihood that individuals would be exactly average. The average value used in these calculations corresponded to the average from which the comparison was made (e.g., average of female responses for the female histogram).

Below Average—We also calculated the corresponding Below Average (i.e., the number of times the subject's Click value was below the average value by at least half the standard deviation of the responses for each Domain question) for each individual. Similarly, the average value used in these calculations corresponded to the average from which the comparison was made.

Difference—Finally, we subtracted Below Average from Above Average to find the Difference.

Social Networking Usage (SNU)—We asked subjects if they use social networking sites like Facebook and Twitter, as well as how often they use these services (Table 2). We are interested to see if there is a correlation between SNU and FTP.

Internet Usage (IU)—We asked subjects how many hours per day, on average, they spend actively using the internet (Table 2). We are interested to see if there is a correlation between IU and FTP.

Reservation Club Card (RCC)—We asked subjects how much they would be willing to pay for a Reservation Club Card if it saved them X amount of time in waiting in lines (e.g., for tickets to a movie), where X ranged from 15 minutes to 90 minutes a month. We are interested to see if individuals value their time, and if so, whether they value more time more.

4.2. Research Design

The study population was drawn from the Wharton Behavioral Lab (WBL) participant list. Participants from the WBL are recruited through on-campus fliers and through email invitations.

Given that we are examining the tradeoff between privacy and customized internet services, the population of interest is adults who use the internet – although they need not be equally experienced users. See Table 3 for demographic information.

The experimental procedure is straightforward. While sitting in a private cubicle within the Wharton Behavioral Lab, participants opened the relevant survey and filled out the informed consent form prior to beginning the survey. Participants came to the Wharton Behavioral Lab once. They completed the tasks and filled out the questionnaires[2]. The main objectives of this research project is to demonstrate that the willingness to accept (WTA) decreased privacy in exchange for better services does exist and does vary across individuals and across Domains.

We hypothesize individuals have different WTAs that are a function of both the individual and the service Domain. To assess both existence and differences, we conducted surveys to quantify the WTA decreased privacy with regards to online services along thirteen Domains. We explored the following ideas:

1. To show that individuals do value their time, and place greater value on more time saved (independent of privacy), we asked subjects how much they would be willing to pay for versions of a Reservation Club Card that would save them varying amounts of time.
2. Additionally, we determined how many hours a day subjects use the internet, whether they participate in services like Facebook or Twitter or other social networks, and how they feel about Domain-independent information like having Google record their search history or their email messages sent.
3. We proceeded to ask participants regarding tradeoffs between enhanced services in search and revelation of private information from thirteen Domains. For each Domain, subjects received a hypothetical situation and then chose between two services – one service tracked the individual's behavior but saved him or her X amount of time (options ranged from 0 to 60 minutes), the other did not track the individual's behavior but did not save him or her time.

Note that we did not ask subjects to reveal private information related to these Domains (e.g., there were no questions related to the subjects' health). However, we did ask them how they feel about revealing information from that Domain to an online service provider.

Individuals' responses were standardized into "Click" equivalents, as described above.

4.3. Instrument Validity

Before conducting the survey, we asked a colleague with expertise in questionnaire design to examine the item for face validity. After the instrument was developed, it was pilot tested on 16 undergraduate

[2] The survey can be found online at http://opim.wharton.upenn.edu/~clemons/PrivacySurvey.pdf.

students at the Wharton Behavioral Lab. Feedback was solicited to correct any question or directions that were confusing or ambiguous, leading to the final survey instrument.

5. Experimental Results

Hypothesis 0: Individuals value their time and they value more time more. Applying linear regression to subjects' responses to the Reservation Club Card (RCC) question, and assuming a Y-intercept of 0, we found a nearly perfect linear relationship (Multiple R = 0.9993, R^2 = 0.9986, Standard Error = 0.5023) between subjects' average amount willing to pay and the minutes saved by the RCC (Figure 1). Applying the same analysis to the individual responses (after Winsorizing one subject's responses ($100, $200, $250, $525, $1000 for 15, 30, 45, 60, and 90 minutes, respectively) by replacing them with the second most extreme values – see Figure 2), we found a similarly strong, though weaker, relationship (Multiple R = 0.7075, R^2 = 0.5005, Standard Error = 10.2388). It is also evident that the range of subjects' willingness to pay increases with the amount of time savings: $0-$30 for 15 minutes, $0-$40 for 30 minutes, $0-$45 for 45 minutes, $0-$60 for 60 minutes, $0-$100 for 90 minutes.

Subjects' responses to how they felt regarding their information being recorded (i.e., ***very willing***, ***willing***, ***indifferent***, ***against***, and ***strongly against*** information being recorded about their internet use) were transformed into dummy variables. We did not use subjects' Likert-valued responses since the Feelings Towards Privacy (FTP) variable is not continuous.

Regression of subjects' average Clicks across all Domains on their FTP responses revealed significant effects of the dummy variables corresponding to indifferent (p = 0.018), against (p = 0.0001), and strongly against (p = 0.0001) on individuals' average Click values (with coefficients 1.7235, 3.3933, and 4.0861 respectively). Similar regression of subjects' average Clicks across all Domains on their Social Networking Usage (SNU) (hours) and Internet Usage (IU) (hours) revealed significant effects from the same FTP dummy variables, but no predictive power from the SNU (p = 0.584) or IU (p = 0.489) variables (see Tables A-B in Appendix[3]). These results suggests that while individuals' stated preferences regarding privacy is predictive of how many Clicks are necessary for them to use a Domain service, their actual usage of privacy-revealing services (e.g., Facebook) does not correspond to their FTP; and thus, revealing a possible disconnect between individuals' revealed and expressed preferences and behavior.

Hypothesis 1: Individuals are not the same (e.g., they do not value their time and/or privacy equally). One-way analysis of variance (ANOVA) comparison of each subject's set of responses to the RCC question showed a significant difference (p = 0.0001, see Table C in Appendix for full ANOVA statistics) between subjects' responses, suggesting a difference among subjects in how they value their time with respect to the RCC. One-way ANOVA of subjects' responses (in Clicks) to the Domain questions showed a significant difference (p = 0.0001, see Table D in Appendix for full ANOVA statistics) among subjects' set of 13 responses.

Hypothesis 2: Individuals are not consistent across Domains (e.g., the standard deviation (SD) of their responses to the Domain questions is not zero (0)). One sample t-test of subjects' average SDs over their Domain question responses (in Clicks) rejected the null hypothesis of the SD=0, with p<0.0001. (Table 4).

Hypothesis 3: Individuals are not consistently more or less sensitive than average across Domains. If subjects are expected to be either hoarders or free providers of information, we would expect a histogram of the Difference to exhibit bimodality. Visual observations of the histograms for males does not indicate bimodality (Figure 3); for females, bimodality may be present given the 7 individuals who were above average in all Domains; and for the combined data, bimodality may also be present, possibly due to the influence of the female dataset. These results imply that subjects cannot easily be grouped into two distinct groups of hoarders and free providers of information – at least now without first separating them by gender (Table 5). A thought experiment may help to clarify this point. If all of our subjects were hoarders, we would expect one peak at 13 in our distribution. If all of our subjects were free providers, we would similarly expect only one peak, but at -13. If individuals are as likely to be hoarders as free providers, half of our subjects would be hoarders and the other half free providers. We would then expect two peaks in our distribution, one at 13 and one at -13. We would also anticipate an expected average of the Difference to be 0. However, regardless of the ratio of hoarders to free providers within our study population, we would expect two peaks at the same locations, but with different magnitudes. As observed in Table 2, this is not the case. We recognize that if all individuals' sensitivities across Domains are not correlated we would expect Difference values to be clustered at 0. However, the presence of Difference values not confined to only 13 or -13 suggests that there are individuals who are more sensitive in certain Domains than in others. Hence, there are individuals who are not consistently more or less sensitive than average across Do-

[3] Wherever possible, supporting analyses have been moved to an online Appendix, which can be found at http://opim.wharton.upenn.edu/~clemons/PrivacyAppendix.pdf

mains.

Bivariate fit of Above Average by Gender for the combined dataset revealed no significant effects of gender (p = 0.2381). However, analysis of each Domain's Click response by Gender showed significant effects of Gender on responses to Domain questions 3 (p = 0.0001, playing military games) and 13 (p = 0.0378, searching for a coffee shop), and almost significant effects on question 5 (p = 0.0554, playing a children's game) (Table 6, also see Appendix Tables E-H for further details). For these Domain questions, males tended to have a lower Click value than females, suggesting that males may have a lower sensitivity to these questions. There was not enough variation in other demographic variables for us to consider them.

Service Domains are not the same with respect to individuals' concerns over privacy. One-way ANOVA analysis of each Domain's Click values revealed a significant difference across Domains (p = 0.0001, see Appendix Table I for more details).

Hypothesis 4: Individuals can be grouped or clustered across collections of Domains with respect to their sensitivity to privacy over those Domains. Post-hoc Bonferroni comparison of each subject's set of responses to the Domain questions revealed four groups of individuals based on their view of the Domain questions (in terms of Clicks). The Neutral Questions group had an average Click of 3.540; the Sex Questions group 10.776; the Personal Questions group 6.924; and the Flippy Questions (occurring in both the Neutral and Personal Groups) group 5.172 (Table 7). In terms of subject *types*, we therefore have those who *hoarders*, *privates* (share only information in the Neutral Questions group), *augmented privates* (share information in the Neutral and Personal Question groups), and free-providers (share information in all Domains).

To identify individuals who may be hoarders of information, we calculated the number of subjects who were unwilling to share information across all 13 Domains (average Click response = 13). Because no subject was willing to share information across all 13 Domains, we calculated the number of potential free-providers by enumerating those with an average Click response less than 2. As seen in Table 8, these individuals make up a relatively small percentage of the total study population, with *slightly* more males in the free-provider group than females.

Separating the data by gender revealed differences in how subjects perceived the Domain questions in the Neutral, Flippy, and Personal groups. Specifically, whereas females tended to view questions 1 (using a web portal to search health symptoms) and 3 (playing military games) as Personal, some males viewed them as Neutral while others see them as Personal; and whereas males viewed 13 (searching for a coffee shop) as Neutral, there were females who viewed it as Neutral and others who saw it as Personal (Table 7). Two-way ANOVA analysis on the effects of gender and question group on the average Clicks of each Question Group revealed significant effects from Question Group (p = 0.0009) but not from Gender (p = 0.1099) (see Table J in Appendix for more details).

6. Discussion

Our present research is focused on understanding consumers' willingness to compromise their privacy for enhanced services, and whether consumers can be neatly categorized as either hoarders or free providers of information. We began by showing that individuals do in fact value their time, that they value more time more, and that there is a difference among individuals in how they value their time (with respect to the Reservation Club Card question). Further, we also showed that individuals do not value their privacy equally, in that their responses to the set of thirteen Domain questions are not equivalent to one another. Additionally, not all Domains are the same with regards to individuals' concerns over privacy – each Domain's set of responses is different, with individuals differing about which Domains they see as sensitive or not.

We next showed that individuals are not consistent across Domains – they may be sensitive in one Domain, but insensitive in another. More specifically, individuals are not uniformly more or less sensitive than average across different Domains – they do not all fall neatly into categories of hoarders or free providers of information. While gender only has near significant effects (p = 0.0642) on how often an individual's Click value is above average for each Domain question, our data showed that gender has a significant effect on individuals' responses to Domain questions 3 (playing military games) and 13 (searching for a coffee shop), and almost significant effects on question 5 (playing a children's game) – with males tending to have a lower Click value than females, suggesting that males may have a lower sensitivity to these questions.

Previous research has shown that individuals' concerns over privacy across Domains are similar (Hann with health, travel, and financial web portals [15]) and willingness to disclose information is not influenced by the sensitivity of the information ([21]). We grouped individuals based on how they perceived the Domain questions, and categorized these questions into either the Neutral, Personal, Sex, or Flippy (both Neutral and Personal) Question groups. Our data showed that individuals consistently viewed Domains related to Travel Maps, Tutoring, Restaurant Searches, and Movie Ratings as if they were non-sensitive (average Click for males 3.147, females 3.559, combined 3.395), and other questions as if they were sensitive, i.e., those related to

Sex (average Click for males 10.395, females 11.028, combined 10.776). Assuming that the Sex questions are more sensitive than the Neutral questions, this set of results show that individuals' willingness to disclose information is influenced by the individual's perception of the sensitivity of the information. Further, our data show that these effects are in part correlated to gender differences. Specifically, whereas females tended to view questions 1 (using a web portal to search health symptoms) and 3 (playing military games) as Personal, males viewed them as both sensitive and non-sensitive; and whereas males viewed 13 (searching for a coffee shop) as non-sensitive, there were females who viewed it as non-sensitive and some as sensitive. It is interesting that there are females whose Click values for question 13 are within the Personal range, since the context of the question is similar to questions 11 (searching for a restaurant) and 12 (movie rating). One possible reason may be that while individuals do not necessarily visit restaurants near their homes and the movie rating service does not reveal their location, they may tend to visit coffee shops near their homes, and thereby would reveal their general location when using the service. Males' possible different habits in coffee shop visits or perception of location data may explain why they do not exhibit this sensitivity.

These results suggest that privacy policies must be constructed differently in the public and market arenas. With respect to public policy, privacy laws written as one-size-fits-all legislations may be optimal when compared to those tailored to specific groups and their privacy sensitivity. Just as mandatory seat belt laws and automobile and food safety standards are established to protect the population as a whole, as opposed to designing them for specific groups of individuals who may or may not care about their safety, one-size-fits-all privacy legislations can serve to protect consumers who may not have any alternatives to the service at hand, have insufficient knowledge to protect themselves, and/or do not recognize the appropriate short term-long term tradeoffs (i.e., providing information for instant gratification [1]).

The implications of our results for marketing and web developers, however, are not as clear-cut. For service Domains that are uniformly and consistently seen as not sensitive, a one-size-fits-all privacy policy is appropriate, but for service Domains whose sensitivity is dependent on the consumer segment, and whose viability is dependent on being able to track consumers' behavior, alternative approaches must be taken. Here, we offer two suggestions. For sensitive service Domains whose services are free to the consumers, but where the tracking of individuals' behaviors is necessary (for example, in order to provide contextually relevant advertisements), anonymity is still possible. Service providers can issue specific and unique usernames and passwords that are associated with each consumer and the records of his or her behavior, but that cannot be associated with individual identities after issued. Of course, since no association with individuals is possible, if the user forgets his username or password, it is not possible to recover either. In the case that the service is not free and the provider requires payment for the service, the establishment of payment intermediaries could still protect anonymity. This would start with mechanism like PayPal, but one where there is anonymity outside the payment firm; the firm would bill individuals and pay service providers, ensuring that the service provider would be unable to associate that account with an individual's identity outside the provider's system[7, 8]. This would of course require an enormous leap of faith in the security of the payment vendor.

Together, our results suggests that future research on privacy will require recognition that individuals' concerns over privacy span a spectrum of areas, and that their attitudes can vary across Domains; this differs from the idea that individuals can be classified as hoarders or free providers. Additionally, we now recognize that gender differences exist for behaviors in certain Domains. Public and corporate policies must be created to encompass gender, individual, and Domain differences, since a one-size-fits-all or a group-specific policy will not necessarily be optimal for the firms.

7. Limitations

There are several limitations to this study. First, indicating that the study concerns online privacy may have led individual subjects at the Wharton Behavioral Lab to be more sensitive to issues of privacy, leading to biased responses not reflective of their true attitudes towards privacy. Second, the dichotomy between what individuals say and what individuals do implies that their responses to the survey may not be indicative of their revealed or true and intrinsic preferences [1, 2, 22]. Third, our study population consisted primarily of college students (in line with previous studies on privacy, e.g., Hann [15], Hui et al. 2004 [21]) and so our findings may not apply to other segments of the population since students fall into a limited range on age, income, education, and awareness of issues in online technology, all of which may influence perception of risk.

Acknowledgements

The authors would like to thank Martin Asher and the Wharton Behavioral Lab for facilitating this study.

Table 1. Feelings towards information being recorded about internet usage (FTP)

Very willing	14 (5.86%)
Willing	24 (10.04%)
Indifferent	57 (23.85%)
Against	102 (42.68%)
Strongly against	42 (17.57%)

Table 2. Social Network and Internet Usage

Social network usage		Internet usage (hours)	
≤ 1x/week	11 (4.60%)	Mean	5.64
2-3x/week	16 (6.69%)	Median	5.00
1x/day	26 (10.88%)	Mode	5.00
2-3x/day	74 (30.96%)	Standard deviation	4.26
4-10x/day	68 (28.45%)		
> 10x/day	44 (18.41%)		

Table 3. Demographics

Gender	
Male	94 (39.33%)
Female	145 (60.67%)
Occupation	
Undergraduate	214 (89.54%)
Graduate	19 (7.95%)
Other students	6 (2.51%)
Age	
18-19	54 (22.59%)
20-23	152 (63.60%)
24-29	27 (11.30%)
>30	6 (2.51%)
Average age	21.5 years

Table 4. Distribution of subjects' average standard deviation (SD) over the Domain questions, 99% confidence interval of mean SD, and test of mean SD = 0.

Quantiles			**Moments**	
100.0%	maximum	6.226	Mean	4.231
75.0%	quartile	5.262	Std Dev	1.517
50.0%	median	4.609	Std Err Mean	0.098
25.0%	quartile	3.479	Upper 95% Mean	4.424
0.0%	minimum	0.000	Lower 95% Mean	4.038
			N	239

Test Mean=value

			t Test	**Signed-Rank**
Hypothesized Value	0			
Actual Estimate	4.2308	Test Statistic	43.1237	13630.50
DF	238	Prob > |t|	<.0001*	<.0001*
Std Dev	1.51672	Prob > t	<.0001*	<.0001*
		Prob < t	1.0000	1.0000

Table 5. Descriptive statistics of Difference (Above Average – Below Average) for males, females, and combined.

	Males	**Females**	**Males & Females**
Mean	-2.547	-1.701	-2.247
Standard Error	0.645	0.558	0.425
Median	-3	-2.5	-3
Mode	-6	-7	-7
Standard Deviation	6.284	6.694	6.569
Minimum	-13	-13	-13
Maximum	13	13	13
Sum	-242	-245	-537
Count	95	144	239

Table 6. Bivariate fit of Domain questions 3, 5, and 13 by Gender (Male=1, Female=0).

Linear Fit	**p-value**
Domain Question 3 Click Response = 7.7083333 - 2.5399123*Gender	0.0001*
Domain Question 5 Click Response = 5.6388889 - 1.175731*Gender	0.0554
Domain Question 13 Click Response = 4.5972222 - 1.2077485*Gender	0.0378*

Table 8. Number and percentage of hoarders and free-providers.

	Hoarders	**Free-Providers**
Males	2 (2.11%)	2 (2.11%)
Females	4 (2.78%)	1 (0.69%)
Combined	6 (2.51%)	3 (1.26%)

Table 7. Post-hoc Bonferroni comparison of each subject's set of responses to the Domain questions allowed for classification of individuals based on perception of Domain questions (in terms of Clicks). See survey[2].

	Males and Females (n=239)		Males (n=95)		Females (n=144)	
Question Group	**Average Clicks**	**Questions**	**Average Clicks**	**Questions**	**Average Clicks**	**Questions**
Neutral	3.540	7, 10, 11, 12, 13	3.196	7, 10, 11, 12, 13	3.559	7, 10, 11, 12
Flippy	5.172	5	4.737	1, 3, 5, 9	5.470	5, 9, 13
Personal	6.924	1, 3, 6, 8, 9	7.511	6,8	7.434	1, 3, 6, 8
Sex	10.776	2, 4	10.395	2, 4	11.028	2, 4

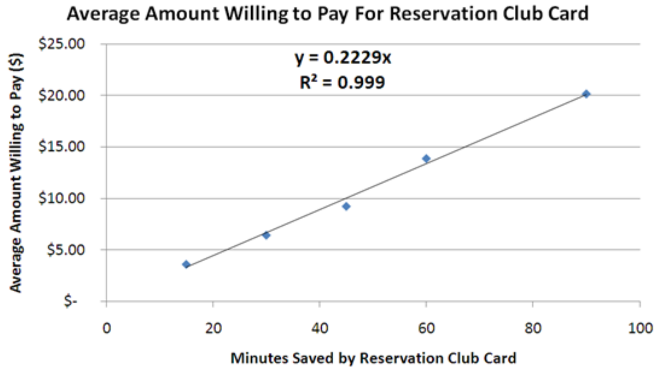

Figure 1. Linear regression; n=239, 5 observations

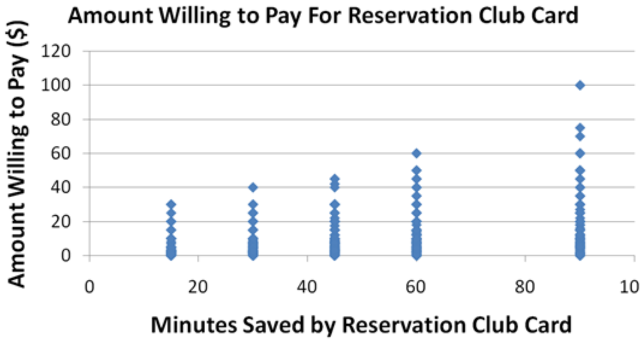

Figure 2. Linear regression; n=239, 1195 observations

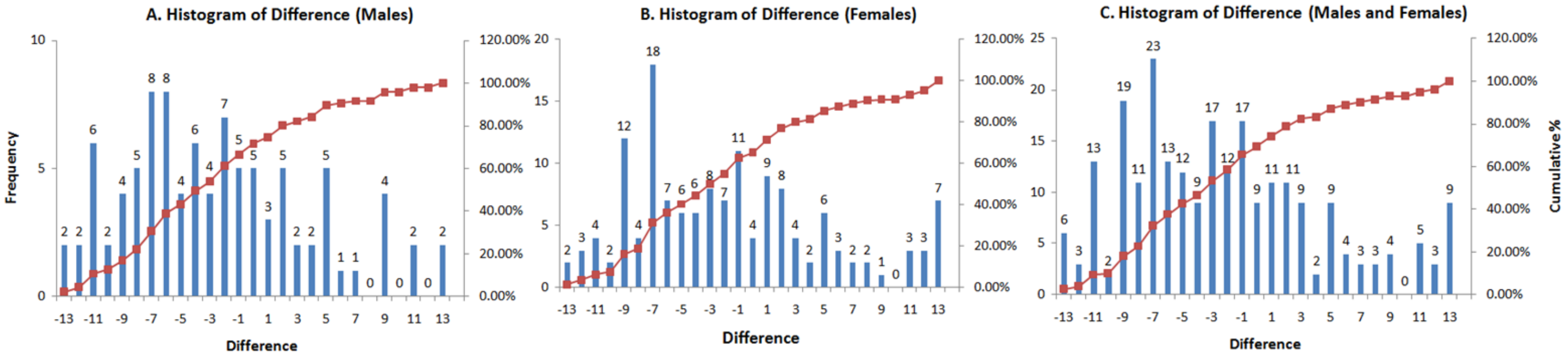

Figure 3. (A) Males n=95. (B) Females n=144 (C) Males & Females n=239.

Bibliography

1. Acquisti, A., *Privacy in electronic commerce and the economics of immediate gratification*, in *Proceedings of the 5th ACM conference on Electronic commerce*. 2004, ACM: New York, NY, USA.

2. Acquisti, A., L. John, and G. Loewenstein, *What is Privacy Worth?* 2010, Carnegie Mellon University: Pittsburgh.

3. Allen, D. *The Great Online Privacy Debate*. 2000; Available from: http://www.ebusinessforum.com/index.asp?doc_id=1785&layout=rich_story.

4. Awad, N.F. and M.S. Krishnan, *The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization.* MIS Quarterly, 2006. **30**: p. 13-28.

5. BBC. *Amazon's old customers 'pay more'*. 2000 [cited 2010 January 6]; Available from: http://news.bbc.co.uk/2/hi/business/914691.stm.

6. Blau, P., *Exchange and Power in Social Life*. 1964, New York: John Wiley & Sons, Inc.

7. Chaum, D. *Security without Identification Card Computers to make Big Brother Obsolete*. [cited 2010 May 30]; Available from: http://www.chaum.com/articles/Security_Wthout_Identification.htm.

8. Chaum, D., *Achieving Electronic Privacy*, in *Scientific American*. 1992, Scientific American. p. 96-101.

9. Chellappa, R., *Consumers' Trust in Electronic Commerce Transactions: The Role of Perceived Privacy and Perceived Security*. 2010.

10. Chellappa, R. and R. Sin, *Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma.* Information Technology and Management, 2005. **6**: p. 181-202.

11. Cvrcek, D., *et al.*, *A study on the value of location privacy*, in *Proceedings of the 5th ACM workshop on Privacy in electronic society*. 2006, ACM: Alexandria, Virginia, USA.

12. Eirinaki, M. and M. Vazirgiannis, *Web mining for web personalization.* ACM Transactions on Internet Technology 2003. **3**(1): p. 1 - 27.

13. European Commission, E. *Justice and Home Affairs - Data Protection*. Available from: http://ec.europa.eu/justice_home/fsj/privacy/overview/index_en.htm.

14. Hann, I.-H., *et al.*, *The Value of Online Information Privacy: An Empirical Investigation*. 2003.

15. Hann, I.H., *et al.*, *Overcoming Online information privacy concerns: An information-processing theory approach.* Journal of Management Information Systems, 2007. **24**(2): p. 13-42.

16. Harris Interactive, H. *First major post-9.11 privacy survey finds consumers demanding companies do more to protect privacy; public wants company privacy policies to be independently verified*. 2002; Available from: http://www.harrisinteractive.com/news/allnewsbydate.asp?NewsID=429.

17. Hoffman, D., T. Novak, and M.A. Peralta, *Information Privacy in the Market-Space: Implications for the Commercial use of Anonymity on the Web.* Information Society, 1999. **15**(4): p. 129-139.

18. Huberman, B.A., E. Adar, and L.R. Fine, *Valuating privacy.* Ieee Security & Privacy, 2005. **3**(5): p. 22-25.

19. Hui, K.-L. and I.P.L. Png, *The Economics of Privacy*, in *Handbook of Information Systems and Economics*, T. Hendershott, Editor. 2007, Elsevier Science.

20. Hui, K.L., *Privacy, Information Presentation, and Question Sequence*. 2004, National University of Singapore.

21. Hui, K.L., H.H. Teo, and T.S.Y. Lee, *The Value of Privacy Assurance: A Field Experiment*. 2004, National University of Singapore.

22. Kobsa, A., *Privacy-enhanced personalization.* Commun. ACM, 2007. **50**(8): p. 24-33.

23. Laudon, K.C., *Markets and Privacy.* Communications of the ACM, 1996. **39**(9): p. 92-104.

24. Laufer, R.S. and M. Wolfe, *Privacy as a Concept and a Social Issue: A Multidimensional Development Theory.* Journal of Social Issues, 1977. **33**(3): p. 22-42.

25. Milne, G.R. and M.E. Gordon, *Direct Mail Privacy-Efficiency Trade-Offs within an Implied Social-Contract Framework.* Journal of Public Policy & Marketing, 1993. **12**(2): p. 206-215.

26. Murthi, B.P.S. and S. Sarkar, *The Role of the Management Sciences in Research on Personalization.* Management Science, 2003. **49**(10): p. 1344-1362.

27. Noam, E., *Privacy in Telecommunications, Part III.* New Telecommunications Quarterly, 1995. **3**(4): p. 51-60.

28. OECD. *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*. Available from: http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html.

29. Posner, R.A., *The Economics of Privacy.* American Economic Review, 1981. **71**: p. 405-409.

30. Rust, R.T., P.K. Kannan, and N. Peng, *The Customer Economics of Internet Privacy.* Journal of the Academy of Marketing Science, 2002. **30**(4): p. 455-464.

31. Rust, R.T. and K.N. Lemon, *E-Service and the Consumer.* **International Journal of Electronic Commerce**, 2001. **5**(3): p. 85-101.

32. Stigler, G.J., *An Introduction to Privacy in Economics and Politics.* Journal of Legal Studies, 1980. **9**(4): p. 623-644.

33. Stone, E.F., *et al.*, *A Field Experiment Comparing Information-Privacy Values, Beliefs, and Attitudes Across Several Types of Organizations.* Journal of Applied Psychology, 1983. **68**(3): p. 459-468.

34. Stone, E.F. and D.L. Stone, *Privacy in Organizations: Theoretical Issues, Research Findings, and Protection Mechanisms*, in *Research Findings in Personnel and Human Resources Management*, K.M. Rowland and G.R. Ferris, Editors. 1990, JAI Press: Greenwich, CT. p. 349-411.

35. Tam, K.Y. and S.Y. Ho, *Web Personalization as a Persuasion Strategy: An Elaboration Likelihood Model Perspective.* Information Systems Research, 2005. **16**(3): p. 271-291.

36. Tseng, M.M. and F.T. Piller, *The customer centric enterprise : advances in mass customization and personalizaton*. 2003, Berlin ; New York: Springer. xii, 535 p.

37. Varian, H., F. Wallenberg, and G. Woroch, *The demographics of the do-not-call list.* IEEE Security & Privacy, 2005. **3**(1): p. 34-39.

38. Westin, A.F., *Privacy and Freedom*. 1967, New York.