

Analyzing Online Information Privacy Concerns: An Information Processing Theory Approach

Il-Horn Hann

*Marshall School of Business,
University of Southern California*

hann@marshall.usc.edu

Kai-Lung Hui

*Department of Information
Systems, Faculty of Business,
City University of Hong Kong*

is_lung@cityu.edu.hk

Sang-Yong Tom Lee

*College of Information and
Communications,
Hanyang University*

tomlee@hanyang.ac.kr

Ivan P.L. Png

*Department of Information
Systems, School of Computing,
National University of Singapore*

ipng@comp.nus.edu.sg

Abstract

The advent of the Internet has made the transmission of personally identifiable information common and often inadvertent to the user. As a consequence, individuals worry that companies misuse their information. Firms have tried to mitigate this concern in two ways: (1) offering privacy policies regarding the handling and use of personal information, (2) offering benefits such as financial gains or convenience. In this paper, we interpret these actions in the context of the information processing theory of motivation. Information processing theories, in the context of motivated behavior also known as expectancy theories, are built on the premise that people process information about behavior-outcome relationships. We empirically validate predictions that the means to mitigate privacy concerns are associated with positive valences resulting in an increase in motivational score. Further, we investigate these means in trade-off situation, where a firm may only offer partially complete privacy protection and/or some benefits.

1. Introduction

Rapid improvements in computing technologies and the advent of e-commerce have amplified public concern about privacy, especially on electronic networks. Website operators can profile browsers to an unprecedented degree and subsequently merge these profiles with other demographic data. Such an enriched data set can then be used by the company or sold to other parties. This information could benefit the customer by more precisely identifying her need. However, it could also be used to her detriment. For example, Amazon.com was suspected of engaging in differential pricing based on prior shopping information and other customer demographics for the

sales of DVDs.¹ Westin concludes: “There has been a well-documented transformation in consumer privacy attitudes..., moving concerns from a modest matter for a minority of consumers in the 1980s to an issue of high intensity expressed by more than three-fourth of American consumers in 2001” [29].

Violation of privacy occurs when an organization collects, stores, manipulates, or transmits personal information unbeknownst to the individual. But, not all of these activities surrounding personal information are necessarily perceived as invading privacy. A person submitting name, e-mail address, residential address and credit card information online for a purchase may not perceive the payment procedure as invasive, but as necessary to benefit from a service. However, the person may feel that her privacy is invaded if that information is then linked to other primary and secondary data such as browsing behavior on the website and demographic information. Yet, other people might welcome these efforts if this leads to price and product promotions. In general, perceptions of privacy infringements vary individually.

Privacy research has shown that this perception can be influenced. Prior literature has documented several effects. Naming the disclosure targets and the purpose of the relationship influences perception of privacy violations [23]. Culnan and Armstrong [6] find that privacy concerns can be addressed by explicitly stating the use of fair procedures for managing private information. In addition, Spiekerman et al. [24] show that in order to reduce product complexity, many participants, even some privacy fundamentalists, willingly share private information with a website. While their study does not measure the cost-benefit trade-off directly, it indicates that perceptions of privacy are context dependent. One important contribution of our study is to analyze such considerations.

¹(www.internetnews.com/ec-news/article.php/4_471541,9/28/00).

Some action of Internet businesses can certainly be interpreted as aiming to mitigate privacy concerns. A firm's promise to adhere to privacy policies regarding the handling and use of personal information may reduce perceptions of privacy violations. One way of reducing privacy concerns has been to offer incentives. Online firms have offered prizes in exchange for personal information. Even more widespread is customization of websites according to customer preferences. For example, Amazon's 'one-click-shopping' greatly simplifies the purchasing process for repeat customers.

Previous privacy research has, perhaps due to the nature of the subject, mostly focused on privacy concerns. We extend this discussion by introducing benefits that the information seeking organization has to offer, namely financial incentives and convenience. In this paper, we are interested in analyzing these means of mitigating privacy concerns. Our research objectives are as follows: First, we analyze privacy mitigation strategies from the viewpoint of information-processing theories of motivation. Specifically, we apply the framework of the expectancy theory of motivation ("expectancy theory"), which assumes that an individual's choice is determined by expectations about attaining desired outcomes. Based on this theory, we empirically validate predictions that the means to mitigate privacy concerns are associated with positive valences resulting in an increase in motivational score. We investigate these means in a trade-off situation, where a firm may only be able to offer partially complete privacy protection and/or promotions and/or convenience.

Previous research has highlighted the possibility that privacy preferences vary across individuals [14, 17]. Our second research objective is to identify distinct segments of Internet users, if any. We apply individual utilities, which we obtain from directly from the trade-off analysis, as a basis for segmentation and explore how individuals' privacy trade-offs depend on personal characteristics.

2. Theory and hypotheses

Information privacy has been defined as the individual's ability to control the collection and use of personal information [26]. This stipulates that privacy is viewed as control of information about the self. Control of personal information requires that an individual manages the outflow of information as well as the subsequent disclosure of that information to other third parties. In many experimental and organizational settings, people are found to perceive privacy invasions when they are not granted sufficient control on the solicitation, storage, use and disclosure of various types of personal information (see, e.g.,

Woodman et al. [30]). Such perceptions may deter them from taking part in transactions that involve personal information solicitation [5, 27].

The growth of electronic commerce has vastly increased the extent to which personal information is acquired, exchanged, and used by vendors. Internet consumers now face additional risk with regard to their personal information. Consequently, the concern about information privacy has increased dramatically [6, 21].

Laufer and Wolfe [17] suggest that individuals perform a "calculus of behavior" to assess the cost and benefit of providing personal information. Individuals explicitly consider the trade-off between the merits of interactions and potential consequences, i.e., maximizing the difference of benefits and costs. Based on this approach, we use an information-processing theory of motivation to analyze the extent of individuals' online information privacy concerns. Information-processing theory focuses on the cognitive process that occurs before an action is undertaken or a choice is made. We employ the expectancy theory framework to give more structure to the question of how individuals make decisions regarding privacy in an online setting. Originally formulated by Vroom [28], expectancy theory is a framework to explain how an individual chooses between alternative forms of behavior. It proposes that an individual considers the outcomes associated with various levels of performance as well as the likelihood of achieving these outcomes. When deciding among alternatives, she selects the option with the greatest *motivational score*.

The motivational force for a behavior or action is a function of three distinct perceptions: expectancy, instrumentality, and valence, i.e., $\text{Motivational Score} = f(\text{Expectancy, Instrumentality, Valence})$. Expectancy is a probability assessment which reflects the individual's belief that a given level of effort will result in a particular level of performance. Instrumentality refers to the subjective assessment that a given performance level will lead to one or more outcomes. Valence refers to the value that an individual places on a given outcome.

To illustrate, we discuss the expectancy theory in the context of a person considering to register at a financial website to trade stocks, to stay current about the value of her stock holdings, and to collect information about the companies in her stock portfolio. For these purposes, the financial website may require the person to submit e-mail address, name, residential address, banking information, social security number, and the names of the stocks and quantities owned.

- Motivation is seen as the force that directs behavior. It deals with the question of choice among competing alternatives. In the case of online information privacy, we investigate which financial

website, after controlling for content, will be chosen given that the individual may choose from various sites with different mixes of privacy policies, convenience, and financial benefits.

- Expectancy is the weight that characterizes the perceived effort-performance relationship. It is the expectancy that one's effort will lead to the desired performance. In the financial website example, she is required to provide detailed information about her stock holdings and specify particular information needs in order to become an account holder.
- Instrumentality is the weight that describes the perceived performance-outcome relationship. It characterizes the belief that if performance expectations are met, then the individual will receive a better outcome. In the example, successfully obtaining an account leads to positive outcomes which may include convenience when checking the stock portfolio and becoming updated on relevant company news, and financial benefits through promotions. Regarding online privacy, an important outcome is the commitment of the financial website to protect personal information according to its privacy policy.
- Valence refers to the value the individual personally places on the outcome. This is a function of his or her needs, goals, and values. Depending on the outcome, the valence can be positive or negative. For the example, positive valences include the convenience of having all relevant information without repeated search and the financial gain for having signed up with this website. In the context of online privacy, positive valence includes the feeling of security due to the specifics of the privacy policy. A website with an incomplete privacy policy may generate negative valences such as the potential to be vulnerable to others or to be exploited by others.

A choice, e.g., a financial website, can typically be characterized by several dimensions that describe the privacy policy, the convenience, and the financial gains. Each of these dimensions is associated with a value for expectancy, instrumentality, and valence. An individual will rank the alternatives and choose the one with the greatest motivational force. More formally, for an alternative with n dimensions, expectancy theory assumes a score that is computed as:

$$\text{Motivational Score} = \sum_{i=1}^n \underbrace{(E \rightarrow P)}_{\text{Expectancy}} \times \underbrace{(P \rightarrow O)}_{\text{Instrumentality}} \times \underbrace{V_i}_{\text{Valence}}$$

Applying this model in the context of online information privacy, we note that expectancies, instrumentalities and valences are specific to each person. However, this model also implies that the outcome (O) variable can be varied to increase the individual's motivational score. If the firm can effectively use outcomes that are associated with positive valences, it

can increase the motivational score and effectively decrease privacy concerns. Important to our approach is that given a certain motivational score for fixed effort, performance, and a manipulated outcome, we can elicit the valences that are associated with the means to mitigate online privacy concerns.

Our first research objective addresses how firms can mitigate privacy concerns by managing outcomes and associated valences within the expectancy model. We first must establish the various outcomes that are associated with valences. Smith et al. [23] identified four specific privacy concern dimensions which represent the cognitive state of consumers towards corporate use of information. These four privacy dimensions are *collection*, *error*, *unauthorized secondary use*, and *improper access*.²

Collection refers to the concern that "extensive amounts of personally identifiable data are being collected and stored in databases". Error refers to the concern that "protections against deliberate and accidental errors in personal data are inadequate". Unauthorized secondary use refers to the concern that "information is collected for one purpose but is used for another, secondary purpose". Improper access refers to the concern that "data about individuals are readily available to people not properly authorized to view or work with this data" [23], Table 2). We use these dimensions (collection, error, unauthorized secondary use, and improper access) as the basis for potential outcome variables, which will determine the instrumentalities.

Consistent with the expectancy theory, businesses can use the dimensions of privacy as an outcome to increase the motivational score of the website. Specifically, individuals link performance (successfully obtaining an account) with outcome (assurance by privacy policy). For example, a person may give a privacy policy that restricts secondary use a higher instrumentality (weight) and hence a greater motivational score than a policy that omits protection from secondary use. Therefore, we hypothesize:

Hypothesis H1a: Specification of privacy policy dimensions increases the motivational scores.

An individual's motivational score may also be affected by extrinsic, positive reinforcements. Resource exchange theory characterizes six categories of interpersonal resources (love, status, information, money, goods and services), and it is well known that people are willing to trade one resource for another [9]. Prior research has shown that this resource framework is quite general, and it can be applied to analyze different types of marketing transactions that involve interpersonal relationships and resource exchanges [2].

² These were subsequently validated by Stewart and Segars [25].

Many e-commerce websites offer monetary reward or added service convenience to customers who disclose certain personal information. Both money and service are primary elements in the resource exchange theory [9], and they may act as positive incentives and resources for online firms to exchange for personal information. Applying this to the expectancy theory, this implies that a firm can offer financial gains and convenience to increase the motivational score. As in the first hypothesis, the firm influences the instrumentalities, by creating outcomes (financial gain, convenience) that are associated with positive valences. Hence, we state:

Hypothesis H1b: Offering financial gains and convenience increases the motivational scores.

Previous research on information privacy was mostly concerned with identifying key dimensions of privacy concerns [23, 25] and how perceptions of privacy infringements can be influenced [6, 15, 27]. Academic research has tended to overlook differences in privacy preferences. However, from a business viewpoint, it is important to identify any systematic differences in preferences and their distribution. Firms could then address the respective privacy concerns appropriately. Hence, we are interested in a characterization of the trade-offs of outcome valences.

Social exchange theory posits that individuals' choice of actions are influenced by their personal experience; the more frequently a person was rewarded by a particular stimulus in the past, the more likely she would be to perform an action that leads to the stimulus [14]. Further, the extent of privacy calculus posited by [17] depends on personal and environmental characteristics. Stone and Stone's [26] expectancy theory-driven privacy model includes individual and social factors such as personality and previous learning. In the context of information privacy, these theories posit that individuals may vary in their judgments towards online privacy. As individuals' expectations about valences are similar, groups may be identified. Past opinion surveys have divided the U.S. population into a majority of "privacy pragmatists" and minorities of "privacy fundamentalists" and "privacy unconcerned" [29]. Therefore, we hypothesize:

Hypothesis H2: Individuals will reveal systematic differences in privacy preferences.

3. Methodology and experiment

Previous research in IS have used judgment models based on within-person based approaches of expectancy theory (e.g., [3, 8]). Typically, an individual is provided with a set of variables, which are used to arrive at a particular decision. The within-person approach requires presentation of multiple cases with unique combinations of variables with each being

individually evaluated. Our approach, the *conjoint analysis* method, shares these characteristics, but is rooted in decision theory. Conjoint analysis grew out of the area of conjoint measurement, which was first developed in economics [7] and psychology [18]. The technique provides a measurement method for decision-making contexts where multiple dimensions must be taken into account.

Conjoint analysis presents test subjects with a set of *alternatives* (stimuli). Each stimulus consists of particular *levels* of various *dimensions* (attributes). In the context of online privacy, *dimensions* of a website include the dimensions of privacy, convenience, and monetary reward. Each *dimension* is represented by two or more *levels*. For example, 'unauthorized secondary use of private information' and 'no unauthorized secondary use of private information' present two levels of the dimension of secondary use. The subject is asked to rank the stimuli according to her own preferences. The conjoint analysis technique decomposes rankings of alternatives into components based on the dimensions of the alternatives. A numerical utility (also called *part-worth*) value is computed for each level of each dimension.

To keep the conjoint tasks to a manageable size, Green and Srinivasan [11] recommend that the number of attributes be limited to six or fewer. Following Green and Krieger [10], we used focus groups prior to the conjoint study. Specifically, we conducted three focus group discussions with upper-division undergraduate and graduate students in the US and Singapore to identify the key benefits that they expected from registration with websites and suitable attribute levels. The focus groups suggested that individuals clearly value direct monetary savings. In addition, they also identified convenience as another important benefit of providing personal information to a website. The focus groups identified two sources of convenience benefits; the explicit time saving per session and the expected visit frequency to the website. Hence, we operationalized convenience by "expected visit frequency/total time savings" in our experiment.

As mentioned before, we considered the four privacy dimensions identified by Smith et al. [23]: collection, error, unauthorized secondary use, and improper access. For our purpose, collection is a necessary antecedent to the three other dimensions. Error, unauthorized secondary use and improper access of information cannot happen without *ex ante* collection of personal information. Further, individuals' concerns on the other three dimensions are a direct function of the amount of information collected; the more information a website collects, the higher should be the concerns with error, unauthorized secondary use, and improper access of information. Therefore, we controlled for the collection of

information and manipulated the other three dimensions.

Taken together, our conjoint study assessed trade-offs among five dimensions; two benefit outcomes and three privacy outcomes. Based on the insight from our focus groups, we created three outcomes for monetary reward (\$5, \$10 and \$20) and visit frequency/time savings (monthly, weekly and daily). The outcomes of the three privacy dimensions (error, unauthorized secondary use and improper access of information) were manipulated by the presence (or absence) of proper information handling and access procedures.

Based on these five dimensions and their treatment levels, there were a maximum of $3 \times 3 \times 2 \times 2 \times 2 = 72$ conjoint stimuli. To avoid asking subjects to rank too many alternatives, we selected 18 stimuli based on an optimal orthogonal design [1]. For example, one particular stimulus was a website that provided a \$5 monetary reward in return for personal information and which the subject visited once a month with a total time savings of 24 minutes per year. Further, the website had no error correction procedure, no policies to prevent unauthorized secondary use, and no policies to prevent improper access to information. Our conjoint analysis asked subjects to rank 18 websites, which represented different combinations of benefits and privacy protection.

The basic model underlying the conjoint analysis is a main-effects ANOVA which computes utilities such that the rank ordering of the sums of each alternative's set of part-worths is the same as the original rank ordering of the alternatives. The basic model is:

$$\begin{aligned} \text{Ranking} = & \alpha + \sum_{j \in \{\$10, \$20\}} PW_{\text{Fin. Rew. } j} * \text{Outcome}_{\text{Fin. Rew. } j} \\ & + \sum_{k \in \{\text{dly, wkly}\}} PW_{\text{Freq. } k} * \text{Outcome}_{\text{Freq. } k} \\ & + PW_{\text{Error}} * \text{Outcome}_{\text{Error}} + PW_{\text{Sec. Use}} * \text{Outcome}_{\text{Sec. Use}} \\ & + PW_{\text{Unauth. Access}} * \text{Outcome}_{\text{Unauth. Access}} + \varepsilon \end{aligned}$$

The part-worth (PW) is the marginal utility of the dimension in the individual's ranking of the conjoint stimuli. To estimate the part-worths, we use least-squares regression with the subjects' rankings (from 1 to 18) as the dependent variable and indicators of the various levels of the two benefit outcomes and three privacy dimension outcomes as the independent variables. The coefficient of each independent variable is interpreted as the part-worth.

The expectancy based theory can be related to the selected research methodology, the conjoint analysis. As previously discussed, the motivational score is a function of expectancy, instrumentalities, and valences. In our research design, the expectancy weights were fixed (set to one), the instrumentality weights were set to zero or one (depending on the outcomes described by the conjoint stimuli), and the motivational score was the actual ranking of the conjoint stimuli (which

represents the set of specific outcomes for the five dimensions). The valences were inferred through the conjoint analysis methodology, given fixed expectancy and instrumentality weights.

The conjoint analysis regressed the subjects' rankings of the conjoint stimuli on dummy variables, which represent the various levels of the two benefit outcomes and three privacy dimension outcomes. The coefficient of each independent variable would be the part-worth corresponding to the outcome of that dimension. In the context of the expectancy theory, the coefficients of the dummy variables represent the valences of the outcomes. In the conjoint analysis, these coefficients are termed part-worths. In order to be consistent with the regression language, we will use the term 'coefficient' to mean part-worth from here on.

The relationship between the expectancy based theory of motivation framework for privacy and the conjoint analysis can be depicted as follows:

$$\begin{aligned} \text{Ranking} = & \alpha + \sum_{j \in \{\$10, \$20\}} \underbrace{\text{Outcome}_{\text{Fin. Rew. } j}}_{\text{Instrumentality for Fin. Rew. } j} * \underbrace{PW_{\text{Fin. Rew. } j}}_{\text{Valence for Fin. Rew. } j} \\ & + \sum_{k \in \{\text{dly, wkly}\}} \underbrace{\text{Outcome}_{\text{Freq. } k}}_{\text{Instrumentality for Freq. } k} * \underbrace{PW_{\text{Freq. } k}}_{\text{Valence for Freq. } k} \\ & + \underbrace{\text{Outcome}_{\text{Error}}}_{\text{Instrumentality for Error}} * \underbrace{PW_{\text{Error}}}_{\text{Valence for Error}} \\ & + \underbrace{\text{Outcome}_{\text{Sec. Use}}}_{\text{Instrumentality for Sec. Use}} * \underbrace{PW_{\text{Sec. Use}}}_{\text{Valence for Sec. Use}} \\ & + \underbrace{\text{Outcome}_{\text{Unauth. Access}}}_{\text{Instrumentality for Unauth. Access}} * \underbrace{PW_{\text{Unauth. Access}}}_{\text{Valence for Unauth. Access}} + \varepsilon \end{aligned}$$

In order to control for industry effects, we posed the conjoint stimuli in three settings – financial, healthcare, and travel. Within each of the three industries, we controlled for the degree of information collection by telling the subjects that all 18 stimuli (that is, hypothetical websites) requested the same set of personal information from the subjects. The personal information consisted of name, home address, phone number, e-mail address, credit card information, and some industry-specific information. The industry-specific information were: for travel websites, occupation, travel purpose, destination and frequency of travel, and frequent flyer numbers; for healthcare websites, medical history, drug allergies, and prescription record; and for financial websites, household income, stock portfolio, and previous stock trading experience.

Each subject was randomly assigned to one of the three industry settings and asked to rank the 18 stimuli (websites) according to her own preferences. In other words, the benefit/privacy dimensions were within-subject factors whereas industry was a between-subject factor. To capture the background of the subjects, we also included demographic questions regarding their gender, age, Internet usage and previous experience with invasion of privacy.

To strengthen the external validity of our study, we conducted the conjoint experiment in the USA and Southeast Asia. The U.S. subjects were senior undergraduate students from a major Eastern university. The Asian sample comprised upper-division undergraduate computing students at a major university. Table 1 presents some descriptive statistics about our subjects.

The experiment proceeded as follows. First, all subjects completed the demographic questions. Then, the experimental task and the meanings of the five dimensions were explained. The subjects ranked the 18 stimuli based on their personal preferences. In the U.S. sample, 84 students completed the experiment. Among them, 35 students received course credit, while the other students were compensated with US\$7 each.³ In Singapore, 184 subjects completed the experiment for course credit. We collected 268 responses in total.

4. Results & discussion: conjoint analysis

Table 2 reports the means of the coefficients (valences or part-worths) for the U.S. and Singapore subjects. Further, we calculated the relative importance of each dimension as the coefficient corresponding to the maximum level of that dimension divided by the sum of the coefficients corresponding to the maximum levels of all five dimensions. The relative importance indicates the impact of a specific outcome relative to other outcomes. We expressed relative importance as a percentage. Note that the coefficients and relative importance for the U.S. and Singapore samples are not directly comparable as the monetary rewards were framed in the respective local currencies.⁴

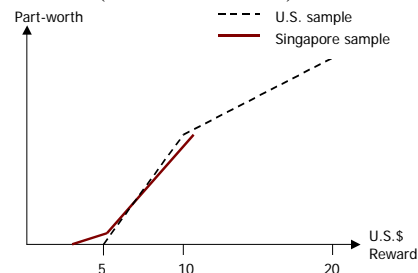
We first examined whether the responses from the subjects differed across the three industries. We conducted one-way analysis of variance (ANOVA) and pairwise t-tests to compare the coefficients of each outcome across the industries. The results suggested that the coefficients were not statistically different across financial, healthcare and travel websites. Accordingly, in all subsequent analyses, we pooled the data across industries.

The coefficients (valences) on the outcomes of the privacy dimensions (error, improper access, and unauthorized secondary use) show strong support for Hypothesis 1a. A positive coefficient for a specific privacy dimension, which differs significantly from zero, indicates that subjects on average prefer a website with this privacy feature. For example, in the US sample, a privacy policy which restricts improper access would raise its motivational score by 3.007 (out of 18). Referring to Table 2, the coefficients for

protection against all three privacy concerns were statistically significant at the 1% level in both samples. Among U.S. subjects, the coefficient for error review and editing was 2.968, while that for disallowing unauthorized secondary use was 2.118. For Singapore subjects, the coefficients for error review and editing, restricting improper access, and disallowing unauthorized secondary use were 1.787, 3.374 and 4.604 respectively.

Consistent with previous research [19], we found that, Singapore subjects were relatively more concerned about improper access and unauthorized secondary use than errors in storing information. However, the U.S. subjects exhibited less concern for unauthorized secondary use than errors in storing information. Despite the discrepancy in relative preferences toward the different privacy protections across the two samples, our conjoint experiment confirmed previous findings that individuals are highly concerned about information privacy [6].

Our results also indicate support for Hypothesis 1b; i.e., that outcomes such as monetary rewards are associated with positive valences and hence increases the motivational score. For the U.S. sample, the coefficient for a US\$20 reward was 3.141 and was statistically significant. This means that a website offering a US\$20 reward for personal information could increase the motivational score by 3.141 as compared to an otherwise identical website offering the base level US\$5 reward. Also, the coefficient for a US\$10 reward was 1.327 and significant. For the Singapore sample, the coefficient for a S\$20 reward was 1.388 and was statistically significant. At the prevailing exchange rate, S\$20 was equivalent to US\$10.80, hence it was not surprising that the coefficient was much less than the US\$20 coefficient in the U.S. sample (3.141). Interestingly, the S\$20 coefficient among Singapore subjects (1.388) was very close to the US\$10 coefficient among U.S. subjects (1.327). This result arose even though the base-level rewards were different in the two samples (S\$5 and US\$5 respectively). The coefficient for a S\$10 reward in the Singapore sample was 0.232 but not statistically significant. The subjects were willing to trade away privacy protection or convenience only when the monetary reward exceeded a threshold, which lay between S\$10-20 (US\$5.40 – 10.80).



³ We found no statistically significant difference in part-worths between those who received course credit and those compensated with US\$7. Hence, we pooled both groups into a single sample.

⁴ At the time, one Singapore dollar = 54 US cents.

Figure 1: Valences for Monetary Reward

Taken together, the results from the U.S. and Singapore samples suggest that a sufficiently large monetary reward did significantly increase the relative attractiveness of a website independent of its privacy policy. Further, when the monetary reward was relatively low (as in the Singapore sample), the marginal utility of the reward was increasing, and when the monetary reward was relatively high (as in the U.S. sample), the marginal utility tended to decrease. These results indicate that the attractiveness of a monetary reward relative to privacy protection or convenience might follow the “S”-shape as shown in Figure 1. The results are consistent with economic analysis that utility functions tend to be non-concave [13].

We also find support for the second part of Hypothesis 1b; i.e., that outcomes such as service convenience, operationalized by visit frequency/time savings, are associated with positive valences and hence increase the motivational score. Referring to Table 2, in the U.S. sample, the coefficient (valence) for weekly visit was significant at the 5% level, but the coefficient for daily visit was significant only at the 10% level. Further, the coefficients for weekly and daily visits were not significantly different. In the Singapore sample, the coefficients for visit frequency/time savings were generally more significant. However, as with the U.S. subjects, the effect due to weekly visit was not significantly different from that due to daily visit.

From the results of both samples, we conclude that there is some evidence that subjects value convenience. The evidence is stronger among Singapore subjects than U.S. subjects. Further, once the subjects expected to visit a particular website sufficiently frequently (at least once a week), more frequent visits did not seem to affect the subjects’ preferences. The coefficients and relative importance associated with visit frequency/time savings among U.S. and Singapore subjects were very close. In both samples, these were much lower than the coefficients and relative importance for the other dimensions. Apparently, among our subjects, convenience was only a minor factor when evaluating websites. By contrast, monetary reward and privacy protection were perceived to be much more important.

5. Results & discussion: cluster analysis

To address our second set of research questions, whether individuals systematically differ in their trade-off between benefits of disclosing personal information and privacy concerns, we applied cluster analysis [10]. This technique groups subjects into distinct segments according to the similarity of their estimated coefficients for the various dimensions. In the present case, we apply cluster analysis to segment the subjects

according to their estimated valences over the various benefits and dimensions of privacy protection.

Specifically, we applied hierarchical cluster analysis using average between-group linkage with (dis)similarity measured by the squared Euclidean distance to both the U.S. and Singapore samples. The hierarchical method was preferred because we had no a priori information on the number of clusters and initial cluster seeds/centers [12]. We used a distance measure for (dis)similarity as all the valences (the inputs to the cluster analyses) were derived from a common scale, viz., the website rankings.

For each sample, we began the analysis with every subject constituting a separate cluster. We then examined the percentage drops in the similarity coefficient as clusters were progressively merged. In both the U.S. and Singapore samples, we stopped at three clusters as further combination of any two clusters resulted in a sharp drop in similarity, a recommended stopping rule [12] (p. 499). Table 3 reports the three clusters, their sample sizes, and the respective mean coefficients.⁵

Overall, we found strong support for Hypothesis 2. Consistent across the two samples, the majority of the subjects formed a cluster that could be characterized by a high value on information privacy. Specifically, 72% of the U.S. subjects and 84% of the Singapore subjects exhibited relatively high coefficients for protection against error, improper access, and unauthorized secondary use of their personal information. By contrast, their coefficients on monetary reward and visit frequency/time savings were relatively low. We label this group of subjects as “privacy guardians” – people who attach a relatively high value to information privacy.

The next largest cluster consisted of subjects who attached a relatively high value to monetary reward. We call them “information sellers”, as they tend to “sell” personal information with little regard for convenience or website privacy policies.

The smallest cluster comprised subjects who focused exclusively on convenience (operationalized by visit frequency/time savings). In fact, their coefficients for visit frequency/time savings were so high that their preferences over alternative websites could almost be predicted by visit frequency/time savings alone. We call these subjects “convenience seekers” – people who prefer convenience with little regard for money or website privacy policies.

Across the three clusters, we observe very different attitudes toward benefits and privacy. The privacy guardians prefer protection, but they still value monetary reward (the mean coefficient for monetary

⁵ We excluded a small number of subjects who could not be classified into any of the three clusters.

reward was significantly different from zero). Only the convenience seekers value convenience; for all other clusters, the coefficients for visit frequency/time savings were insignificant. Among the three privacy concerns, only unauthorized secondary use was significant in all three clusters.

Based on opinion surveys, 12% of the U.S. population has been characterized as “privacy unconcerned” [29]: “for 5 cents off, they will give you any information you want about their family, their lifestyle, their travel plans, and so forth”. Interestingly, we found that 12.5% of the U.S. sample were “information sellers”. However, our evidence is that information sellers demand a great deal more than “5 cents off.” Indeed, this point distinguishes our analysis from opinion surveys: we can estimate the dollar amount that information sellers must be paid for their information.

Our analysis revealed a cluster not previously identified -- convenience seekers, people who would “sell” their personal information for convenience rather than money. For the remainder of the U.S. population, Westin [29] differentiated between privacy pragmatists (63%) and privacy fundamentalists (25%). Our cluster analysis did not find such a distinction. We did detect some evidence among U.S. subjects that the privacy guardians could be further segmented, with each sub-segment placing relatively greater weight on one of the three privacy concerns.

Having identified three clusters, we investigated whether cluster membership depended systematically on particular demographic variables. We first sought systematic differences between information sellers and privacy guardians. Among the U.S. subjects, we found that information sellers had significantly more prior experience of providing personal information to websites than privacy guardians ($t = 3.115$, $p < 0.01$). The information sellers’ greater prior experience was consistent with their relatively high coefficients for money. Among the Singapore subjects, there was no significant difference between information sellers and privacy guardians in terms of prior experience of providing personal information to websites.

We next investigated systematic differences between convenience seekers and privacy guardians. Among the U.S. subjects, convenience seekers were much more accepting of cookies than privacy guardians ($t = 4.282$, $p < 0.001$). Specifically, convenience seekers were less concerned about cookies, and they typically accepted all cookie manipulations from websites without warning. By contrast, the majority of privacy guardians requested to be warned about cookies. Many of them even configured their browsers to reject all cookies. The convenience seekers’ greater acceptance of cookies was consistent with their relatively high coefficients for

visit frequency/time savings. Among the Singapore subjects, the convenience seekers were also less concerned about the use of cookies than the privacy guardians ($t = 6.954$, $p < 0.001$). This result was consistent with the preferences of the U.S. sample.

Overall, we found some evidence that information sellers had more prior experience of information provision than privacy guardians, and strong evidence that convenience seekers were more accepting of cookies than privacy guardians.

6. Concluding remarks

In this paper, we have analyzed strategies that might mitigate online information privacy concerns. To that end, we applied the expectancy based theory of motivation to define our research questions and hypotheses. Further, we linked the expectancy base theory framework to the chosen methodology, the conjoint analysis. We empirically validated hypotheses based on the expectancy based theory framework that stipulates that individuals have positive valences for privacy protection policies, which increases the motivational score. Similarly, we confirmed the hypotheses that benefits such as financial rewards and convenience have a positive valence and increase the motivational score.

One important implication of this research is that organizations have the means to actively manage the privacy concerns of Internet users. Our results distinctly show that privacy policies are valued by users. Hence, organizations can capitalize on this, by stating their privacy policy more prominently. Similarly, investing in well-known privacy policy standards such as BBBOnLine or TRUSTe may increase the willingness to part with private information. One benefit of increasing convenience is increased value offering through personalization [4]. In addition, it appears that convenience also has a benefit that has been overlooked, namely mitigating privacy concerns. Perhaps the least surprising result, given the subject group (students), is that financial incentives are also a persuasive means to elicit private information. This finding is consistent with anecdotal evidence that has shown that people are willing to disclose personal information for gifts and catalogs [20], and even a \$100 drawing [16].

Our secondary set of research questions investigated the differences in privacy preferences. By applying cluster analysis to the subjects’ marginal rankings of the various benefits and concerns, we find that our subjects can be categorized into three distinct segments – privacy guardians, information sellers, and convenience seekers. The majority of subjects were relatively sensitive to online information privacy concerns (“privacy guardians”). By contrast, a smaller

proportion was relatively willing to provide information in exchange for money (“information sellers”), and an even smaller proportion was relatively willing to provide information in exchange for convenience (“convenience seekers”). All of the preceding results were robust in the sense that they held in both the U.S. and Singapore samples.

The immediate implication is that firms with online presence must differentiate their services to serve these distinct segments to best meet the needs of segments with differing trade-offs among money, convenience, and privacy concerns. Convenience seekers will be the first to register with a website if it simplifies website navigation or enables personalized content. Businesses can exploit this by offering them the opportunity to provide personal information to customize the website and simplify the shopping experience. Information sellers are distinguished from privacy guardians by prior experience of information provision. This customer type cannot be lured to provide personal information by offering them convenience. To the extent that businesses cannot observe an individual’s prior experience, they must use indirect methods to induce segmentation by self-selection [22]. Businesses could use monetary rewards to attract information sellers to provide personal information. Preferably, businesses would seek convenience seekers first before enticing information sellers. By elimination, the individuals who do not respond to either monetary reward or convenience would be privacy guardians.

7. References

- [1] Addelman, Sidney “Orthogonal Main-Effect Plans for Asymmetrical Factorial Experiments,” *Technometrics*, 4(1), February 1962.
- [2] Brinberg, David and Ronald Wood “A Resource Exchange Theory Analysis of Consumer Behavior,” *Journal of Consumer Research*, vol. 10, no. 3, December 1983, pp. 330-338.
- [3] Burton, F Greg, Yi-Ning Chen, Varun Grover, and Kathy Stewart “An Application of Expectancy Theory for assessing user motivation to Utilize an Expert System,” *Journal of Management Information Systems*, Vol. 9, No. 3, Winter 1992-1993, pp. 183 – 199.
- [4] Chellappa, Ramnath and Sin, Raymond “Personalization versus Privacy: New Exchange Relationships on the Web,” Working Paper, USC, May 2002.
- [5] Culnan, Mary J. “How Did They Get My Name? An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use,” *MIS Quarterly*, 17(3):341-363, Sept 1993.
- [6] Culnan, Mary J. and Pamela K. Armstrong “Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation,” *Organization Science*, vol. 10, no. 1, January-February 1999, pp. 104-115.
- [7] Debreu, G. “Topological methods in cardinal utility theory.” In S. Karlin K. J. Arrow and P. Suppes, editors, *Mathematical methods in the Social Sciences*, pages 16{26. Stanford University Press, Stanford, 1960.
- [8] DeSanctis, G. (1983). Expectancy theory as an explanation of voluntary use of a decision-support system. *Psychological Reports*, 52, pp. 247-260.
- [9] Foa, Uriel G. “Interpersonal and Economic Resources,” *Science*, vol. 171, 1971, pp. 345-351.
- [10] Green, Paul E. and Abba M. Krieger “Segmenting Markets with Conjoint Analysis,” *Journal of Marketing*, 55(4), Oct. 1991, pp. 20-31.
- [11] Green, Paul E. and V. Srinivasan “Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice,” *Journal of Marketing*, 54(4), 1990, pp.3-19.
- [12] Hair, Joseph F., Ronald L. Tatham, Rolph E. Anderson and William C. Black. *Multivariate Data Analysis with Readings*. Prentice Hall, 1998.
- [13] Hartley, Roger, and Lisa Farrell “Can Expected Utility Theory Explain Gambling?” *American Economic Review*, vol. 92, no. 2, June 2002, pp. 613-624.
- [14] Homans, George Caspar. *Social Behavior: Its Elementary Forms*. Harcourt Brace Jovanovich, Inc. 1974.
- [15] Hui, K.L., H.H. Teo, and S.Y.T. Lee “The Value of Privacy Assurance: A Field Experiment,” Working paper, NUS, 2004.
- [16] Jupiter Media Metrix “Seventy Percent of US Consumers Worry About Online Privacy, But Few Take Protective Action,” Press Release, June 3, 2002.
- [17] Laufer, Robert S. and Maxine Wolfe “Privacy as a Concept and a Social Issue: A Multidimensional Developmental Theory,” *Journal of Social Issues*, 33(3), 1977, pp. 22-42.
- [18] Luce, R. D. and J. W. Tukey “Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement,” *Journal of Mathematical Psychology*, No. 1, 1964, 1-27.
- [19] Milberg, Sandra J., Sandra J. Burke and H. Jeff Smith “Values, Personal Information Privacy, and Regulatory Approaches,” 1995, *Comm.s of the ACM*, 38(12):65-74.
- [20] Oberndorf, Shannon “Registering for Success,” *Catalog Age*, vol. 16, no. 13, 1999, pp. 47-48.
- [21] Phelps, Joseph, Glen Nowak and Elizabeth Ferrell “Privacy Concerns and Consumer Willingness to Provide Personal Information,” *Journal of Public Policy and Marketing*, vol. 19, no. 1, Spring 2000, pp. 27-41.
- [22] Png, Ivan. *Managerial Economics*, Malden, MA: Blackwell, 2002.
- [23] Smith, H. Jeff, Sandra J. Milberg and Sandra J. Burke “Information Privacy: Measuring Individuals’ Concerns About Organizational Practices,” *MIS Quarterly*, vol. 20, no. 2, June 1996, pp. 167-196.
- [24] Spiekermann, Sarah, Jens Grossklags, and Bettina Berendt “E-privacy in 2nd Generation E-Commerce: Privacy Preferences vs. Actual Behavior.” Proceedings of the 3rd *ACM Conference on Electronic Commerce*, October 2001, Tampa, FL., pp. 38-47.
- [25] Stewart, Kathy A. and Albert H. Segars “An Empirical Examination of the Concern for Information Privacy Instrument,” *Information Systems Research*, 13(1):36-49, Mar 2002.
- [26] Stone, Eugene F. and Dianna L. Stone “Privacy in organizations: theoretical issues, research findings, and protection mechanisms,” *Research in Personnel and Human Resources Management*, vol. 8, 1990, pp. 349-411.
- [27] Stone, Eugene F., Hal G. Gueutal, Donald G. Gardner and Shepherd McClure “A Field Experiment Comparing Information-Privacy Values, Beliefs, and Attitudes Across Several Types of Organizations,” *Journal of Applied Psychology*, vol. 68, no. 3, 1983, pp. 459-468.
- [28] Vroom, V. H. *Work and Motivation*. NY: Wiley, 1964.
- [29] Westin, Alan. Testimony before U.S. House of Representatives, Committee on Energy and Commerce, Subcommittee on Commerce, Trade, and Consumer Protection, Hearing on “Opinion Surveys: What Consumers Have To Say About Information Privacy,” May 8, 2001.
- [30] Woodman, Richard W., Daniel C. Ganster, Jerome Adams, Michael K. McCuddy, Paul D. Tolchinsky and Howard From-

Table 1: Descriptive Statistics

	U.S.	Singapore
Number of subjects	84	184
Percentage of females	42%	44%
Average age	24	23.1
Average Internet experience (years)	6.8	5.9
Percentage of subjects having online purchase experience	95%	61%
Subjects' country of origin (number of subjects)	U.S. (48), India (13), 10 other countries (each less than 5)	Singapore (145), Malaysia (12), 9 other countries (each less than 5)

Table 2. Coefficients and Relative Importance

Instruments	Level	U.S.		Singapore	
		Coefficient ⁺	Relative Importance	Coefficient ⁺	Relative Importance
Monetary Reward	\$5 [#]	0	26.24%	0	11.69%
	\$10 [#]	1.327 ^{***} (0.341)		0.232 (0.165)	
	\$20 [#]	3.141 ^{***} (0.534)		1.388 ^{***} (0.281)	
Visit Frequency/Time Savings	Monthly	0	6.13%	0	6.02%
	Weekly	0.568 ^{**} (0.260)		0.432 ^{***} (0.153)	
	Daily	0.734 [*] (0.411)		0.715 ^{***} (0.254)	
Error	No Review	0	24.80%	0	15.06%
	Review	2.968 ^{***} (0.355)		1.787 ^{***} (0.194)	
Improper Access	No restriction	0	25.12%	0	28.43%
	Restriction	3.007 ^{***} (0.529)		3.374 ^{***} (0.349)	
Unauthorized Secondary Use	Allowed	0	17.70%	0	38.80%
	Not allowed	2.118 ^{***} (0.324)		4.605 ^{***} (0.297)	

⁺ Standard errors are presented in parentheses. The control stimulus consisted of the lowest levels of each of the included dimensions. Because the control was represented by a least squares intercept, we label all lowest level coefficients as zero. The mean intercept is not reported for brevity. [#] US dollars for U.S. subjects and Singapore dollars for Singapore subjects.

^{***} significant at 1% level; ^{**} significant at 5% level; ^{*} significant at 10% level.

Table 3: Clusters

Segment (no. of observations)		Average coefficient				
		Monetary reward	Visit Freq./ Time Savings	Error	Unauthorized Secondary Use	Improper Access
U.S. (78) ⁺	Privacy guardians (56)	1.637 ^{***} (0.385)	0.027 (0.316)	4.040 ^{***} (0.434)	2.576 ^{***} (0.448)	5.116 ^{***} (0.519)
	Information sellers (16)	10.865 ^{***} (0.330)	-0.781 (0.753)	0.245 (0.458)	1.255 ^{**} (0.483)	-0.099 (0.462)
	Convenience seekers (6)	1.445 (0.781)	11.028 ^{***} (0.613)	1.500 ^{**} (0.348)	0.750 [*] (0.371)	0.542 (0.945)
Number of outliers/unclassifiable observations: 6						
Singapore (165) ⁺	Privacy guardians (138)	0.464 ^{**} (0.195)	0.089 (0.166)	2.234 ^{***} (0.183)	5.734 ^{***} (0.318)	4.973 ^{***} (0.314)
	Information sellers (14)	11.286 ^{***} (0.360)	-0.714 (0.855)	0.107 (0.263)	1.768 ^{***} (0.434)	0.446 (0.470)
	Convenience seekers (13)	1.127 (0.862)	10.512 ^{***} (0.682)	0.404 (0.372)	1.077 ^{**} (0.484)	0.173 (0.382)

⁺ Numbers excluding outliers. ^{***} significant at 1% level; ^{**} significant at 5% level; ^{*} significant at 10% level. Standard are errors in parentheses.