Would You Sell Your Mother's Data? Personal Data Disclosure in a Simulated Credit Card Application

Miguel Malheiros, Sacha Brostoff, Charlene Jennett, M. Angela Sasse
University College London
Gower Street
London WC1E 6BT, UK
{m.malheiros, s.brostoff, c.jennett, a.sasse}@cs.ucl.ac.uk

ABSTRACT

To assess the risk of a loan applicant defaulting, lenders feed applicants' data into credit scoring algorithms. They are always looking to improve the effectiveness of their predictions, which means improving the algorithms and/or collecting different data. Research on financial behavior found that elements of a person's family history and social ties can be good predictors of financial responsibility and control. Our study investigated how loan applicants applying for a credit card would respond to questions such as *"Did any of your loved ones die while you were growing up?"* 48 participants were asked to complete a new type of credit card application form containing such requests as part of a "Consumer Acceptance Test" of a credit card with lower interest rates, but only available to "financially responsible customers." This was a double-blind study – the experimenters processing participants were told exactly the same. We found that: (1) more sensitive items are disclosed less often - e.g.: friends' names and contact had only a 69% answer rate; (2) privacy fundamentalists are 5.6 times less likely to disclose data; and (3) providing a justification for a question has no effect on its answer rate. Discrepancies between acceptability and disclosure were observed – e.g.: 43% provided names and contact of friends, having said they found the question unacceptable. We conclude that collecting data items not traditionally seen as relevant could be made acceptable if lenders can credibly establish relevance, and assure applicants they will be assessed fairly. More research needs to be done on how to best communicate these qualities.

INTRODUCTION

To lend money responsibly, as well as protect their own business, lenders assess the risk of applicants not repaying their loans. For the assessment process, lenders collect personal data items directly from applicants, and from organizations such as credit reference agencies, and feed the data collected into their credit scoring algorithms. The lenders will reject loan requests from applicants who fall above a certain risk threshold. The goal is to ensure that the lending business remains profitable, but it also prevents applicants who would not be able to afford the loan from getting into financial hardship.

Lenders are continuously looking to improve the accuracy of their risk assessments, either by improving the algorithms used, or by collecting new types of data. Based on the literature on credit scoring and interviews with experts in personal finance and credit risk, we identified factors that seem to be associated with financial behavior, but are not widely used (and if they are used, the general public is not aware of it). These include a person's relationship with parents while growing up (Hunt & Fry, 2009; Pine & Gnessen, 2009), social links (Glaeser et al., 1999), bill payment history (Belsky & Calder, 2004, Microbilt, 2011) among others. Such data is clearly sensitive, but using it in this way is no different from how health data is used by insurance companies, and psychometric and drug tests data by some companies to assess job applicants. But such data could also be beneficial for some loan applicants: new types of data with predictive value could help those with 'thin' credit histories, who currently find themselves excluded from many financial services because they cannot prove their creditworthiness.

We first review the literature on credit scoring, and present results from interviews with experts in personal finance and credit risk; we then discuss factors known to influence privacy perceptions of individuals. We then present a study in which participants were asked to complete a credit card application in which they had to disclose data commonly requested in this process, and some alternative data. The results from the experiment and the post-experiment questionnaire show that providing justifications for questions has no effect on disclosure rates. Surprisingly, participants did disclose some data they rated as *"unacceptable for lenders to request",* but were less likely to disclose such information about people other than themselves. We conclude that lenders should avoid collecting indices of social capital for the time being, and should keep in mind the potential mismatch between the perceived relevance of a data request and its actual relevance in an empirically based credit scoring algorithm.

BACKGROUND

Credit Scoring

Credit can be a force for good: it can be an investment - for example, buying a car might enable someone to obtain job which they otherwise might not be able to get to, or it can help to manage unexpected expenses, such as emergency repairs. However, individuals obtaining loans they cannot repay has serious consequences on their lives, as well as the lenders' balance sheets: in the UK, for instance, 331 people

are declared insolvent or bankrupt every day (Credit Action, 2011).

To minimize the number of loan defaults and maximize profit (not giving a loan to applicants who could repay it equals lost profit), lenders assess the likelihood that an applicant will repay a loan. This process is known as credit scoring, first used in the 1940s, when it relied on human judgement: credit analysts read an application form and making a decision based on the 5 C's (Thomas, 2000): "the *character* of the person (do you know the person or their family?); the *capital* (how much is being asked for?); the *collateral* (what is the applicant willing to put up from their own resources?); the *capacity* (what is their repaying ability. How much free income do they have?); the *condition* (what are the conditions in the market?)".

Today, credit scoring is based on automatic statistical algorithms which are fed data from the applicant's application form, data related to past dealings with the lender, and their credit report – obtained from a credit bureau (see, for example: RBS, 2011). The risk of an applicant defaulting is inferred from the performance of borrowers whose data profile is similar (Collard and Kempton, 2005; Jentzsch, 2010). Credit scoring algorithms are faster, more consistent and less prejudiced than human decision makers, and there is evidence that these algorithms are better predictors of which applicants would be "good" or "bad" customers (Thomas, 2000).

But credit scoring algorithms are not perfect. Mistakes occur in the classification of some applicants (good risks classified as bad risks and vice-versa), because of limitations in the building of the algorithms themselves (data used to develop predictive models sometimes has poor quality and is based only on samples of accepted borrowers; Hand, 2001), interactions between variables that become outdated (people's behavior changes over time), and because some factors that are the cause of bankruptcy are difficult to predict – e.g. divorce, health problems or unemployment. (Expert 1, 2010; Jentzsch, 2007). To improve the accuracy of their credit scoring, lenders can improve the way their algorithms are built - by adjusting how variables are transformed - or by collecting more data. The latter is seen as a more promising approach because the statistical methods underlying credit scoring are well understood, and no imminent breakthroughs in improving their performance is expected (Expert 1, 2010).

Alternative Indicators

Based in our review of literature on personal credit (Brostoff et al., 2011), and interviews with experts on credit risk and financial behavior, we identified several types of data that are potential indicators of financial behavior, but which are not currently requested in loan applications. These types of data include: bill payments (other than utilities), tax payments, employer recommendations, health condition, stability in life, and social relationships.

Utility payments, for example, are considered to be a measure of willingness to pay debts. There have been initiatives in the US for applicants with no traditional credit history to use their history of utility payments as a measure of their willingness to pay, and these data have been incorporated into credit report products offered by mainstream credit reference agencies – for example the "PRBC credit report with FICO expansion score" from Fair Isaac. Some utility payments are now part of the UK credit bureau data, but it is not clear whether applicants realize this, or how they would perceive an explicit request for this data. Data such as TV license payments are not yet collected, and it is not clear to what extent applicants consider them to be utilities (as opposed to less socially acceptable categories of expenditure), and how this personal classification might be reflected in perceptions of requests for the data.

The same applies to accommodation-related payments. Rent (Microbilt, 2011) and Council Tax payments indicate that the applicant makes regular payments and demonstrates responsible behavior. A larger number of insurance claims might also indicate that you're a riskier person (resulting in higher insurance premiums). Too many may indicate a propensity for fraud.

Sometimes employers vouch for new employees, so that they can get bank accounts (Expert 2, 2009). A recommendation from the employer could therefore function as signal for creditworthiness.

Health condition may also be linked to ability to repay. Body-mass index (BMI), for example, has been linked to some aspects of self-control (Junger & Kampen, 2010), which can be perceived as being related to ability to pay. Also, some lenders purchase insurance to recover the loan in the case of the borrower's death, and these policies require declarations of health and pre-existing conditions. Health checks can reveal lifestyle choices that correlate with responsibility and self-control, and ability to pay back loans. Moreover, mental illness, disability, and physical illness are large risk factors for borrowers not paying back debts (Expert 3, 2009).

Stability in applicants' lives is a key predictor for creditworthiness. One way to assess stability is by asking whether the applicant lives with a partner or spouse (Expert 4, 2009). Kirchler et al. (2008) suggest that relationship dynamics can have an impact on credit decisions, with mutual social influence of the partners potentially changing their behavior.

Stability and attitudes to money are also corrected with experiences while growing up. Analysis of case studies of over-spenders found that these often have a family background where money was used as a method of control, where the relationships with fathers were problematic and distant and mediated by money (Pine & Gnessen, 2009),

and where the patient had experienced major and unresolved loss (Hunt & Fry, 2009).

In a study that examined the performance of listings in a peer-to-peer lending service (*Prosper*), the structural component "degree centrality" of the applicant's social network was related to their probability of being granted a loan: applicants who had more friends and were more central in their social networks were more likely to receive loans. Lin et al. (2009) found that the number and type of friends an applicant had was related to how likely they were to receive a loan – with likelihood increasing with the number of friends who were lenders on *Prosper*. Friend lists may therefore be used as a way of estimating social capital – if an applicant has friends who are rich, powerful and trustworthy, then then s/he is seen as trustworthy and less risky to lend to. It is also seen to assist fraud prevention because such connections facilitate tracing of a defaulting borrower who has changed address. Similarly, the names, addresses and phone numbers of people that know you well could be obtained. This is already done by some sub-prime lenders (Jones, 2001). Although Lin et al. (2009) did not study it, it is plausible that some measure of message flow between an applicant and their social network is an indication of the strength of ties between that applicant and their network, and so could be used as an index of social capital, and therefore trustworthiness to receive loans.

Privacy Factors

Even though the data items discussed above could potentially improve the assessment, their use by lenders raises the number of questions. The key one – which we address in this study - is whether requesting them would raise privacy concerns. Past research has identified 3 criteria that are like to impact applicants' privacy perceptions: sensitivity, transparency, and privacy values.

Sensitivity

Adams and Sasse (2001) investigated privacy perceptions from a user-centric perspective, and found that users' assessment of privacy risks depends on three main factors: (1) information receiver; (2) information usage; and (3) information sensitivity. The first factor refers to how much the user trusts the person or people who will have access to their data. The second factor addresses the way users think receivers use their data in the present, and are going to use it in the future. When individuals perceive that they have some degree of control over future usage of their personal data, they react in a more positive manner to its collection (Culnan, 1993). The third factor consists of the users' perceptions of the data being disclosed and how others (e.g. the receivers) will interpret it. Believing that data portrays individuals in a fair and accurate manner is an important acceptance factor – from a privacy perspective - of technologies and processes that collect personal data (Culnan, 1993; Malheiros et al., 2011). Metzger (2007) investigated the effect of sensitivity on disclosure and found individuals were more likely to withhold items they found more sensitive.

We believe that the different data sensitivities of the various items requested will have an impact on disclosure rates. Consequently, we propose that:

H1: The proportion of participants disclosing each data item will be correlated with the sensitivity of the data items.

Transparency

Relevance or legitimacy of the data request in the context of the interaction has also been identified as an important privacy factor (Culnan, 1993; Hine & Eve, 1998). Annacker et al. (2001) identify legitimacy of a data request as a significant driver for privacy costs, i.e., the lower the perceived legitimacy of the data request the more privacy individuals felt they were giving away. Drawing from the concept of "contextual integrity" (see Nissenbaum, 2004), O'Hara & Shadbolt (2008) describe examples in which there is a negative reaction to a type of data request in one context, but not another: e.g.: collecting data about one's marital status may be appropriate during a date, but is inappropriate in the context of a job interview. In a previous study, Jennett et al. (2011) suggested that transparency of purpose of data requests, in the context of credit applications, could make individuals feel more comfortable with answering questions. Thus, in the current study we advance the following hypothesis:

H2: Participants will disclose more data when a reason for the data request is given, compared to when no reason is given.

Privacy Values

Individual differences may also contribute to different privacy perceptions of specific data requests: some individuals are more sensitive to privacy issues than others. There have been several attempts to develop ways to measure privacy concern (see Buchanan et al., 2007 for a review). One of the most widely used privacy scales is Westin Privacy segmentation (Harris & Associates Inc. & Westin, 1998), which requires participants to rate three statements on a 4-level scale. Based on their answers participants are assigned to one of three groups: (1) privacy fundamentalists, who have strong feelings about privacy and are very defensive of their personal data; (2) privacy unconcerned, who don't have many concerns about privacy or disclosing personal data; and (3) privacy pragmatists, the majority of people, who are willing to disclose personal data when they see a legitimate use for it and see the benefits of doing so (Taylor, 2003). In our study, we expect participants categorized as privacy fundamentalists to be more protective of their personal data, therefore, our third hypothesis states that:

H3: "Privacy fundamentalists" (according to Westin Privacy segmentation) will disclose less data than "privacy unconcerned" or "privacy pragmatists."

Privacy Attitudes *vs.* Privacy Behavior

Privacy research has identified a discrepancy between stated privacy attitudes and concern and actual disclosure behavior (see Acquisti, 2004; and Berendt & Spiekermann, 2005). Most privacy research has relied on data collection techniques such as questionnaires and interviews to capture privacy perceptions and attitudes. In the past two decades, several surveys have identified privacy as a serious concern for consumers and citizens in general (Federal Trade Commission, 1998; Business Week/Harris Poll, 1998; Pew Internet & American Life Project, 2000; Jupiter Research, 2002); yet there are many documented examples of individuals surrendering their personal data for seemingly small rewards (Eskenzi, 2008; Kourti, 2009). Thus, it is important to observe how people act in situations where they are confronted with real trade-offs involving their personal data, rather than just ask them about hypothetical scenarios. Our study explores the difference between the stated acceptability of some questions, and the actual disclosure behavior of the same participants on those questions.

Actual privacy behavior is guided by cost-benefit considerations. When organizations providing a service request personal data from individuals, these assess the potential economic or social benefits that will result from the exchange, and weigh them against the costs of providing the data (Milne & Gordon, 1993; Phelps et al., 2000). If the benefits are perceived to outweigh the costs, individuals will agree to the exchange; if they do not, they will withhold or falsify data to reduce the privacy costs, while still obtaining the benefits of the exchange (Horne, 2007; Metzger, 2007).

Some studies have investigated disclosure behavior when economic rewards (such as money, future convenience or time savings) are offered in exchange for personal data (Grossklags & Acquisti, 2007; Hann et al. 2002a; 2002b; Hui et al. 2007). Results indicate that there is a point – albeit variable from context to context - at which individuals will trade their data for material benefits. When individuals apply for credit there is also a potential economic reward that can be obtained through the disclosure of personal data. However, to our knowledge, no empirical research has been conducted on privacy perceptions and decision-making in the context of credit application forms. It is not clear whether privacy decision-making when individuals apply for credit follows the same rules as in other contexts. The research described here tries to address this gap in the literature by simulating an application process for a credit card that requires different types of personal data to be disclosed.

In the following section we describe our experimental design.

STUDY DESIGN

Demographics

There were 48 participants in the study. Ages ranged from 19 to 31 years, average age 20 years old (s=1.97). Thirty five (72.9%) participants were female and 13 (27.1%) were male. Thirty six (75%) participants were UCL psychology students; 8 (16.7%) were students in other degrees at UCL; 2 (4.2%) were students at other universities; and 1 (2.1%) was not a student.

Procedure

Participants were told that they would be helping to test *"the acceptability of the application process for a new Super Credit Card that beats all other cards on the market. Because the deal is so good it can only be offered to people who are very reliable at repaying. The bank (we cannot reveal which one because of commercial sensitivity) thinks it has discovered a better way of assessing financial responsibility, but it requires more and also different information than is used in the standard credit reference reports."*

The application process consisted of an online application form with 24 questions. Participants were asked to complete and submit the form. They could submit once they had answered at least 20 out of the 24 questions), and were paid £5 (approx. $8) regardless whether they were able to submit the form or not. Participants were told that no actual credit card would be awarded, but that the person who was found to be most creditworthy would receive a £50 (approx. $80) prize. One factor that could potentially affect the way personal data disclosure decisions are made in the context of credit applications is the large value of the credit service being offered compared to the privacy cost of disclosing sensitive data. Thus, this reward was meant to create a real trade-off between disclosing personal data and obtaining an economic benefit similar to what happens in real life credit applications.

To disincentivize submission of false data, participants were told that *"the card can only be offered to people that are completely honest during the application procedure, if you lie on a single item you are not eligible. […] all application data is being sent to a credit reference agency for validation… [using a] … sophisticated combination of cross-comparisons between data in the application form, the individual's current credit record, and also comparison to the Agency's most advance customer profiling system."* Again, the goal was to simulate as realistically as possible a real application process for obtaining credit.

After filling in the application form, participants answered a short questionnaire on privacy values, and were interviewed about the acceptability if the form's questions and whether they had engagaed in any privacy protection behaviors (such as lying). Participants were told that this questionnaire and interview were not part of the evaluation of the bank's application form, but instead part of the

research group's own investigation into the acceptability of the data requests. They were further reassured that the experimenters would not share the interview data with the bank.

To prevent bias, the study was conducted "double-blind". The experimenters who processed the participants – three psychology students - were told the same story as participants. The experimenters were told that the research group was conducting a consumer acceptance trial of the new application process for the bank, and also wanted to determine if people would be inclined to lie on those forms.

The study design was submitted to the university's ethics approval process, and received approval before the study commenced. After the study participants' had been processed, experimenters and participants were informed (face-to-face and by email respectively) that the bank did not really exist. The £50 reward was given to a participant selected at random out of those who did submit.

Application Form

In a past study we generated 53 hypothetical questions which are thought to have relevance for assessing creditworthiness, but which are not normally collected in loan application processes (Brostoff et al. 2012). These include "internet payment history", "any insurance claims", "list of friends from your social networking sites". For each item, participants (N=285) were asked to rate on a 6-point scale to what extent they were comfortable with giving a lender this data.

After an initial principal components analysis (PCA) with Cattell's scree plot method, we identified five main factors that the 53 questions varied on. These five factors accounted for 57% of the total variance. The varimax rotation provided a far more interpretable solution than the direct oblimin rotation. Therefore the varimax rotation was interpreted. The five factors produced were seen to have common themes in the items they contained and as such were given the tags: (1) *Personal/sensitive*, (2) *Bills*, (3) *Attitudes*, (4) *Social network*, and (5) *Partners and Children*. We selected 14 items for use in the current study (see Table 2 below) that were representative of each of the factors identified in the previous study – but that could also be changed into a question that could be "responded to" by a participant.

The application form in the current study began with 10 *Basic* questionnaire items that are present on existing credit application forms (see **Error! Not a valid bookmark self-reference.** below). These were included to make participants believe that the data was really going to be checked against credit reference agency data, and be used by the bank to identify e. We also assumed that - given how the study was advertised - participants would expect that they had to provide these items – giving a baseline to compare the more sensitive items with.

Table 1 - List of *Basic* Items

	Items
1.	Full name
2.	Gender
3.	Date of birth
4.	Current Home Address
5.	Mobile phone number
6.	Home phone number
7.	Nationality
8.	Employment status
9.	Have you had a credit card before?
10.	What is the name of your bank?

Table 2 - List of *Novel* Items

	Items
1.	Did any of your loved ones die while you were growing up? Please give their relation to you (e.g. mother, brother, friend, etc.)
2.	Do you suffer from any medical conditions? Please list...
3.	Did you live with both your mother and father while you were growing up?
4.	Could you list the names and either phone numbers or email addresses of three of your closest friends?
5.	Do you give us permission to contact your local council to get a copy of your council tax payment history?
6.	Do you give us permission to obtain a copy of your TV licence payment history?
7.	Do you give us permission to obtain a copy of your gas or electricity payment history?
8.	Please provide the name and address (or other contact details) of a previous employer so that we can request a copy of the last recommendation from him / her about you...
9.	What is the job of your partner / spouse? Please describe...
10.	What are the names of 3 people that you are friends with on a social networking site (facebook, twitter) whose profiles you would be happy share with us? Please list...
11.	What are the names of 3 people that you are friends with on a professional networking site (LinkedIn, Orkut) whose profiles you would be happy share with us? Please list...
12.	Will you allow us to measure the typical number and length of messages between you and your friends on social networking sites?
13.	What is the length of the longest relationship you have had with a partner / spouse? (years/ months/ weeks)
14.	May we obtain a copy of your insurance claims (e.g. car, house)?

These were followed by the 14 *Novel* items. Responses were either textual data, or required the participant to tick a box to state that he/she consented to their data being looked up by the bank.

A progress bar was set up so that participants had to give a certain amount of data before they could submit their application to the bank (see Figure 1).

Figure 1 - Application Form Progress Bar

We deliberately set the threshold high (20/24) to increase the likelihood of participants providing some *Novel* data items: even if participants gave all *Basic* data (10 items), they needed to provide 10 of the 14 *Novel* items. If participants tried to submit their answers before the progress bar reached 100%, they received an error message (see Figure 2). If participants chose "not applicable" (N/A), this did not contribute towards the tally, since in a real credit application, an applicant would have to submit alternative data items if s/he was unable to answer a question.

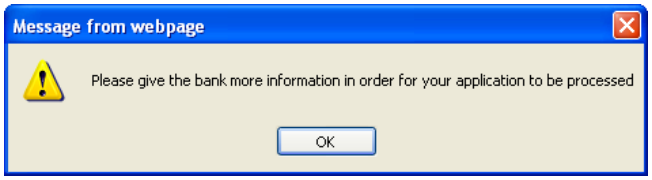

Figure 2 - Insufficient Information Error Message

This was part of the deception: when a participant clicked "submit", their data was not sent anywhere, but deleted instead, i.e. no record of the content of participants' responses was kept. Instead, experimenters' kept notes on which questions participants answered in the form. Experimenters did record audio of the post-scenario interviews for later analysis.

Different Versions of the Application Form

Past research suggests that individuals are more comfortable with disclosing personal data when they understand and agree with the purpose of its collection and how it going to be used (e.g.: Culnan, 1993; Hine & Eve, 1998; Adams & Sasse, 2001). To test this, we set up two versions of the form:

- *Explanation* condition: Participants were given a brief explanation of why each item was needed by the bank (small text that was presented below the item)
- *No Explanation* condition: Participants were not given an explanation of why each item was needed by the bank.

For example, for half of the participants the question *"Did any of your loved ones die while you were growing up? "* was accompanied by the following explanation: *"We need this information to help judge how your early experiences might shape your behavior as an adult – early loss has been related to later financial behavior."*

For each of these conditions, we created a *Normal Order* version and a *Reverse Order* version to control for item order. In both versions the 10 *Basic* items were always presented first. In the normal order the *Novel* items were presented as above in Table 2 – and in the reverse order the Novel items were presented in reverse.

Privacy Values Questionnaire and Follow-Up Interview

As noted in the Background section, there is evidence that some people are more privacy-sensitive than others. Thus, as well as controlling for *age* and *gender*, we also collected level of privacy concern as assessed by the Westin privacy segmentation (Harris and & Westin, 1998). In the Westin scale participants are asked to rate three statements on a 4-point Likert type scale, where 1 = strongly disagree and 4 = strongly agree. The three statements are:

- Consumers have lost all control over how personal information is collected and used by companies
- Most businesses handle the personal information they collect about consumers in a proper and confidential way
- Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today

Privacy fundamentalists are respondents who agreed (strongly or somewhat) with the first statement and disagreed (strongly or somewhat) with the second and third statements. *Privacy unconcerned* are respondents who disagreed with the first statement and agreed with the second and third statements. All other respondents are categorized as *privacy pragmatists*.

A short interview followed, where participants were invited to discuss the acceptability of each of the 24 questions in the application form. If they had decided not to submit the form they were asked about their reasons. They were also asked about whether they had lied on or exaggerated any of their answers.

RESULTS

Submission and Answer Rates

Twenty eight (58.3%) participants submitted the application form, which means they answered at least 20 questions out of the 24 asked. All participants answered at least one question however, even participants that did not submit the form, and the answer rate across all participants for each question is shown below in Table 3.

Of the ten *Basic* information items, 6 were answered by all participants for whom they were applicable (100%), 3 were not answered by one participant each for who they were

applicable (98%), 1 was not answered by two participants for whom it was applicable (92%), giving an average answer rate of 99% of people for the *Basic* items.

The answer rate for *Novel* items was lower, averaging 85% and ranging from 100% to 44% of participants answering data items that applied to them. Only one of the *Novel* data items was answered by all respondents – "*Grew up with both mother and father"*.

Testing the Hypotheses

Hypothesis 1 predicted that the answer rate for each data item request would be correlated with the sensitivity of the item as measured in a previous study (see Methodology section). This hypothesis was supported: the percentage of participants who answered an item (excluding N/A answers) was significantly correlated with the sensitivity of that item (measured on a 5-point comfort scale) $\rho = 0.624$, $p<0.01$.

Hypothesis 2 stated that participants would be more willing to disclose personal data in the version of the form where a justification was given for each question. The data did not support this hypothesis. There was no association between the presence of explanations for the questions and whether participants submitted the form or not: $\chi^2(1) = 0.34$, which is below the critical value of 3.84 ($p=0.05$). There was also no association between the presence of explanations and the number of questions participants answered: t value was not significant ($p=0.05$). Finally, there was no association between the presence of explanations and whether participants had answered a particular question: Pearson's Chi Square or Fisher's Exact Tests were conducted for each item and none were significant ($p=0.05$).

Hypothesis 3 stated that participants categorized as *privacy fundamentalists* according to Westin's privacy scale would be less willing to disclose data. This hypothesis was supported by the data, but only when *privacy unconcerned* and *privacy pragmatists* were blocked. When comparing the behavior of privacy fundamentalists against that of privacy pragmatists and unconcerned separately no statistically significant relationship was found. We believed that we did not have enough participants for the test to have enough power to detect the difference. In order to increase the power of the test, we attempted to sharpen the differences in predicted behavior between the segmentations by contrasting fundamentalists with the other two Westin segmentation groups, using the statistical

Table 3 - Answer Rates

Item	N	Answered	Not Answered	Not Applicable (N/A)	% Answered	% Answered (excluding N/A)
Grew up with both mother and father	48	48	0	0	100.0%	100.0%
Current home address	48	48	0	0	100.0%	100.0%
Employment status	48	48	0	0	100.0%	100.0%
Gender	48	48	0	0	100.0%	100.0%
Mobile phone number	48	48	0	0	100.0%	100.0%
Nationality	48	48	0	0	100.0%	100.0%
Full name	48	48	0	0	100.0%	100.0%
Date of birth	48	47	1	0	97.9%	97.9%
Ever had a credit card	48	47	1	0	97.9%	97.9%
Loved ones passed away while growing up	48	45	3	0	93.8%	93.8%
Name of your bank	48	45	1	2	93.8%	97.8%
Copy of TV licence payment history	48	28	1	19	58.3%	96.6%
Medical conditions	48	45	3	0	93.8%	93.8%
Copy of gas / electricity payment history	48	38	3	7	79.2%	92.7%
Home phone number	48	24	2	22	50.0%	92.3%
Length of longest relationship	48	34	3	11	70.8%	91.9%
Copy of council tax payment history	48	24	3	21	50.0%	88.9%
Previous employer contact details	48	26	4	18	54.2%	86.7%
Social networking profiles of 3 friends	48	37	6	5	77.1%	86.0%
Copy of insurance claims	48	23	4	21	47.9%	85.2%
Job of partner / spouse	48	17	3	28	35.4%	85.0%
Number and length of mobile text messages	48	33	13	2	68.8%	71.7%
Name and phone number / email of 3 friends	48	33	15	0	68.8%	68.8%
Professional networking profiles of 3 friends	48	4	5	39	8.3%	44.4%

technique of blocking. There was a significant association between whether participants were *privacy fundamentalists* and whether they submitted the form $\chi^2(1) = 4.39$, $p < 0.05$. Based on the odds ratio, the odds of a person submitting the form were 5.6 times higher if they were non-*fundamentalists*.

Acceptability of Data Requests

We transcribed the recordings of the post-session interviews and analyzed participants' comments using thematic analysis (Braun & Clarke, 2006). We identified several factors that influence the acceptability of a data request (see Figure 3, the frequencies of participants that mentioned each theme are between parentheses). These factors help clarify why some data requests are considered acceptable while others are not.

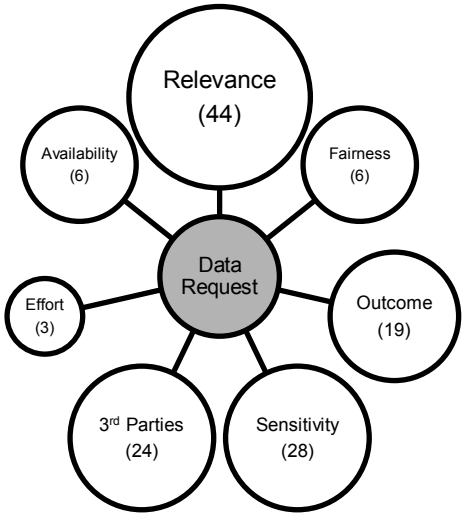

Figure 3 - Acceptability Themes

The acceptability of a data request is related to its perceived **relevance**. A relevant data request is one where the data item is perceived to be related to financial behavior, personality of the applicant, or probability of debt repayment. Relevant data requests were perceived more positively than irrelevant ones:

"I don't think it's acceptable, it's got nothing to do with my credit status" P6

"Yeah it's good, because the bank needs to know how much income you've got" P13

Some participants questioned the **fairness** of using certain items to assess an applicant. Fairness perceptions are associated with relevance perceptions; however, while perceived relevance seems to be related to how acceptable it is to use an item to draw conclusions from a statistical point of view, perceive fairness is related to how acceptable it is to use the item from an ethical perspective.

Perceptions of the consequences of disclosing a data item had an influence on acceptability as well. When participants thought that a data disclosure would result in a positive or neutral **outcome,** they saw it as more acceptable. On the other hand, participants perceived data disclosures they thought could harm them in future as less acceptable

"I did reply, I answered, but only because I don't suffer from a medical condition. Probably if I did I might have reacted differently." P17

"I did disclose it on the answers because again I had nothing to hide, it would all go in my favor" P29

"I know that because I have medical conditions it could be used to discriminate against me" P40

The **sensitivity** of a data request has an influence on how acceptable it is perceived to be. When participants considered a request too personal, sensitive, or invasive, they perceived it as less acceptable.

"I found that very intrusive. I don't think that's acceptable." P48

Requests for data related to **third parties**, such as colleagues or friends of the participants, were perceived as less acceptable.

"[S]haring other people's details is always something I find like quite hard to do" P48

Participants said they feared their friends might be hassled by the bank, that disclosing their data would be a privacy invasion, that it was not their data to give, and that their friends had not consented for their data to be disclosed.

"I wouldn't really want them to impose on my friends' personal space without them giving consent to that." P25

The **effort** required to answer a data request may also impact how it is perceived with request that are involve more work being seen more negatively.

"[I]t would be difficult to get hold of the information, so again I was less inclined to provide it" P30

"Depending on how long the form is, I wouldn't mind doing it" P36

Asking for data that was already **publicly available** from other sources was perceived by some participants as more acceptable. In fact, a couple of participants even disclosed items they thought were unacceptable because they believed the data was already public (see next section).

"Yes I thought this was acceptable, insofar that social networking sites are sort of publicly accessible, and so giving the details of people with whom I have connections on these sort of sites is a reasonable thing to ask" P23

Acceptability and Disclosure

As expected, the acceptability ratings of items correlated significantly with their previously measured sensitivity ratings, $\rho = 0.607$, $p<0.01$. However, the association between participants finding an item acceptable and

disclosing it was only significant for 3 questions: *insurance claims* $\chi^2(2) = 10.44$, $p<0.05$, *council tax* $\chi^2(2) = 10.10$, $p<0.05$, and *emails and phone numbers of friends* $\chi^2(2) = 8.42$, $p<0.05$

For some items there was no association between acceptability and disclosure rate because every participant (or almost every participant) found the item acceptable and disclosed it. There were several items which a large proportion of participants found unacceptable, but still disclosed (see Table 4).

Participants who answered data requests they considered unacceptable were asked why they did. Fourteen participants said that, **on reflection, they did not mind disclosing** the data:

"I did, though I felt I shouldn't... they don't need to know that [...] Although I did answer the question, because then I thought it might not be that bad." P17

Ten participants answered that even though they considered a question **generally unacceptable,** they personally had no problem with answering it.

"Again I did disclose it, but I don't think the general public would be happy […] Because I see myself as quite an open person, so I would be happy." P28

Five participants said they did disclose because they **wanted to complete the form**.

"I did disclose some things mainly just to complete the questionnaire. But it didn't seem a great question." P27

Other reasons for answering unacceptable data requests included:

1. Answering is **not harmful** to me (4 participants);
2. The data is **publicly available** anyway (2 participants);
3. The **bank will not actually look** at the data (2 participants);
4. **Wasn't thinking** about it when I answered (1 participant);
5. I felt safe answering because I was **part of a study** (1 participant).

Privacy Protection Behaviors

All participants were asked during the post-survey interview if they had lied or exaggerated on some items when completing the form : 22.9% of participants said that they had. Examples include saying that the bank could check on their electricity bills when they actually do not pay any, and using friends' initials instead of their own names. One reason mentioned to do this was to increase the amount of data disclosed to the minimum required to be able to submit the form. Another reason given was to protect the privacy of friends.

DISCUSSION

Our study investigated the role of sensitivity, transparency, and privacy values in decision-making about disclosure in the context of a simulated credit card application form. We also wanted to explore the interaction between stated acceptability of a data request, and disclosure behavior regarding the same data request.

Sensitivity, Transparency, and Privacy Values

Hypothesis 1 stated that the number of participants sharing each data item would be inversely correlated with the sensitivity of the data items. In fact, the answer rates for each question showed a significant negative correlation with the sensitivity rating of the question (as measured in a previous study), thus supporting the hypothesis. Past research found that more sensitive items were more likely

Table 4 - Acceptability *vs.* Disclosure

Item	N[1]	Found unacceptable but disclosed	% found unacceptable but disclosed	% found unacceptable but disclosed (excluding N/A)
Loved ones passed away while growing up	46	26	56.5%	56.5%
Social networking profiles of 3 friends	47	25	53.2%	61.0%
Name and phone number / email of 3 friends	47	20	42.6%	42.6%
Number and length of mobile text messages	46	19	41.3%	43.2%
Length of longest relationship	47	18	38.3%	50.0%
Grew up with both mother and father	44	18	40.9%	40.9%
Medical conditions	46	11	23.9%	23.9%
Professional networking profiles of 3 friends	45	3	6.7%	33.3%
Job of partner / spouse	46	3	6.5%	15.8%
Copy of insurance claims	41	2	4.9%	7.1%
Previous employer contact details	46	2	4.3%	6.7%
Copy of TV license payment history	45	2	4.4%	7.1%
Copy of gas / electricity payment history	45	1	2.2%	2.8%
Copy of council tax payment history	46	1	2.2%	3.8%

[1] Participants who, in the interview, did not answer clearly whether they found an item acceptable or not were deleted pairwise

to be withheld (Metzger, 2007). The importance of our finding is that it can be used to estimate *a priori* how an application or registration form will fare, before actually deploying it. Knowing how sensitive certain data items are perceived to be in general makes it possible to predict the likelihood of applicants withholding such items, and weigh the impact of missing data on the lender's business processes to determine whether it is actually worth requesting it.

H3 stated that *privacy fundamentalists* would disclose less data than *privacy unconcerned* or *privacy pragmatists*. As expected, participants who were categorized as privacy fundamentalists on Westin's scale were significantly less likely to submit the form than non-fundamentalists. Privacy fundamentalists are generally more concerned about the risks of their personal data falling into the wrong hands and of the harmful effects that disclosing personal data can have on their lives (cf. Westin, 2003). This would explain their reluctance in submitting their personal data to an unknown party for an uncertain reward, i.e., the reward would have to be larger to offset the perceived cost of answering and submitting the form.

H2 predicted that participants would disclose more data when a reason for the data request was given than when no reason was given. However, even though previous studies identified lack of transparency and legitimacy as promoters of negative reactions (Annacker et al., 2001; Jennett et al., 2011), in our study the presence of explanations for the questions being asked had no significant effect on participant behavior. Thus, this hypothesis was not supported. One possible explanation is that participants did not notice the explanations positioned below each question. Another possibility is that they saw the explanations, but did not read them. Past research on privacy policies found that people rarely read them, or other terms online, because of the time cost, which has been estimated as an average of 10 minutes per policy (McDonald and Cranor, 2008) - so our participants may not have wanted to spend time reading the explanation. If participants read the explanations they may not have understood, or believed them – we did not ask our participants about this. In future studies, user behavior, such as mouse and eye movements, should be tracked to check whether participants are noticing and reading the explanations.

Disclosure and Acceptability of *Novel* Items

Overall, the disclosure rates for the *Novel* items (excluding N/A answers) can be considered high: 85% or more for all but three items. Items related to family history had surprisingly high disclosure rates (100% and 93.8% respectively for *"Grew up with both mother and father"* and "*Loved ones passed away while growing up"*), as did "Medical conditions" and "Length of longest relationship". These are all items generally considered to be very sensitive.

One possible explanation for the high disclosure rates is that no relationship was found between acceptability of a question and its disclosure rate. Even though acceptability and sensitivity ratings were significantly correlated, the acceptability and disclosure rates for individual questions were not with many participants both rating questions as unacceptable and answering them. For example, 56.5% of participants considered the question about "*Loved ones passed away while growing up"* unacceptable, but still answered it. The only exceptions were *"Copy of insurance claims*", "*Copy of council tax payment history"*, and "*Name and phone number/email of 3 friends"*

The thematic analysis provides some insights into why this happens. Several participants said that - even though they found a particular question generally unacceptable - they personally did not mind answering it. This suggests that the assessment of the acceptability of a data request precedes the actual individual cost-benefit evaluation of the disclosure. Participants may believe that it is wrong for a lender to ask for particular data items, but feel that in their personal case it is beneficial (or not costly) to answer. This is further supported by some participants saying they answered "unacceptable" questions so they could submit the application form. They weighed the effort already invested plus the benefit of entering the prize draw against the costs of disclosure, and decided for disclosure. This suggests that when individuals answer surveys about privacy, they may be answering according to the perceived abstract acceptability of certain data practices which may differ from their personal cost-benefit assessment in a real situation. This would help explain why a difference between privacy attitudes and behaviors has been observed in the literature (Acquisti, 2004; Berendt and Spiekermann, 2005).

Items relating to social networks had among the lowest disclosure rates. These items included:

- Number and length of mobile text messages,
- Name and phone number / email of 3 friends,
- Professional networking profiles of 3 friends, and
- Social networking profiles of 3 friends.

All of these can be taken as indexes of participants' social capital. We have already noted that social capital is related to trustworthiness (cf. Lin et al., 2009). However, these data items are about individuals other than the participant and with whom the participant is friends. The thematic analysis revealed that participants were not comfortable revealing data about their friends without their permission. Similarly, "Partner's job" was among the least disclosed items.

Items related to bill payment history, such as utility, TV license, and council tax payment history had high disclosure rates, and a low proportion of participants found them unacceptable. This gives support to the current trend

for lenders to use some types of bill payment history as indicators of creditworthiness, especially when applicants have "thin" credit histories, to make credit scoring more accurate.

Several factors identified in the thematic analysis confirm previous results. In a previous study on applicants' perceptions of loan application forms (Jennett et al., 2011), participants similarly raised issues with: perceived lack of relevance of data requests; level of detail needed to reply to some requests; potential negative outcome of a disclosure; and perceived unfairness of application process. Relevance of a data request, sensitivity of data, and disclosure outcome are all also identified by Culnan (1993) when reviewing factors which impact perceptions of secondary use of information. Culnan (1993) argues that individuals are less likely to perceive that their privacy was invaded when the collected personal data is considered to be relevant for the interaction taking place and will be used to draw reliable conclusions about them. Sensitivity of data is generally considered to be related to privacy perceptions (see Adams and Sasse, 2001 for a privacy model in multimedia communications, or Metzger, 2007 for findings in e-commerce). In a study focused on privacy perceptions in serious games, Malheiros et al. (2011) also identified perceived outcome of sharing data as an important factor. The emergence of factors in our thematic analysis which have been identified in studies focused on different types of contexts suggest that the process through which individuals assess data requests may be context-independent, which does not mean the assessments themselves are.

Privacy Protection Behaviors

23% of participants admitted to falsifying some of the data they submitted as a way to obtain the benefits of submitting the form (and the chance to get a £50 prize) while minimizing the data actually disclosed. Metzger's (2007) study found an almost identical correlation between item sensitivity and disclosure (0.61) as this study (0.62), but a higher proportion of participants that falsified (40% of participants that falsified at least one data item). Metzger's participants were asked about falsification in a self-administered questionnaire, whereas ours were asked face-to-face by the experimenter. Survey work to estimate the prevalence of socially undesirable behavior (for example sexual infidelity in marriage) has found that more people admit to these behaviors to self-administered questionnaires than to experimenters. The difference can be large - six times as many admitted infidelity when asked by a form than by interview (Whisman and Snyder, 2007). We hypothesize that social desirability effects due to the presence of experimenters may have led to under reporting of falsification in our study, and encourage other researchers to address this source of bias more effectively when designing their studies, by employing methods that are more resistant to this bias, for example: self-report questionnaires, or random response techniques (such as participants flipping a coin to answer truthfully or answer yes; Barnett, 1998) that make it impossible to tell if each individual respondent's answer is truthful, but allow an accurate assessment of the true proportion in the sample as a whole.

No data was collected in this study on the rate of falsification per item (we made sure participants' data was not saved to comply with ethics guidelines), but if a relationship could be found between sensitivity of an item and its falsification rate (as in Metzger, 2007), then the data quality impact for lenders of asking for certain items could be bounded.

Limitations

Our participants were college students with an average age of 20. We acknowledge that this limits the generalizability of our results, and plan to repeat the study with a larger, more representative sample. We would, however, argue that the findings of our study have face validity when considered in the context of previous results. Westin's Privacy Segmentation has been repeatedly given across many different samples in different years. A consistent finding is that approximately 25% of respondents fall into the Privacy Fundamentalist category (Kumaraguru and Cranor, 2005). Our participants had a smaller proportion, with 16.7% being Privacy Fundamentalists. This agrees with Tsarenko and Tojib's (2009) finding that young people were more pragmatic in their privacy concerns *viz* financial institutions than other segments of the population. We argue that by being more pragmatic and unconcerned than the general population, the disinclination shown by our participants for disclosing certain data items can be expected in the general population, and that our results would form an upper bound for disclosure of these items in the general population. Also, the previous study on sensitivity ratings, was conducted with a larger (N=285), nationally representative sample, and the sensitivity ratings correlated significantly with this study's disclosure rates.

CONCLUSIONS

From a methodological point of view, this study breaks with common practice by deceiving participants into thinking the data they submitted was actually going to be used to assess financial reliability. A monetary reward for the most creditworthy participant was also offered to nudge participants into submitting their form and answering questions in a truthful manner. Furthermore, experimenters were under the same deception as the participants, to minimize bias. Since privacy decision-making and disclosure behavior are highly contextual it is important to capture and observe them in conditions as realistic as possible.

A goal of this study was to discover which novel data items could potentially be used as alternatives for evidence of credit worthiness for applicants who do not have conventional credit histories, and so could not otherwise participate in and receive the benefits of low cost credit. Among the most sensitive of the novel data items studied in

this research (as measured by sensitivity score and disclosure rates) were those relating to people other than the participant. Although the results need to be validated with a wider socio-demographic (where we estimate individuals to be less pragmatic), we consider this study to be a warning that use of indices of social capital as signs of creditworthiness may currently not be acceptable. The explicit collection of items associated with bill payment history, on the other hand, seem to be less sensitive. Items such as TV license and council tax payment history, which are not currently collected, could be used for credit scoring in situations where applicants have "thin" credit histories.

In the context of applying for credit, we found a direct relationship between an item's sensitivity, and its likelihood of being disclosed, and that this relationship might be employed in a cost/benefit analysis during the design phase for credit application procedures. However, care must be taken in choice of language when assessing sensitivity using survey methods: we found that similar language can tap quite distinct constructs that relate very differently to observed behavior. We found no relationship between items' "acceptability" and their disclosure; many people disclosed information whilst reporting that the antecedent information requests were unacceptable. We hypothesize that there are two separate but related tests employed by credit applicants for assessing information requests – one for testing the requests' general acceptability (that has little impact on disclosure behaviors), and one with respect to the individual's costs and benefits (with much greater impact on disclosure).

A growing body of privacy research is starting to look at privacy decision-making as outcome-oriented: individuals assess the costs and benefits of trading their personal data for some kind of reward. Our research provides some insights into the factors that guide this decision-making process.

The impact of perceived relevance and fairness in particular should be of note to any organization that collects personal data and uses it for profiling purposes. Empirical score-carding (Hand et al., 2008), for example, may find a relationship between a data item and likelihood of default which, while statistically sound, may not be understood by applicants. In fact, these relationships are usually kept hidden from applicants to prevent gaming of the application process, which makes it more difficult for applicants to perceive the relevance of certain data requests. Furthermore, even if the collection of certain types of data is seen as statistically relevant, applicants may still consider the practice unfair or unethical.

We detected no effect of request transparency on disclosure – participants were just as likely to disclose data whether or not an explanation was given for the request. This suggests that, in contexts where there is a low perceived relevance of data requests, organizations should explore new ways to assure individuals that their data collection and data use practices are actually relevant and fair.

We also found that 23% of our participants admitted to falsifying, exaggerating or omitting information when completing our simulated application form. We have no data with which to compare an item's sensitivity to its falsification rate in the context of applying for credit – a topic that requires further studies in which participants' responses are retained and verified through more robust processes.

ACKNOWLEDGEMENTS

We would like to thank Madalina Vasilache, Diana Franculescu and Jessica Colson for testing and interviewing the study's participants. We would also like to thank Conor Fisk for helping to transcribe the interview data.

REFERENCES

Acquisti, A. 2004. Privacy in Electronic Commerce and the Economics of Immediate Gratification. Proceedings of ACM Electronic Commerce Conference (EC '04). New York, NY: ACM Press, 21-29.

Annacker, D., Strobel, M. & Spiekermann, S., 2001. E-Privacy: Evaluating a New Search Cost in Online Environments. *SSRN eLibrary*.

Barnett, J., 1998. Sensitive questions and response effects: an evaluation. *Journal of Managerial Psychology*, 13(1/2), pp.63 – 76

Belsky, E. & Calder, A., 2004. Credit matters: Low-income asset building challenges in a dual finanacial service system. BABC 04-1 Harvard University, Joint Center for Housing Studies. Available at http://www.jchs.harvard.edu/sites/jchs.harvard.edu/files/babc_04-1.pdf

Berendt, B., Günther, O. & Spiekermann, S., 2005. Privacy in e-commerce: stated preferences *vs.* actual behavior. Commun. ACM, 48(4), 101-106.

Buchanan, T., Carina, P., Joinson, A. N. & Reips, U., 2007. Development of measures of online privacy concern and protection for use on the Internet. *Journal of the American Society for Information Science and Technology*, 58(2), pp.157-165.

Braun, V. & Clarke, V., 2006. Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77-101.

Brostoff, S., Jennett, C., Malheiros, M., Sasse, A., 2012. Investigating Loan Applicants' Perceptions of Alternative Data Items. Unpublished Manuscript.

Brostoff, S., Lacoheé, H., Sasse, M.A., 2011. Privacy Value Networks Project: Financial Services Case Study. Unpublished report.

Business Week/Harris Poll, 1998. Online Insecurity, Business Week/Harris Poll. Available at:

http://www.businessweek.com/1998/11/b3569107.htm [Accessed February 24, 2012].

Collard, S. & Kempton, E., 2005. *Affordable credit: The way forward*. Bristol, UK, The Policy Press.

Credit Action, 2011. *UK Debt Statistics from Credit Action*. Available at: http://www.creditaction.org.uk/helpful-resources/debt-statistics.html [Accessed 11 August 2011]

Culnan, M.J., 1993. "How Did They Get My Name?": An Exploratory Investigation of Consumer Attitudes toward Secondary Information Use. *MIS Quarterly*, 17(3), pp.341-363.

Culnan, M.J. & Armstrong, P.K., 1999. Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. Organization Science, 10, pp.104–115.

Eskenzi, 2008. Woman 4 Times More Likely Than Men To Give Passwords For Chocolate. Press release for Infosecurity Europe 2008 Avilable at: http://www.eskenzipr.com/page.cfm/T=m/Action=Press/PressID=202

Expert 1, 2010. Discussion on Consumer Finance Statistics. (Personal Communication with university professor with background in consumer finance statistics research, 27 April, 2010).

Expert 2, 2010. Discussion on financial behavior of immigrant populations in London (Personal Communication with an academic specializing in migration and immigration, 28th April 2010)

Expert 3, 2009. Discussion on Peer-to-Peer Lending. (Personal Communication with executive from a peer-to-peer lending company, 19 August, 2009).

Expert 4, 2009. Discussion on Risk Management in Lending. (Personal Communication with a risk management consultant for a financial services authority, 28 September, 2009).

Federal Trade Commission, 1998. Privacy Online: A Report to Congress , Federal Trade Commission.

Glaeser, E., Laibson, D., Scheinkman, J.A. & Soutter, C.L., 1999. *What Is Social Capital? The Determinants Of Trust And Trustworthiness*. Working Paper 7216: National Bureau Of Economic Research, Cambridge, MA. http://www.nber.org/papers/w7216.pdf?new_window=1

Grossklags, J. & Acquisti, A., 2007. When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *Workshop on Economics of Information Security*.

Hand, D.J., 2001. Modelling consumer credit risk. IMA *Journal of Management Mathematics*, 12(2), p.139.

Hand, D., Brentnall, A. & Crowder, M., 2008. Credit scoring: A future beyond empirical models. *Journal of Financial Transformation, 23, pp.121–128.*

Hann, I., Hui, K., Lee, S. T. & Png, I. P. L., 2002a. Online information privacy: Measuring the cost-benefit trade-off. In *Proceedings of the Twenty-Third International Conference on Information Systems*. Barcelona: L. Applegate, R. D. Galliers, & J. I. DeGross, pp. 1-10.

Hann, I., Hui, K., Lee, S. T. & Png, I. P. L., 2002b. The Value of Online Information Privacy: Evidence from the USA and Singapore. *International Conference on Information Systems*.

Harris and Associates Inc., & Westin, A., 1998. E-commerce and privacy: What net users want. Privacy and American Business and Pricewaterhouse Coopers LLP.

Horne, D.R., Norberg, P.A. & Ekin, A.C., 2007. Exploring consumer lying in information-based exchanges. *Journal of Consumer Marketing*, 24(2), pp.90 - 99.

Hui, K., Teo, H.H. & Lee, S.T., 2007. The Value of Privacy Assurance: An Exploratory Field Experiment. *MIS Quarterly*, 31(1), 19-33.

Hunt, J., & Fry, B., 2009. *Spendsmart*. London: Piatkus Books.

Jennett, C., Malheiros, M., Brostoff, S. & Sasse, M. A., 2011. Privacy for applicants versus lenders' needs for predictive power: Is it possible to bridge the gap? In *Proceedings of the 4th International Conference on Computers, Privacy & Data Protection (CPDP 2011),* 25-27 January 2011, Brussels, Belgium.

Jentzsch, N., 2007. *Financial privacy: an international comparison of credit reporting systems*. Springer.

Jones, P. A., 2001. Access to credit on a low income.

Junger, M. & van Kampen, M., 2010. Cognitive ability and self-control in relation to dietary habits, physical activity and bodyweight in adolescents. *International Journal of Behavioral Nutrition and Physical Activity*, 7(22).

Jupiter Research, 2002. Security and Privacy Data. Available at: http://www.ftc.gov/bcp/workshops/security/020520leathern.pdf

Kirchler, E., Hoelzl, E. & Kamleitner, B., 2008. Spending and credit use in the private household, *Journal of Socio-Economics* 37(2): 519–532.

Kourti, I., 2009. Project FLAME Social Study Report. London School of Economics and Political Science. Available at: http://tnc2009.terena.org/core/getfile2f59.pdf [Accessed February 20, 2012].

Kumaraguru, P. & Cranor, L.F., 2005. Privacy indexes : a survey of Westin's studies. In *[Technical report] /*

Carnegie Mellon University School of Computer Science Institute for Software Research International CMU-ISRI-5-138 Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Pittsburgh, Pa.

Lin, M., Prabhala, N. R., & Viswanathan, S., 2009. Judging Borrowers by the Company They Keep: Social Networks and Adverse Selection in Online Peer-to-Peer Lending. SSRN eLibrary.

McDonald, A.M. & Cranor, L.F., 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society* 4 (3).

Malheiros, M., Jennett, C., Seager, W., & Sasse, M.A., 2011. Trusting to Learn: Trust and Privacy Issues in Serious Games. In J. M. McCune et al., eds. *Trust and Trustworthy Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 116-130.

Metzger, M.J., 2007. Communication Privacy Management in Electronic Commerce. *Journal of Computer-Mediated Communication*, 12(2), p.335–361.

MicroBilt, 2011. PRBC Alternative Credit Data. http://www.microbilt.com/nontraditional-credit-report.aspx. Last viewed 03/02/2012

Pew Internet & American Life Project, 2000. Trust and privacy online: Why Americans want to rewrite the rules, The Pew Internet & American Life Project.

Phelps, J., Nowak, G. & Ferrell, E., 2000. Privacy Concerns and Consumer Willingness to Provide Personal Information.

Pine, P.K.J. & Gnessen, S., 2009. *Sheconomics*, Headline.

Royal Bank of Scotland, 2011. *Your Loan Application*. Available at: http://www.rbs.co.uk/personal/loans/loan-application-info.ashx [Accessed 11 August 2011]

Thomas, L. C., 2000. A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers. *International Journal of Forecasting, 16*, 149-172.

Taylor, H., 2003. Most people are“ privacy pragmatists” who, while concerned about privacy, will sometimes trade it off for other benefits. *The Harris Poll*, 17, p.19.

Tsarenko, Y. & Tojib, D.R., 2009. Examining customer privacy concerns in dealings with financial institutions. *Journal of Consumer Marketing,* 26 (7), 468-476.

Westin, A.F., 2003. Social and political dimension of privacy. *Journal of Social Issues,* 59 (2), 431-453.

Whisman, M.A., Snyder, D.K. (2007) Sexual infidelity in a national survey of American women: Differences in prevalence and correlates as a function of method of assessment. *Journal of Family Psychology*, 21(2), 147-154