

Potoglou, Dimitris; Palacios, Juan; Feijoo, Claudio; Gómez Barroso, Jose-Luis

Conference Paper

The supply of personal information: A study on the determinants of information provision in e-commerce scenarios

26th European Regional Conference of the International Telecommunications Society (ITS):
"What Next for European Telecommunications?", Madrid, Spain, 24th-27th June, 2015

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Potoglou, Dimitris; Palacios, Juan; Feijoo, Claudio; Gómez Barroso, Jose-Luis (2015) : The supply of personal information: A study on the determinants of information provision in e-commerce scenarios, 26th European Regional Conference of the International Telecommunications Society (ITS): "What Next for European Telecommunications?", Madrid, Spain, 24th-27th June, 2015, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/127174>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

The supply of personal information: A study on the determinants of information provision in e-commerce scenarios

Dimitris Potoglou^{1*}, Juan Palacios², Claudio Feijoo², Jose-Luis Gómez Barroso³

¹School of Planning and Geography, Cardiff University, Cardiff, United Kingdom

²Technical University of Madrid, Madrid, Spain

³Universidad Nacional de Educación a Distancia, Madrid, Spain

* Corresponding author: potoglou@cardiff.ac.uk

Suggested Citation:

Potoglou, D., Palacios, J., Feijoo, C. and Gómez Barroso, J.-L. " The supply of personal information: A study on the determinants of information provision in e-commerce scenarios" *26th European Regional Conference of the International Telecommunications Society*, 24th – 27th June 2015, San Lorenzo de El Escorial, Spain

Abstract

Disjoint research efforts have so far considered latent constructs (e.g. privacy concerns) either as an independent variable to explore consumers' actual or stated intentions or - to a less extent - as dependent variable explained through a number of antecedents (e.g. privacy awareness, demographic differences). However, there has not been a formal link across antecedents, latent constructs and behavioural (or stated) intentions in the context of e-commerce or information disclosure. This paper aims to fill this gap in the literature in order to establish a better understanding of the role of attitudes in respondents' willingness to engage in online purchases. We employ a stated preference discrete choice experiment to collect respondents' choices across online retailers, conventional store and opt-out options under different levels of personal-information requirements. Personal information in the experiment is presented across three dimensions (attributes): amount/type of information collected, duration of storage and the likelihood of this information being shared with third parties. These dimensions are introduced in order to be able to capture risks involved in online transactions according to consumer perceptions. Using the Privacy Calculus as a guiding conceptual framework, the experiment also offers respondents trade-offs between benefits such as faster check-out, detailed reviews and priority shipping of the purchased product. The choice data are complemented with a set of attitudinal indicators (psychometric scales) describing individuals' attitudes toward information privacy protection. The data comes from 502 participants representing the online-user population in the UK. We report results from Integrated Latent Variable models, which test the influence of these latent constructs in the consumers' decision to purchase a product online and their sensitivity upon attributes describing online retailers. Preliminary model estimation results show that the higher an individual's concern, general caution and technical protection the less likely a consumer is to purchase a product online. In a joint model, the influence of privacy concern is found to outweigh the influence of general caution and technical protection. Also, consumers with increased general caution are less sensitive in the case an online retailer shares their personal data with third parties.

1. Introduction

The increasing uptake of broadband, a more technologically aware population, and improved digital infrastructure have facilitated an impressive rise of global e-commerce services (OECD, 2012). According to the Centre for Retail Research (2015), a higher-education research institute, online sales in Western Europe are expected to grow by 18.4%¹ in 2015 and reach £185.44bn (€219.44bn) in 2016.

Purchase data and personal information collected from e-consumers are nowadays highly valuable assets for e-commerce stakeholders. The use of data mining tools enables online retailers to uncover purchase patterns, which help them tailor products, identify consumer segments and thus better predict consumer behaviour. Advertisement revenues are also at the heart of e-commerce business models. E-commerce platforms sell personal information about their visitors to advertisers; where the latter are willing to pay a premium for increasing the efficiency of their ads by targeting them to individuals with certain characteristics or behaviours (Godel, Litchfield, & Mantovani, 2012). Companies without access to consumer data generally face significant barriers, for example, to enter in certain industries, and thus are at competitive disadvantage against those firms that hold this asset (Acquisti, 2010).

Normally, information on consumers' online activity such as the number of visits to the e-retailer's site or what products they buy are monitored, collected, traded and retained by online companies. When accessing e-commerce services, e-shoppers benefit from lower prices and personalised consumer experiences, such as purchase recommendations. On the other hand, access to such services requires the provision of personal data by Internet users raising threats in the form of, for example, adverse price discrimination or invasive advertising (Goldfarb & Tucker, 2011).

A large body of work has focused on understanding the benefits and drawbacks (trade-offs) of personal data being disclosed or protected by the two sides of the market. Information privacy is commonly defined as "the desire of individuals to control or have some influence over data about themselves" (Bélanger & Crossler, 2011). In their decision to disclose information to retailers, customers weigh out perceived risks and benefits of disclosing their personal information to retailers, a form of cost-benefit analysis. Following this rationale, experimental studies show that respondents tend to search for online retailers, which require the minimum amount of personal information (Hui, Teo, & Lee, 2007) and those committed to limit future use and access to/use of information by third parties (Hann, Hui, Lee, & Png, 2007). Smith et al. (2011) note the existence of heterogeneity in preferences, which can be determined by observable (e.g. gender, age), and especially by latent psychological characteristics of consumers such as attitudes and perceptions.

In a parallel though disjoint, stream of research, a number of studies have used self-reported scales to examine the role of information-privacy concern as a central construct in shaping consumer behaviour (Smith et al., 2011). Information privacy concerns are measured through psychometric scales (indicators) based on respondents' opinions about privacy-related issues or past behaviour (for a collection of psychometric measures, see Preibusch 2013). In a review of the articles by Smith et al. (2011) and Bélanger & Crossler (2011), Pavlou (2011) notes that there is a 'general consensus among researchers that information privacy concerns correspond to a person's willingness to render personal information, transaction activity and government regulation'. Yet, with the exception of Dinev & Hart (2006), there have been very few studies on how consumers' willingness to disclose their

¹ Prediction based on data from in the UK, Germany, France Sweden, The Netherlands. Italy Poland and Spain.

personal information in e-commerce relates to latent constructs (e.g. attitudes) and how these constructs are influenced through observed individual characteristics (e.g. age).

This paper is aimed at obtaining a better understanding of consumers' decision to engage in an online purchase where the provision of personal information is needed to complete the transaction. Therefore, the paper adds evidence to a growing body of literature that has used different experimental survey and laboratory approaches to study e-shopper behaviour and their willingness to engage and disclose their personal information in online transactions (see, for example, Beresford, Kübler, & Preibusch, 2012; Hann et al., 2007; Hui et al., 2007; Jentzsch et al., 2012; Tsai et al., 2010). More precisely, we explore: (a) how consumers react to different levels of information requirements based on the type/load of information they have to disclose, for how long this information may be retained by retailers and whether their personal information may be shared with third parties; and (b) how privacy related attitudes in the form of, for example, privacy concerns influence e-shoppers choices for online retailers. We argue that a person's willingness to disclose their personal information is dependent upon their preferences against personal data requirements and latent constructs (attitudes) such as privacy concerns, which simultaneously influence these preferences. The objective of this paper is therefore to test the following hypotheses: (a) 'Higher levels of an consumers' privacy concern, caution and technical protection are related to a lower level of willingness to engage in online transactions and thus provide personal information when purchasing products from online retailers' and (b) 'Consumers with different level of Concern, General Caution and Technical protection may exhibit different levels of sensitivity to personal information disclosure in online transactions'.

Our study is based on a Stated Preference Discrete Choice Experiment (SPDCE) – an advanced, survey-based preference elicitation instrument, which simulates an e-commerce scenario of online retailer choice with varying levels of information collection, handling and storage. The data comes from a nationally representative sample of 502 respondents representing the Online User Population in the UK. Using the respondents' preference data for online retailers along with a set of psychometric scales relevant to privacy concerns we develop a model known as Integrated Choice and Latent Variable model (Daly, Hess, Patruni, Potoglou, & Rohr, 2012) that enables the use of information provided by psychometric indicators measuring respondents' attitudes towards privacy concern, general caution and technical protection when using the internet. This modelling framework also provides us with a better understanding of what drives e-shopper choices of retailers. In this study, we use the Integrated Choice and Latent Variable (ICLV) model to explore the link between such choices and self-reported concerns, gauging to what extent information concerns drive individuals to opt-out from their e-commerce transactions.

The remainder of the paper is organised as follows. Section 2 presents a summary of current efforts on privacy concerns and identifies the current gap in the literature. Section 3 provides details on the sample, the experimental design and the psychometric indicators used. Section 4 describes the analytical approach based on ICLV and finally, the paper closes with the discussion of results and conclusions.

2. Privacy concerns, disclosure intentions and the privacy calculus

Empirical studies on 'Privacy Concerns' in the subject areas of Information Systems and related Social and Behavioural Science (e.g. marketing, organisational behaviour) can be consolidated under the 'Antecedents → Privacy Concerns → Outcomes' (APCO) model in which 'measurement of privacy concerns is the central construct', (see, Figure 1 and Smith et al., 2011) for a detailed description of the APCO modelling framework).

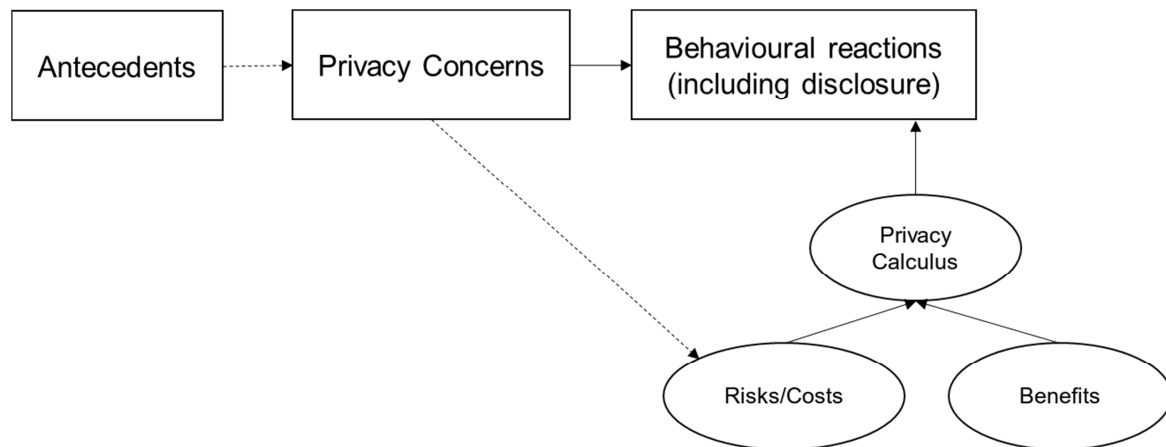


Figure 1. A simplified version of the APCO Macro Model (adapted from Smith et al. (2011))

APCO reflects two streams of research efforts. The first stream is captured on the left hand side in Figure 1 (i.e., Antecedents → Privacy Concerns) in which Privacy Concern is the dependent variable. In this conceptual model, the level of individuals' concerns with regard to their privacy is determined by a number of antecedents. These antecedents may include socioeconomic characteristics – e.g. age (Goldfarb & Tucker, 2012), gender (Bartel Sheehan, 1999); culture/climate (Smith, Milberg, & Burke, 1996); personality differences – e.g. social awareness (Dinev & Hart, 2006); past distressing experiences related to disclosing personal information or privacy awareness (Smith et al., 1996). Privacy concerns are indirectly captured using psychometric scales across different dimensions; for example the concern for information privacy (CFIP; Smith, Milberg, & Burke, 1996) and Internet users' information privacy concerns (IUIPC; Malhotra, Kim, & Agarwal, 2004) scales.

The second stream of research considers privacy concern as determinant of several online-related behavioural reactions or 'stated intentions' (see, right-hand side in Figure 1) – i.e., Privacy Concern is regarded as the independent variable and behavioural reactions are hypothesised to be influenced by the so-called Privacy Calculus. The concept of Privacy Calculus reflects the assumption that individuals' behavioural reactions (and stated intentions) are the result of a calculus between costs and benefits as a consequence of providing personal information (Dinev & Hart, 2006). In the area of e-commerce, several studies explore the role of privacy concerns on consumers' intention to engage in e-commerce transactions, the selection of online retailers (e.g. Tsai et al. 2010) or their willingness to provide information to retailers under certain privacy conditions/scenarios (e.g. Hui et al. 2007; Preibusch et al. 2013).

What is missing from the current literature is an analytical framework that enables antecedents, privacy concerns and stated intentions. It is the aim of this paper to fill this gap in the literature both in substantive and methodological terms. By conducting a stated preference discrete choice experiment and relevant survey, we effectively employ a privacy calculus framework to study consumers' intentions to engage in e-commerce and simultaneously analyse these intentions using an integrated latent variable choice model (ICLV), which as explained later in the paper, enables linkage of stated intentions with latent constructs (i.e., privacy concerns) and its antecedents (e.g. demographic differences) and vice versa. Therefore, the proposed framework enables to capture the effects of privacy concerns in terms of both influencing antecedents and stated intentions.

3. Data and Methods

3.1 Data

The data comes from a UK-representative sample of online users collected through an Internet Panel in August 2012. Sample quotas are specified in order for the sample to match the profile of the 2011 (Q4) Internet-user population in the UK (ONS, 2011). The total sample size includes 502 respondents (see, Table 1). Participants' ages range from 18 to 82 years old (with mean years of age 43). The sample includes 248 male (49.6%) and 269 female (52.0%) respondents. More than half of the participants in the sample work full or part time, 7% of individuals are students, and the rest of participants are unemployed or retired. Most participants report an individual income below £50,000. Most respondents are frequent internet users; 50% of respondents in the sample spend more than 14 hours per week and the majority are regular e-shoppers. Sixty-six percent of the sample search for products at least once a month and 59% report that they buy products online monthly.

Table 1. Sample characteristics vs. the 2011 UK online user population (n=502)

Variable	Sample (%)	Internet users in UK (2011 Q4, %)	Variable	Sample (%)	Internet users in UK (2011 Q4, %)
Gender (female)	52.0	49.6	<i>Region</i>		
<i>Age group</i>			East of England	10.1	7.2
18-24	13.9	17.1	East Midlands	7.2	9.5
25-34	21.5	19.6	London	12.8	13.3
35-44	19.3	19.5	North East	3.7	4.0
45-54	18.4	18.8	North West	11.6	11.0
55-64	15.9	14.0	Northern Ireland	2.3	2.5
65-74	7.9	7.9	Scotland	8.5	8.3
75 and over	3.1	3.2	South East	13.7	14.1
<i>Annual individual income</i>			South West	9.3	8.7
Less than £10,399	27.8	20.9	Wales	4.5	4.7
£10,400 - £15,599	14.1	15.2	West Midlands	8.3	8.3
£15,600 - £20,799	12.6	15.9	Yorkshire / Humberside	8.1	8.4
£20,800 - £25,999	9.3	12.9			
£26,000 - £31,199	6.6	10.4	<i>Occupational status</i>		
£31,200 - £36,399	6.6	7.3	Working full time	41.0	
£36,400 - £41,599	4.1	4.6	Working part time	17.2	
£41,600 - £46,799	2.5	3.8	Student	7.2	
			Retired	16.1	
£46,800 - £51,999	2.7	2.7	Not in paid work because of long term illness or disability	7.0	
£52,000 - £77,999	2.9	4.1	Seeking work	5.8	
£78,000 - £103,999	1.2	1.8	Other	5.8	
£104,000 or higher	0.0	0.3			
Not reported	9.7	20.9			

3.2 Choice of online retailers involving varying levels of information privacy

Respondents' choice for online retailers and subsequently, their decision to engage in online transactions is captured through a stated preference discrete choice experiment, a core component of the online survey. This method is a multi-attribute survey-based approach for eliciting consumer's preferences and has been applied in a wide range of fields such as health economics, transport or marketing (Louviere, Hensher, & Swait, 2000).

The advantage of the SPDCE approach is that it can successfully capture the Privacy Calculus concept (Dinev & Hart, 2006) as it can incorporate both potential risks to privacy and benefits. With regard to risks, Smith et al. (1996) identify four dimensions of individuals' concerns about organizational information privacy practices: collection of personal information, handling errors, unauthorized secondary use and improper access. Therefore,

we include an attribute reflecting the amount of information required in order to complete the transaction. As shown in Table 2, there are four possible versions based on the information collected ranging from a single identifier (email) to a complete range of purchase activity including purchase history, browsing, navigation history, email and additional personal details.

Table 2. Attributes and levels in the choice of online-retailer

Attribute	Levels
<i>Premium/discount against security costs</i>	(1) Discount £4.00 (2) Discount £2.00 (3) No charge (4) Premium £2.00 (5) Premium £4.00
<i>Additional information saved and linked to your account</i>	(1) Only email [Reference Level] (2) Purchase history and email (3) Purchase history, browsing and navigation history and email (4) Purchase history, browsing, navigation history, email and additional personal details
<i>Permission of sharing personal information with third parties</i>	(1) Yes (2) No [Reference Level]
<i>Time your personal information is stored for</i>	(1) 1 year [Reference Level] (2) 2 years (3) 5 years (4) Without an explicit temporal limit
<i>Availability of product or service at a conventional store/outlet</i> (Only available in the Conventional store/outlet alternative)	(1) This item can also be easily purchased in your neighbourhood at a conventional retailer (2) This item can also be purchased from a conventional retailer, but it would require from you to make a special effort because of day/hour of purchase, distance to reach the merchant, etc.) (3) This item is not available to purchase from a conventional retailer in your neighbourhood [Reference Level]
<i>Additional services offered by the product provider at the same price</i>	(1) None [Reference Level] (2) Faster checkout (one-click order) (3) Detailed reviews of products/seller (4) Priority shipping of product

Online retailers also vary in the amount of restrictions they have on the information collected from individuals, namely (1) sharing of costumers' data with other entities and (2) the duration of storage of such information. In the most restricted privacy-protective case, retailers are not allowed to transfer the information collected to any other organisation. Also, they may be obliged to remove customers' information following a specific time period (i.e. 1, 2 or 5 years). These two restrictions describe the perceived degree of risk involved in the online purchase. Retaining personal information for longer of time or sharing it with third parties increases the risk of adverse price discrimination (Acquisti & Varian, 2005; Taylor, 2004) or invasive advertising (Tucker, 2012).

On the other hand, consumers may receive benefits in the form of discounts or additional services for the same product price. Monetary incentives have been previously found to positively affect individuals' willingness to disclose information (Hui et al., 2007); cited in (Bélanger & Crossler, 2011). Also, the additional-services attribute in Table 2 reflects potential benefits offered by online retailers in exchange of the personal information provided by respondents. These attributes are combined such that participants weigh up between restricting the conditions under which retailers can collect and handle their information, and the access to benefits such as discount or additional services (e.g. priority shipping, etc.).

Finally, retailers are also presented with a cost attribute which is defined as either an extra-payment to be paid by participants to acquire the product or discount off the market price of the product.

In the survey, respondents are firstly asked to recall an online purchase they have recently completed.² This strategy of linking the experiment to a real-life experience helps respondents to frame the choice scenario (Hess & Rose, 2009). As shown in Figure 2, respondents are then offered the choice to buy the product from three different online retailers, a Conventional Store and an Opt-out option.

In the previous questions you indicated that you purchased DVDs/Games online most recently. Now thinking about the next purchase of this item please choose from one of the options below.

Description	Online Retailer A	Online Retailer B	Online Retailer C	Conventional store/outlet	
Time your personal information is stored for	1 year	5 years	2 years		
Cost per transaction	Discount £4	£2	£2		
Additional information saved and linked to your account	Only email	Purchase history, browsing, navigation history, email and additional personal details	Purchase history, browsing and navigation history and e-mail	This item can also be purchased from a conventional retailer, but it would require from you to make a special effort because of day/hour of purchase, distance to reach the merchant, etc.	I'll not purchase this item
Permission of sharing this additional information with third parties	No	Yes	Yes		
Additional services offered by the service provider	Detailed reviews of products/seller	Faster checkout (one-click order)	Priority shipping of product at the same price		
Please, indicate which of the option you would choose:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 2. An example choice exercise

In the context of e-commerce, geographic proximity to conventional stores is a determinant aspect to complete e-commerce transactions (Dinev & Hart, 2006). An individual might decide that the need for a highly desired product that is not in close proximity overrides privacy risk and concerns. In our experiment, we explore this dimension by including a conventional-retailer alternative, which is described by one attribute reflecting the ease of access to a physical store (conventional store) with the desired product in stock. Finally, an opt-out option allows participants not to buy the product under any of the conditions offered in the choice exercise. The inclusion of the Opt-out option does not force respondents to choose an undesired alternative and prevents not-participation (Kjaer, 2005).

The combinations of the attribute levels of each attribute are combined to form the Online Retailer and Conventional Store alternatives using a D-efficient experimental design based on the Multinomial Logit (MNL) model with zero priors (Rose & Bliemer, 2009). The final design contains 60 rows (scenarios) which are divided into 12 blocks so that each respondent is presented with a set of five choice tasks. The experimental design matrix is generated using the software Ngene (ChoiceMetrics, 2014). The 60 scenarios are further divided into 12 orthogonal blocks so that each respondent is presented with a set of five scenarios corresponding to one block. Each orthogonal block is randomly assigned to respondents to ensure that any heterogeneity retrieved in the parameter estimates is not affected by the attribute levels presented to respondents (Arentze, Borgers, Timmermans, & DelMistro, 2003). In this study, we use responses from 502 individuals, which result into 2,510 observations.

3.3 Privacy and information protection psychometric scales

As part of the survey questionnaire, respondents are also asked to indicate their strength of preference to a series of statements relating to attitudes toward privacy and information protection. As shown in Table 3, these statements of three sets of attitudinal scales

² The choice experiment is introduced to the subjects by the following statement "In the previous questions you indicated that you purchased <PRODUCT> online most recently. Now thinking about the next purchase of this item please, choose from one of the options below."

previously validated by other studies (Buchanan, Paine, Joinson, & Reips, 2007; Robinson, Potoglou, Kim, Burge, & Warnes, 2010).

Table 3. Psychometric scales for privacy concern, general and technical protection

Latent Construct	Description	Never (1)	Rarely (2)	Sometimes (3)	Very often (4)	Always (5)
General Caution Source: Buchanan et al. 2007	Do you shred/burn your personal documents when you are disposing of them?	5.2%	5.6%	17.0%	21.1%	51.1%
	Do you hide your bank card PIN number when using cash machines/making purchases?	1.6%	3.5%	13.8%	16.83%	65.4%
	Do you only register for websites that have a privacy policy?	4.6%	9.1%	31.1%	27.7%	27.5%
	Do you read a website's privacy policy before you register your information?	11.4%	20.3%	32.5%	18.0%	17.8%
	Do you look for a privacy certification on a website before you register your information?	12.2%	18.2%	27.1%	21.5%	21.1%
	Do you read license agreements fully before you agree to them?	12.4%	23.2%	29.6%	19.3%	15.5%
Technical Protection Source: Buchanan et al. 2007	Do you watch for ways to control what people send you online (such as check boxes that allow you to opt-in or opt-out of certain offers)?	4.3%	5.6%	17.8%	28.6%	43.7%
	Do you remove cookies?	7.5%	13.5%	35.2%	24.4%	19.3%
	Do you use a pop up window blocker?	4.1%	4.6%	18.4%	26.7%	46.2%
	Do you check your computer for spy ware?	4.3%	4.3%	17.4%	27.5%	46.6%
	Do you clear your browser history regularly?	5.8%	11.0%	26.3%	22.0%	34.8%
	Do you block messages/emails from someone you do not want to hear from?	6.0%	9.1%	21.2%	20.7%	41.0%
		Not at all important (1)	Not very important (2)	Neither (3)	Somewhat important (4)	Very important (5)
Privacy Concern Source: Kamaguru and Cranor 2005	Protecting the privacy of my personal information is...	0.6%	1.9%	1.6%	13.3%	82.6%
	Taking action against important security risk (e.g. international terrorism, organised crime) is...	1.0%	2.7%	2.1%	15.9%	78.3%
	Defending current liberties and human rights is...	1.0%	6.2%	2.5%	24.6%	65.8%

Table 4 shows the rotated component matrix arising from the three factor solution. We obtained a similar structure as Buchanan et al. (2007) and Daly et al. (2012), where each of the significant instruments loads in the expected factor. Thus, the results from the explanatory factor analysis provide evidence on the existence of three latent variables capturing respondents': (a) General Caution (alpha = 0.877), (b) Technical Protection (alpha = 0.804) and (c) Privacy Concern (alpha = 0.611). The explanatory factor analysis used is a principal-axis factor analysis with promax rotation (Tabachnick & Fidell, 2007). Eigenvalue scores greater than 0.9 and the percentage of variance explained being smaller than 5% are the criteria used for selecting the number of components to extract from the analysis obtaining a three factor solution.

Table 4. Rotated factor loadings against psychometric indicators

Psychometric Indicator	Factor1	Factor2	Factor3
Do you shred/burn your personal documents when you are disposing of them?	0.244	0.135	0.317
Do you hide your bank card PIN number when using cash machines/making purchases?	0.242	0.193	0.248
Do you only register for websites that have a privacy policy?	0.625		0.118
Do you read a website's privacy policy before you register your information?	0.883		
Do you look for a privacy certification on a website before you register your information?	0.784		
Do you read license agreements fully before you agree to them?	0.812		
Do you watch for ways to control what people send you online (such as check boxes that allow you to opt-in or opt-out of certain offers)?		0.383	0.298
Do you remove cookies?	0.103	0.707	-0.137

Psychometric Indicator	Factor1	Factor2	Factor3
Do you use a pop up window blocker?	-0.213	0.656	0.107
Do you check your computer for spy ware?		0.646	0.107
Do you clear your browser history regularly?	0.103	0.751	-0.185
Do you block messages/emails from someone you do not want to hear from?		0.612	
Protecting the privacy of my personal information is...		-0.123	0.680
Taking action against important security risks (e.g. international terrorism, organised crime) is...			0.589
Defending current liberties and human rights is...			0.602
Technology has almost got out of control		0.143	-0.167
Government can generally be trusted to look after our interests		-0.114	
The way one votes has no effect on what the Government does	0.102	-0.118	
In general business helps us more than it harms us		-0.111	

The first two factors reflect two latent variables which measure individuals attitudes toward privacy protection in online contexts based on privacy-protective measures. These sets of scales were derived and validated by Buchanan et al. (2007). The latent variable *General Caution* reflects general concern with regard to protection of privacy in online scenarios in different ways. Finally, the latent variable *Technical Protection* comprised a series of questions specific to the use of technology for the protection of information privacy (Buchanan et al., 2007). The final set of statements asked respondents to indicate their strength of preference on a five-point Likert scale against an adapted version of the *Privacy Concern Index* (Kumaraguru & Cranor, 2005) consisting of a series of questions about their attitudes towards privacy, security and liberty.

4. Analytical Approach: The Integrated Latent Variable Choice Model

The analytical framework is based on the ICLV model (Bolduc, Ben-Akiva, Walker, & Michaud, 2005; Walker & Ben-Akiva, 2002). This model structure involves the simultaneous estimation of a choice model and attitudinal-structural equation (SEM) model. The ICLV model assumes that both choice and responses to attitudinal indicators are influenced by latent constructs (e.g. privacy concerns) modelling together both outcomes to give more insights into the process that motivate respondents' choices (Daly et al., 2012). ICLV updates the standard choice models based on characteristics of the alternatives and socio-economic characteristics of the respondents by incorporating latent variables describing attitudes and perceptions of individuals. At the same time, these attitudinal variables are used to explain respondents' answers to attitudinal indicators used to quantify the psychological characteristics of individuals. The ICLV model can be then used to explicitly explore the role of the latent constructs in the context of e-commerce transactions – i.e., the willingness of individuals to render their personal information online and while they engage in an online transaction.

In this study, the choice model is employed to explain how respondents make choices between the three online retailers, the conventional-store and the opt-out options given the different attributes describing each option and respondents' latent (unobserved) attitudinal constructs. In the specification of the utilities of the choice model, therefore, we incorporate the three latent variables (latent constructs) identified in Table 4. These factors cannot be directly observed, but can be inferred by respondents' socio-economic characteristics and

their responses to the corresponding psychometric scales for privacy concern, general and technical protection, respectively (see, Latent Variable Model section below).

To incorporate attitudinal concerns in a discrete choice model without introducing the risk of measurement errors and endogeneity bias, the ICLV approach treats the answer to attitudinal instruments as dependent rather than explanatory variables (for a comprehensive overview of the methodological framework, see Daly et al. (2012)). Thus, the ICLV model requires the specification of two components and the associated set of structural equations: (1) a Latent Variable Model (LVM) describing respondents' responses to psychometric indicators, and (2) a Choice Model describing the choices by individuals in the choice tasks forming the DCE (see Figure 3).

On the right-hand side of the graph in Figure 3, there is a set of psychometric indicators capturing respondents' attitudes. These indicators are explained using the three latent variables General Caution, Technical Protection and Concern, which are in turn a function of socio-demographic characteristics of individuals (see, top-left part in Figure 3). These latent variables are then used as predictors, along with the other attributes in the choice model, in order to explain the choices of respondents across the Online retailer, Conventional Store and Opt-out alternatives.

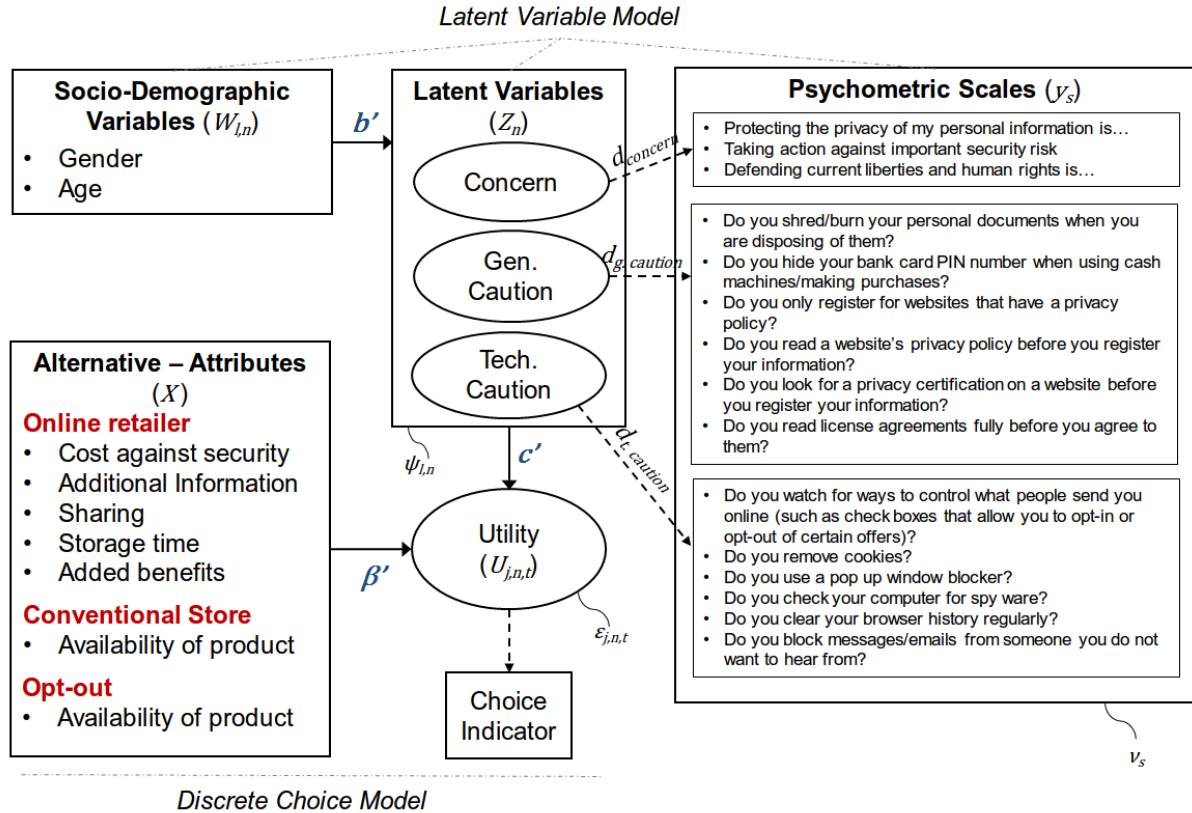


Figure 3. The Integrated Choice and Latent Variable Modelling approach (adapted from: Walker & Ben-Akiva (2002); Bolduc et al. (2005).

Choice Model

Five utility functions are specified in the choice model: one for each of the three Online Retailers, one for the Conventional Store (CS) and one the opt-out (OO) alternative.

The utility of individual n when considering the online-retailer j in the choice task t is described as follows:

$$U_{j,n,t} = \beta_1 Cost_{j,n,t} + \beta'_2 Add.Information_{j,n,t} + \beta_3 Sharing_{j,n,t} + \beta'_4 StorageTime_{j,n,t} + \beta'_5 Add.Services_{j,n,t} + \varepsilon_{j,n,t} \quad \forall 1 \leq j \leq 3 \quad (1)$$

Where *Cost*, *Add.Information*, *Sharing*, *StorageTime*, *Add.Services* correspond to the attributes describing the online retailer j and β s are unknown coefficients to be estimated. Specifically, β_1 is the coefficient representing the influence of variation in the cost attribute (premium or discount) in selecting an online retailer.

All privacy-related attributes (i.e., except the cost attribute) are dummy coded. For each attribute, we set a “base level” (see, Table 2), which represents maximum level of information protection by retailers, thus implying minimum levels of: information requirements, length of storage and restricted access to such information by third parties. The remaining attribute levels describe retailers with higher information requirements in terms relative to the corresponding base level - i.e., collecting additional details of respondents, having the right to share the information with third parties or storing respondents' information for longer periods of time.

β'_2 - β'_5 are vectors of coefficients capturing the effect of the associated attribute level (against its base level) on the likelihood of choosing a specific online retailer. These parameters capture the effect of relaxing the restrictions to retailers on data collection and data handling. For instance, the coefficient β_3 captures the effect of unrestricted access to respondents' personal information on the odds of choosing a retailer, relative to the base level, in which the retailer guarantees that the information will not be shared with third parties.

The random error $\varepsilon_{j,n,t}$ is i.i.d. Type I Extreme Value (Gumbel) distributed and captures the effect of unobservable attributes, unobserved taste variations and measurement or specification errors (Ben-Akiva & Lerman, 1985).

The utilities associated with the conventional store and opt-out alternatives are as follows:

$$U_{CS,n,t} = ASC_{CS} + \beta_6 Spec.Effort_{CS,n,t} + c'Z_n + \varepsilon_{CS,n,t} \quad (2)$$

$$U_{OO,n,t} = ASC_{OO} + c'Z_n + \varepsilon_{OO,n,t} \quad (3)$$

where ASC_{CS} and ASC_{OO} are the alternative specific constants and capture the average effect of unobserved factors in the utility of the conventional retailer and opt-out options, respectively, and relative to online retailers (Train 2003, p.20).

The conventional-store utility function includes the attribute *Spec.Effort*, which describes the availability of the product at a conventional store. This variable takes the value of zero when the product can be easily found at a nearby store and takes the value of one when special effort is required by respondents to purchase it.

Z_n reflects the latent variables considered in this study, namely Privacy Concern, General Caution and Technical Protection and the coefficients c' to be estimated indicate the marginal probability of selecting the Conventional Store alternative or the Opt-out relative to the probability of selecting an online retailer option.

The probability for the sequence of T choices $k = \{k_1, \dots, k_T\}$ made by respondent n , conditional on Z_n is given by:

$$P(k|Z_n, \xi_n) = \prod_T \left(\frac{\exp(\beta' X_{k,n,t} + c'_k Z_n)}{\sum_j \exp(\beta' X_{j,n,t} + c'_j Z_n)} \right) \quad (4)$$

Where $X_{j,n,t}$ corresponds to the vector of attributes describing alternative j , and β' is the vector of unknown coefficients to be estimated.

Latent Variable Model

As mentioned in the previous section, the analysis focuses on three latent variables: (a) General Caution, $Z_{Gen_caution,n}$, (b) Technical Protection, $Z_{Tech_protection,n}$ and (c) Privacy Concern, $Z_{Concern,n}$. These variables are measured by the set of psychometric scales shown in Table 4. The responses to psychometric scales are modelled via 15 equations known as *measurement equations*, which are specified as follows:

$$y_s = \delta_s + d_s Z + v_s \quad (5)$$

where y_s is the observed response to the s th psychometric scale (out of $S=15$). The effect of the latent variables on the value of the scale is given by the vector of parameters d_s which are specific to a given attitudinal indicator. Finally, v_s is the random component of the response to the psychometric indicator.

The three latent variables are assumed to be determined by a series of linear ‘structural’ relationships:

$$Z_{l,n} = b_l W_{l,n} + \sigma_l \psi_{l,n} \quad (6)$$

For $l = \text{General Concern, General Caution, Technical Protection}$

where $b_l W_{l,n}$ represents the deterministic part of $Z_{l,n}$, with $W_{l,n}$ corresponds to a set of socioeconomic variables (e.g. age and gender) and $\sigma_l \psi_{l,n}$ is a normally-distributed error term - with mean zero and standard deviation σ_l .

Following Daly et al. (2012), we take into consideration the ordinal character of the S scales included in our survey by using an ordered logit model specification. Thus, the probability that individual n gives the observed response q in the psychometric indicator s is described by:

$$Pr\{y_{s,n} = q|Z_n\} = \Lambda(\mu_{y_{s,n}} - d_s Z_n) - \Lambda(\mu_{y_{s-1}} - d_s Z_n) \quad (7)$$

The likelihood of the series of answers of respondent n to the psychometric indicators can be described as follows:

$$Pr\{y_n = q|Z_n\} = \prod_s \left(\Lambda(\mu_{y_{s,n}} - d_s Z_n) - \Lambda(\mu_{y_{s-1}} - d_s Z_n) \right) \quad (8)$$

Where $y_{s,n}$ is the observed response to the s th attitudinal indicator by respondent n ; Z_n is a vector of latent variables; and d_s is the vector of parameters describing the impact of each latent variable on the value of the s th indicator, which takes the value of zero when a latent factor does not have any impact on a specific indicator. We follow Ben-Akiva et al. (1999) for the normalisation of the scale for the measurement equations, which consists of normalising one of the parameters associated with each of the latent variables d_s to one.

As part of the ICLV specification shown in Figure 3, the utility functions in the choice model component incorporate a vector of latent variables Z_n as an additional predictor for the choices made by the respondents. Where c' is the vector of coefficients capturing the effect of the latent variables on the utility of selecting alternative j – specifically here the probability of selecting the conventional store and opt-out options. In this study, we normalise the value of c_1, c_2 and c_3 to zero such that c_{CS} and c_{OO} correspond to the effect of variation on the level of privacy concern on the probability of opting-out an e-commerce transaction.

The likelihood function of the ICLV model is the product of the likelihood for the measurement model (Eq. 8) and the likelihood of the series of choices (Eq. 4). For details on the estimation of maximum likelihood estimation of the ICLV model, see Daly et al., (2012). The estimated parameters are estimated by maximising the simulated likelihood function given by the product of the likelihoods of each observation – i.e. the product of the series of five choice tasks and the set of responses to the S attitudinal indicators. As a base model, we also estimated a simple Multinomial Logit (MNL) model without latent variables.

5. Results

This section discusses the results of the estimated models based on the collected stated choice, individual characteristics and psychometric-scale response data.

Tables 6, 7 and 8 present the estimated parameters of two ICLV models split into their corresponding components, namely, the choice, structural equation and the measurement equations models, respectively. The Joint ICLV model includes the three latent variables Concern, General Caution and Technical Protection. The Joint ICLV and Interaction model also shows how latent constructs (i.e., General Caution) may reflect different sensitivities to potential information disclosure risks through the interaction of online retailer attributes and latent constructs. Also, Table 6 presents the estimated parameters of a multinomial logit model (MNL), a model specification which only includes the estimates of the attributes describing online retailers, the conventional store and the opt-out alternatives and does not include any latent variables. The AIC and BIC indicators also show that the Joint ICLV and Interaction models offer a much better fit compared to the Joint ICLV model.

Table 5 presents a summary of each model in terms of: the number of respondents, the number of observations, the number of Halton draws used in the simulations (where applicable), the number of parameters in the discrete choice model (DCM) component and the ICLV overall, and the values of the corresponding likelihood functions. All models reported in the following sections are estimated in Python-Biogeme (Bierlaire, 2003; Bierlaire & Fretschler, 2009) and these estimates are also confirmed by re-estimating the models in Ox (Doornik, 2001). As shown in Table 5, the log-likelihood for the choice component of the model is substantially improved with the inclusion of the latent variables and the corresponding latent-variable model components. The AIC and BIC indicators also show that the Joint ICLV and Interaction models offer a much better fit compared to the Joint ICLV model.

Table 5. Models statistics

	MNL	Joint ICLV	Joint ICLV and Interaction
Number of individuals	502	502	502
Number of observations	2,510	2,510	2,510

Number of Halton draws	-	1,000	1,000
Log-likelihood (DCM)	-3,497.16	-3,177.76	-2826.38
Number of parameters (DCM)	13	16	15
Number of parameters (ICLV)	-	82	54
Log-likelihood (ICLV)	-	-10764.7	-6748.968
AIC (DCM)	7020.3	6387.5	5682.8
BIC(DCM)	7038.5	6409.9	5770.2

Choice Model Component

Initial model specifications separately tested the hypotheses that the higher an individual's Concern, General Caution and Technical Protection the more likely these consumers were to opt-in in an online transaction, respectively. All else being equal, the estimated coefficients of Concern, General Caution and Technical Protection each in the corresponding models, satisfied the above hypotheses (for details, see Palacios, Potoglou, & Feijoo, 2015). The Joint ICLV model tests the above hypothesis when all three latent variables are introduced in the Conventional Store and Opt-out alternatives. In the Joint model, it is shown that the effect of the Concern latent variable remains is strongly significant (t-ratio = 4.43) whereas influence of the General Caution and Technical Protection diminishes. This is an indication that the latent variable Concern outweighs the other two latent variables, General Caution and Technical Protection.

The Joint ICLV and Interaction model shows how latent constructs may interact with privacy-related attributes incorporated in the SPDCE experiment rather than the decision to purchase a product from a conventional store or opt-out. After extensive hypothesis testing, it emerged that only the valuation of sharing of personal data was influenced by attitudes towards General Caution when using the Internet. Therefore, this interaction was incorporated in the simultaneous model structure; no other statistically significant interactions were identified. In particular, if $Sharing_{j,n,t}$ is equal to 1 level applies, which means that the online retailer *shares information with third parties*, in the MNL and Joint ICLV models, the contribution of this attribute to the utility would then be given by $\beta_3 * Sharing_{j,n,t}$ as shown in Eq. 1 while in the specification of the Joint ICLV and Interaction model it will be given by $(\beta_3 + \beta * General_Caution_n)$ where gives the increased General Caution latent variable for respondent n .

Further, the parameter estimates in the MNL, Joint model and Joint ICLV & Interaction models in Table 6 also reveal the effect of each attribute describing the online retailers and the Conventional Store alternatives without the specification of any latent attitude variables in the utilities. The effect of premium/discount (*cost per transaction*) included in the observed part of the utility of the online-retailer is specified with a linear term. As expected, the parameter of the premium/discount has negative sign implying that the higher the premium the less likely a consumer is to choose an online retailer presenting that premium.

The coefficients in the attribute levels of the *additional information saved and linked to your account* attribute were dummy coded. 'Only e-mail' corresponds to the minimum information required to complete the purchase, and is set as the reference level. Again, the negative sign of the coefficients indicates that consumers are more likely to choose retailers that collect the minimum amount of personal information (i.e., only email) to complete the purchase. Therefore, consumers are less likely to select retailers that, besides email address, record and store the 'purchase history' or the 'purchase history and browsing history' of their customers.

Preferences in favour of restricting access to consumer information beyond the online retailer are also reflected in the negative effect of *sharing information with third parties* attribute. At the reference level, retailers compromise to restrict the access of any recorded information

within the retailer and not share that information with any other public or private institution. While this strategy reduces the potential profitable strategies by companies, it increases the odds of being selected by individuals. It is worth noting that the absolute value of the coefficient describing the effect of unbounded access to personal information is the highest among the privacy-related (dummy coded) coefficients.

To complete the results of privacy-related attributes, we present the effect of increasing the timeframe within which retailers store respondents' information. Setting 'one year' as the reference level, the estimated coefficients show that as retailers increase the timeframe during which the collected information is stored, the likelihood of a retailer being selected decreases. Preliminary estimations showed no statistical difference storing personal information between one and two years; hence the reference level in Table 6 is set at 1-2 years. Further, the significantly negative coefficient associated with 'without an explicit temporal limit' indicates that participants tend to avoid retailers who do not compromise to remove consumers' information after a specified period of time. However, the (negative) effect between a 5-year storage limit and a 'without an explicit temporal limit' are valued (statistically) equally negative – i.e., there is no significant difference between the 5-year and 'no explicit temporal limit' coefficients.

Online retailers may persuade individuals to select them by reducing the price of the product, protecting their personal data, and offering additional services. To capture the latter, we included three different additional services that online retailers may offer free of charge. We set the reference level as the situation in which retailers do not offer any additional service. As shown in Table 6, results indicate that – all else being equal, additional services free of charge effectively increase the likelihood of an online retailer being selected by respondents. These services are equally valued as there are no significant differences in the coefficients the free services on offer, namely *faster check out*, *detailed reviews* of the product or the seller and *priority shipping* of the purchased product.

Finally, the utility of the fourth and fifth alternatives are given by two alternative specific constants (ASC). The negative coefficients of the ASC for Conventional Store and Opt-out options show that, all else being equal, respondents are less likely to buy products from a conventional retailer or opt out from the choice context presented to them when compared to the reference levels describing an online retailer. As shown in Table 2, the reference levels describing an online retailer include the maximum level of information protection that is collecting the minimum amount of information, restricting third parties from access to such information and storing customers' information just for one-two years and offer no additional services. The above patterns of attribute-level coefficients remain fairly consistent across the DCM parts of the ICLV models reported in Table 6.

Latent Variable Model Component

As mentioned at the beginning of this section, we explore the impact of the latent constructs (Concern, General Caution, Technical Protection) on respondents' choices of retailers by adding the three latent variables as regressors in the utilities for Conventional Store and Opt-out options in the choice-model component in a joint model and a joint model with interaction. The choice- and latent-variable model components are estimated simultaneously thus achieving consistent and efficient parameter estimates (Daly et al., 2012). The repeated-response (panel) nature of the data is also taken into considerations across all models.

The parameter estimates of the latent variable are better understood in conjunction with the structural and measurement equations in Tables 7 and 8. Table 7 shows the estimation results for the structural equations. The estimated coefficients indicate that respondents' age

and gender are positively associated with each of the latent constructs, except in the last latent construct, Technical Protection, in which it is shown that women are less likely to use available technology to protect the privacy of their information when they are online. Other socio-economic effects tested included respondents' familiarity with the Internet measured as the 'number of hours spent on the Internet' and social class, a composite index of education and occupation status, but these were not statistically significant.

Table 6. Parameter estimates in the choice model

	MNL		Joint ICLV		Joint ICLV & Interaction	
	Est.	t-ratio	Est.	t-ratio	Est.	t-ratio
Online Retailer Alternatives						
<i>Cost per transaction (in £)</i>	-0.128	-14.49	-0.130	-14.58	-0.129	-14.42
<i>Additional information saved and linked to your account</i>						
Only mail			Reference level			
Purchase history and email	-0.165	-2.42	-0.181	-2.59	-0.186	-2.65
Purchase history, browsing, navigation history and email	-0.470	-6.49	-0.478	-6.47	-0.475	-6.40
Purchase history, browsing, navigation history, additional personal details and email	-0.403	-5.58	-0.411	-5.54	-0.416	-5.59
<i>Information is shared with third parties</i>	-0.704	-13.66	-0.723	-13.79	-0.565	-7.75
<i>Time your personal information is stored for</i>						
1-2 years			Reference level			
5 years	-0.311	-4.83	-0.353	-5.36	-0.366	-5.53
Without an explicit temporal limit	-0.340	-5.27	-0.363	-5.53	-0.363	-5.52
<i>Additional services free of charge</i>						
None			Reference level			
Faster checkout (one-click order)	0.369	4.92	0.385	5.06	0.372	4.86
Detailed reviews of products/seller	0.438	5.82	0.435	5.70	0.427	5.56
Priority shipping of product	0.441	5.84	0.469	6.10	0.464	6.02
Conventional Store Alternative						
Alternative Spec. Constant - Conventional Store	-0.481	-4.31	-3.11	-5.92	-3.05	-5.81
This item can also be purchased in your neighbourhood at a conventional retailer			Reference Level			
This item can also be purchased from a conventional retailer, but it would require from you to make a special effort (because of opening hours, distance to reach the merchant, etc.)	-0.637	-4.51	-0.834	-5.09	-0.830	-5.05
Opt-out Alternative						
Alternative Spec. Constant – Opt-Out	-0.869	-9.64	-3.70	-7.08	-3.64	-6.97
Influence of Latent Variables						
Concern x (Conventional Store; Opt-out Alternatives)			2.97	4.43	3.26	4.26
General Caution x (Conventional Store; Opt-out Alternatives)		-	0.095	0.99		
Technical Protection x (Conventional Store; Opt-out Alternatives)			-0.119	-1.31		-
Information is shared with third parties x General Caution					-0.128	-4.00

Table 7. Parameter estimates in the structural equation model component of ICLV

		ICLV Concern		Joint ICLV		Joint ICLV & Interaction	
		Est.	t-ratio	Est.	t-ratio		
Increased Concern	Age	0.008	2.19	0.00852	2.27	0.00749	2.17
	Gender (female)	0.348	2.89	0.339	2.78	0.314	2.75
	Standard deviation	0.864	4.55	0.926	4.82	0.843	4.46
Increased General Caution	Age			0.0268	4.43	0.0287	4.86
	Gender (female)	-		0.362	2.03	0.418	2.26
	Standard deviation			1.87	11.97	1.84	12.06
Increased Technical Protection	Age			0.00972	1.43		
	Gender (female)	-		-0.670	-3.04	-	
	Standard deviation			2.13	11.05		

Table 8 shows that the latent variable Concern is positively associated with the privacy, security and liberty indicators. The highest rating is attributed to the privacy indicator when compared with the security and liberty indicators. Similarly for the other two latent variables, General Caution, Technical Protection, these are positively associated with all their indicators. Individuals with high levels of 'Concern' are those who attach high relevance to the indicators 'security', 'liberty' and especially 'privacy'. High values of the latent variables 'General Caution' and 'Technical Protection' are both associated with those individuals who more regularly implement measures to protect the privacy of their personal information in online scenarios. Therefore, an increased value of any of the abovementioned latent variables implies a positive association with all their corresponding indicators indicating 'strong agreement' or 'always' in the psychometric scales presented in Table 3.

Table 8. Parameter estimates in the measurement model component of ICLV

		ICLV Concern		Joint ICLV		Joint ICLV & Interaction	
		Est.	t-ratio	Est.	t-ratio	Est.	t-ratio
Concern indicators (d parameters)							
Privacy (Indicator 10)		1	-	1	-	1	-
Security (Indicator 11)		0.613	3.21	0.607	3.39	0.620	3.17
Liberty (Indicator 12)		0.543	3.36	0.544	3.56	0.563	3.35
General Caution indicators (d parameters)							
Register only if privacy policy (Indicator 1)				1	-	1	-
Read privacy policy (Indicator 2)				2.650	6.63	2.88	6.10
Privacy certification (Indicator 3)		-		1.450	9.64	1.41	9.75
License agreements (Indicator 4)				1.500	8.95	1.49	9.11
Technical Protection indicators (d parameters)							
Cookies (Indicator 5)				1	-		
Pop-up blocker (Indicator 6)				0.718	7.54		
Spy ware (Indicator 7)		-		1.190	7.22		
Clear browsing history (Indicator 8)				1.240	8.58		
Block messages (Indicator 9)				0.747	7.88		
Constants (δs)							
Privacy		6.110	6.02	6.17	6.07	6.12	6.04
Security		4.950	8.43	4.96	8.45	4.96	8.45
Liberty		4.950	8.46	4.97	8.48	4.96	8.47
Register only if privacy policy				3.10	8.23	3.01	7.89
Read privacy policy				2.80	3.44	2.69	3.00
Privacy certification				1.84	4.08	1.68	3.77
License agreements				1.83	3.95	1.69	3.60
Cookies		-		4.05	9.04	-	
Pop-up blocker				4.22	11.03		
Spy ware				5.34	9.25		
Clear browsing history				5.03	8.58		
Block messages				3.77	10.61		

	ICLV Concern		Joint ICLV		Joint ICLV & Interaction	
	Est.	t-ratio	Est.	t-ratio	Est.	t-ratio
Thresholds ($\mu_{y,s,n}$)						
Concern						
<i>Privacy Thresholds</i>						
Threshold 1	0	-	0	-	0	-
Threshold 2	2.350	2.47	2.350	2.46	2.350	2.46
Threshold 3	2.919	2.99	2.926	2.99	2.916	2.98
Threshold 4	4.819	4.79	4.856	4.81	4.806	4.77
<i>Security Thresholds</i>						
Threshold 1	0	-	0	-	0	-
Threshold 2	1.720	3.27	1.720	3.27	1.720	3.27
Threshold 3	2.240	4.07	2.242	4.07	2.239	4.06
Threshold 4	3.900	6.70	3.902	6.71	3.889	6.69
<i>Liberty Thresholds</i>						
Threshold 1	0	-	0	-	0	-
Threshold 2	2.480	4.47	2.490	4.48	2.490	4.48
Threshold 3	2.826	5.01	2.838	5.02	2.836	5.02
Threshold 4	4.556	7.81	4.578	7.83	4.566	7.82
General Caution						
<i>Register only if privacy policy</i>						
Threshold 1			0	-	0	-
Threshold 2	-		1.650	6.99	1.660	6.97
Threshold 3			4.160	13.70	4.15	13.66
Threshold 4			6.030	17.81	6.020	17.79
<i>Read privacy policy</i>						
Threshold 1			0	-	0	-
Threshold 2	-		3.810	7.68	4.090	7.04
Threshold 3			8.270	11.22	8.850	10.19
Threshold 4			11.360	13.40	12.190	12.21
<i>Privacy certification</i>						
Threshold 1			0	-	0	-
Threshold 2	-		2.140	9.82	2.100	9.91
Threshold 3			4.420	15.04	4.320	15.21
Threshold 4			6.490	18.68	6.350	18.61
<i>License agreements</i>						
Threshold 1			0	-	0	-
Threshold 2	-		2.640	11.05	2.620	11.05
Threshold 3			5.160	16.32	5.130	16.37
Threshold 4			7.400	19.27	7.37	19.31
Technical Protection						
<i>Cookies</i>						
Threshold 1			0	-		
Threshold 2	-		1.780	8.56	-	
Threshold 3			4.340	15.19		
Threshold 4			6.460	18.76		
<i>Pop-up blocker</i>						
Threshold 1			0	-		
Threshold 2	-		1.090	5.05	-	
Threshold 3			2.860	10.22		
Threshold 4			4.440	14.41		
<i>Spy ware</i>						
Threshold 1			0	-		
Threshold 2	-		1.170	4.74	-	
Threshold 3			3.340	9.76		
Threshold 4			5.540	14.01		
<i>Clear browsing history</i>						
Threshold 1			0	-		
Threshold 2	-		2.000	7.46	-	
Threshold 3			4.340	12.61		
Threshold 4			6.210	15.87		
<i>Block messages</i>						
Threshold 1			0	-		
Threshold 2	-		1.350	7.07	-	
Threshold 3			3.070	12.61		
Threshold 4			4.310	16.01		

7. Discussion and Conclusion

This study explores the role of privacy concern, general caution and technical protection, three latent constructs captured through corresponding sets of psychometric privacy-scales, on consumers' willingness to engage in online transactions and their sensitivity towards different dimensions of privacy-related requirements in the online world. From a methodological perspective, this paper demonstrates how consumer choices collected through an SPDCE and subsequent analyses using an ICLV model can be employed to simultaneously capture different effects across the Antecedents Privacy Concern Outcomes framework (Smith et al., 2011). This framework allows to obtain a holistic understanding of the latent constructs that drive consumers' choices for retailers with the latter involving varying requirements for personal information among other possible benefits and risks. By specifying the latent attitudinal constructs in the discrete choice models we do find that the explanatory power of these models substantially improve relative to a conventional discrete choice model without latent attitudinal variables.

Respondents face trade-offs between restricting the conditions under which retailers can collect and handle their information, and the access to benefits such as discount or additional services (e.g. priority shipping, etc.). Overall, respondents are more likely to buy products from retailers that only require consumers' email, store their personal information for one year and do not share that information with third parties. This finding is in agreement with previous research such as Hui et al., (2007) who also find that consumers try to minimise the amount of information that they provide to online retailers. On the other hand, they are also against paying premiums in order to introduce control of their personal data. Our study further shows that the effect of Privacy Concern outweighs the effect of General Caution and Technical Protection, which reflect consumers' increased awareness of privacy threats and adoption of technical measures to mitigate these risks when using the Internet. These results reflect a comparison of the relative strength across privacy concern, general caution and technical protection.

Also, we find that consumers exhibit different sensitivity to the potential sharing of their personal information with third parties through online retailers. Specifically, all else being equal, consumers with increased General Caution are less sensitive to their information being shared with third parties (e.g. smaller negative observed utility).

While individuals' concern with regard to their privacy have been included in several models attempting to explain willingness to provide information in online transactions (Dinev & Hart, 2006) or willingness to accept for providing personal data (Motiwalla, Li, & Liu, 2014); this is the first time (to the knowledge of the authors) that this relationship is explicitly incorporated into a model describing choices for online retailers and how these are simultaneously influenced by respondents'/consumers' attitudes such as privacy concern, general caution and technical protection. The adopted analytical framework simultaneously determines which antecedents influence Privacy Concerns, General Caution and Technical Protection. We primarily focus on demographic differences, namely age and gender and find that women and older consumers are more likely to be privacy concerned, take general caution when it comes to their personal information compared to men and younger consumers. These findings agree with previous studies; for example, Sheehan (1999) report that women are more concerned about the impact of personal information disclosure may have upon their privacy. Also, Culnan (1995) reports that young consumers are less likely to be concerned about their privacy. On the other hand, women are less likely to seek ways to enhance their technical protection when using the Internet than men.

One of the strengths of our paper is that the empirical study was undertaken on consumer preferences from a nationally representative sample of the Online User population in the UK. Therefore, this paper directly addresses previous calls for research on personal information to focus on representative and particularly, non-student samples (Bélanger & Crossler, 2011, pp.1026). Another critique on previous studies is that they are primarily US-based (Bélanger & Crossler, 2011, pp.1019). Our study draws findings from a nationally representative sample of online users in the UK, ensuring the representativeness of our results for the British population of internet users and not specific to the nature of the subjects studied (e.g. university students). The British population is one of the most active in terms of internet adoption with 76% of adults (38 million) accessing the Internet on a daily basis; purchasing goods or services online is of the most popular activities (ONS, 2014). Thus, this paper widens the existing empirical evidence departing from narrow socio-geographic based samples generally found in the literature (Bélanger & Crossler, 2011)³ by gathering nationally representative data.

Finally, we have included a set of psychometric scales in the survey in order to capture a latent variable reflecting an individuals' general trust, but not trust specifically related to the online retailer. This is a limitation of our study as we are unable to develop a latent construct out of existing scales to capture trust to the online retailer and thus we are unable to test hypotheses between willingness to disclose personal information online and retailer trust. The latter has been previously found as having a mediating relationship between privacy concerns and willingness to transact (Dinev & Hart, 2006; Van Slyke, Shim, & Jiang, 2006).

Acknowledgements

The authors are grateful to Dr Patil Sunil (Rand Europe) for his advice on the design of the experiments and Prof. Stephane Hess (ITS Leeds), Prof. Andrew Daly (ITS Leeds and RAND Europe) and Bhanu Patruni (Rand Europe) for their suggestions in model estimations. Any errors are our own.

References

- Acquisti, A. (2010). The economics of personal data and the economics of privacy. Background Paper for OECD Joint WPISP-WPIE Roundtable 1. Retrieved from <http://www.oecd.org/sti/ieconomy/46968784.pdf>
- Acquisti, A., & Varian, H. R. (2005). Conditioning Prices on Purchase History. *Marketing Science*, 24(3), 367–381.
- Arentze, T., Borgers, A., Timmermans, H., & DelMistro, R. (2003). Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transportation Research Part E: Logistics and Transportation Review*, 39 (3), 229–244.
- Bartel Sheehan, K. (1999). An investigation of gender differences in on-line privacy concerns and resultant behaviors. *Journal of Interactive Marketing*.
- Bélanger, F., & Crossler, R. E. (2011). Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems. *MIS Quarterly*, 35(4), 1017–1041.

³ For an existing review of empirical studies in the literature , see among others Bélanger & Crossler (2011); Smith et al. (2011); or Pavlou (2011).

- Ben-Akiva, & Lerman. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.
- Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D. A., Morikawa, T., & Polydoropoulou, A. (1999). *Integration of Choice and Latent Variable Models*. Massachusetts Institute of Technology, Cambridge.
- Beresford, A. R., Kübler, D., & Preibusch, S. (2012). Unwillingness to pay for privacy: A field experiment. *Economics Letters*, 117(1), 25–27.
- Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. In *Proceedings of the 3rd Swiss Transportation Research Conference*. Ascona, Switzerland.
- Bierlaire, M., & Fethiarison, M. (2009). Estimation of discrete choice models : extending BIOGEME Estimation of discrete choice models : extending BIOGEME. *Swiss Transport Research Conference*, (September).
- Bolduc, D., Ben-Akiva, M., Walker, J., & Michaud, A. (2005). Hybrid choice models with logit kernel: applicability to large scale models. In Elsevier (Ed.), *Lee-Gosselin, M., Doherty, S. (eds.) Integrated land-use and transportation models: behavioural foundations*, (pp. 275–302). Oxford.
- Buchanan, T. of measures of online privacy concern and protection for use on the I., Paine, C., Joinson, A. N., & Reips, U. D. (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), 154–165.
- Centre for Retail Research. (2015). Online Retailing: Britain, Europe, US and Canda 2015. Retrieved February 2, 2015, from <http://www.retailresearch.org/onlineretailing.php>
- ChoiceMetrics. (2014). Ngene 1.1.2. User Manual and Reference Guide: The cutting edge in experimental design, ChoiceMetrics, Sydney.
- Culnan, M. J. (1995). Consumer Awareness of Name Removal Procedures: Implication for Direct Marketing. *Journal of Interactive Marketing*, (9), 10–19.
- Daly, A., Hess, S., Patruni, B., Potoglou, D., & Rohr, C. (2012). Using ordered attitudinal indicators in a latent variable choice model: A study of the impact of security on rail travel behaviour. *Transportation*, 39(2), 267–297.
- Dinev, T., & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research* (17:1).
- Doornik, J. A. (2001). Ox: An Object-Oriented Matrix Language. *Timberlake Consultants Press, London*.
- Godel, M., Litchfield, A., & Mantovani, I. (2012). The Value of Personal Information Evidence from Empirical Economic studies. In *Communications & Strategies*, No. 88, 4th Quarter 2012 (pp. 41–60).
- Goldfarb, A., & Tucker, C. (2011). Online Display Advertising: Targeting and Obtrusiveness. *Marketing Science*, 30(3), 389–404.

- Goldfarb, A., & Tucker, C. (2012). Shifts in Privacy Concerns. *American Economic Review*.
- Hann, I.-H., Hui, K.-L., Lee, S.-Y. T., & Png, I. P. (2007). Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach. *Journal of Management Information Systems*, 24, 13–42.
- Hess, S., & Rose, J. M. (2009). Should reference alternatives in pivot design SC surveys be treated differently? *Environmental and Resource Economics*, 42(3), 297–317.
- Hui, K.-L., Teo, H. H., & Lee, S.-Y. T. (2007). The Value of Privacy Assurance: An Exploratory Field Experiment. *MIS Quarterly*, 31(1), 19–33. <http://doi.org/Article>
- Jentzsch, N., Preibusch, S., & Harasser, A. (2012). Study on monetising privacy An economic model for pricing personal information. *Agenda*, 76.
- Kjaer, T. (2005). A review of the discrete choice experiment- with emphasis on its application in health care. *Denmark: Syddansk Universitet*.
- Kumaraguru, P., & Cranor, L. F. (2005). Privacy Indexes: A Survey of Westin's Studies.
- Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: analysis and applications*. *Analysis and Applications* (Vol. 12). Cambridge University Press.
- Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*.
- Motiwalla, L. F., Li, X. (Bob), & Liu, X. (2014). Privacy paradox: does stated privacy concerns translate into the valuation of personal information? In *PACIS 2014 Proceedings*.
- OECD. (2012). *Internet Economy Outlook*. Retrieved from <http://dx.doi.org/10.1787/9789264086463-en>
- ONS. (2011). Internet Access - Households and Individuals. Retrieved February 19, 2015, from <http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcn%3A77-226727>
- ONS. (2014). *Internet Access-Household and Individuals*. Retrieved from <http://www.ons.gov.uk/ons/rel/rdit2/internet-access---households-and-individuals/2014/stb-ia-2014.html>
- Palacios, J., Potoglou, D., & Feijoo, C. (2015). Does privacy concern influence e-shoppers choices for online retailers? An application of the Integrated Latent Variable Choice Model. In *International Choice Modelling Conference 2015, 10-13 May, Austin, Texas, USA*. Retrieved from <http://repositories.lib.utexas.edu/handle/2152/29805>
- Pavlou, P. A. (2011). State of the Information Privacy Literature : Where Are We Know and Where Should We Go? *MIS Quarterly*, 35, 977–988.
- Preibusch, S. (2013). Guide to measuring privacy concern: Review of survey and observational instruments. *International Journal of Human Computer Studies*, 71(12), 1133–1143.

- Preibusch, S., Kübler, D., & Beresford, A. (2013). Price versus privacy: an experiment into the competitive advantage of collecting less personal information. *Electronic Commerce Research*, 13(4), 423–455.
- Robinson, N., Potoglou, D., Kim, C. W., Burge, P., & Warnes, R. (2010). Security at What Cost? Quantifying People's Trade-Offs across Liberty, Privacy, and Security. Retrieved October 28, 2014, from http://www.rand.org/pubs/technical_reports/TR664.html
- Rose, J. M., & Bliemer, M. C. J. (2009). Constructing Efficient Stated Choice Experimental Designs. *Transport Reviews*.
- Sheehan, K. B. (1999). Toward a Typology of Internet Users and Online Privacy Concerns. *Information Society*, (18:1), 21–32.
- Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: an interdisciplinary review. *MIS Quarterly*, 35(4), 989–1015.
- Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information Privacy: Measuring Individuals' Concerns about Organizational Practices. *MIS Quarterly*, 20(2), 167–196.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. (Pearson International, Ed.) *Using multivariate statistics* (5th edit.). Boston.
- Taylor, C. R. (2004). Consumer privacy and the market for customer information. *RAND Journal of Economics*, 35(4), 631–651.
- Train, K. E. (2009). *Discrete choice methods with simulation* (2 edition). Cambridge university press.
- Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2010). The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. *Information Systems Research*, 22(June), 254–268.
- Tucker, C. E. (2012). The economics of advertising and privacy. *International Journal of Industrial Organization*, 30(3), 326–329.
- Van Slyke, C., Shim, J. T., & Jiang, J. J. (2006). Concern for Information Privacy, Risk Perception and Online Consumer Purchasing. *Journal of the Association for Information Systems*, (7:6), pp. 415–444.
- Walker, J., & Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical Social Sciences*, 43(3), 303–343.